

**Tipología textual y cohesión
en la traducción biomédica inglés - español:
un estudio de corpus**

Clara Inés López Rodríguez

Tesis doctoral
Departamento de Traducción e Interpretación
Universidad de Granada, 2000

**Tipología textual y cohesión
en la traducción biomédica inglés - español:
un estudio de corpus**

Tesis doctoral presentada por Clara Inés López Rodríguez
y dirigida por la Dra. Pamela Faber

Granada, julio de 2000

AGRADECIMIENTOS

Mi primer agradecimiento va dirigido a la Dra. Pamela Faber por haber sido una maestra excepcional durante el desarrollo de esta tesis, por sus sabias orientaciones, su dedicación y todo el aliento y la amistad que me ha brindado.

Agradezco la ayuda de la Dra. Chantal Pérez Hernández, que me proporcionó las primeras lecturas sobre lingüística de corpus y al Dr. Christopher Butler, que guió mis comienzos en esta disciplina y me animó a ponerme en contacto con la lista de distribución *CORPORA*. A esta comunidad virtual también estoy agradecida, en especial al profesor Jeremy Ellman de la *University of Sunderland* (Reino Unido), quien me permitió utilizar el generador de cadenas léxicas del que es creador antes de su comercialización.

A mis compañeros del Grupo de Investigación "Lexicografía contrastiva: aplicaciones a la traducción" también debo cualquier acierto o provecho derivado de esta tesis. Ellos han compartido conmigo inquietudes terminológicas y gratos momentos de trabajo.

Doy también las gracias a la Consejería de Educación y Ciencia de la Junta de Andalucía y al Departamento de Traducción e Interpretación de la Universidad de Granada, que han proporcionado recursos materiales para mi formación como investigadora y docente. No menos importantes en mi formación han sido los profesores de dicho departamento y mis alumnos de la Facultad de Traducción e Interpretación de la Universidad de Granada.

Quiero que este sentimiento de gratitud alcance con la misma intensidad a mi familia y a mis amigos por su comprensión, cariño y paciencia sin límites. Ellos también me han inyectado optimismo cuando todo parecía una maraña indomable de ideas. Gracias por haber vivido conmigo las venturas y desventuras que han acompañado esta tesis.

SIGLAS

FL: función léxica

PDQ: *Physician Data Query*, un servicio de información sobre el cáncer del *National Cancer Institute*

QDT: Textos sobre tratamiento del cáncer para profesionales de la salud (*Physician Data Query* del *National Cancer Institute*)

QPT: Textos sobre tratamiento del cáncer para pacientes (*PDQ*)

QDC: Textos para profesionales de la salud sobre cuidados médicos (*PDQ*)

QPC: Textos para pacientes sobre cuidados médicos (*PDQ*)

TO: Texto origen

TST: Teoría Sentido Texto (*Meaning Text Theory*)

TT: Texto término

UF: Unidad fraseológica

USC: Unidad de significación clave

USE: Unidad de significación especializada

CONVENCIONES

H + número: Cadena léxica identificada por *Hesperus*, programa que genera cadenas léxicas

MAYÚSCULA: lema

VERSALITAS: concepto

VERSALITAS EN CURSIVA: cadena léxica

VERSALITAS EN CURSIVA CON SUBRAYADO: marco conceptual

(véase también pp. 303-305)

ÍNDICE

1. Introducción	1
1.1. Presentación	1
1.2. Hipótesis de trabajo y objetivos	3
1.3. Ubicación del trabajo en un marco interdisciplinar	4
1.4. Breve resumen de los capítulos	6
1.4.1. Primera parte: revisión bibliográfica	6
1.4.2. Segunda parte: metodología	7
1.4.3. Tercera parte: resultados	7
1.4.4. Cuarta parte: discusión y conclusiones	8
PRIMERA PARTE: REVISIÓN CRÍTICA DE LOS CONCEPTOS DE TIPOLOGÍA TEXTUAL Y COHESIÓN	
2. Tipologías textuales en los estudios de traducción	9
2.1. La traducción como proceso: el modelo cognitivo de Bell (1991)	10
2.1.1. Memoria, procesamiento de la información y traducción	11
2.1.2. Etapa de análisis	15
2.1.2.1. Reconocimiento visual de palabras y análisis sintáctico	15
2.1.2.2. Análisis semántico	16
2.1.2.3. Análisis pragmático	18
2.1.2.4. Representación semántica independiente de lengua	19
2.1.2.5. Organizador de ideas	20
2.1.2.6. Planificador	20
2.1.3. Etapa de síntesis	21
2.1.3.1. Síntesis pragmática	21
2.1.3.2. Síntesis semántica	22
2.1.3.3. Síntesis sintáctica	22
2.1.3.4. Escritura	22
2.2. Tipologías textuales y traducción	24
2.2.1. Del concepto de <i>equivalencia</i> a la elaboración de taxonomías textuales	24
2.2.2. Las taxonomías textuales y el comienzo de las aproximaciones funcionales	28

2.2.2.1. La aproximación funcional de Katharina Reiss	29
2.2.2.2. El enfoque retórico de Werlich	31
2.2.2.3. Las dimensiones situacionales de Juliane House	32
2.2.2.4. La primacía del texto término: La teoría del <i>skopos</i>	33
2.2.3. La relativización de las dicotomías. Hacia un enfoque interdisciplinar de la traducción	34
2.2.3.1. La aproximación cognitivo-textual de Albrecht Neubert	34
2.2.3.2. La aproximación integradora de Mary Snell-Hornby	40
2.2.3.2.1. Tipos textuales y criterios de traducción	40
2.2.3.2.2. Tipos textuales y cultura	45
2.2.3.2.3. Análisis textual	46
2.2.3.2.4. El papel de los <i>marcos</i> y las <i>escenas</i> (Fillmore 1977)	47
2.2.3.2.5. Los actos de habla	49
2.2.4. El funcionalismo de Christiane Nord: su modificación a la teoría del <i>skopos</i>	50
2.2.5. La tipología de traducciones según Juan C. Sager	55
2.2.5.1. Las estrategias de traducción	59
2.2.5.2. Una tipología de traducciones	60
2.2.6. Paul Kussmaul: convención, tipos textuales y la codificación lingüística de los actos de habla	63
2.2.6.1. Convención cultura y tipo textual	63
2.2.6.2. Macroestructuras	64
2.2.6.3. Microestructuras	65
2.2.7. La aproximación pragmático-semiótica de Ian Mason y Basil Hatim	68
2.2.7.1. Contexto situacional y <i>registro</i>	68
2.2.7.2. Las tres dimensiones del contexto	69
2.2.7.2.1. La dimensión institucional-comunicativa	69
2.2.7.2.2. La dimensión pragmática	70
2.2.7.2.3. La dimensión semiótica	71
2.2.7.3. Tipos textuales y textura	76
2.3. Recapitulación	77
3. La variación textual en la lingüística y la lengua para fines específicos	79
3.1. El texto ¿estructura u organización?	80
3.1.1. El texto como estructura	80
3.1.2. El texto como organización	81

3.1.2.1. El análisis relacional de cláusulas (Winter y Hoey)	81
3.1.2.2. Los estudios sobre <i>género</i>	86
3.1.2.3. Recapitulación	89
3.2. Registro, tipo textual y género	89
3.3. La variación textual en el inglés con fines médicos	91
3.3.1. Géneros en la comunicación entre especialistas	94
3.3.2. Propuesta simplificada de tipos textuales médicos	99
4. Revisión del concepto de <i>cohesión</i>	101
4.1. Coherencia y cohesión	101
4.2. Halliday y Hasan (1976)	102
4.2.1. La conjunción	102
4.2.2. La referencia	102
4.2.3. La sustitución	103
4.2.4. La elipsis	103
4.2.5. La cohesión léxica	103
4.2.5.1. La reiteración	104
4.2.5.2. La colocación	104
4.3. Hasan (1984)	106
4.4. Winter (1974, 1979, 1986)	108
4.5. Metodología de Hoey (1991, 1995)	110
4.5.1. Cohesión y tipos de repetición	111
4.5.1.1. Repetición léxica simple	111
4.5.1.2. Repetición léxica compleja	111
4.5.1.3. Paráfrasis simple	112
4.5.1.4. Paráfrasis compleja	112
4.5.1.5. Repetición superordinada, hiponímica y correferencial	114
4.5.1.6. Otras formas de repetición que no son léxicas	114
4.5.2. Conexiones, vínculos y redes	115
4.5.3. Cómputo y representación gráfica de las conexiones y vínculos	117
4.6. Aplicaciones derivadas de la metodología de Hoey	122
4.6.1. Fragmentación de textos y didáctica de las lenguas	122
4.6.2. Vinculación e intertextualidad (Hoey 1995)	123
4.6.2.1. La percepción de vínculos por parte del lector	124
4.6.2.2. Creación de vínculos y procesamiento del lenguaje	125

4.6.2.3. Intertextualidad	129
4.7. Morris (1988), Morris y Hirst (1991)	131
4.8. Aplicaciones informáticas para el reconocimiento automático de la cohesión	133
4.8.1. Aproximaciones basadas en la repetición léxica	133
4.8.2. Aproximaciones que incluyen relaciones conceptuales	135
4.9. La cohesión desde la <i>Teoría Sentido Texto</i>	138
4.9.1. Consideraciones generales	138
4.9.2. Aproximaciones a la cohesión	145
4.9.3. Implicaciones	151
4.10. Cohesión y tipo textual	152
4.10.1. Cohesión <i>referencial</i> y <i>género</i>	152
4.10.2. Cohesión <i>léxica</i> y conocimiento científico en textos especializados y divulgativos	153
4.10.2.1. Textos especializados: requisitos epistemológicos y cohesión	155
4.10.2.2. Textos de divulgación: recursos cohesivos y adquisición de conocimiento	158
4.10.2.3. Recapitulación	159
4.10.3. Cohesión y dificultad comunicativa	160
4.11. Aproximaciones a la cohesión desde la traducción	164
4.11.1. Cambios cohesivos en la traducción de textos de divulgación científica	165
5. La lingüística de corpus al servicio de la traducción y la terminología	169
5.1. Introducción	169
5.2. Evolución de la lingüística de corpus	170
5.3. Corpus paralelos y corpus comparables	175
5.4. Aplicación de la lingüística de corpus a la traducción	177
5.4.1. La utilización de corpus en la traducción biomédica	180
SEGUNDA PARTE: METODOLOGÍA	
6. Fundamentos metodológicos para el estudio de la cohesión	185
6.1. Introducción	185
6.2. La cohesión como fenómeno interoracional e intraoracional	185
6.3. La cohesión léxica como principal recurso cohesivo	186
6.4. Papel de la cohesión léxica en la activación de marcos	188
6.5. Cohesión y repetición en el eje paradigmático y sintagmático	190

6.6. Funciones léxicas en la generación de textos cohesivos	197
6.6.1. Propuesta de funciones léxicas cohesivas	197
6.6.1.1. Repetición exacta	197
6.6.1.2. Repetición paradigmática	197
6.6.1.2.1. Repetición paradigmática simple	198
6.6.1.2.2. Repetición paradigmática sintáctica	198
6.6.1.2.3. Repetición paradigmática semántica	199
6.6.1.3. Repetición sintagmática	205
6.6.1.3.1. Funciones léxicas nominales	206
6.6.1.3.2. Funciones léxicas adjetivales / adverbiales	207
6.6.1.3.3. Funciones léxicas verbales	208
6.6.1.4. Repetición mixta de unidades fraseológicas	215
6.6.1.5. Repetición mixta de unidad fraseológica y lexema	216
7. Esquemas de representación para el análisis intensivo de la cohesión	217
7.1. Modificaciones necesarias a las matrices de repetición de Hoey (1991)	217
7.1.1. Matriz de repetición de relaciones cohesivas primarias (Tabla A)	220
7.1.2. Matriz de repetición de relaciones cohesivas primarias (Tabla B)	221
7.2. Esquemas de representación complementarios	224
7.2.1. Datos estadísticos sobre el texto/ corpus	225
7.2.2. Listas de frecuencia	226
7.2.3. Matriz de repetición	228
8. Esquemas de representación para el análisis extensivo de la cohesión	233
8.1. Descripción de los resultados obtenidos con el programa informático <i>Hesperus</i>	233
8.1.1. Perspectiva general	233
8.1.2. Perfil del documento	234
8.1.3. Versión del texto con hipervínculos hacia las cadenas léxicas	236
8.1.4. Representación de una cadena léxica	237
8.2. Limitaciones de los resultados aportados por <i>Hesperus</i>	239
8.3. Etapas para el análisis extensivo de la cohesión	242
8.3.1. Índice del texto	242
8.3.2. Áreas conceptuales reconocidas a partir de listas de frecuencia lematizadas	242
8.3.3. Representación modificada de los resultados de <i>Hesperus</i>	244
8.3.4. Propuesta definitiva de cadenas léxicas	246
8.3.5. Cuadro contrastivo (<i>Hesperus</i> – Análisis combinado)	247

9. Selección del corpus	251
9.1. Criterios de selección del corpus	251
9.2. Limitaciones del corpus utilizado	252
9.3. Composición del corpus	253
9.4. Tratamiento del corpus: eliminación de <i>palabras forma</i> y lematización	257
TERCERA PARTE: RESULTADOS	
10. Análisis intensivo de una noticia sobre oncología en español	259
10.1. Datos estadísticos del texto	260
10.2. Listas de frecuencia	261
10.3. Matrices de repetición de las relaciones cohesivas primarias	264
10.3.1. Cómputo de las conexiones en las que participa cada oración	267
10.3.2. Tipos de repetición	268
10.3.2.1. Repetición mixta	269
10.3.2.2. Repetición exacta	274
10.3.2.3. Otros tipos de repetición	275
10.3.3. Funciones léxicas	276
10.4. Matriz de repetición de relaciones cohesivas secundarias	282
10.5. Hacia un metodología empírica para la representación de marcos	285
11. Análisis extensivo comparativo de textos para especialistas y textos para pacientes	291
11.1. Textos para especialistas sobre tratamiento del cáncer de pulmón	291
11.1.1. Mesothelioma (QDT1)	291
11.1.1.1. ÍNDICE DEL TEXTO	291
11.1.1.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	292
11.1.1.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	293
11.1.1.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	302
11.1.1.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS	310
11.1.1.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	329
11.1.2. Non-small cell lung cancer (QDT2)	334
11.1.2.1. ÍNDICE DEL TEXTO	334
11.1.2.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	334
11.1.2.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	336

11.1.2.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	344
11.1.2.4.1. DESARROLLO DE LAS CADENÁS LÉXICAS DEFINITIVAS	349
11.1.2.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	370
11.1.3. SMALL CELL LUNG CANCER (QDT3)	375
11.1.3.1. ÍNDICE DEL TEXTO	375
11.1.3.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	375
11.1.3.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	376
11.1.3.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	383
11.1.3.4.1. DESARROLLO DE LAS CADENÁS LÉXICAS DEFINITIVAS	387
11.1.3.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	404
11.2. Textos para pacientes sobre tratamiento del cáncer de pulmón	408
11.2.1. Mesothelioma (QPT1)	408
11.2.1.1. ÍNDICE DEL TEXTO	408
11.2.1.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	408
11.2.1.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	410
11.2.1.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	414
11.2.1.4.1. DESARROLLO DE LAS CADENÁS LÉXICAS DEFINITIVAS	417
11.2.1.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	431
11.2.2. Non-small cell lung cancer (QPT2)	434
11.2.2.1. ÍNDICE DEL TEXTO	434
11.2.2.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	434
11.2.2.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	436
11.2.2.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	440
11.2.2.4.1. DESARROLLO DE LAS CADENÁS LÉXICAS DEFINITIVAS	444
11.2.2.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	453
11.2.3. Small cell lung cancer (QPT3)	455
11.2.3.1. ÍNDICE DEL TEXTO	455
11.2.3.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS	455
11.2.3.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE <i>HESPERUS</i>	456
11.2.3.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS	462
11.2.3.4.1. DESARROLLO DE LAS CADENÁS LÉXICAS DEFINITIVAS	465
11.2.3.5. CUADRO CONTRASTIVO (<i>HESPERUS</i> – ANÁLISIS COMBINADO)	474
11.3. Datos estadísticos relacionados con el estudio de la cohesión y sus limitaciones	477
11.3.1. Representación de las cadenas léxicas y tipos de repetición	477
11.3.2. Estadísticas de textos sobre tratamiento del cáncer de pulmón	479
11.3.3. Estadísticas generales para el resto del corpus	481
11.3.4. Recapitulación	485

CUARTA PARTE: DISCUSIÓN Y CONCLUSIONES

12. Cadenas léxicas y activación conceptual	487
12.1. Diferencias conceptuales y terminológicas entre textos para especialistas y textos para pacientes	490
12.1.1. <i>TREATMENT</i>	493
12.1.2. <i>PATHOLOGY</i>	497
12.1.3. <i>RESEARCH</i>	498
12.1.4. <i>OUTCOME OF TREATMENT</i>	499
12.1.5. <i>LOCATION IN THE HUMAN BODY</i>	502
12.1.6. <i>DIAGNOSIS</i>	504
12.1.7. <i>PATIENT</i>	505
12.1.8. <i>STAGING</i>	505
12.1.9. <i>SIZE / QUANTITY</i>	505
12.1.10. Resto de las macrocadenas	507
12.2. Recapitulación	508
13. El papel de la cohesión en la construcción del conocimiento: recursos cohesivos en textos de divulgación	509
13.1. Expresiones <i>punte</i> entre el lenguaje del especialista y el del lector lego	511
13.1.1. Explicación + (término especializado)	511
13.1.2. Término especializado + (explicación / definición)	512
13.1.3. Explicación / lexema de la lengua general (término superordinado) + <i>called</i> + + término especializado	512
13.1.4. Oración de relativo con <i>that / which</i>	515
13.1.5. Término especializado + <i>is a</i> + definición	516
13.1.6. <i>Type / kind / form</i>	516
13.1.7. <i>This / these</i>	517
13.1.8. Otros recursos	518
13.2. Recapitulación	519
14. Frecuencia léxica y cohesión	521
14.1. Textos para especialistas	521
14.1.1. Correlación entre frecuencia léxica y organización textual	523
14.2. Textos para pacientes. Diferencias con respecto a los textos para especialistas	527
14.3. Frecuencia léxica y grado de especialización	530
14.4. El cotexto de los lemas más frecuentes como restrictor semántico	533

14.4.1. Elaboración de una ficha con información cotextual	536
14.4.2. Unidades de significación más frecuentes del subdominio de la biomedicina y la oncología	538
14.4.2.1. PATIENTS	538
14.4.2.2. CANCER	541
14.4.2.3. TREATMENT	542
14.4.2.4. SYMPTOMS	544
14.4.2.5. THERAPY	545
14.4.2.6. CHEMOTHERAPY	546
14.4.3. Unidades de significación de la lengua general <i>terminologizadas</i> en el subdominio de la oncología	548
14.4.3.1. USE, STAGE, COMBINATION-COMBINED	548
14.4.3.2. PROSODIA SEMÁNTICA EN LA EXPRESIÓN DE LA CAUSA Y EL EFECTO Y DE LA APARICIÓN DE CÁNCER/ SÍNTOMAS / EFECTOS SECUNDARIOS	553
14.4.3.2.1. EFFECT, CAUSE, RESPONSE, PRODUCE, FACTOR	553
14.4.3.2.1.1. EFFECT	553
14.4.3.2.1.2. CAUSE	558
14.4.3.2.1.3. RESPONSE	559
14.4.3.2.1.4. PRODUCE	559
14.4.3.2.1.5. FACTOR	561
14.4.3.2.2. APARICIÓN DE SÍNTOMAS / EFECTOS SECUNDARIOS / ENFERMEDAD: DEVELOPMENT, ASSOCIATED, OCCUR, EXPERIENCE, EFFECT	562
14.5. Observaciones en torno a las <i>palabras forma</i> más frecuentes del corpus	567
15. Cambios en la cohesión derivados de la traducción	573
15.1. La cohesión en textos originales en español	573
15.2. Diferencias cohesivas entre textos originales en inglés y su traducción al español	576
15.3. Recapitulación	579
16. Conclusiones	581
BIBLIOGRAFÍA	595

APÉNDICES	621
Apéndice I: El proceso cognitivo de la traducción (Bell 1991: 59)	621
Apéndice II: Textos analizados	622
Apéndice III: Tipos de repetición	693
Tabla 1. Detección de posibles instancias de cohesión léxica	693
Tabla 2. Repetición léxica simple	694
Tabla 3. Repetición léxica compleja	695
Tabla 4. Paráfrasis	696
Apéndice IV: Fragmento inicial de <i>Masters of Political Thought</i> (Hoey 1991: 246-247)	697
Apéndice V: Matriz de repetición del texto del Apéndice IV	698
Tabla 1. Desarrollo de la primera línea de la matriz de repetición en Tabla 2	699
Tabla 2. Matriz de repetición del texto del Apéndice IV	700
Apéndice VI: Matrices de repetición en forma de pares de coordenadas para las oraciones 3-12 del texto 6 del Apéndice II	701
Apéndice VII: Resultados de Hesperus	706
VII a: Perspectiva general de los 36 textos del <i>PDQ</i> (<i>Cancer Treatment Information summaries, Supportive Care and Advocacy Issues</i>)	707
VII b: Perfil del texto QDT2 (TEXTO 10)	710
VII c: Versión de QDT2 con hipervínculos hacia las cadenas léxicas	712
VII d: Representación de la cadena 1 del texto QDT2	721
VII e: Representación tabular de las cadenas reconocidas por <i>Hesperus</i>	731
Malignant mesothelioma (QDT1)	731
Non-small cell lung cancer (QDT2)	734
Small cell lung cancer (QDT3)	739
Malignant mesothelioma (QPT1)	743
Non-small cell lung cancer (QPT2)	745
Small cell lung cancer (QPT3)	746
Apéndice VIII: Lista de frecuencia del Corpus <i>Brown</i> con los 75 lexemas más frecuentes	749
Apéndice IX: Lista de frecuencia del <i>Corpus de Referencia del Español Actual (CREA)</i> con los 75 lexemas más frecuentes	750
Apéndice X: Palabras clave identificadas con la aplicación <i>Keywords</i> de <i>Wordsmith Tools</i> en los textos sobre tratamiento para profesionales de la salud del <i>PDQ</i> (QDT)	751

ÍNDICE DE TABLAS Y DIAGRAMAS

TABLAS

Tabla 1: Funciones del lenguaje y tipos textuales	29
Tabla 2: El tenor en interacción con el campo y el modo. (Hatim 1997: 26)	70
Tabla 3: Movimientos y funciones discursivas presentes en los artículos experimentales (Nwogu 1997: 125)	93
Tabla 4: Taxonomía textual según el tenor y el canal de comunicación basada en Lévy-Leblond (1996)	98
Tabla 5: Colocaciones de la palabra <i>enseñanza</i>	105
Tabla 6: Categorías de cohesión léxica	106
Tabla 7: Reemplazos y elementos constantes del ejemplo (9)	109
Tabla 8: Tabla de conexiones del texto del Apéndice IV	119
Tabla 9: Tabla de vínculos del texto del Apéndice IV	121
Tabla 10: Tabla de vínculos (adaptación de Hoey 1995: 77)	122
Tabla 11: Tipos de Funciones léxicas (Mel'čuk 1996: 41)	140
Tabla 12: Ejemplos de corpus monolingües y multilingües especializados	174
Tabla 13: Tipos de textos paralelos	176
Tabla 14: Definición de los verbos de apoyo (Mel'čuk 1996: 64).	209
Tabla 15: Propuesta de funciones léxicas paradigmáticas	215
Tabla 16: Modificación de las categorías de cohesión propuestas por Hoey (1991)	218
Tabla 17: Fragmento de <i>Tabla tipo A</i>	220
Tabla 18: Fragmento de <i>Tabla tipo B</i>	222
Tabla 19: Datos estadísticos sobre el <i>abstract</i> en inglés	226
Tabla 20: Frecuencia de lexemas aislados	227
Tabla 21: Combinaciones de dos palabras	228
Tabla 22: Combinaciones de tres palabras	228
Tabla 23: Combinaciones de cuatro palabras	228
Tabla 24: Primera columna de la matriz de repetición del texto 6 (desarrollo de Hoey 1991)	229
Tabla 25: Cohesión con respecto a la primera oración (<i>Tabla A</i>)	230
Tabla 26: Cohesión con respecto a la segunda oración (<i>Tabla A</i>)	231
Tabla 27: Cómputo de las conexiones en las que participa cada oración	231
Tabla 28: Tabla de las conexiones en las que participa cada oración	232
Tabla 29: Perspectiva general correspondiente a los textos QDT1 Y QDT2	234

Tabla 30: Fragmento del perfil del texto QDT2.	235
Tabla 31: Versión HTML del texto QDT2.	236
Tabla 32: Fragmento de una de las 14 cadenas léxicas del texto QDT2.	237
Tabla 33: Activación de términos que designan partes del cuerpo con frecuencia mayor al 0,1% en QDT2 (<i>Microsoft Excel</i>).	243
Tabla 34: Representación de las cadenas reconocidas por <i>Hesperus</i> (Cadena 4 de QDT2)	244
Tabla 35: Representación de los conceptos identificados por <i>Hesperus</i> en las cadenas 4, 7 y 11 de QDT2: porcentaje de activación.	245
Tabla 36: Casilla perteneciente a una de las <i>cadenas léxicas definitivas</i> de QDT2	247
Tabla 37: Fragmento del cuadro contrastivo correspondiente al texto QDT2.	248
Tabla 38: Composición del corpus	256
Tabla 39: Datos estadísticos del texto "El cáncer de pulmón se asocia a alteraciones del cromosoma 9"	260
Tabla 40: Lista de frecuencia lematizada	262
Tabla 41: Combinaciones dos palabras	263
Tabla 42: Combinaciones de tres palabras	263
Tabla 43: Combinaciones de cuatro palabras	263
Tabla 44: Formas más frecuentes del texto traducido y el texto original	263
Tabla 45: Matriz de repetición de las relaciones cohesivas primarias (<i>Tabla A</i>)	265
Tabla 46: Matriz de repetición de las relaciones cohesivas primarias (<i>Tabla B</i>): tipos de repetición y funciones léxicas	266
Tabla 47: Conexiones en las que participa cada oración	267
Tabla 48: Tabla de vínculos	268
Tabla 49: Tipos de repetición	269
Tabla 50: Funciones léxicas activadas	278
Tabla 51: Matriz de repetición de las relaciones cohesivas secundarias	284
Tabla 52: Conceptos más activados en el texto	286
Tabla 53: Identificación preliminar de áreas conceptuales activadas en QDT1 a partir de lista de frecuencia lematizada	292
Tabla 54: Cadenas 1, 2a y 3 a	295
Tabla 55: Cadena 7 (QDT1)	296
Tabla 56: Cadenas 0 y 8 (QDT1)	296
Tabla 57: Cadenas 2b, 4 y 9 (QDT1)	297

Tabla 58: Cadena 3b (QDT1)	298
Tabla 59: Cadena 5 (QDT1)	298
Tabla 60: Cadenas 12, 13 y 14 (QDT1)	299
Tabla 61: Cadenas difusas: 6, 10 y 11 (QDT1)	300
Tabla 62: Multidimensionalidad	307
Tabla 63: Cadenas léxicas en QDT1	308
Tabla 64: Grado de especialización de los términos que designan el concepto MALIGNANT MESOTHELIOMA y su alternancia para crear cohesión.	317
Tabla 65: Tipos de relaciones cohesivas a partir del término <i>cytology</i> .	320
Tabla 66: Descripción cuantitativa de las cadenas léxicas reconocidas en QDT1	329
Tabla 67: Identificación preliminar de áreas conceptuales activadas en QDT2 a partir de lista de frecuencia lematizada	335
Tabla 68: Cadena 0, 1 y 2a (QDT2)	337
Tabla 69: Cadena 3 (QDT2)	338
Tabla 70: Cadenas 4, 7 y 11 (QDT2)	339
Tabla 71: Cadenas 5 y 9 (QDT2)	340
Tabla 72: Cadena 6 (QDT2)	341
Tabla 73: Cadena 8 (QDT2)	342
Tabla 74: Cadenas difusas: 10 y 12 (QDT2)	342
Tabla 75: Cadena difusa: 2b	343
Tabla 76: Cadena 13 (QDT2)	344
Tabla 77: Cadenas léxicas en QDT2	346
Tabla 78: Descripción cuantitativa de las cadenas léxicas reconocidas en QDT2	371
Tabla 79: Identificación preliminar de áreas conceptuales activadas en QDT2 a partir de lista de frecuencia lematizada	375
Tabla 80: Cadena 0, 2, 3 a (QDT3)	378
Tabla 81: Cadena 1 (QDT3)	379
Tabla 82: Cadena 3b: Difusa (QDT3)	380
Tabla 83: Cadena 4 (QDT3)	381
Tabla 84: Cadena 5b y 6 (QDT3)	382
Tabla 85: Cadena 5a (QDT3)	382
Tabla 86: Cadena 7 (QDT3)	383
Tabla 87: Cadenas léxicas en QDT3	384
Tabla 88: Descripción cuantitativa de las cadenas léxicas reconocidas en QDT3	405

Tabla 89: Identificación preliminar de áreas conceptuales activadas en QPT1 a partir de lista de frecuencia lematizada	409
Tabla 90: Cadena 0 (QPT1)	411
Tabla 91: Cadena 1 (QPT1)	411
Tabla 92: Cadena 2 (QPT1)	412
Tabla 93: Cadena 3 (QPT1)	412
Tabla 94: Cadena 4 (QPT1)	413
Tabla 95: Cadena 5 (QPT1)	413
Tabla 96: Cadena 6 (QPT1)	414
Tabla 97: Cadenas léxicas en QPT1	415
Tabla 98: Descripción cuantitativa de las cadenas léxicas reconocidas en QPT1	432
Tabla 99: Identificación preliminar de áreas conceptuales activadas en QPT2 a partir de lista de frecuencia lematizada	434
Tabla 100: Cadena 0 (QPT2)	437
Tabla 101: Cadena 1 (QPT2)	437
Tabla 102: Cadena 2 (QPT2)	438
Tabla 103: Cadena 3 (QPT2)	438
Tabla 104: Cadena 4 (QPT2)	439
Tabla 105: Cadena 5 (QPT2)	439
Tabla 106: Cadena 6 (QPT2)	439
Tabla 107: Cadenas léxicas en QPT2	441
Tabla 108: Descripción cuantitativa de las cadenas léxicas reconocidas en QPT2	453
Tabla 109: Identificación preliminar de áreas conceptuales activadas en QPT3 a partir de lista de frecuencia lematizada	455
Tabla 110: Cadena 0 (QPT3)	458
Tabla 111: Cadenas 1, 3 y 6 (QPT3)	458
Tabla 112: Cadena 2 (QPT3)	459
Tabla 113: Cadena 4 (QPT3)	459
Tabla 114: Cadena 5 (QPT3)	460
Tabla 115: Cadena 7 (QPT3)	460
Tabla 116: Cadena 8 (QPT3)	461
Tabla 117: Cadena 9 (QPT3)	461
Tabla 118: Cadenas léxicas en QPT3	463
Tabla 119: Descripción cuantitativa de las cadenas léxicas reconocidas en QPT3	475
Tabla 120: Subcadena <i>TREATMENT</i> del texto QDT1	478

Tabla 121: Estadísticas correspondientes al corpus analizado en el capítulo 11	483
Tabla 122: Cadenas léxicas correspondientes a textos para especialistas (QDT)	488
Tabla 123: Cadenas léxicas correspondientes a textos para pacientes (QPT)	489
Tabla 124: Lexemas encadenados en las macrocadenas de los textos para especialistas sobre TRATAMIENTO	492
Tabla 125: Lexemas encadenados en las macrocadenas de los textos para pacientes sobre TRATAMIENTO	492
Tabla 126: Unidades de significación especializada dentro de la subcadena <i>RADIOTHERAPY</i>	495
Tabla 127: Cuadro contrastivo con la activación conceptual y terminológica de la macrocadena <i>LOCATION IN THE HUMAN BODY</i>	503
Tabla 128: Macrocadenas léxicas que activan un 15% de los conceptos	507
Tabla 129: Lemas con frecuencia relativa superior al 1% en los textos sobre tratamiento para especialistas	522
Tabla 130: Lemas con frecuencia relativa superior al 1% en los textos sobre tratamiento para pacientes	529
Tabla 131: Palabras clave en QDT1 (<i>Treatment for malignant mesothelioma</i>)	531
Tabla 132: Palabras clave en los textos sobre tratamiento para especialistas (QDT)	532
Tabla 133: Lemas más frecuentes en textos sobre tratamiento y cuidados paliativos para profesionales de la salud	535
Tabla 134: Ficha con el radio colocacional de <i>administ*</i>	538
Tabla 135: Radio colocacional de <i>patient*</i>	539
Tabla 136: Radio colocacional de <i>cancer*</i>	541
Tabla 137: Radio colocacional de <i>treat*</i>	543
Tabla 138: Radio colocacional de <i>symptom*</i>	544
Tabla 139: Radio colocacional de <i>therapy*</i>	545
Tabla 140: Radio colocacional de <i>chemotherap*</i>	547
Tabla 141: Radio colocacional del lema USE	548
Tabla 142: Radio colocacional de <i>stag*</i>	549
Tabla 143: Radio colocacional de <i>combin*</i>	551
Tabla 144: Radio colocacional de <i>combination</i>	552
Tabla 145: Radio colocacional de <i>combined</i>	553
Tabla 146: Colocaciones de <i>effect*</i> en el corpus de textos para especialistas	554
Tabla 147: UF formadas sobre la base <i>effect*</i>	554
Tabla 148: Radio colocacional de <i>effect*</i>	555

Tabla 149: Radio colocacional de <i>cause</i> *	558
Tabla 150: Radio colocacional de <i>respons</i> *	559
Tabla 151: Radio colocacional de <i>produc</i> *	560
Tabla 152: Radio colocacional de <i>factor</i> *	561
Tabla 153: Radio colocacional de <i>develop</i> *	562
Tabla 154: Radio colocacional de <i>associated</i>	564
Tabla 155: Radio colocacional de <i>experience</i> *	565
Tabla 156: Formas léxicas en el radio colocacional (-5, +5) de <i>occur</i> *	566
Tabla 157: Radio colocacional de <i>occur</i> *	567
Tabla 158: Posición de los verbos modales de <i>PDQ</i> con respecto al corpus de referencia <i>Brown</i>	569
Tabla 159: Lexemas para articular matices de certeza más usuales: frecuencia por cada 10 000 palabras (Hyland 1998: 149)	570

DIAGRAMAS

Diagrama 1: El proceso cognitivo de la traducción (Bell 1991: 46. Cfr. Apéndice I)	14
Diagrama 2: Tipo de texto y criterios relevantes en la traducción (Snell-Hornby 1988: 32)	41
Diagrama 3. Jerarquía <i>top-down</i> de decisiones a la hora de traducir	53
Diagrama 4: Las tres dimensiones del contexto	72
Diagrama 5: Relación jerárquica entre género, discurso y texto (Hatim 1990)	74
Diagrama 6: Triángulo de conexiones	113
Diagrama 7. Diagrama topológico del texto del Apéndice V	120
Diagrama 8: Porcentaje de los principales tipos de conexión en 6 libros (Ellman 1999)	137
Diagrama 9: Referencia anafórica directa e indirecta	147
Diagrama 10: La organización del generador (Lee y Evens 1996: 301)	149
Diagrama 11: Cohesión y dificultad comunicativa (Tanskanen 1995: 537)	163
Diagrama 12: Prototipo de estación de trabajo para la traducción humana asistida por ordenador (Peters y Picchi 1997: 270)	179
Diagrama 13: Redes colocacionales (Phillips 1985: 131-132)	187
Diagrama 14: Elementos del <i>EVENTO ONCOLÓGICO</i>	189
Diagrama 15: Verbos de apoyo que se combinan con la palabra clave <i>analysis</i> . Relaciones sintácticas en la estructura profunda	211
Diagrama 16: Propuesta de representación del marco conceptual activado en el texto	287
Diagrama 17: Realizaciones del concepto ALTERACIONES EN EL CROMOSOMA 9	288

Diagrama 18: Realizaciones del concepto CÁNCER	289
Diagrama 19: Representación preliminar de áreas conceptuales activadas en QDT1	293
Diagrama 20: Porcentaje de activación de las cadenas léxicas definitivas de QDT1	309
Diagrama 21: Comparación de las áreas conceptuales reconocidas por <i>Hesperus</i> y las reconocidas por un humano asistido por medios informáticos (QDT1)	332
Diagrama 22: Representación preliminar de las áreas conceptuales activadas en QDT2	335
Diagrama 23: Porcentaje de activación de las cadenas léxicas definitivas de QDT2	348
Diagrama 24: Comparación de las áreas conceptuales reconocidas por <i>Hesperus</i> y las reconocidas por un humano asistido por medios informáticos (QDT2)	374
Diagrama 25: Representación preliminar de áreas conceptuales activadas en QDT3	376
Diagrama 26: Porcentaje de activación de las cadenas léxicas definitivas de QDT3	385
Diagrama 27: Comparación de las áreas conceptuales reconocidas por <i>Hesperus</i> y las reconocidas por un humano asistido por medios informáticos (QDT3)	407
Diagrama 28: Representación preliminar de áreas conceptuales activadas en QPT1	409
Diagrama 29: Porcentaje de activación de las cadenas léxicas definitivas de QPT1	416
Diagrama 30: Representación preliminar de áreas conceptuales activadas en QPT2	435
Diagrama 31: Porcentaje de activación de las cadenas léxicas definitivas de QPT2	443
Diagrama 32: Representación preliminar de áreas conceptuales activadas en QPT3	456
Diagrama 33: Porcentaje de activación de las cadenas léxicas definitivas de QPT3	464
Diagrama 34 : Macrocadenas léxicas en los textos para especialistas (QDT)	491
Diagrama 35: Macrocadenas léxicas en los textos para pacientes (QPT)	491
Diagrama 36: Líneas de dispersión del lema STAGE	524
Diagrama 37: Líneas de dispersión de los lemas TREATMENT, RADIOTHERAPY, CHEMOTHERAPY y SURGERY	525
Diagrama 38: Líneas de dispersión de los lemas TUMOR, DISEASE, CANCER, SMALL, CELL, LUNG Y MESOTHELIOMA	526
Diagrama 39: Líneas de dispersión de los lemas PATIENT y SURVIVAL	527

1. INTRODUCCIÓN*

* Este trabajo se ha elaborado dentro del marco del proyecto de investigación *OncoTerm: Sistema bilingüe de información y recursos oncológicos*, financiado por la DGICYT (código PB98-1342). Su objetivo es el de poner al alcance de pacientes y profesionales de la medicina y la traducción especializada información relacionada con el cáncer. El sistema consta de una ontología y una base de datos terminológica y textual en inglés y español, recursos que estarán integrados entre sí en *Internet*.

1. Introducción

1.1. Presentación

En nuestra época, la fascinación por nuevos artefactos audiovisuales convive con un atenuado culto hacia la palabra escrita, encarnada en una infinidad de textos y géneros que las sociedades humanas han legitimado como vehículos de comunicación apropiados para un fin concreto. Si bien es verdad que la historia del hombre se ha construido sobre textos orales y escritos, sin embargo, no todas las disciplinas cuyo objeto de estudio son los textos han estudiado de una forma sistemática los distintos tipos textuales reconocidos por una comunidad lingüística.

De hecho, de las disciplinas que han llegado más lejos en su análisis del texto, la crítica literaria, la lingüística y la traducción, sólo la primera saltó de planteamientos programáticos a la elaboración de taxonomías textuales con anterioridad al siglo XX. Podemos encontrar una explicación de esto en el carácter poco tangible de esta unidad, el texto, que Brown y Yule (1983:6) definen como “the verbal record of a communicative act”.

Los estudios de traducción comenzaron a considerar al texto como unidad de traducción y de equivalencia con el desarrollo de la lingüística textual en la década de los setenta. No obstante sólo en la década de los noventa se ha producido el salto de modelos traductológicos programáticos y muy generales, a modelos que pueden explicar el ejercicio concreto de la traducción en un dominio determinado.

Esto se debe en parte a que, durante muchos años, los estudios de traducción no han sabido aprovechar las aportaciones de otras disciplinas que han explorado el área de la variación textual: la lingüística, el inglés para fines específicos, la psicolingüística, la inteligencia artificial, la retórica contrastiva, la sociología de la ciencia, la lingüística de corpus y la terminología. Estas disciplinas han aportado términos como el de *modo de discurso*, *superestructura*, *macroestructura*, *género*, *tipo textual*, *registro*, *discurso* y *sublenguaje* y han intentado identificar los parámetros que los definen. Buena prueba de esta escasa colaboración interdisciplinar lo constituye el hecho de que el concepto de *género* se ha introducido de un modo débil en el área de los estudios de traducción.

Por tanto, partimos de la premisa de que los estudios sobre género pueden ayudar al traductor a desarrollar estrategias para reconocer las opciones y las limitaciones que imponen determinados tipos textuales (Trosborg 1997: viii).

En la traducción de textos biomédicos, el traductor tiene que estar familiarizado con las opciones y limitaciones relacionadas con una serie de recursos morfosintácticos, con unas convenciones retóricas y ortotipográficas y con una terminología constantemente en expansión. En estas tres facetas subyace la identificación de tipos textuales biomédicos, sin la cual, el texto traducido perdería su función comunicativa y se podrían ocasionar consecuencias indeseables. Entre estas, podemos mencionar el rechazo de un artículo por el comité editorial de una revista prestigiosa o la creación de expectativas de curación falsas, ocasionadas por la traducción incorrecta de un verbo modal.

El traductor debe conocer los recursos morfosintácticos asociados a determinados géneros. Por ejemplo, para traducir un folleto explicativo de cómo realizar una autoexploración mamaria, es necesario que el traductor sepa que, en inglés, los folletos de instrucciones utilizan el imperativo mientras que la lengua española impone una mayor variación sintáctica, y así, es frecuente encontrar además del imperativo, el presente de indicativo, el impersonal con *se* o incluso un uso incorrecto del infinitivo. También es preciso que reproduzca las convenciones tipográficas y retóricas establecidas por la comunidad médica internacional. Entre estas se pueden citar las normas de publicación de manuscritos enviados a publicaciones científicas, que especifican las secciones de que consta un artículo experimental.

El conocimiento de la terminología biomédica comprende el conocimiento de las entidades y procesos implicados en la investigación biomédica y una competencia *documentalista* para acceder a bases de datos, glosarios y a organismos de normalización dependientes de la ISO (*International Standards Organization*) como AENOR, ANSI y BSI. Asimismo requiere un conocimiento de cómo la terminología cambia en distintos registros como el propio de una publicación periódica especializada o el de una revista de divulgación o el que utiliza un médico en su consulta.

La utilización de terminología médica supone la activación de redes conceptuales de significado que hacen del texto una unidad gracias al fenómeno que se

denomina *cohesión léxica*. La cohesión léxica contribuye a la textualidad y es un parámetro sensible al tipo textual en el que se establece. De ahí la necesidad de que el traductor tome consciencia de los mecanismos cohesivos presentes en el texto origen (en adelante, TO) y el texto término (TT), de las redes conceptuales activadas mediante las unidades léxicas y terminológicas del texto, al mismo tiempo que evalúa si se ha mantenido el valor discursivo de la cohesión en el TT.

1.2. Hipótesis de trabajo y objetivos

En este trabajo se parte de tres hipótesis. En primer lugar, existe una correlación entre tipo textual y los recursos cohesivos que lo materializan. En segundo lugar, existen diferencias con respecto a la cohesión entre las lenguas inglesa y española que tienen una importante repercusión en la traducción. Por último, la cohesión léxica activada en los textos médicos refleja en gran medida la compleja red conceptual que sustenta ese campo del saber.

Para validar estas hipótesis, se proponen una serie de objetivos que se materializarán en un corpus de textos sobre oncología:

- a) Revisar la literatura sobre tipología textual en traducción incidiendo en los aspectos cognitivos, lingüísticos y sociales implicados en la producción y comprensión de tipos textuales.
- b) Revisar los estudios sobre cohesión léxica en cuanto que esta puede ser un parámetro sensible al tipo textual en el que se establece
- c) Diseñar un método de análisis de la cohesión donde se computen las relaciones cohesivas en textos completos con la ayuda de recursos informáticos.
- d) Estudiar qué diferencias cohesivas se pueden encontrar al contrastar dos tipos textuales que difieren con respecto al destinatario y a los conocimientos que este tiene sobre medicina.
- e) Estudiar si existen diferencias con respecto a la cohesión entre un texto y su traducción a otra lengua y discernir cuáles son el resultado de las convenciones retóricas inherentes al inglés y al español.
- f) Establecer un marco teórico-práctico que ayude al traductor, principalmente, al que trabaja con el inglés y el español, a sistematizar los recursos cohesivos y asociarlos a

tipos textuales diferentes. El traductor debe ser consciente de las decisiones implicadas en la traducción de distintos tipos textuales relacionados con el ámbito de la medicina de manera que consiga traducir de una forma lo más comunicativa posible. También debe ser sensible al hecho de que en algunas ocasiones la selección de terminología médica en los textos no siempre responde a la organización conceptual de ese dominio. Conociendo cómo se articula la cohesión en los textos y los factores retóricos que inciden sobre la misma, el traductor puede discernir por qué se han activado unos términos, y no otros de esa organización conceptual.

- g) Aplicar la metodología de la lingüística de corpus y el concepto de cohesión en la detección de la representación conceptual subyacente a los textos y estudiar las implicaciones que esto puede tener en terminología, documentación y traducción.

1.3. Ubicación del trabajo en un marco interdisciplinar

Todos estos objetivos se van a contemplar desde una perspectiva interdisciplinar que aúne las aportaciones sobre tipologías textuales provenientes de los *estudios de traducción* y la *lingüística*. También recogerá los estudios sobre *género* realizados desde la *lingüística aplicada* y algunas de sus subdisciplinas como la lengua para fines específicos (Bhatia 1993; Eggins y Martin 1997; Nwogu 1997; Swales y Feak 1994), la retórica de contrastes (Connor 1996; Connor y Kaplan 1987) y la textología contrastiva. (Hartmann 1996).

El inglés para fines específicos arroja luz sobre cómo se estructuran los principales géneros médicos en inglés y sus rasgos lingüísticos. La retórica de contrastes identifica las diferencias lingüísticas y retóricas entre textos escritos en la lengua materna y en una segunda lengua. La textología contrastiva propone una integración de la lingüística de contrastes, el análisis de discurso y la lingüística de corpus en el estudio de textos paralelos en más de una lengua. Relacionada con estas disciplinas está también la *sociología de la ciencia* (Bazerman y Paradis 1991; Myers 1990), que indaga sobre el conocimiento cultural y social que determina y estructura retóricamente los textos científicos, a pesar de que estos aluden a una realidad objetiva independiente del investigador.

Asimismo, se revisarán algunas nociones de *psicolingüística* (Baddeley 1992, Bell 1991; Faber 1998a) e intentaremos explicar el papel de la cohesión en el procesamiento de la información y del lenguaje.

La *terminología* (Budin y Wright 1997; Cabré 1999a; Somers 1996; Sonneveld y Loening 1993) también será un claro referente porque es la disciplina que estudia los términos de una materia especializada, cómo recopilarlos, los mecanismos de normalización de los términos y cómo estructurar las relaciones conceptuales de ese campo de especialidad. Dentro de la terminología se adoptará una aproximación descriptiva en la que se gestione la información de una forma orientada a la traducción. Esto es lo que Wright (1999) denomina *translation-oriented terminology management*. Dentro de esta, no nos interesa tanto la terminología puntual que suscita la traducción de un texto especializado sino el intento de sistematizar recursos y aplicaciones útiles para el traductor. Esta aproximación es la que Cabré (1999b) denomina *terminología para la traducción*.

Por último, como se pretende la aplicación de recursos informáticos al estudio de la cohesión, nos vamos a beneficiar de la metodología de la *lingüística de corpus* y de algunas aportaciones provenientes de la *lingüística computacional* y la *ingeniería de conocimiento* (Morris y Hirst 1991, Ellman 1998). Esta última elabora ontologías, es decir, repositorios estructurados de conocimiento, en las que se organizan los conceptos que articulan un determinado campo del saber. La lingüística de corpus (Aijmer y Altenberg 1991; Baker *et al.* 1993; Biber 1993, 1998; Biber y Finegan 1991; Butler 1992; McEnery y Wilson 1996; Sinclair 1991; Stubbs 1996; Svartvik 1992) aporta respuestas a algunos de los dilemas que se plantean a la hora de recopilar un corpus en soporte electrónico *representativo*:

- criterios de selección del corpus (fecha de publicación, textos completos, variaciones dialectales, géneros, tema, etc.)
- tamaño
- formato de anotación más apropiado para que el corpus sea reutilizable
- programas informáticos que elaboren listas de frecuencia y líneas de concordancia y que etiqueten los corpóra para poder extraer de ellos información sintáctica, semántica, pragmática y suprasegmental.

- distinción entre *córpore comparables* y *córpore paralelos* en soporte electrónico

1.4. Breve resumen de los capítulos

1.4.1. Primera parte: revisión bibliográfica

La primera parte, dedicada a la revisión de bibliografía, girará en torno a las tipologías textuales, la cohesión y la lingüística de corpus. En el capítulo 2 se parte de un modelo cognitivo de la traducción para poder explicar los elementos de la competencia traductológica que intervienen en la comprensión y producción de tipos textuales. Se recalca la idea de que en la traducción participan los mismos sistemas neurológicos que intervienen en el procesamiento de la información. Estos ejecutan procesos interactivos y no-lineales que suponen una continua toma de decisiones. A continuación, se aborda el concepto de tipología textual, por una parte, desde aproximaciones funcionales y textuales como las de Reiss, House, la teoría del *skopos* (Vermeer), Sager, Nord y Hatim y Mason, y por otra, desde aproximaciones más eclécticas que, aún reconociendo la importancia de los aspectos pragmáticos en la traducción, hacen un gran hincapié en la cognición, como la de Werlich, Neubert, Snell- Hornby y Kussmaul.

En el capítulo 3, se aclara el significado de los términos *tipo textual*, *registro* y *género* y se incide en la idea de que el texto es un todo organizado. Se describen algunos de los tipos textuales médicos más usuales.

En capítulo 4, se revisan los estudios sobre cohesión y se presta una especial atención a la metodología propuesta por Hoey (1991, 1995) y, de todos los recursos cohesivos, a la *cohesión léxica*. También se incluye una aproximación proveniente del campo de la lingüística computacional, la de Morris y Hirst (1991), que hablan de la cohesión en términos de *cadena léxica*. Se comentan algunas herramientas informáticas para la detección automática de la cohesión y se destacan aquellas que reconocen relaciones conceptuales. Nuestro interés por el aspecto semántico de la cohesión nos llevará a revisar la propuesta de la *Teoría Sentido Texto* (Mel'cuk *et al.* 1984) en relación con las representaciones semánticas de los textos. Por último, se sugieren algunos elementos que interactúan con la cohesión, tal y como el tipo textual, el conocimiento que comparten el emisor y el receptor de un texto, las dificultades

comunicativas que impone el contexto situacional, el grado de explicitación y convenciones retóricas que dominan las lenguas, y el proceso traductor en sí.

El capítulo 5 trata sobre la relevancia que está adquiriendo la lingüística de corpus en la traducción y en terminología y de la necesidad de distinguir entre corpora paralelos y comparables. Se citan algunas aproximaciones que se pueden tener en cuenta en la traducción biomédica.

1.4.2. Segunda parte: metodología

En la segunda parte se conjugan la mayoría de los modelos sobre cohesión presentados en la primera parte y se propone una doble metodología para el estudio de la cohesión que tiene en cuenta la naturaleza del fenómeno cohesivo, que sintetizamos en el capítulo 6. En este capítulo también se proponen categorías para describir las instancias de cohesión léxica y se sistematizan las posibles relaciones entre elementos cohesionados en una taxonomía de funciones léxicas cohesivas. La primera metodología, llamada *análisis intensivo*, es la que se sugiere en el capítulo 7 para el análisis de textos breves o fragmentos de texto y destaca la faceta interoracional de la cohesión. Se aplicará en el capítulo 10. El capítulo 8 se centra en el *análisis extensivo* de la cohesión, que parte de una concepción de la cohesión como un fenómeno tanto interoracional como intraoracional. En este capítulo se presentan los módulos de la herramienta informática que nos va a ayudar a detectar las *cadena léxicas* del texto y cómo se plasmarán en papel los datos que se extraen del ordenador. Por último, en el capítulo 9 se detalla el corpus de textos sobre oncología que hemos manejado en esta investigación.

1.4.3. Tercera parte: resultados

Los resultados se exponen en la tercera parte de la tesis. Los obtenidos con un análisis intensivo van en el capítulo 10 y dejan entrever algunas diferencias cohesivas entre textos en inglés y español, y entre textos originales y traducidos. El capítulo 11 presenta los resultados del estudio de la cohesión léxica de forma extensiva en dos tipos textuales diferentes sobre exactamente el mismo tema: el tratamiento del cáncer de pulmón. El primer tipo textual lo constituyen textos escritos por oncólogos para profesionales de la salud y publicados en Internet; el segundo es la explicación del mismo tema al paciente y a sus familiares. En el análisis, después de describir los resultados obtenidos

automáticamente, se modifican estos para ajustarlos mejor a las relaciones cohesivas que detecta el analista humano. En todo caso, se comparan los resultados obtenidos de una y otra forma.

1.4.4. Cuarta parte: discusión y conclusiones

En el capítulo 12 comienza la parte de *Discusión y conclusiones* de la tesis y se establece la relación que existe entre las cadenas léxicas reconocibles en un texto y la activación de conceptos que se produce durante la lectura. Se comentan las diferencias conceptuales y terminológicas más significativas en las cadenas léxicas correspondientes a textos para especialistas y para pacientes. Las cadenas analizadas están vinculadas principalmente con los conceptos TREATMENT, PATHOLOGY, RESEARCH, OUTCOME OF TREATMENT, LOCATION IN THE HUMAN BODY, DIAGNOSIS, PATIENT, STAGING y SIZE / QUANTITY.

En el capítulo 13 se buscan las expresiones que contribuyen a crear cohesión en textos para lectores legos y, gracias a la cohesión, aumentan los conocimientos científicos de dichos lectores. Estas expresiones permiten que términos ininteligibles para el lector lego aparezcan en el texto cerca de lexemas de la lengua general que explican su significado.

En el capítulo 14 se indaga sobre el papel que la frecuencia de determinadas unidades léxicas en un texto o en un corpus puede jugar, por una parte, en la configuración de cadenas léxicas, y por otra, en el reconocimiento de unidades fraseológicas. Los datos que se comentan se obtienen mediante listas de frecuencia y líneas de concordancia. Se reflexiona sobre cómo la distribución de *palabras forma* (*function words*) en un corpus nos puede orientar sobre las características del mismo.

En el capítulo 15, se detectan las diferencias cohesivas entre textos originales en inglés y su traducción al español y se señalan aquellas que responden a las características del inglés y del español, y al proceso traductor, tal y como aparecen en traducciones al español de textos sobre oncología en inglés. Las conclusiones quedan reflejadas en el capítulo 16.

**PRIMERA PARTE:
REVISIÓN BIBLIOGRÁFICA**

2. Tipologías textuales en los estudios de traducción

La didáctica de la traducción enfatiza hoy más que nunca la necesidad de que los alumnos se familiaricen con una gran variedad de *tipos textuales*, de modo que adquieran una serie de estrategias que les permitan producir traducciones que cumplan los criterios de textualidad propuestos en Beaugrande y Dressler (1981: 182 y ss.). Estos autores (*ibid*: 3) definen el *texto* como:

A communicative occurrence which meets seven standards of textuality. If any of these standards is considered not to have been satisfied, the text will not be communicative. Hence, non-communicative texts are treated as non-texts.

Estos criterios de textualidad que debe cumplir cualquier texto, ya sea original o una traducción, son los siguientes:

- a) *Cohesión*: relaciones internas entre los elementos lingüísticos del texto que hacen del texto una unidad en la que las oraciones están vinculadas. Esta unidad se consigue mediante recursos léxicos y sintácticos. (Cfr. Capítulo 4)
- b) *Coherencia*: continuidad de significado entre el mundo interno del texto y sus referencias externas. La coherencia se consigue cuando la secuencia de proposiciones del texto se acomoda a nuestro conocimiento del mundo.
- c) *Intencionalidad*: la finalidad que guía la producción de un texto o, siguiendo la terminología de la teoría de los actos de habla (Austin 1962; Searle 1969, 1979), la fuerza ilocucionaria del mismo.
- d) *Aceptabilidad*: el reconocimiento por parte del receptor de que el texto está cumpliendo una finalidad. Corresponde a la fuerza perlocucionaria de un texto.
- e) *Informatividad*: cantidad de información que aporta el texto en función del grado de expectativa por parte del lector y la mayor o menor novedad temática y formal del texto.
- f) *Situacionalidad*: factores que hacen que un texto sea relevante en una situación dada

- g) *Intertextualidad*: relación entre un texto y la tradición textual de la que depende, que hace que reconozcamos un texto como una instancia de un determinado género o tipo textual.

Para conseguir textos con estas características, son necesarias una serie de aptitudes que constituyen lo que Bell (1991: 42) denomina *translator communicative competence*. Esta competencia es la que posibilita el proceso cognitivo que llamamos traducción.

2.1. La traducción como proceso: el modelo cognitivo de Bell (1991)

El modelo propuesto por Bell (1991) nos parece muy adecuado para explicar la complejidad de los problemas y las decisiones a las que se enfrenta el traductor porque, en primer lugar, aprovecha los avances ocurridos en la ciencia cognitiva, la inteligencia artificial, la lingüística del texto y la lingüística sistémica, y en segundo lugar, concede un lugar preferente al significado, que, según Neubert (1991: 19), se encuentra siempre como trasfondo del proceso de traducción.

Según Bell (*ibid*: 36), la competencia traductológica comprende una serie de conocimientos y habilidades. El traductor necesita poseer competencia lingüística y comunicativa (Hymes 1972) en las lenguas origen y término, conocimiento de distintos tipos textuales, conocimiento del mundo y conocimiento contrastivo. En lo que se refiere a habilidades, el traductor ha de ser capaz, en primer lugar, de descodificar textos, lo cual implica la lectura y comprensión de los mismos; y en segundo lugar, de codificar o escribir textos.

Teniendo en cuenta que algunos de estos elementos están presentes en la comunicación monolingüe, Bell ubica la traducción dentro de la actividad humana que denominamos “procesamiento de la información” (*human information processing*). Asimismo, pretende dotar a su modelo de validez neurológica adaptando los estudios sobre la memoria realizados por Baddeley (1990, 1992)¹.

2.1.1. Memoria, procesamiento de la información y traducción

Bell (ibid: 231) describe las tres etapas que comprende el procesamiento de la información. En la primera de ellas, el cerebro filtra, de todas las sensaciones recibidas por los sentidos, sólo aquellas que le resultan relevantes y las almacena en el sistema de información sensorial (*sensory information system*), el cual convierte estas sensaciones caóticas en unidades de información, en percepciones.

En una segunda etapa, estas percepciones se almacenan temporalmente en la memoria a corto plazo. Dentro de la memoria a corto plazo, Baddeley reconoce una sección llamada memoria de trabajo (*working memory*), que es la responsable de focalizar nuestra atención y es la que se encarga de ejecutar las operaciones mentales, manipular datos, números y acontecimientos y conectar la memoria a corto plazo con la memoria a largo plazo. Pues bien, la memoria de trabajo analiza los rasgos constitutivos de esas percepciones, los organiza y los relaciona con conceptos y experiencias personales almacenadas en la memoria a largo plazo. En una tercera etapa, la información nueva queda integrada en la memoria a largo plazo. Esta operación la realiza también la memoria de trabajo y se consolida durante el sueño.

El proceso que convierte sensaciones caóticas en percepciones y las percepciones, en representaciones conceptuales almacenables en el sistema de memoria a largo plazo es común a todos los humanos. Por este motivo, es posible la comunicación entre los miembros de una misma comunidad que tienen un compendio de conceptos comunes almacenados en la memoria a largo plazo. También es posible la comunicación interlingüística porque el hecho de que muchos de nuestros conceptos tengan una misma base sensorial y cognitiva, los convierte en universales. Gracias al carácter universal de muchos de nuestros conceptos, el traductor puede convertir el significado codificado en el texto origen en una representación semántica universal independiente de lengua. Esta representación semántica le servirá de puente hacia las palabras de su lengua materna.

En cuanto al tipo de conocimiento que se almacena en nuestra memoria, Bell sostiene que es de dos tipos, que dan nombre a las dos divisiones funcionales de la memoria:

¹ En la explicación que sigue, también hemos seguido a Givón (1997, 1998a y 1998b) y Faber (1998a).

(a) *Conocimiento factual o declarativo (memoria declarativa)*

Es el conocimiento que tenemos sobre las cosas. La memoria declarativa comprende dos divisiones:

- *la memoria semántica*, que es un apartado de la memoria a largo plazo que almacena datos, números y todos los conceptos que componen el lexicón mental. Gran parte del contenido de la memoria semántica es conocimiento compartido con los demás hablantes de la comunidad humana a la que pertenecemos, lo cual posibilita la comunicación. El conocimiento del significado de las palabras y de las convenciones retóricas asociadas a un género está almacenado en la memoria semántica.
- *la memoria episódica*, que tiene registrados recuerdos de sucesos y experiencias de nuestra vida. Una parte de la memoria episódica pertenece a la memoria a corto plazo y es la que detecta acontecimientos, estados o situaciones nuevas.

(b) *Conocimiento procedimental (memoria procedimental)*

Conocimiento inconsciente y automatizado de cómo hacer las cosas. Los traductores expertos tienen algunas de las tareas implicadas en el proceso traductor como conocimiento procedimental, lo cual ahorra esfuerzo a la memoria a corto plazo. Por ejemplo, un traductor que trabaje en una compañía es capaz de escribir una carta comercial que se amolde a este tipo de género, sin pensar conscientemente de qué partes se compone la carta y cuál es su estructura.

La traducción consiste en la transformación del texto origen (TO) en el texto término (TT) mediante, básicamente, dos procesos que implican a la memoria a corto plazo y a la memoria a largo plazo. El primero es el *análisis del TO* para obtener una representación semántica universal, es decir, desvinculada de cualquier sistema lingüístico. El segundo es la *síntesis* de esa representación semántica para obtener el texto término.

Se estima que en el proceso de traducción, el sistema de la memoria cuenta con los siguientes componentes (Cfr. Apéndice I):

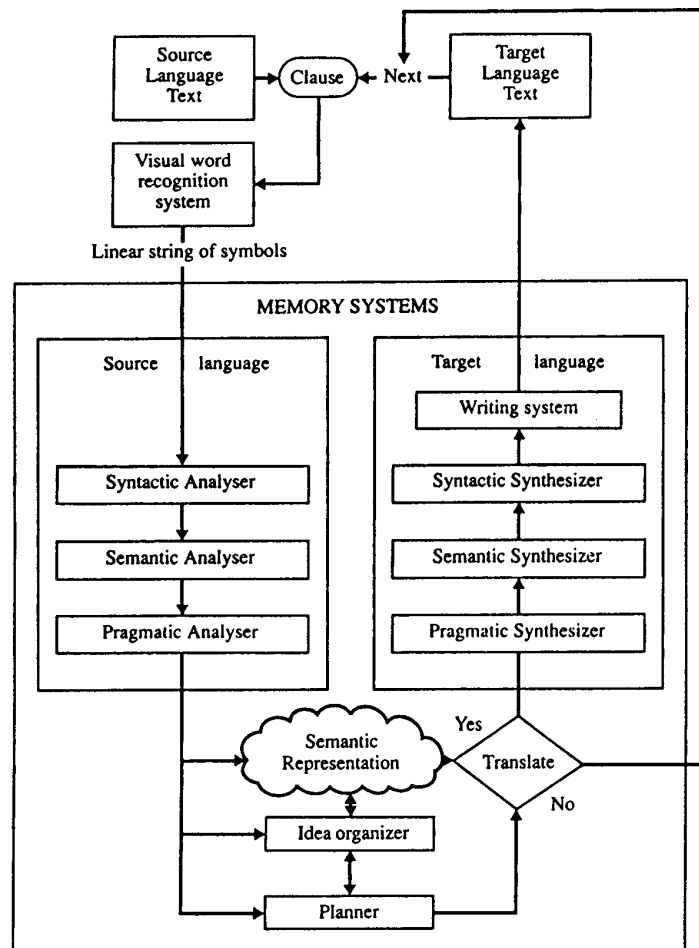
- un sistema visual de reconocimiento de palabras y un sistema de escritura
- un procesador sintáctico que se ocupa de las opciones disponibles en el sistema de la *modalidad*² y que cuenta con:
 - una base de datos con las palabras más frecuentes (*frequent lexis store*) y un inventario con las estructuras sintácticas más frecuentes (*frequent structure store*), que reducen el esfuerzo que tiene que realizar la memoria de trabajo
 - un mecanismo de búsqueda léxica (*lexical search mechanism*)
- un analizador sintáctico (*parser*), por el que pasan estructuras sintácticas *marcadas* que, o bien provienen de, o bien van dirigidas hacia otro elemento llamado procesador semántico
- un procesador semántico (*semantic processor*), que analiza las opciones disponibles en el sistema de la *transitividad* e intercambia información con un procesador pragmático
- un procesador pragmático (*pragmatic processor*), que se ocupa de las opciones disponibles en el sistema de *tematización*
- un organizador de ideas (*idea organizer*), que articula la progresión de los actos de habla en el texto, de acuerdo con los patrones de tipos textuales que el traductor tiene almacenados en la memoria. Cuando estos patrones son insuficientes para la comprensión o la producción de un texto, un planificador indica las inferencias que han de hacerse para lograrlo.
- un planificador (*planner*), que diseña las estrategias para alcanzar todo tipo de objetivos comunicativos

Todos estos elementos intervienen en un proceso que no se desarrolla de una forma lineal, etapa por etapa, sino de un modo interactivo que permite volver de nuevo a una misma etapa, hacer revisiones y saltar entre etapas en las que actúan distintos

² Estamos utilizando esta palabra con el sentido que le da la lingüística sistémica (Halliday 1985). La *modalidad* junto con la *transitividad* y la *tematización* son los tres sistemas que codifican las tres macrofunciones del lenguaje: la *interpersonal*, la *ideacional* y la *textual*, respectivamente.

componentes. Asimismo, la traducción implica el procesamiento *bottom-up*, que posibilita el reconocimiento de los caracteres tipográficos como patrones mentales ya existentes y un procesamiento *top-down*, que permite hacer inferencias basadas en experiencias pasadas, expectativas y otros esquemas mentales para interpretar los signos lingüísticos.

DIAGRAMA 1: El proceso cognitivo de la traducción (Bell 1991: 46. Cfr. Apéndice I)



Veamos las dos etapas en las que se divide la traducción (análisis y síntesis) y los distintos mecanismos (procesamiento sintáctico, semántico y pragmático) que se desarrollan en ambas etapas.

2.1.2. Etapa de análisis

2.1.2.1. Reconocimiento visual de palabras y análisis sintáctico

En esta etapa se activa el sistema visual de reconocimiento de palabras, que transforma un estímulo visual en una cadena lineal de símbolos discretos. Estos símbolos son el material que va a ser analizado por el procesador sintáctico.

El procesador sintáctico descompone cada cláusula en estructuras sintácticas y le asigna una de las opciones funcionales disponibles en el sistema de la *modalidad* de la lengua de partida. En el caso del inglés y el español, estas funciones son las de sujeto, predicado, objeto, atributo y complemento circunstancial. Estas opciones funcionales son realizadas por sintagmas, a nivel de la cláusula y por palabras, a nivel del sintagma. Posteriormente, en función del orden en que se disponen estos elementos funcionales en la cláusula y sus rasgos suprasegmentales, el procesador sintáctico reconoce cómo se codifica la actitud del emisor con respecto a un estado de cosas, es decir, si la oración está en modo indicativo, imperativo o subjuntivo y si es declarativa o interrogativa.

Bell (1991: 136, 227) considera la *cláusula*³, como la unidad de comunicación, del sistema de la modalidad, e incluso de traducción, una opinión que pocos traductores suscribirían, teniendo en cuenta que la equivalencia pocas veces se consigue en el nivel de la oración. Para el traductor la unidad de traducción es el enunciado. Bell argumenta su posición sosteniendo que la cláusula codifica contenidos proposicionales, universales mediante formas que son específicas a una determinada lengua y a un determinado *cotexto*, entendido este como el contexto lingüístico que rodea a una palabra. Además, Bell sostiene que la cláusula tiene un grado de abstracción apropiado para explicar los enunciados y los textos.

Para poder llegar a estas representaciones sintácticas, la cláusula pasa por dos bases de datos: una que contiene los lexemas más frecuentes (*frequent lexis store*) y otra que contiene las estructuras sintácticas más frecuentes (*frequent structure store*). Estas bases de datos liberan el esfuerzo impuesto sobre la memoria a corto plazo en el sentido de que permiten el procesamiento automático de palabras y combinaciones sintácticas que son muy usuales a la hora de comunicarnos. Si en las bases de datos no están los

³ Término que este autor utiliza para designar oraciones

lexemas o las estructuras sintácticas de la cláusula que se está analizando, esta pasa por un mecanismo de búsqueda léxica o por un analizador sintáctico.

En todo caso, lo que queda claro es que existe una íntima relación entre el léxico y la sintaxis (Dik 1989; Martín Mingorance 1998). En el significado de cada palabra están implícitas las funciones sintácticas que puede desempeñar en estructuras superiores y sus posibles roles semánticos. Entre las funciones sintácticas que puede desempeñar, mencionaremos la de sujeto, predicado, objeto, atributo, complemento circunstancial, núcleo de sintagma nominal, verbal, preposicional, adjetivo y adverbial, determinante y modificador. En cuanto a los roles semánticos, los detallaremos en el siguiente apartado.

Al final de esta etapa, la cadena de símbolos se ha transformado en una estructura en la que cada palabra lleva asociado un significado léxico provisional y unas etiquetas sintácticas.

2.1.2.2. Análisis semántico

El analizador semántico detecta las relaciones lógicas que se han verbalizado en cada cláusula, es decir, los roles semánticos de las entidades de las que se predica algo para presentar un *estado de cosas* estático o dinámico. A la representación mental de un estado de cosas se le denomina *proposición*. La cláusula es la representación verbal de una o varias proposiciones y presenta las relaciones lógicas que el emisor ha decidido codificar de entre todas las relaciones que pueden existir entre dichas entidades en el mundo real o imaginario.

Mientras que las cláusulas son dependientes de la lengua, el contenido de las proposiciones es el único que puede considerarse universal y es el que hace posible la traducción. De la codificación de estas relaciones lógicas en la cláusula se encarga el sistema de la *transitividad*.

La Gramática Funcional (Dik 1978, 1989) ha formalizado las proposiciones mediante una estructura llamada *predicación nuclear*. Esta consta de un *predicado verbal* que designa las propiedades y relaciones entre entidades (*argumentos*). Según Dik (1989: 101-104), estos argumentos pueden tener uno de los siguientes roles semánticos: *agent*, *positioner*, *force*, *processed*, *zero*, *zero experiencer*, *processed*

experiencer, goal, recipient, location, direction, source, reference, goal experiencer y recipient experiencer.

A continuación, vamos a presentar una predicación nuclear que pueda ilustrar el tipo de operaciones que ha de realizar el analizador semántico para extraer el contenido al que se refiere una estructura sintáctica como, por ejemplo, *Escucha esta canción*. El contenido es el conjunto de relaciones lógicas que existen entre participantes, procesos y circunstancias.

(1) escuchar V (x_1 :< humano> (x_1))_{Agente} (x_2 :<sonido emitido por concreto> (x_2))_{Meta}

En este marco predicativo, aparece un predicado verbal (*escuchar*) que tiene dos argumentos, representados por las variables x_1 y x_2 . Estos argumentos tienen la función de Agente y Meta (*Goal*) respectivamente. Las etiquetas <humano> y <sonido emitido por concreto> especifican las restricciones de selección que el verbo *escuchar* impone sobre los posibles argumentos que pueden ocupar la posición de *Agente* y *Meta*. Estas restricciones hacen que, por ejemplo, el sustantivo *piedra* no pueda aparecer como agente del verbo *escuchar*.

En definitiva, gracias a este analizador, cuando leemos cualquier oración, somos capaces de descomponerla en fragmentos que responden a las siguientes preguntas:

¿Quién/qué hace algo?

¿A quién/qué se lo hace?

¿Cuándo, dónde y por qué lo hace?

Aunque Bell no afirma nada al respecto, creemos que en la detección de roles semánticos entre formas léxicas que son nuevas para el lector juega un papel crucial la cohesión. Como se verá con posterioridad, los recursos cohesivos vinculan esos nuevos términos con otros relacionados cuyos roles semánticos ya ha analizado el lector, con lo que se facilita el análisis semántico.

En esta etapa se ha puesto en juego el sistema de la *transitividad*, que codifica principalmente la función *ideacional* del lenguaje y se ha obtenido el contenido

proposicional de la cláusula, pero no su fuerza ilocucionaria. De esto se encargan los mecanismos que participan en el análisis pragmático.

2.1.2.3. Análisis pragmático

Las principales funciones del procesador pragmático son las de aislar la estructura temática, detectar los recursos cohesivos que han enlazado las cláusulas para formar un texto y analizar el registro para inferir el propósito o fuerza ilocucionaria de cada acto de habla.

La estructura temática distingue entre información conocida e información nueva. Las organizaciones de la información más frecuentes (estructuras no marcadas) están almacenadas en un inventario con las estructuras sintácticas más frecuentes (*frequent structure store*). Según Bell, la estructura temática crea vínculos *dentro* de la oración, mientras que los que existen *entre* las oraciones se establecen mediante recursos cohesivos.

En cuanto al análisis de registro, supone la consideración de los tres parámetros situacionales del discurso (Halliday 1978, 1985; Hatim y Mason 1990), el *campo*, el *tenor* y el *modo*, con el fin de desentrañar la función de cada acto de habla. El campo es el asunto sobre el que tratan el emisor y el receptor. Bell utiliza el término *dominio* para referirse al campo, lo cual no nos parece adecuado porque puede confundirse con el que se utiliza en semántica para referirse a los dominios léxicos (Faber y Mairal 1999: 59) o en terminología, como sinónimo de *field*, es decir, para denominar un campo del saber⁴. El tenor es la relación social entre los participantes en un intercambio verbal, que está condicionada por el estatus y el grado de intimidad entre los mismos. Por último, el modo es la forma en que se articula el lenguaje para convenir significados y el canal mediante el que se efectúa la comunicación (oral, escrito, escrito para ser leído, etc.).

Este análisis, que se produce inconscientemente en la mente del traductor experto, proporciona al traductor la función del acto de habla que aparece en el texto (representativo, directivo, de compromiso, expresivo y declarativo⁵), la visión de la realidad presentada, qué información es nueva en el mismo, la relación entre el emisor y

⁴ Meyer y Mackintosh (1996: 4) definen *dominio* como "the area of specialized knowledge that is described for a particular terminology project, given the resources and target audience of that project".

el receptor y el modo en que se transmite del texto para que la comunicación sea más efectiva.

A partir de este momento, el traductor ya puede empezar a asignar de un modo provisional el tipo textual al que pueden pertenecer esos actos de habla. El traductor empieza a identificar los artilugios retóricos que ha utilizado el autor del texto, a partir del inventario de tipos textuales básicos que, como miembro de una comunidad de discurso, tiene almacenados en la memoria a largo plazo.

Si el traductor ha realizado las opciones correctas, tendrá almacenada en su memoria a corto plazo una representación semántica abstracta y universal del contenido transmitido por todas y cada una de las oraciones del texto origen. En la memoria a corto plazo también permanecerá hasta que se complete la traducción alguna información sintáctica para que la estructura temática no quede desvirtuada en el texto término.

2.1.2.4. Representación semántica independiente de lengua

Bell postula que esta representación semántica es el resultado de un proceso de abstracción sobre la base de los elementos que hemos mencionado con anterioridad. El problema de su aproximación es que considera la cláusula, en lugar del enunciado, como unidad de comunicación y de traducción. Por este motivo, donde Bell utiliza la palabra cláusula u oración nosotros utilizaremos el término *enunciado*. A continuación presentamos los elementos de los que se abstrae la representación semántica universal:

- a) Estructura sintáctica de los enunciados: patrones que se han elegido de entre los disponibles en el sistema de modalidad de cada lengua, por ejemplo, *Sujeto + Predicado + Objeto*, y el significado léxico de las opciones léxicas, o en su defecto, una etiqueta con información sintáctica y léxica sobre su posible significado.
- b) Contenido proposicional: opciones en el sistema de la transitividad. Como dijimos en la etapa de análisis semántico, representa las relaciones lógicas entre

⁵ Esta tipología de actos de habla es la propuesta en Searle (1979).

entidades codificadas por los enunciados. Este contenido es el único que es totalmente universal.

- c) Estructura temática y cohesiva⁶: organización de la información según esta sea nueva o conocida y progresión de unidades léxicas vinculadas semánticamente.
- d) Rasgos propios del registro: campo, tenor y modo del discurso
- e) Fuerza ilocucionaria, que, al combinarse con el contenido proposicional produce un acto de habla
- f) Acto de habla que se produce a partir de los enunciados.

Este tipo de representación semántica es esencial porque no traducimos oraciones de una lengua a otra sino textos. Al extraer de una serie de oraciones su representación semántica, podemos construir en la lengua término un texto con el mismo significado que contenga el mismo número de oraciones, o que prescinda de alguna de ellas, o que presente la información en más oraciones, según las convenciones textuales de la lengua término.

Es necesario ahora que un mecanismo evalúe si esa representación semántica se adecua al cotexto y al contexto situacional y que decida qué hacer con esa representación semántica.

2.1.2.5. Organizador de ideas

Su función es la de integrar el análisis realizado con la estructura global del texto, con la presentación tipográfica del mismo, etc. Esto supone evaluar periódicamente la información que se tiene y volver a las etapas previas para modificar la representación semántica inicial conforme aparece nueva información.

2.1.2.6. Planificador

Diseña las estrategias para obtener los objetivos que se marcaron antes de leer el texto. Este elemento decide en función de esos objetivos iniciales, si es conveniente leer el texto de nuevo, proceder a un nuevo análisis, consultar obras de referencia, o empezar a traducir directamente.

Esta fase de análisis tiene lugar en la mente del traductor y de cualquier lector monolingüe. Para poder traducir hay que pasar a la segunda gran etapa: la síntesis

2.1.3. Etapa de síntesis

El traductor comienza a construir el texto en la lengua término llevando al procesador pragmático de la lengua término la representación semántica que obtuvo tras el análisis del texto origen. Como vimos anteriormente, esa representación semántica aporta pistas sobre el tipo textual al que pertenecen esas oraciones.

2.1.3.1. Síntesis pragmática

En esta etapa, la mente del traductor decide qué elementos mantener o eliminar del texto original, de acuerdo con la naturaleza de la traducción que el cliente solicita, factor que, a pesar de su importancia, Bell nunca menciona. Al tomar estas decisiones se plantean tres problemas:

- cómo transmitir la intención del texto original
- cómo presentar la estructura temática del original
- cómo presentar las dimensiones contextuales de campo, tenor y modo del texto original en el texto meta.

Para poder llegar a una decisión final sobre cuál va a ser la función del texto término, el traductor intenta encontrar el equilibrio entre la intención del texto original, codificada mediante recursos léxicos, sintácticos y pragmáticos, y el encargo de traducción. Esta decisión será fundamental a la hora de decidir qué elementos de la lengua origen se han de reproducir en la traducción y cuáles se han de adaptar a los conocimientos previos del receptor, expectativas, necesidades comunicativas, restricciones del medio y convenciones de la cultura origen o de la cultura término.

⁶ Elemento que hemos añadido en función de la asociación que Bell (ibid: 154) establece entre ambas.

2.1.3.2. Síntesis semántica

Sabiendo la función que ha de tener el texto término, el traductor elige las proposiciones que va a transmitir en cada oración del texto término.

2.1.3.3. Síntesis sintáctica

El procesador sintáctico tiene que transformar las proposiciones de la fase anterior (síntesis semántica) y, aunque no lo especifique Bell, tiene que transformar el significado pragmático de la etapa de síntesis pragmática, en estructuras sintácticas o cláusulas. Para ello, se escanean las bases de datos de lexemas y estructuras sintácticas más frecuentes en busca de las unidades léxicas y las estructuras apropiadas. Si no se encuentran, las proposiciones pasan por el mecanismo de búsqueda léxica o por el analizador sintáctico (*parser*), que también puede funcionar como sintetizador sintáctico.

2.1.3.4. Escritura

Es la última etapa que permite la representación del mensaje del texto origen mediante los caracteres tipográficos de la lengua término.

Como conclusión podemos decir que la traducción es un proceso cognitivo interactivo en el que se parte de una actividad fisiológica, el reconocimiento de caracteres escritos en la lengua origen, se continúa con una serie de actividades psicológicas implicadas en la lectura y comprensión y se culmina con las actividades psicológicas y fisiológicas que posibilitan la escritura del texto término. En los procesos cognitivos, la memoria a corto plazo procesa casi simultáneamente información de índole sintáctica, semántica y pragmática en dos lenguas. También interviene en ese proceso la memoria a largo plazo, que aporta bases de datos con información sobre el mundo, cada una de las lenguas y de las culturas implicadas en el proceso de la traducción.

No obstante, es preciso señalar algunos puntos débiles del modelo de Bell. En primer lugar, este modelo asume que los procesos de comprensión y producción de mensajes son simétricos. Sin embargo, cualquier traductor puede recordar alguna

ocasión en la que, a pesar de comprender un texto perfectamente, le resultó muy difícil traducirlo a la lengua término. Desde la psicolingüística, algunas voces (López García 1991) han criticado modelos de este tipo argumentando que la comprensión y la producción no son fenómenos equivalentes y contrarios, aunque sin duda, están estrechamente relacionados.

Por otra parte, nos ha sorprendido que, en la etapa final, Bell haya olvidado el significado pragmático y consideraciones en torno a los tipos textuales que cumplen mejor un determinado fin. Quizá esto se deba a su pretensión de formular un modelo simplificado y a que se sobreentiende que, por ser interactivo, el proceso de la traducción tiene en cuenta todos los aspectos presentados de una forma simultánea. También se puede deber a que el proceso de traducción, según Bell, opera al nivel de la oración. En esto no coincidimos con él y postulamos que la memoria de trabajo puede procesar más que una oración a la vez y, en consecuencia, en traducción se opera con enunciados (Faber 1998a; Givón 1997, 1998a y 1998b).

Asimismo, Bell menciona la cohesión aunque no en las explicaciones del proceso de análisis y síntesis. Como se verá, la cohesión juega un papel crucial en la traducción, algo que Bell no ha reflejado.

Por último, es evidente que todos los mecanismos descritos en el apartado anterior intervienen en el procesamiento de textos y en el reconocimiento y presentación de los aspectos culturales y textuales que consideramos esenciales en una tipología de textos aplicada a la traducción. No obstante, algunos mecanismos juegan un papel más importante como el sistema de reconocimiento de caracteres, que detecta aspectos tipográficos propios de determinados tipos textuales, los mecanismos que se ocupan del léxico y del contenido proposicional, el analizador y el sintetizador pragmático, el organizador de ideas y el planificador. Junto a estos, estamos convencidos de que existe en la memoria a largo plazo una base de datos que recoge la organización de tipos textuales propios de la lengua origen y la lengua término. Esta base de datos va a guiar las decisiones concernientes a la elección de tipos textuales.

2.2. Tipologías textuales y traducción

Tradicionalmente, los estudios de traducción se han caracterizado por la parquedad de investigaciones sobre el texto. De hecho, hubieron de esperar al desarrollo de la lingüística textual (Van Dijk 1972; Petöfi y Rieser 1973; Werlich 1976; Dressler 1978) para empezar a considerar como unidad de traducción y de equivalencia al texto. A partir de ese momento, la colaboración entre la traducción, la lingüística, y una serie de disciplinas derivadas de esta ha producido aproximaciones interesantes a la hora de identificar *tipos textuales*.

Estas aproximaciones, que revisaremos en esta sección, hubieran alcanzado un mayor grado de consistencia en cuanto al uso de términos (*género, tipo textual, registro, discurso*) y de concreción si la colaboración interdisciplinar hubiera sido mayor. De hecho, el concepto de *género*, que ha suscitado en la última década un creciente interés en áreas distintas a la literatura, se ha introducido de un modo muy débil en el área de los estudios de traducción (Trosborg 1997: vii). Esto es sorprendente porque, como apunta Trosborg (ibid: viii), “text typology involving genre analysis can help the translator develop strategies that facilitate his/her work and provide awareness of various options as well as constraints”.

2.2.1. Del concepto de *equivalencia* a la elaboración de taxonomías textuales

La importancia de estudiar el lenguaje desde el fenómeno de la variación lingüística ya fue señalada por Catford (1965: 83)

The concept of a ‘whole language’ is so vast and heterogeneous that it is not operationally useful for many linguistic purposes, descriptive, comparative, and pedagogical. It is therefore desirable to have a framework of categories for the classification of ‘sub-languages’ or varieties within a total language.

Este mismo autor también hablaba de la dicotomía que existe en traducción entre equivalencia textual y correspondencia formal. Su aproximación es bastante ilustrativa del metalenguaje articulado en torno a dicotomías que impregnaba la lingüística y la traducción, que empezaba a delimitarse como disciplina autónoma.

En el caso del metalenguaje de la traducción, el término *equivalencia* era, y aún es, demasiado frecuente. De hecho, desde que Jakobson (1959) lo utilizara, la equivalencia fue considerada como la panacea que iba a permitir la traducción de cualquier texto. Sin embargo, se empezó a percibir que una teoría de la traducción basada en la equivalencia producía textos demasiado centrados en la lengua origen que no se amoldaban a la cultura de la lengua término (Toury 1980). Frente a la tradicional dicotomía entre *source-text oriented translation* y *target-text oriented translation*, había una clara preferencia por la primera opción.

Esta aproximación favorecía que se produjera una cierta incompatibilidad entre el texto traducido y el cuerpo de textos que existen en la cultura término, algo así como el rechazo que puede experimentar un receptor de órganos cuando le realizan un trasplante. Neubert (1981: 133) lo expresa acertadamente con la siguiente metáfora: “Translating texts, then, is equivalent to transplanting an ‘organ’ from a whole ‘body of texts’ into another ‘body of texts’ ”.

Katharina Reiss (1971) también defiende la idea de equivalencia, entendida como la permanencia de la función del texto origen (TO) en la cultura del texto meta (TM), esto es, conseguir que el TO siga perteneciendo al mismo tipo textual en la lengua de llegada. A partir de una concepción de la equivalencia textual y funcional se empieza a reconocer la importancia de los tipos textuales. De todas formas, señala que existen dos contextos en los que el traductor puede prescindir de la equivalencia. Uno es cuando el texto término tiene una función distinta a la del texto origen. El otro caso es cuando la audiencia del TM es distinta de la del TO.

Del escepticismo con el que hay que enfrentarse a la equivalencia entendida como algo estático que hay que mantener a toda costa hay suficientes muestras en Rabadán (1991). En este libro se nos presenta la falta de validez de una teoría de la traducción basada exclusivamente en la equivalencia y el excesivo desgaste que ha sufrido el término. Estos dos factores han contribuido a un desplazamiento del texto origen, en cuanto a determinante de la traducción, hacia el texto meta, su función y la situación comunicativa en que cobra existencia.

En suma, el concepto de equivalencia perdió vigor frente al de función del texto meta o *skopos* (Vermeer 1978). Y así, desde el reconocimiento de que la equivalencia

no es posible en infinidad de casos, se abrieron las puertas al estudio de la tipología de textos. Por este motivo, en muchas corrientes traductológicas contemporáneas, el término equivalencia ha sido sustituido por otros términos como la “fidelidad y lealtad” (*fidelity and loyalty*) de Nord (1991) o la *equivalencia pragmática* de Baker (1992). Algunos estudiosos como Wilss (1995: 35) apuntan a la vaguedad del término y sostienen que la traducción intenta alcanzar “some sort of equivalence”.

Los estudios de tipología de textos aplicados a la traducción aparecen en la década de los setenta y están influenciados en gran medida por las ideas de la *lingüística textual alemana* (Beaugrande 1980; Beaugrande y Dressler 1981) y en la misma línea Van Dijk (1972), que, a partir del primer simposio sobre lingüística del texto celebrado en Constanza (Petöfi y Rieser 1973), se ocupan del estudio lingüístico del discurso de un modo sistemático. Este estudio culmina en una antología de estudios del discurso (Dressler 1977). Además de esta influencia, también es muy destacada la de la *lingüística anglosajona* que derivó de los postulados de Firth y cuyo más destacado representante es M.A.K. Halliday.

Firth (1951, 1968) apunta hacia una concepción del texto como un elemento inmerso en el contexto de situación. Firth, siguiendo al antropólogo Malinowski, sostiene que el significado es la razón de ser de la lingüística, entendiendo el significado como lo que se pretende conseguir mediante los enunciados. Asimismo, según Firth, el significado ha de ser contemplado desde la función del lenguaje en contexto, ya sea el contexto situacional o el contexto cultural.

Halliday, en la misma línea de Firth, también hace una distinción entre *contexto situacional* y *contexto cultural*. Dentro del contexto de la cultura, Halliday (1973: 481) define la lengua como un conjunto indefinido de opciones de comportamiento disponibles para el hablante. El entorno en el que se selecciona una de esas opciones de entre todas las disponibles en el contexto cultural se denomina contexto situacional (ibid: 71). Es decir, el contexto cultural define el potencial de la lengua, mientras que el contexto situacional define la realización.

En la propuesta *sistémico-funcional* de Halliday, el significado aparece como paso intermedio entre el *poder hacer* y el *poder decir*. El potencial de significado es una serie de opciones a disposición del hablante-oyente que hacen posible codificar

mediante el lenguaje lo que no es lenguaje. Este potencial comprende una serie de variaciones que tienen significado para los hablantes que tienen competencia comunicativa (Hymes 1972). La competencia comunicativa es un mecanismo que indica a los hablantes cuándo tienen que hablar y cuándo tienen que permanecer en silencio, qué código utilizar, dónde y con quién.

La idea de que existen variaciones en el lenguaje lleva a Halliday, McIntosh y Strevens (1964: 87) a formular la teoría del registro:

Language varies as its function varies: it differs in different situations. The name given to a variety of a language distinguished according to use is 'register' [...] It is by their formal properties that registers are defined.

Estos autores reconocen dos dimensiones de variación lingüística. Una de ellas es la variación según el hablante, a la que denominan *dialecto*. Entre los tipos de dialecto, mencionan el dialecto geográfico, temporal, social, estándar y el idiolecto. La otra dimensión es la variación según el uso en una situación determinada, a la que denominan *registro*. La categoría situación no designa simplemente a una serie de circunstancias sino también la convención según la cual un enunciado es apropiado para un cierto uso. Halliday (1978) distingue tres tipos de variación según el registro: el campo, el tenor y el modo.

Halliday (1978) considera que la característica esencial de un texto es la *interacción*. Cada texto tiene un significado social en un *contexto situacional* determinado, lo que implica un proceso continuo de selección y actualización semántica. Por este motivo, para poder definir un texto, es necesario caracterizar el contexto situacional de forma que sea posible sistematizar las relaciones y dependencias entre la realización social y el entorno social. El contexto situacional se manifiesta en el texto en tres niveles: el *campo*, que corresponde a la función ideacional; el *tenor*, relacionado con la función interpersonal; y el *modo*, que actualiza la función textual. La combinación de estos tres factores determina el *registro*. El registro se convierte entonces en el puente entre el texto y el entorno socio-semiótico (ibid: 145).

Tanto la lingüística textual alemana como la lingüística derivada de Firth, surgen en un terreno que ya estaba abonado por las aportaciones que se habían producido en áreas afines a la lingüística:

- a) la etnografía (Malinowski 1923, Hymes 1972)
- b) la filosofía del lenguaje (Austin 1962; Searle 1969; Grice 1975)
- c) la sociolingüística (Labov 1972a, 1972b)
- d) la inteligencia artificial y la psicolingüística (Schank y Abelson 1977, Minsky 1975, Rosch 1975)

Las investigaciones de Hatim y Mason (1990, 1997) (véase 2.2.7.) y Mayoral Asensio (1997) aplican el enfoque sistémico de Halliday a la traducción.

2.2.2. Las taxonomías textuales y el comienzo de las aproximaciones funcionales

El desarrollo de las nuevas teorías textuales en la década de los setenta se refleja en la aparición de una serie de taxonomías textuales enfocadas al estudio de la traducción, en parte, debido a que, desde la lingüística ya se habían elaborado *text typologies* (Gülich y Raible 1972; Petöfi 1979). Estas tipologías, al igual que se hacía en el análisis componencial de lexemas de distintas lenguas, pretenden aislar tipos discursivos que indiquen variaciones específicas a una lengua. Después, estos tienen que aplicarse de una forma contrastiva tanto a los textos originales como a los textos meta y a textos paralelos. Este planteamiento será el germen de la *retórica de contrastes* (Connor 1995, 1996; Connor y Kaplan 1987) y la *textología contrastiva* (Hartmann 1980, 1996). La primera identifica las diferencias lingüísticas y retóricas entre textos escritos en la lengua materna y en una segunda lengua, mientras que la segunda propone una integración de la lingüística de contrastes, el análisis de discurso y la lingüística de corpus en el estudio de textos paralelos en más de una lengua.

Desafortunadamente, las aproximaciones iniciales cometen el error de utilizar la estructura de la oración para explicar la estructura del texto y de adoptar una postura prescriptiva, incapaz de explicar textos que se producen en la *actuación* de los hablantes. Esta inoperatividad práctica hace que se empiece a cuestionar la validez de la

aproximación lingüística (en el ámbito del *sistema*) y que se tome el texto como sistema de referencia. De la interlingüística se produce un salto al marco de la intertextualidad. A partir de este momento, los textos no se considerarán como una cadena de oraciones separadas que contienen una serie de elementos gramaticales y léxicos, sino como una estructura compleja y multidimensional que consiste en algo más que la mera suma de las oraciones que lo componen.

Los estudios de traducción centrados en la tipología de textos, que, como se ha mencionado previamente, derivan del desarrollo de la lingüística textual, se remontan a Reiss (1968) y sostienen que la unidad de traducción es el texto.

2.2.2.1. La aproximación funcional de Katharina Reiss

Reiss (1968) clasifica los textos de acuerdo con las funciones del lenguaje de Bühler (1934): la representativa, la expresiva y la apelativa. En Reiss (1971) estos tipos textuales reciben el siguiente nombre⁷:

- a) *Inhaltsbetonte texte*: Textos en lo que lo importante es el contenido, por ejemplo, los textos científicos y técnicos.
- b) *Formbetonte texte*: Textos en los que prima la forma, que serían los textos literarios.
- c) *Appellbetonte texte*: Textos que cumplen una función conativa, como los anuncios.

TABLA 1: Funciones del lenguaje y tipos textuales (Reiss 1971: 33)

Función del lenguaje	Darstellung	Ausdruck	Appell
Dimensión del lenguaje	Logisch	Ästhetisch	Dialogisch
Textyp	Inhaltsbetont	Formbetont	Appellbetont

⁷ *Apud* Nord (1996a: 82-83).

Reiss incluye estas tres categorías dentro de lo que se denomina en alemán *texttyp*, una distinción funcional diferente a la indicada por otro término relacionado con el concepto de *tipo textual: textsorte*. *Textsorte* hace referencia a los textos que se producen en situaciones estereotipadas: manuales, cartas comerciales, contratos, oraciones, etc. Esta categoría entronca con la de *género*, que surgió dentro de la literatura.

Una vez que se ha reconocido el tipo textual, Reiss sostiene que el traductor debe identificar las *instrucciones semánticas, léxicas, gramaticales y estilísticas* (Nord 1996a: 84) asociadas a cada tipo textual, lo cual puede proporcionar criterios objetivos para la traducción.

Posteriormente, Reiss (1976: 8) y Reiss y Vermeer (1984: 204) modifican la terminología a textos informativos, expresivos y operativos, con lo cual cambian el énfasis, de las funciones del lenguaje, a las funciones del texto.

Reiss (1977) añade una nueva categoría, los *audiomediale texte* (textos audio-media), para dar cuenta de aquellos tipos textuales que han sido escritos para ser hablados o cantados como obras de teatro, óperas, etc. Estos textos dependen de medios no lingüísticos para materializarse. Reiss (1980) modifica esta categoría para incluir textos como los cómics, que contienen elementos visuales en lugar de elementos acústicos y sustituye el término de textos *audiomedia* por textos *multimedia*. Al reconocer este nuevo tipo textual, incluye el parámetro de *medio* (Crystal y Davy 1969: 65) en su taxonomía textual y enriquece sobremanera su tipología textual.

No obstante, como indican Nord (1996a: 83) y Snell-Hornby (1997), para Reiss los textos multimedia no constituyen una categoría aparte, sino que son una superestructura que puede englobar a los otros tres. En otras palabras, un texto multimedia puede ser informativo, expresivo u operativo.

Según Rabadán (1991: 70), no tardan en aparecer críticas a Reiss (1976) porque su tipología presupone que en un texto existe una única constante funcional abstracta (House 1977). Las críticas se basan en una lectura superficial de Reiss, que reconoce que pueden existir tipos textuales mixtos. Podemos intuir esto en las categorías iniciales de textos: textos *centrados* en la forma, *centrados* en el contenido y *centrados* en la función. Reiss (1979: 50) escribe:

[...] il existe également des types mixtes, dans lesquels les fonctions de communication découlant des trois situations de base (la transmission de connaissances, l'expression artistique originale et le déclenchement de stimuli comportementaux) se présentent alternativement ou simultanément.

Es decir, Reiss no espera a 1981, para reconocer que existen formas mixtas y que la función del texto origen no siempre se mantiene en el texto meta. De todos modos, es verdad que en 1981, su formulación de los tipos textuales denominados *multimedia* es mucho más concreta.

Su postura inicial se relativiza al afirmar que, cuando el tipo textual y su función son diferentes en la lengua meta, la tipología previamente establecida no es relevante para determinar el tipo de equivalencia y es preciso tomar como criterio dominante otros factores extralingüísticos (*pragmatische Kategorie*) como la referencia temporal, referencias al mundo real, relaciones afectivas, etc. para obtener un texto apropiado a las exigencias del *polisistema* meta (Toury 1980). Y así en Reiss y Vermeer (1984: 204 y ss.), sólo existe una relación entre tipo textual y estrategia de traducción cuando existe una equivalencia funcional entre texto origen y texto meta.

2.2.2.2. El enfoque retórico de Werlich

Casi simultáneamente a las aportaciones de Reiss, Werlich (1976: 21) define los tipos textuales como “idealized norms of distinctive text structuring” y propone una tipología con las categorías de *descripción, narración, exposición, argumentación e instrucción*. Lo interesante es que este autor aspira a que su propuesta tenga una adecuación psicológica (*psychological adequacy*) que justifique la existencia de esos tipos textuales. Es decir, cada tipo textual está asociado a un proceso cognitivo concreto, lo cual queda expresado claramente cuando Werlich (1976: 21) afirma:

Texts, conceived of as assignable to *text types* primarily derive their structural distinctions (text structuring) from innate cognitive properties. Accordingly the five text types correlate with *forms and ranges of human cognition*. They reflect the basic cognitive processes of contextual organization.

Y así, en la *descripción* se presentan entidades, estados, acontecimientos y relaciones en el espacio. En la *narración*, estos se presentan en el tiempo. En la *exposición*, el emisor los organiza mediante la descomposición (análisis) o la composición (síntesis) de sus elementos. En la *argumentación*, se evalúan las relaciones entre conceptos mediante la identificación de similitudes y diferencias. Por último, en la *instrucción* se planifican los comportamientos futuros del lector.

Esta tipología es importante porque recoge la idea de propósito retórico, que Werlich (1976: 19) denomina *foco contextual dominante*. Asimismo, esta tipología ha sido adoptada por Zydatis (1983) y Hatim y Mason (1990), aunque estos últimos fusionan los tres primeros tipos para formar una sola categoría, la de los textos *expositivos*.

2.2.2.3. Las dimensiones situacionales de Juliane House

Otra aproximación con un mayor grado de concreción es la de House (1977: 38 y ss.), que se basa en las *dimensions of situational constraint* de Crystal y Davy (1969: 66). Al igual que Crystal y Davy, House sostiene que los tipos textuales están definidos por una serie de dimensiones situacionales. Éstas se dividen en dos: la del usuario (origen geográfico, clase social, época) y la del uso lingüístico (medio, participación, relaciones emisor-receptor, el estilo, etc.). Teniendo en cuenta en todo momento el horizonte de la traducción, House añade una nueva categoría que indica el grado de confianza que hay entre los participantes y la denomina *social attitude*. Con esta categoría, House pretende determinar el conocimiento que comparten los participantes y así inferir la información que se ha de hacer explícita en el texto término.

En estas dimensiones el material lingüístico se organiza en tres niveles: sintáctico, léxico y textual. El análisis situacional y lingüístico conduce a un perfil textual del TO que funciona como norma y criterio básico de la equivalencia. Según las características de este perfil se aplicará uno de los dos tipos de traducción que propone: a) *overt*, en la que el TO está vinculado de una forma específica a los destinatarios de la lengua origen; b) *covert*, en la que el texto meta adapta la función textual del TO al polisistema meta (House 1977:194). Estos tipos de traducción adquieren un nuevo

nombre con Newmark (1981, 1988), *communicative translation* y *semantic translation*; y con Sager (1997), *dependent translations* y *derived translation*.

Al señalar la existencia de *overt translations*, plantea que existen límites a la hora de traducir que no son meramente lingüísticos sino de origen cultural, lo cual no es una novedad con respecto a Reiss (1971), que también reconoce que existen traducciones en las que la equivalencia no es posible. Asimismo, sostiene que, para alcanzar la equivalencia textual en este tipo de traducciones, se ha de proceder al análisis del TO, que contiene los elementos que determinan su función o “the application or use which the text has in the particular context of a situation” (House 1977:37). No obstante, aquí nunca especifica cómo se organizan los rasgos que definen el perfil textual del TO.

House (1977), siguiendo a Halliday, clasifica los textos en dos *macrofunciones*, la ideacional y la interpersonal y las subdivide en categorías *nocionales: técnico/no técnico, ficcional/ no ficcional*. El problema de esta clasificación es que obvia el hecho de que en todos los tipos textuales, ambas macrofunciones están presentes y que no concede espacio para la traducción literaria, en la que la función estética predomina más que las otras dos.

Por último, podemos decir que la principal aportación de House se encuentra en su incorporación de parámetros situacionales y en el reconocimiento de las relaciones emisor-receptor, que más tarde cristalizarán en su formulación de la interacción comunicativa en traducción como *marco de negociabilidad* (House 1986).

2.2.2.4. La primacía del texto término: La teoría del *skopos*

Vermeer (1978) y Reiss y Vermeer (1984) son los artífices de la *Skopostheorie*, en la que sostienen que la traducción es una forma de interacción humana y, como tal, está determinada por su función o *skopos*. La función del texto está dispuesta en gran medida por el destinatario de un texto ya sea el TO o el TT. De ahí que tanto el TO como el TT estén conformados a las convenciones vigentes de la cultura origen y de la cultura término, respectivamente. De todas formas, el destinatario que determina las decisiones que han de tomarse a la hora de traducir es en última instancia el lector de la traducción.

Al igual que el modelo de equivalencia, la teoría del *skopos* es un modelo general que reconoce que puede existir una relación de analogía o similitud entre TO y TT, que Vermeer (1978: 59) denomina *intertextual coherence* o *fidelity* y que, en todo caso, los procedimientos traductológicos están subordinados al *skopos* o función del texto traducido en la cultura de llegada. La teoría del *skopos* subyace las propuestas iniciales de Nord (1987).

2.2.3. La relativización de las dicotomías. Hacia un enfoque interdisciplinar de la traducción

2.2.3.1. La aproximación cognitivo-textual de Albrecht Neubert

También en la línea de la lingüística del texto encontramos la propuesta de Neubert, uno de los representantes de la Escuela de Leipzig. Neubert enarbola con firmeza la idea de que la lingüística del texto y el análisis del discurso pueden ser muy útiles para la traducción. Esto no es de extrañar, teniendo en cuenta su afirmación de que la unidad de traducción es el texto. Su aproximación es multidisciplinar e intenta integrar posiciones cognitivas y pragmáticas. De hecho, sostiene que la Lingüística del Texto *explora la sutil relación entre la cognición y el discurso*. (Neubert, 1983: 132). En otras palabras, la Lingüística del Texto se encarga de ver el efecto comunicativo de un texto en función de la intención del autor y las expectativas del lector.

La necesidad de echar mano a otras disciplinas adyacentes como la sociología (Beaugrande 1980), psicología (Bobrow y Collins 1975) y la inteligencia artificial (Kintsch 1977; Schank y Abelson 1977) radica en el hecho de que los textos son unidades superiores a la oración que no se construyen simplemente con unidades lingüísticas sino que son “explicit models of the world and our ways of coping with it” (Neubert 1981: 133).

Los textos son el resultado de un proceso *top-down* por el que, tras un proceso de decisión continua entre distintas opciones formadas por secuencias textuales, se reconstruye el texto origen bajo las condiciones textuales de la audiencia término. En palabras de Neubert (1983: 161):

[...] the translation process is a holistic activity always keeping the whole textual product in mind when making each individual choice of words and structures.

Esta última cita es ilustrativa de cómo Neubert relativiza la importancia del texto origen y, de esta forma, se encuentra en la línea de la recién formulada *teoría del skopos*. En efecto, el lector de una traducción no lee el texto que escribió su autor sino un producto secundario que se construye a partir de un texto (*text-induced text*). Otra coincidencia con la teoría del *skopos* se encuentra en su opinión de que, en traducción, el aspecto más decisivo es la finalidad de la traducción (Neubert, 1981: 135). Para que la traducción sea efectiva tiene que provocar una sensación de inmediatez. Esto conlleva que no exista servidumbre respecto al texto origen para que los receptores de los textos traducidos los perciban como *codificaciones naturales de la experiencia* (Neubert 1981: 131).

En cuanto a la tipología de textos, su primera clasificación de los textos es bastante modesta. Neubert (1968) establece una clasificación tripartita según el grado de traducibilidad potencial de los elementos lingüísticos del TO, esto es, según la mayor o menor correspondencia sintáctica, semántica y pragmática: textos de traducibilidad relativa, parcial y óptima. De ahí que decida abandonar su intento de establecer taxonomías textuales que no tengan una base psicológica y comience a investigar el proceso cognitivo de la traducción, un interés presente en sus últimas publicaciones (Neubert 1997). También abandona cualquier intento de establecer una total equivalencia entre texto origen y el texto término, ya que ambos son diferentes textualmente. El texto origen está enraizado en una determinada comunidad lingüística desde el punto de vista sintáctico, léxico, estilístico y contextual.

Para minimizar la incompatibilidad entre el texto traducido y el cuerpo de textos que existe en la cultura término, toma como punto de partida la noción de *macroestructura* y de *superestructura* de Van Dijk (1977, 1981). Las macroestructuras representan la organización semántica global del discurso, el tema o el asunto del discurso. Las superestructuras constituyen la forma global del discurso que define la organización general del mismo y las relaciones jerárquicas de sus respectivos fragmentos. Tanto las macroestructuras como las superestructuras subyacen los textos y

por eso, conforme el lector lee varios textos con una superestructura y una macroestructura similar, este comienza a asociarlas a determinados tipos textuales y las almacena en la memoria.

En consecuencia, los tipos textuales disponibles en una lengua los tiene interiorizados el hablante como un paradigma con opciones que permite, por una parte, la activación de un texto concreto y, por otra parte, la distinción entre este y otros tipos textuales. Todos estos tipos textuales constituyen un conjunto llamado *background texts*, que son *una realidad objetiva de la experiencia textual previa de los hablantes que tiene una base psicológica* (Neubert 1981: 134).

Los *background texts* son tipos discursivos recurrentes que no son estructuras fijas, sino *modos* de organizar el material textual que están asociados a una determinada actuación textual. Según Neubert (ibid: 134), los textos pertenecen a determinados paradigmas textuales: “One text shares a textual paradigm. It is produced and understood in terms of this larger paradigm of discourse”.

También se muestra en la línea de lo que se propugna desde otras áreas, como la lingüística de corpus y la textología contrastiva. Neubert habla de la necesidad de estudiar distintos tipos de texto y sus repertorios gramaticales y léxicos tanto en la lengua origen como en la lengua término con el fin de detectar cuantitativamente preferencias en el nivel macroestructural, morfosintáctico y léxico de una lengua. Esto es, Neubert opina que, aunque existen una serie de opciones macroestructurales, morfosintácticas y léxicas abiertas al hablante de una lengua para codificar un tipo concreto de texto, existen opciones que son preferentes y que se pueden determinar cuantitativamente. Neubert (1981: 134) escribe al respecto:

To a certain extent macrostructures show quantitative dominances, i.e. certain words or constructions occur more often than in other contexts which makes the clear-cut delineation of the linguistic setup of certain textual macrostructures even harder.

Neubert también identifica elementos variables e invariables dentro de una traducción. Las superestructuras son invariables. Y así, por ejemplo, un libro de instrucciones, al ser traducido, tiene que seguir siendo un libro de instrucciones. Por eso afirma que *la mediación nunca debe producir una mezcla de tipos discursivos* (ibid:

139). Esta opinión no parece confirmarse en la traducción de nuevos tipos textuales que han aparecido en la sociedad moderna, como, por ejemplo, los resúmenes de cartas comerciales y los *abstracts* de artículos científicos.

También puede haber invarianza en el nivel macroestructural, aunque esto, a veces supone ordenar de nuevo el contenido semántico en la lengua término. De aquí se deduce que el gran problema en traducción es localizar y explicar la estructura textual del texto origen y los cambios que ha de sufrir en el texto término. Y en eso, según Neubert, la lingüística del texto arroja luz sobre este proceso por su análisis de cómo las proposiciones están internamente estructuradas y cómo se combinan para formar macroproposiciones relacionadas. Estas relaciones semánticas se pueden objetivar analizando los lexemas actualizados en un texto, que son los que nos llevan a detectar la superestructura y la macroestructura del texto. De ahí se deduce que en la comprensión de un texto intervienen tanto procesos *bottom-up* como *top-down*. Como se verá en los capítulos 10 y 11, estos lexemas semánticamente relacionados son los responsables de la cohesión del texto.

Vamos a ilustrar la interacción entre los lexemas activados y la superestructura y macroestructura mediante un ejemplo. Al leer la primera página de un artículo de la revista *Cancer* (Véase Texto 1 del Apéndice II), la presentación tipográfica y los lexemas contenidos en el título (*Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary Chromosome 9 Aberrations*) y en los títulos de las secciones (*Abstract, Background, Methods, Results, Conclusion, Keywords*) hacen que el lector detecte que ese texto se amolda a la *superestructura de un artículo biomédico*, que como dijimos sigue el formato *IMRAD* y a la macroestructura de textos sobre *Pronóstico del cáncer de pulmón*. Este proceso *bottom-up* es seguido por la activación en la memoria a largo plazo de las macroestructuras normalmente asociadas a cada uno de los elementos de la superestructura, por ejemplo, la organización semántica de la sección *Methods*, que gira en torno a los conceptos *PATIENTS, HYPOTHESIS, MATERIALS* y *METHODS*. De esta forma, los esquemas mentales crean unas expectativas que ayudan a los lectores a interpretar los lexemas de estos textos (proceso *top-down*). Y de nuevo, estos lexemas, a su vez, construyen la macroestructura concreta del texto (proceso *bottom-up*).

Desde la convicción de que tanto el texto origen como el texto término comparten una totalidad semántica (*semantic whole*), que permanece invariable, se puede llegar a alcanzar la equivalencia textual. Neubert (ibid: 142) escribe:

Textual equivalence is achieved through precisely those textual elements which play the same role in the target text that receivers experience as the 'respective' ones in the source language.

En relación con esa totalidad semántica encontramos el concepto de *intertextualidad*, es decir, el hecho de que relacionemos unos textos con otros. La intertextualidad permite que hablantes de distintas lenguas *sintonicen* para compartir una experiencia codificada en las lenguas origen y término, gracias a la labor del traductor o el intérprete, que realiza su cometido “in a receiver-based way” (Beaugrande 1980: 292).

En 1985, según Rabadán (1991: 183-184), Neubert pone el énfasis en los textos como nexos entre la *lengua-sistema* y la *lengua-instrumento de interacción social*. Como tales los textos no pueden ser analizados según las unidades *cerradas* de la lingüística oracional. Es necesario pues, ampliar el campo abstracto hasta poder establecer unidades más ajustadas a la realidad de la comunicación y al contexto social: los textos. Su definición de los textos es la siguiente: “socially effective, efficient, and appropriate moulds into which the linguistic material available in the system of a language is recast” (Neubert 1985: 125).

Para que una traducción sea comunicativa, tiene que tener como finalidad proporcionar a la audiencia del polo meta un texto que pertenece a una comunidad cuyos hábitos textuales son distintos. Esto se consigue mediante la incorporación de la traducción al espectro textual propio del polisistema meta. Parece como si la equivalencia textual se pudiera alcanzar siempre.

Si a esto añadimos que la producción textual depende de condiciones sociales, que son variables, también lo serán los tipos textuales. De aquí se deriva una nueva concepción de las tipologías aún más relativa que la del *background text*, la noción de *texto prototipo*. Neubert (1985: 127) lo define como el modo de organizar el

conocimiento en el discurso oral y escrito que está condicionado histórica y socialmente.

Con esta definición se hace eco de la teoría semántica de los prototipos (Rosch 1975, 1977)⁸, aunque la aplica a la *categorización* de textos, en lugar de a la categorización de la realidad. También permite dar cuenta de las dimensiones histórica (las normas) y social (los parámetros sociolingüísticos) de los textos. Esta aplicación de la teoría de prototipos a los textos no es sorprendente si tenemos en cuenta que Neubert (1991: 19) concede una gran importancia al significado:

[...] behind the translation process, whether macro- or microcontextual, lies the problem of the genuine reconstruction of utterance meanings.

En suma, podemos decir que Neubert incide sobre la necesidad de integrar una postura interdisciplinar que combine estructuras textuales y significado discursivo, macroestructuras textuales y macroproposiciones textuales, y que sea flexible para dar cabida a la variedad de tipos textuales que responden a una situación histórica y social determinada.

La postura de Neubert marca la frontera entre una serie de aproximaciones en que las tipologías establecen compartimentos estancos que no responden a la realidad. Estas aproximaciones tienen una orientación normativa en la que las clasificaciones se establecen *a priori*, sobre el sistema textual de un solo polo (suponemos que el del TO) y que no tiene por qué ser equiparable a la situación en el polo meta. Asimismo, al igual que se hacía en el análisis componencial de lexemas, pretenden aislar tipos discursivos que indiquen las variaciones específicas de una lengua. Sin embargo, en la práctica, los tipos textuales no se dan de una forma tan clara y homogénea como pretenden los que elaboran estas taxonomías.

⁸ La teoría de prototipos sostiene que el significado de las palabras no tiene unos límites claros y que los hablantes de una lengua no encuentran ninguna dificultad a la hora de decidir si una serie de entidades pueden incluirse dentro de la categoría designada por una palabra. Para demostrar esta teoría, Eleanor Rosch diseñó un experimento en el que presentaba a hablantes nativos de inglés americano una lista de palabras que hacían referencia a una determinada categoría semántica y les pedía que eligieran qué palabra era la que mejor representaba esa categoría (*prototipo*) y que ordenaran el resto de las palabras en relación a este. Por ejemplo, ante una lista de nombres de pájaro (*sparrow, robin, penguin, peacock, etc.*), los hablantes señalaron que *robin*—tordo, en los Estados Unidos—era un pájaro más prototípico que el resto.

En consecuencia, hay que clasificar los textos conforme a funciones o dimensiones situacionales, y no adscribirlos a una categoría establecida de antemano. También, es conveniente, como hace Snell-Hornby (1988), hablar de *prototipologías*.

2.2.3.2. La aproximación integradora de Mary Snell-Hornby

Este escepticismo hacia modelos de análisis basados en dicotomías y en taxonomías textuales de carácter estático es lo que lleva a Snell-Hornby (1988, 1997) a plantear una prototipología de textos y a adoptar una perspectiva ecléctica.

Como en otros muchos estudios centrados en los textos, Snell-Hornby señala lo inadecuado del concepto de equivalencia y propone formas de categorización que se plantean en términos del principio de *gestalt* y el de *prototipo*. Su metodología es integradora en el sentido de que intenta aunar aspectos textuales, culturales y cognitivos de la traducción. Esta perspectiva nos puede resultar útil a la hora de proponer tipos de texto.

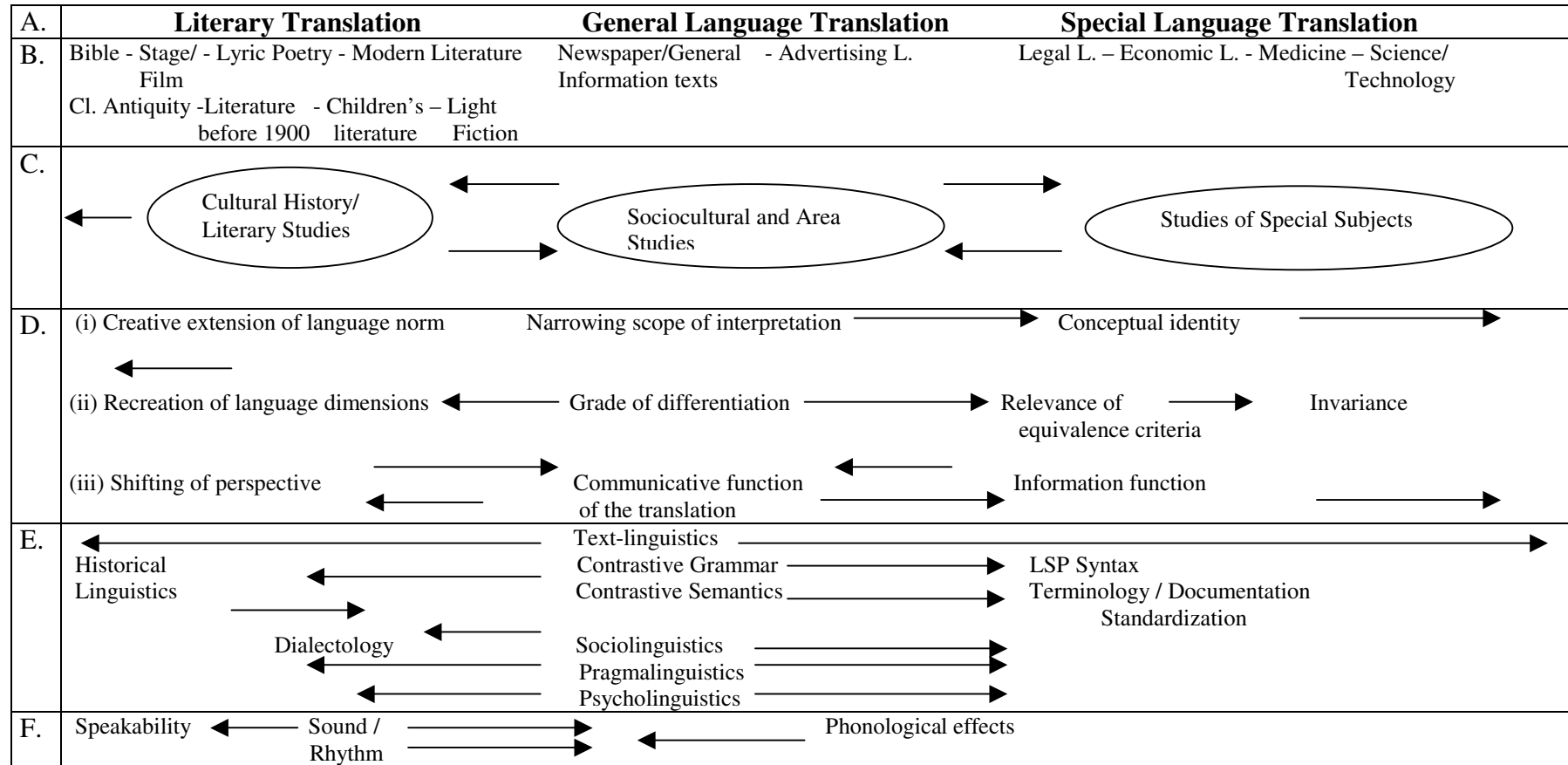
La propuesta de tipos de texto de Snell-Hornby es una aplicación del principio de *gestalt* (Wertheimer 1959), que sostiene que la totalidad se compone de algo más que la mera suma de sus partes y del concepto de prototipo (Rosch 1977, 1979), que tan fructífero ha resultado para explicar el proceso cognitivo de la categorización (Lakoff 1987). Estos dos principios se amoldan más al hecho de que la mayoría de los textos son formas híbridas, estructuras multidimensionales con una mezcla de características. Por este motivo sugiere que el concepto de prototipología es más apropiado que el de tipología, que adolece de una excesiva prescripción y un abuso de compartimentos estanco. El concepto de prototipología es dinámico y no pretende hacer distinciones entre categorías estrictas, sino detectar diferencias sutiles.

Pasemos ahora a presentar su propuesta de tipos de texto y cómo los relaciona con los aspectos más importantes de la traducción.

2.2.3.2.1. Tipos de texto y criterios de traducción

Con el fin de sistematizar los textos prototípicos que se pueden encontrar en la práctica de la traducción, Snell-Hornby (1988) elabora un diagrama de seis niveles en el que se aglutinan no sólo disciplinas asociadas a la lingüística, sino también disciplinas extralingüísticas.

DIAGRAMA 2: Tipo textual y criterios relevantes en la traducción (Snell-Hornby 1988: 32)



En el plano horizontal del diagrama, aparece la idea de escala gradual en la que las transiciones y los límites entre categorías no están perfectamente delimitados. En el plano vertical hay una representación estratificada que, de acuerdo con el principio de *gestalt*, funciona de un nivel más general (A) o *macronivel* a uno más particular (F) o *micronivel*.

El nivel A presenta las áreas convencionales de la traducción: traducción literaria, traducción general (*general language translation*) y traducción de lengua de especialidad (*special language translation*). La separación entre estos tres tipos de texto no es absoluta.

El nivel B presenta una prototipología de los tipos de texto básicos desde textos bíblicos hasta el lenguaje de la tecnología moderna, aunque realmente opinamos que estos no son tipos de texto, sino distinciones de acuerdo con el campo del discurso. Mientras que los tipos de traducción que se encuentran en la izquierda (la Biblia, la antigüedad clásica, la traducción lírica, etc.) han sido el objeto de la teoría de la traducción tradicional, la literatura infantil y la traducción teatral (*stage-translation*) no han recibido atención hasta la década de los ochenta. En el otro extremo están los tipos de texto de lenguajes de especialidad, que es la principal actividad del traductor profesional actual y de los académicos de distintas disciplinas (el derecho, la economía, la medicina, etc.). Esta prototipología sólo presenta algunos tipos de texto básicos o prototípicos.

En el nivel C aparecen disciplinas no-lingüísticas. Estas presentan la realidad extralingüística, lo cual indica que en traducción es esencial el conocimiento del contexto sociocultural tanto en la cultura origen como en la cultura término.

El nivel D menciona los principales criterios que gobiernan el proceso de traducción. El primer subnivel (Di) se centra en el texto origen y su comprensión. Al acercarse más a lenguajes de especialidad se trabaja con términos y conceptos estandarizados y las opciones en la interpretación de los mismos se reducen bastante. En el segundo subnivel, D (ii), se aplica la noción de *invarianza* a casos de identidad conceptual y también el concepto de *equivalencia*. Para Snell-Hornby, la equivalencia es algo dinámico (Hönig y Kussmaul 1982) y el criterio dominante en cualquier traducción es la función comunicativa del texto meta (subnivel D iii). En el D(iii) se

presentan de una forma críptica los cambios que puede experimentar el texto término según la función del texto origen. Por ejemplo, ante textos que llevan consigo la extensión creativa de la norma lingüística, la traducción implica recrear las dimensiones lingüísticas, y esto resulta en un cambio de perspectiva en el texto término.

En el nivel E se nombran áreas de la lingüística que son relevantes para la traducción; entre ellas, se destaca el papel de la lingüística del texto en todos sus aspectos, desde el análisis de la macroestructura, la progresión temática, la perspectiva de la oración, la coherencia y la cohesión. La traducción de lenguajes especiales presupone una cierta familiaridad con la terminología y el acceso a bancos de datos. La lingüística contrastiva tiene un papel importante en el estudio de la sintaxis y la lexicología, aunque se deberían explorar otras áreas como la retórica contrastiva, un área que Snell-Hornby no menciona. Otras disciplinas pertinentes son la sociolingüística, la pragmalingüística y la psicolingüística.

Y por último en el nivel F, se tratan aspectos fonológicos de relevancia en algunos tipos de texto y traducción. Estos son importantes en la traducción *audiomedia*, que, como afirma Snell-Hornby (1997: 288), están adquiriendo gran relevancia en la era audiovisual.

Con este marco prototípico, Snell-Hornby (1988: 35-36) pone los cimientos a la traducción entendida como una disciplina integradora y al mismo tiempo independiente, que cubre todos los tipos de traducción, desde la traducción técnica a la literaria. También propone cuatro postulados:

- La traducción debe tener un enfoque interdisciplinar.
- La traducción ha de seguir un enfoque distinto al de la lingüística tradicional, en la que se va desde el micronivel hasta el macronivel. Snell-Hornby concibe los textos en función de su contexto situacional y cultural, coincidiendo de este modo con la psicología de la *gestalt*: el todo ha de ser analizado *top-down*.
- La *tipología* se sustituye por la idea de *prototipología*, que admite tipos textuales mixtos y límites difusos, mientras que la *dicotomía* se sustituye por el concepto de *escala gradual*.

- Más que aislar fenómenos y palabras, la traducción aísla redes de relaciones y analiza la relevancia de cosas concretas en relación con el contexto más amplio del texto, la situación y la cultura.

En cuanto a las características de los prototipos de texto que van desde la traducción de lenguajes de especialidad a la traducción literaria, también plantea una serie de hipótesis que posteriormente intenta validar con el análisis de cuatro tipos de texto distintos (ibid: 115):

- Mientras más especializado o pragmático sea el texto origen, más vinculado está a una situación específica, y resulta más fácil definir la función de la traducción.
- Mientras más específica sea la situación y más claramente esté definida la función, la traducción estará más orientada hacia el polo meta.
- Mientras más literario un texto (original o traducción) es, más depende la situación y la función de la activación que hace el lector
- Mientras más literaria es una traducción, el estatus del texto origen es mayor.

En estas hipótesis aparecen destacados tres aspectos que determinan el papel que el tipo textual al que pertenece el texto original va a tener en la traducción. Estos aspectos son la *situación del texto origen* (el emisor, el receptor, el medio), la *función del texto meta* y el *estatus del texto origen*. Estamos en el macronivel del análisis de los textos.

Además de estos aspectos, existen una serie de convenciones que la comunidad de hablantes ha asociado a determinados tipos de texto. Estas convenciones afectan a los microniveles de la presentación formal del texto, de la sintaxis, de la semántica, del léxico y de la grafología / fonología. También, aunque Snell-Hornby no lo incluye, también afecta al nivel pragmático, a la sucesión de actos de habla que aparecen en un determinado tipo textual.

Las convenciones lingüísticas de la comunidad de discurso en que cobran existencia los textos se suelen identificar con la idea de estilo, que ha sido ampliamente explorada desde la traducción, la lingüística y la crítica literaria (Reiss 1971; Crystal y

Davy 1969; House 1977; Leech 1969; Leech y Short 1981). Snell-Hornby los considera como dos conceptos distintos pero relacionados y define estilo como el sistema de opciones en el uso de la lengua que realiza un escritor como individuo. Para conciliar la idea de estilo y de convención de grupo afirma que, cuanto más especializado sea un texto y más concreta la situación en que se produce, menos se manifiesta el estilo individual y más se manifiestan las convenciones de grupo. Además, las opciones estilísticas también afectan a los microniveles que hemos mencionado en el párrafo anterior.

Si al hablar de los tipos de texto es ineludible hablar de convenciones de la comunidad lingüística, al hablar de estas, es ineludible hablar de la cultura. Estas convenciones no se derivan de los textos como instancias estáticas de la lengua sino de la función que adquieren en una situación extralingüística para conseguir un objetivo concreto (Snell-Hornby 1997: 287).

2.2.3.2.2. Tipos textuales y cultura

Antes de relacionar estos conceptos, Snell-Hornby (1998: 40), en consonancia con las ideas de Hönig y Kussmaul (1982) y Reiss y Vermeer (1984), menciona las claves que encierra el término *cultura*:

- La cultura es una totalidad de conocimientos, habilidades y percepciones.
- La cultura está en estrecha conexión con el comportamiento y la acción.
- La cultura es dependiente de la norma, ya sea de comportamiento social o de uso lingüístico.

Los tipos de texto que Snell-Hornby presenta en su modelo, o al menos, sus variantes y las formas lingüísticas que los materializan, no son universales, ya que la traducción es un *cross-cultural event* que tiene que mantener el equilibrio entre la lengua como *sistema*, la *norma* y el *habla*. (Coseriu 1977). La norma impone un determinado comportamiento social y un determinado uso lingüístico. En nuestra opinión, la norma también impone una forma concreta de organizar los textos para que sean efectivos en una comunidad de hablantes con una idiosincrasia propia.

Como en traducción se trabaja con dos culturas, se plantea el problema de decidir entre mantener los aspectos culturales codificados en el texto origen o adaptar el texto a la cultura que va a recibir la traducción. Ante este dilema tan antiguo como la traducción misma, Snell-Hornby toma partido por la teoría del *skopos*, es decir, prefiere que la traducción esté orientada hacia la función del texto meta, en lugar de hacia las normas que impone el texto origen. No obstante, para evitar un excesivo énfasis en la función del texto meta, esta autora llama la atención sobre el estatus del texto origen y reclama un estudio de la relación entre el estatus del texto origen en relación al tipo textual. También destaca la relación entre la situación y el tipo de texto origen, por una parte, y la función y el tipo de texto término, por otra parte.

2.2.3.2.3. Análisis textual

El análisis textual propuesto en Snell-Hornby (1988) concede importancia a los tipos textuales porque procede *top-down*, de un macronivel a un micronivel, del texto al signo. Sin embargo, como vimos en el modelo de Bell (1991), en la traducción hay procesos tanto *top-down* como *bottom-up*. La única explicación de por qué no ha incluido los últimos es que ella no utiliza estas palabras en el sentido cognitivo de Bell. Para ella, los procesos *top-down* suponen ir de lo más general a lo más particular, siguiendo el principio de *gestalt*. Y así, sugiere que el análisis textual debe comenzar con la identificación del texto en términos de la cultura y la situación (Vermeer 1983). El paso siguiente es el análisis de la estructura del texto, yendo de la macroestructura al nivel de la cohesión léxica e incluyendo la relación entre el título y el cuerpo principal del texto. Por último, se deben desarrollar estrategias para traducir el texto, basadas en las conclusiones que se derivan del análisis. En todo caso, las conclusiones derivadas del análisis han de estar relacionadas y proporcionar una interpretación coherente del texto.

Aunque esta forma de análisis puede ser útil desde el punto de vista pedagógico, pensamos que tiene algunos errores desde el punto de vista teórico. Esta forma de análisis pasa por alto los procesos *bottom-up* y el papel que juegan, entre otros, la comprobación de que el texto está siendo interpretado de una forma correcta. Por otra parte, nos parece que, además de partir de la macroestructura a la cohesión léxica, se

debe también proceder a la inversa, de la misma manera que funciona la cognición. De hecho, a la hora de interpretar los textos, los lexemas y las relaciones cohesivas que se establecen entre ellos son los que nos permiten llegar a la macroestructura y no al contrario. Coincidiendo con Wotjak (1998) y Jiménez (1994), creemos que cada lexema activa una serie de relaciones con los otros lexemas del texto (relaciones *en presencia*), y con otros lexemas que, aunque no aparecen en el texto, pertenecen al mismo *marco (frame)*⁹ que el lexema en cuestión (relaciones *en ausencia*). Por este motivo, conforme se avanza en la lectura del texto, podemos reconocer una macroestructura del discurso distinta a la que reconocimos en un primer momento. Por ejemplo, un texto que aparece en la sección de opinión de un periódico puede comenzar con elogios a una persona, lo cual nos hace pensar en un texto laudatorio, y, conforme seguimos leyendo, descubrir que se trata de una crítica acérrima basada en la ironía.

2.2.3.2.4. El papel de los *marcos* y las *escenas* (Fillmore 1977)

Por último, en su visión integradora, Snell-Hornby no olvida aspectos cognitivos de la traducción que pueden explicar el proceso de la traducción y sostiene, siguiendo a Fillmore, que los textos son portadores de *marcos* y *escenas (scenes)* (Fillmore 1977). Tanto unos como otros han de ser recreados e interpretados por el lector o por el traductor. Fillmore (1977: 63), artífice de la *Gramática de Casos*, define *marco* como:

any system of linguistic choice—the easiest being collections of words, but also including choices of grammatical rules or grammatical categories—that can get associated with prototypical instances of scenes.

En su definición, se percibe que Fillmore concibe los marcos como constructos lingüísticos asociados a un esquema mental vinculado con experiencias, creencias y prácticas que él denomina *escena*. Uno de los ejemplos que propone es la relación lingüística de base cognitiva que existe entre los verbos *buy*, *sell* y *cost*. Las diferencias

⁹ Aquí estamos entendiendo el concepto *marco* como una red de conocimientos que relaciona los distintos dominios asociados a un lexema (Taylor 1995: 87), de forma que este queda asociado a otros lexemas que suelen aparecer en su cotexto. En el apartado siguiente, presentaremos la acepción que le da Fillmore (1977).

semánticas y sintácticas entre ellos radica en que focalizan una perspectiva cognitiva diferente. En cuanto al término *escenas*, Fillmore (ibid) lo emplea para incluir

familiar kinds of interpersonal transactions, standard scenarios, familiar layouts, institutional structures, enactive experiences, body image; and in general, any kind of coherent segment, large or small, of human beliefs, actions, experiences, or imaginings.

Snell-Hornby utiliza estos conceptos para explicar el proceso cognitivo de la traducción. En nuestra mente tenemos almacenadas experiencias o *escenas* que están codificadas mediante *formas lingüísticas relacionadas o marcos*. Al leer un texto, las *escenas* interactúan con los marcos y evocan una serie de asociaciones que activan otras formas lingüísticas y otras *escenas*. Por eso, la comprensión de los textos y, por ende, la traducción son procesos dinámicos que culminan con la obtención de un *todo* con significado. Asimismo, la forma lingüística queda relacionada con la experiencia y la respuesta que da el lector y el traductor como lector. En esto Snell-Hornby destaca la importancia de los lectores en la construcción de los textos al más puro estilo de la estética de la recepción (*Rezeptionsästhetik*) de Iser (1976). Al fin y al cabo, como se verá, es el lector el que forma en su mente una unidad conceptual y coherente a partir de signos lingüísticos separados en el papel pero unidos mediante la cohesión.

De acuerdo con todo esto, la traducción es un acto complejo de comunicación en el que interactúan el autor del texto origen, el traductor como lector del mismo y autor del texto meta y el lector del texto meta. El traductor parte de los marcos presentados en el texto (componente lingüístico), que fue compuesto por un autor de acuerdo con su propio repertorio de *escenas*, en parte prototípicas. Basándose en ese marco, el lector-traductor construye sus propias *escenas* dependiendo de su propia experiencia y de su conocimiento del tema del texto. Si las *escenas* activadas en la mente del traductor difieren de la intención del autor, se puede llegar a una traducción errónea. Posteriormente, con las *escenas* que ha activado, el traductor debe encontrar marcos adecuados en la lengua término, lo cual supone un proceso constante de toma de decisiones (*decision making*), que, en todo caso, depende del conocimiento de la lengua término. En esta etapa, el traductor da forma al nuevo texto.

Con esta visión de la traducción basada en marcos y escenas, Snell-Hornby supera la idea simplista y estática de que la traducción consiste en buscar la equivalencia e incorpora un principio holístico que incluye elementos textuales interrelacionados, la experiencia, la percepción y el contexto situacional. Sin embargo, vamos a hacer una serie de puntualizaciones.

En primer lugar, la definición de *marco* de Snell-Hornby como forma lingüística, no coincide con la definición dada por otros autores como Minsky (1975), para quien un marco es “la estructura mental evocada por la palabra, que representa una situación estereotípica en la que se usa” (Muñoz 1995: 168). Tampoco coincide con formulaciones posteriores de Fillmore (Fillmore y Atkins 1998) en las que el énfasis no está en las relaciones paradigmáticas y sintagmáticas entre los signos lingüísticos, sino en la estructuración conceptual de las palabras en dominios de la experiencia y de las creencias. En segundo lugar, creemos que su aproximación cognitiva no presta la suficiente atención a otras estructuras psicológicas esenciales en la comprensión de los textos: las *macroestructuras*, que, como dijimos, indican el contenido global del discurso, y las *superestructuras*, que definen la ordenación global del discurso y las relaciones jerárquicas de sus respectivos fragmentos (Van Dijk 1977b, 1980). Tampoco menciona los guiones (*scripts*), planes (*plans*) y metas (*goals*) de Schank y Abelson (1977), que son fundamentales a la hora de organizar el discurso. Estas estructuras mentales representan “experiencias textuales” anteriores que cumplieron de forma efectiva un fin concreto y que, en parte, son universales, y en parte, tienen características establecidas convencionalmente por una comunidad de discurso (*discourse community*).

2.2.3.2.5. Los actos de habla

Snell-Hornby concede bastante importancia a los actos de habla de Austin y Searle, sobre todo, en cuanto que, para cada lengua, están asociados a una o varias formas lingüísticas que el traductor debe conocer. Uno de los recursos más útiles al traductor para familiarizarse con las formas naturales del lenguaje asociadas a una función comunicativa son los textos paralelos. Aquí especificaremos que Snell-Hornby utiliza el término de *texto paralelo* en el sentido más habitual en traducción, que es distinto al

más aceptado recientemente en lingüística de corpus. Define los *textos paralelos* como “two linguistically independent products arising from an identical (or very similar) situation” (ibid. 1988: 86). Estos se denominan en lingüística de corpus, *textos comparables*.

Aunque estamos de acuerdo con esta autora, creemos que ha olvidado mencionar que existe una correlación entre *tipos textuales* y secuencias de actos de habla, estas últimas denominadas por Hatim (1995) *actos textuales*. Esta correlación podría ser estudiada con la ayuda de la lingüística de corpus y con la elaboración de *textos comparables* en dos lenguas que pertenezcan al mismo tipo textual.

2.2.4. El funcionalismo de Christiane Nord: su modificación a la teoría del *skopos*

La propuesta de Nord (1997) se encuentra en la línea de la teoría del *skopos* (Vermeer 1978; Reiss y Vermeer 1984) y constituye una crítica al concepto de equivalencia en traducción y a la tradicional primacía que esta disciplina ha dado al texto origen esgrimiendo que no proporciona soluciones a una serie de problemas traductológicos que son muy frecuentes en el mundo profesional del traductor.

It is no longer the ST which sets the standards for the translator's decisions in the translation process, but the intended receiver of the translation, whose reception will be entirely guided by target culture expectations, conventions, norms, models, real-world knowledge, perspective, etc. (Nord 1997: 46)

Otros argumentos que Nord (1997: 44-45) aporta en contra de la teoría de equivalencia se podrían resumir del siguiente modo:

- a) El modelo de equivalencia se centra en cualidades estructurales del texto origen y descuida en cierto modo aspectos pragmáticos y culturales.
- b) No existe consistencia en las definiciones que se dan de equivalencia, lo cual no aclara si algunos textos pragmáticos se han de traducir de una forma no literal.
- c) Al basarse en una supuesta universalidad de conceptos, el modelo de equivalencia no da respuesta a las diferencias interculturales entre conceptos, normas y convenciones.

- d) Este modelo excluye algunos textos meta que no satisfacen el criterio de equivalencia, por ejemplo, las versiones interlineales, la traducción filológica, etc.
- e) El texto original y los valores que se encarnan en el mismo tienen primacía sobre cualquier otra consideración o decisión que se haya que tomar a la hora de traducir. Siendo esto así, el traductor se convierte en una figura infravalorada socialmente.

Como alternativa al modelo de equivalencia, Nord adopta un modelo que ella caracteriza como pragmático, orientado hacia la cultura entendida como comportamiento verbal y no verbal, antiuniversalista, capaz de englobar todos los tipos textuales por su carácter práctico y respetuoso del papel del traductor como profesional experto. Su modelo propone que, aparte de la funcionalidad del TT, el traductor ha de tener en cuenta factores extratextuales e intratextuales que ayudan a entender la función del TO y el TT. De entre ellos un factor esencial es la lealtad (*loyalität*) o “responsabilidad del traductor hacia los participantes en la interacción translaticia” (ibid: 48). A la hora de traducir, el traductor ha de comparar el *escopo* con las funciones del texto en la lengua origen y considerar hasta qué punto se puede cumplir el encargo sin violar el principio de lealtad y qué problemas pueden surgir y qué técnicas de transferencia (*transfer procedures*) se pueden emplear para solucionarlas. De este compromiso entre el *escopo* y la *lealtad* hacia los participantes en la comunicación, nace una buena traducción.

En otras palabras, la traducción se sustenta en dos pilares, la *funcionalidad* del TT y la *lealtad* (Nord 1991a: 28 y ss.; Nord 1991b) hacia las expectativas de los participantes implicados en la traducción: el autor del texto, el cliente y el lector de la traducción. De esta forma Nord evita las interpretaciones simplistas del proceso traductor a las que se habían llegado por seguir un modelo de equivalencia y un modelo de *escopo* extremos.

En cuanto a su propuesta de tipología de traducciones, la podemos calificar como funcional, aunque entendiendo función no como algo inherente al texto, sino como una cualidad pragmática que el receptor del mismo le asigna en una situación

determinada y que no tiene por qué coincidir con la función del TO. El marco de referencia de su tipología de textos lo constituye una combinación de los archiconocidos modelos propuestos por Bühler (1934) y Roman Jakobson (1960), autor que Nord califica erróneamente como funcionalista, dejando de lado el papel que tuvo en la articulación del formalismo ruso. De estos dos modelos, Nord selecciona sólo cuatro funciones del lenguaje: la *referencial*, la *expresiva*, la *apelativa* y la *fática*.

La clasificación básica de tipos de traducción que propone identifica dos tipos de traducciones en los que se pueden encontrar distintos subtipos:

- *Documentary translation*: traducción que, según Nord, tiene una función metatextual porque reproduce las condiciones comunicativas y distintos aspectos del TO. Como ejemplos, menciona la traducción literal, la traducción filológica y la traducción exotizante (*exoticizing translation*). Esta última reproduce de un modo muy fiel las costumbres de la cultura origen y es muy común en la traducción de literaturas postcoloniales.
- *Instrumental translation*: traducción que, aun utilizando como modelo el TO, cumple una función que se adapta a la situación comunicativa del TT, de modo que el lector no se da cuenta de que está leyendo una traducción. La traducción instrumental tiene una o varias de las siguientes funciones: referencial, expresiva, apelativa o fática. Dentro de este tipo, menciona la traducción *equifuncional* (textos técnicos, manuales de instrucciones, etc.), la traducción *heterofuncional* y la que denomina *homóloga*, usual en la traducción poética. La traducción *homóloga* pretende conseguir un efecto similar al que produce el texto original, reproduciendo en el contexto literario del TT la función que el original tiene en su contexto literario.

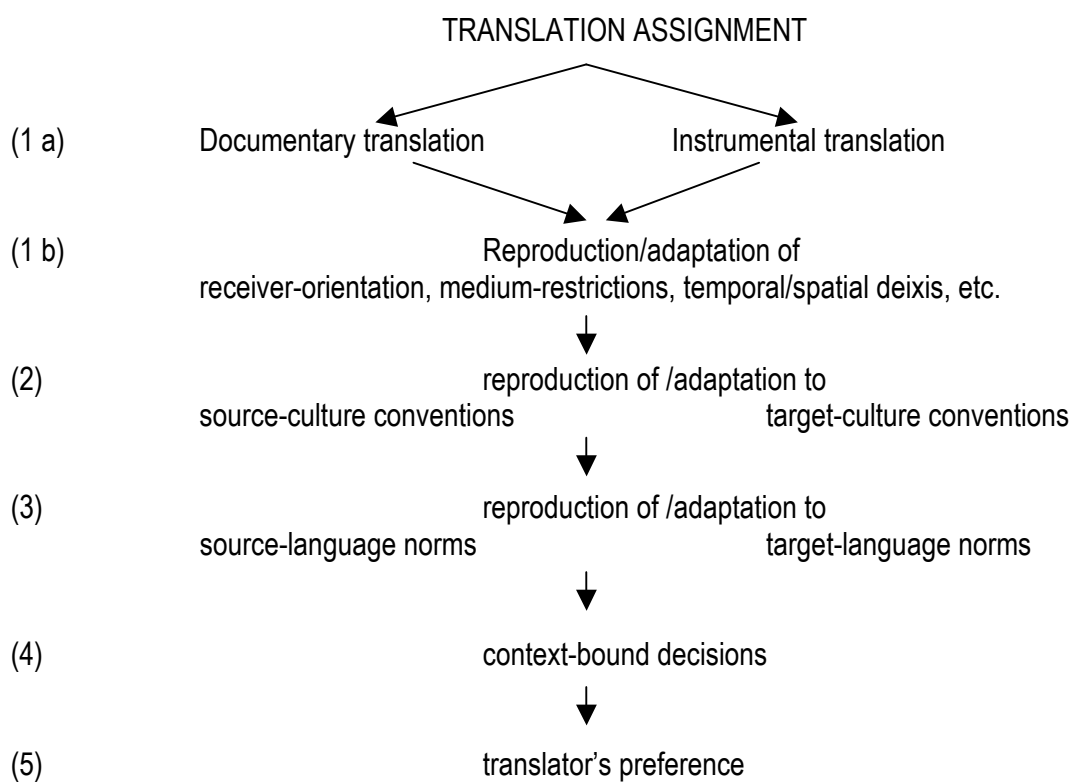
Asimismo, su modelo aporta una clasificación de los problemas de traducción en el que, de las cuatro categorías propuestas en Nord (1987), tres de ellas son relevantes al hablar de tipología de textos:

- a) problemas de traducción pragmáticos (*pragmatic translation problems*), como decidir la función de la traducción

- b) problemas de traducción específicos al texto (*text-specific translation problems*) como la traducción del humor y de figuras retóricas, entre otros, que normalmente están asociadas a un determinado tipo textual
- c) problemas de traducción interculturales (*intercultural translation problems*) como la traducción de convenciones textuales y estilísticas.

Para solucionar estos problemas de traducción de tipo pragmático, intercultural y específicos a un tipo textual, Nord incluye un esquema con los pasos que se han de seguir a la hora de traducir. En el esquema se recomienda que, en lugar de proceder de forma *bottom-up*, es decir, partiendo de las palabras individuales en una primera etapa y posteriormente, prestando atención a los factores pragmáticos, se ha de proceder *top-down*. El primer paso consiste en decidir la función que va a tener la traducción.

DIAGRAMA 3: Jerarquía *top-down* de decisiones a la hora de traducir



En la aproximación funcional de Nord, el traductor ha de comenzar en un nivel pragmático decidiendo cuál es la función asignada a la traducción y si va a realizar una traducción de tipo documental (*documentary*) o instrumental (*instrumental*). En función de esta decisión, el traductor debe considerar qué elementos funcionales de la lengua origen se han de reproducir tal cual en la traducción y cuáles se han de adaptar a los conocimientos previos del receptor, expectativas, necesidades comunicativas y restricciones del medio. El tipo de traducción elegida determina si ha de seguir las convenciones de la cultura origen o de la cultura término en cuanto a traducción y estilo. Sólo en este instante el traductor ha de tomar en consideración diferencias en los sistemas lingüísticos. Si se presentan distintas posibilidades a la hora de traducir, el traductor ha de decidir, tomando en cuenta, o bien aspectos contextuales, o bien, en textos menos convencionalizados como los textos literarios, las preferencias personales del traductor. Todo esto siempre queda supeditado a la función de la traducción.

Al igual que decíamos al comentar el modelo de análisis de Snell-Hornby, este esquema resulta muy pedagógico pero no es realista con los procesos mentales que se producen a la hora de traducir que, como dijimos, no son unidireccionales y son de tipo *bottom-up* como *top-down*. Por ejemplo, no se menciona por ninguna parte qué procesos de tipo *bottom-up* como el reconocimiento de elementos visuales o tipográficos pueden guiar la decisión sobre el tipo textual del original y el tipo de traducción a realizar (documental o instrumental). También hemos detectado que elementos esenciales en la organización de cada tipo textual como los marcos, las escenas y los guiones brillan por su ausencia. Ayudar a los futuros traductores y a los traductores profesionales a comprender el proceso cognitivo de la traducción también resultaría bastante pedagógico.

Por otra parte, nos ha resultado interesante para cualquier estudio de tipología de textos, la información que Nord sugiere que se ha de incluir junto con cualquier traducción que se asigna a los alumnos de traducción. Esta información que Nord denomina *translation brief* (Nord 1996) puede ayudar a los alumnos a inferir información sobre las funciones que se esperan ha de cumplir el texto término. Los puntos que ha de incluir la especificación del *encargo de traducción* son bastantes útiles a la hora de determinar una tipología de textos:

- a) La función que el emisor espera que cumpla el texto: referencial, expresiva, apelativa y/o fática
- b) El/los receptor/es
- c) El tiempo y el lugar de la recepción del mensaje
- d) El medio de transmisión de la traducción
- e) El motivo que ha llevado a la producción o recepción del texto.

Para que se cumpla la función que el emisor espera en los destinatarios de la traducción, el texto término ha de conformar las convenciones textuales y de estilo de la cultura término y mantener el registro adecuado. También se han de tener en cuenta las presuposiciones culturales asumidas por los lectores del texto término.

Por último, Nord concluye diciendo que su aproximación funcional se apoya en un concepto de textualidad orientado hacia la comunicación y un concepto de traducción basado en el principio de lealtad hacia los *cooperation partners*. A nuestro parecer, aunque Nord mencione algunas interpretaciones del concepto de equivalencia que no son estáticas¹⁰, su negación del concepto de equivalencia pasa por alto el hecho de que en algunos tipos de traducción, como la traducción para el aprendizaje de lenguas, es muy recomendable la equivalencia. Asimismo, pensamos que, al acuñar el término de *lealtad*, Nord subsume las concepciones dinámicas de equivalencia arriba mencionadas. Por último, nos gustaría manifestar nuestra sorpresa al no encontrar en un modelo que se denomina funcional ninguna mención a una de las aproximaciones más funcionales de la lingüística moderna: el funcionalismo sistémico de Halliday.

2.2.5. La tipología de traducciones según Juan C. Sager

Sager (1997) no rechaza la idea de *equivalencia* porque, al fin y al cabo, la vaguedad del concepto de equivalencia está también presente en la teoría del *skopos* o función. Por eso, Sager relativiza el concepto de equivalencia y la entiende como un proceso dinámico en el que intervienen estrategias que pueden modificar en el texto traducido,

¹⁰ Nord (1997: 44) hace alusión a la *equivalencia dinámica* (Nida 1964: 159), la equivalencia textual (*text-bound equivalence*) de Neubert (1984:68) y la *equivalencia pragmática* de Koller (1979: 187 y ss.).

el contenido y la intención del texto original, en función de elementos como la cultura, el emisor, el receptor y el traductor. Este último es un mediador en el proceso de traducción y sigue distintos tipos de equivalencia según el cliente que solicita la traducción. Uno de estos tipos de equivalencia es la equivalencia de *tipo textual* (ibid: 25), que responde al reciente interés por los tipos textuales y al hecho de que hayan aparecido nuevos tipos de traducción que no se acomodan al tradicional concepto de equivalencia y a los tipos de texto existentes en la lengua meta.

La postura de Sager en relación con la tipología textual y los tipos de traducción se deriva de su familiaridad con la traducción automática y las ciencias de la información. De las ciencias de la información, extrae su distinción entre *lectores primarios* y *secundarios*, por una parte, y *texto, documento y mensaje*, por otra.

En primer lugar, los lectores primarios son la audiencia ideal a la que el autor pretende dirigirse. Bajo el término lectores secundarios se incluye al resto de los lectores. La segunda distinción, que completa la tradicional dicotomía establecida desde la lingüística entre texto-discurso¹¹, se basa en la intención del escritor. Esta intención es uno de los parámetros que Sager distingue para determinar el tipo de traducción, aunque la llama *función*. Según Sager, si el traductor sabe, en primer lugar, si está traduciendo un texto, un documento o un mensaje y, en segundo lugar, si debe mantener la intención del texto original en el texto término, la toma de decisiones inherentes al proceso de traducción estarán simplificadas.

Según Sager (ibid: 27), un *texto* es “a unit of form and content only, which can be given a new intention by an information mediator like an abstractor or indeed a translator when it is addressed to someone the writer did not have in mind”. El texto se convierte en *documento* si la intención del autor está claramente codificada en el

¹¹ Las acepciones de estos dos términos en la lingüística son de lo más variado. Stubbs (1987) hace un breve resumen de las diferencias entre estos dos términos. En primer lugar, se suele hablar de texto escrito frente a discurso hablado (Coulthard 1985). También, *discurso* suele implicar un discurso interactivo, mientras que *texto* implica un monólogo no interactivo. En segundo lugar, discurso implica longitud, mientras que un texto puede ser muy corto. Desde esta perspectiva, Halliday y Hasan consideran ‘Salida’ o ‘No fumar’ como textos completos. En tercer lugar, Van Dijk (1977) utiliza *texto* para indicar un constructo teórico y abstracto que se actualiza en el *discurso*. Halliday (1978: 40) hace esa misma distinción pero utiliza los términos en el sentido contrario: el lenguaje se actualiza en el texto. Por último, Widdowson (1979) distingue entre *cohesión textual*, que aparece en el léxico superficial, la gramática y el desarrollo proposicional, y la *coherencia del discurso*, que opera entre actos de habla subyacentes.

mismo. Y así, los *documentos* son unidades de texto que tienen una función en la comunicación pero que están fuera de la situación comunicativa que los originó. Pues bien, el traductor ha de enfrentarse tanto a *textos*, en los que la intención del destinatario no está claramente definida, como a *documentos*, cuyos lectores primarios son desconocidos por el traductor. El resultado de su traducción tiene el potencial de convertirse en un *mensaje*, que es un documento que cobra existencia mediante la interacción comunicativa entre escritor y lector en una situación concreta. Un texto traducido puede considerarse un mensaje si la intención codificada en el documento original o en la traducción y las necesidades del lector de la traducción coinciden. En otras palabras, un texto traducido será un instrumento de comunicación o un *mensaje* si el traductor pone al lector de la traducción en la posición de un lector primario. Para ello, el traductor ha de tener muy presente la siguiente máxima: “when there is a match between a writer’s presuppositions about a reader’s expectation and the reader’s assumption about the writer’s intention, communication is maximally effective” (ibid: 28).

Al hacer estas distinciones, Sager resalta el papel tan importante que juega la *intención* a la hora de distinguir entre estos vehículos de expresión que denominamos *tipos textuales*. En cualquier caso, el traductor ha de ver si esa intención se va a mantener en el texto término porque existen nuevos tipos de traducción en los que la intención puede cambiar. Esto es bastante habitual entre los profesionales de la traducción especializada, que cuentan con clientes fijos que no solicitan una traducción completa, sino *abstracts* de artículos científicos, resúmenes de cartas comerciales, etc. Estos nuevos documentos, surgidos en una sociedad en la que la información y la familiaridad con los últimos avances científicos es vital, tienen que encontrar cabida en una teoría de la traducción.

La propuesta que hace Sager para traducir estos tipos textuales es la de buscar una solución de compromiso que tenga en cuenta:

- la intención del autor con respecto al lector primario
- la intención del cliente, que, al fin y al cabo, es el que indica la intención que ha de aparecer en el texto término y las expectativas de los lectores de la traducción

- las necesidades del lector secundario.

También ha de tener presente la máxima de que la comunicación es más efectiva cuando hay una correspondencia entre la intención del autor y las expectativas del lector.

Si el traductor decide mantener la intención del autor en el texto término, el traductor sólo tiene que preocuparse por ver el conocimiento previo que tienen los lectores en comparación con los lectores de la lengua origen. En el caso de que se pretenda que el texto término tenga una intención distinta al texto original, el traductor tiene que negociar los cambios necesarios con el cliente.

Sager parte de la idea de *intención* para empezar a hablar de tipos textuales porque la elección de un tipo textual convencional es el medio más efectivo y económico (en términos lingüísticos) para expresar la intención y para que el lector primario la detecte. Los tipos textuales surgieron a partir de situaciones comunicativas convencionales. En estas situaciones comunicativas específicas empezaron a utilizarse de una forma regular unos determinados patrones lingüísticos. En palabras de Sager (ibid: 30): “Regular repetitions of messages in particular circumstances have created expectations of recognisable structural and rhetorical features which condition our modes of reading a message”. En definitiva, los tipos textuales son el resultado de relaciones comunes tanto sociales como cognoscitivas entre el escritor y el receptor, y por este motivo, son tan efectivos como instrumento de comunicación.

Después de explicar el origen de los tipos textuales, Sager apunta hacia las distintas perspectivas respecto a las cuales se pueden clasificar. Normalmente se suele asociar un tipo textual con su contenido, por ejemplo, el informe médico, la reseña de un libro, la receta de cocina, etc. Sin embargo, este criterio no es suficiente porque muchas veces el mismo contenido se puede presentar en una variedad de tipos textuales y, a la inversa, distintos contenidos pueden ser presentados con la misma estructura. Por eso añade que los tipos de texto también se caracterizan por el tema, el modo de expresión, la situación y los rasgos de composición del texto.

A partir de estos elementos, el lector puede, si tiene en sus manos una buena traducción, identificar un determinado tipo textual, el cual le va a indicar quién es el emisor del mensaje y cómo se estructura la información *grosso modo*. Por ejemplo,

cuando leemos un artículo académico, sabemos el tipo de información que vamos a encontrar en la sección de *Métodos* y, si quisiéramos, podríamos leer sólo el *Abstract* para extraer la idea principal del artículo.

De aquí se deduce que un buen traductor ha de asegurarse de que la traducción haga que los lectores reconozcan de inmediato el tipo textual de forma que puedan ajustar sus expectativas lo antes posible. En esta tarea al traductor se le presenta el dilema de qué estrategia de traducción debe seguir.

2.2.5.1. Las estrategias de traducción

Sager añade al conjunto de estrategias de traducción que se han identificado tradicionalmente, tres nuevas que intervienen cuando el tipo textual del TO no se va a mantener en la traducción. Y así, agrupa las estrategias de traducción en cuatro tipos:

- a) **Las estrategias de traducción tradicionales, que intentan preservar tanto el contenido como la intención.** Sirven para traducir un manual técnico, o una carta, o un artículo científico/académico en el área de la ciencia o la tecnología
- b) **Las estrategias de traducción que cambian el *contenido del texto*.** Por ejemplo, las traducciones comentadas, los *translingual abstracts* y las traducciones para extraer el contenido esencial de un texto (*gist translation*).
- c) **Las estrategias que cambian la *intención del texto*.** Cuando esto ocurre, a menos que la intención del documento a traducir sea evidente, el traductor debe definir e identificar la nueva situación para poder elegir el tipo textual apropiado. El resultado de estas estrategias puede ser o bien un tipo textual conocido en la lengua término, o bien tipos textuales que aparecen en la cultura de la lengua término como una traducción, como un nuevo acto traductivo. Esto ocurre, por ejemplo, cuando las actas de un simposio se convierten en un informe, y el informe en un resumen. Los cambios de intención van acompañados normalmente de un cambio del tipo textual, lo cual puede acarrear un cambio en el contenido y de la estructura del mismo.
- d) **Las estrategias de traducción en las que se produce un cambio en el *contenido y la intención*.** Una ilustración de esto la encontramos al elaborar un

abstract, un comentario o una adaptación del texto original. El cambio en el contenido lleva aparejado un cambio en la intención y como consecuencia, una adaptación a las convenciones textuales de la cultura meta.

En resumen, Sager establece que una aproximación dinámica al proceso traductor puede requerir no sólo un cambio en el código lingüístico, sino que también puede requerir un cambio en el contenido (por adición o supresión) y/o un cambio de finalidad para dar a los lectores de la traducción las mismas ventajas que las proporcionadas a los lectores primarios.

2.2.5.2. Una tipología de traducciones

Sager propone una tipología de traducciones basada en las estrategias de traducción que se siguen para obtener el texto término:

- a) **según el modo de producción:** traducción *humana* y traducción *automática*
- b) **según se produzca un cambio en el contenido:** traducción *íntegra* y traducción *modificada* por reducción o por adición (*full/modified by reduction or addition*). Estos tipos de traducción van normalmente acompañados de un cambio de intención y, por tanto, de tipo textual.
- c) **según se produzca un cambio en la intención.** Sager concede gran énfasis a esta faceta. Al producirse un cambio en la intención, cambia el tipo textual. Según esto tenemos, por una parte, traducciones en las que la intención no cambia como las traducciones literarias, de textos de instrucciones, de directivas, etc.; y por otra parte, traducciones que se ajustan a satisfacer las expectativas del lector. En un artículo posterior, Sager (1998) denomina al primer tipo de traducción *dependent texts*, y al segundo, *derived texts*. Los cambios en la intención se pueden conseguir introduciendo en la traducción cambios en el código lingüístico o modificaciones del contenido y/o tipo textual del documento. Ejemplos de este tipo de traducción son los resúmenes, las *gist translations* y los *texts for "information only"*.

Esta clasificación de tipos de traducción se deriva de tomar una perspectiva en la que no sólo el código lingüístico puede cambiar, sino también el contenido y la intención del texto original. Desde esta perspectiva, desaparece la posibilidad de evaluar una traducción en función de criterios como el de fidelidad o exactitud.

A esta clasificación de las traducciones según las estrategias de traducción, hay que añadir una nueva distinción basada en los tipos textuales. Existen traducciones que están plenamente integradas en la cultura término (*existing target language text types*) y traducciones que no se acomodan a ninguno de los tipos textuales que existen en la lengua término y cuya lectura es suficiente para percibir que son el resultado de una traducción (*translation-specific text types*) como la traducción en prosa de un poema o la traducción para obtener información (*translation “for information only”*). Con el tiempo, el segundo tipo de traducción puede llegar a consolidarse en la lengua término como un nuevo tipo textual. Esto fue lo que ocurrió con los artículos académicos traducidos del inglés.

Sager (1998) vuelve a incidir en las estrategias que guían la secuencia de opciones a las que se tiene que enfrentar el traductor a la hora de traducir. En esta ocasión, objetiva estas estrategias mediante una serie de parámetros y demuestra que estos parámetros no sólo sirven para distinguir entre *tipos de texto* o *documentos*, sino que también justifican la distinción tradicional entre los siguientes tipos de traducción: traducción literaria, bíblica y técnica (o traducción de textos con fines específicos).

Estos parámetros pueden guiar las decisiones que se plantea el traductor y son los siguientes:

1. *Antecedentes situacionales en la cultura meta (situational antecedents)*

El traductor ha de constatar si la situación comunicativa que produce la traducción ya existe en la cultura meta.

2. *Tipo de documento en la lengua término (target-language document type)*

El traductor tiene que ver si la lengua término dispone de modelos de documento iguales o parecidos al texto origen, que puedan servir de base para traducir. Como vimos antes, si estos tipos de texto no existen, se creará un

translation-specific document type, es decir, un tipo textual que sólo se da como resultado de una traducción.

3. *Función del texto (purpose)*

La función del texto está marcada por el tipo de documento y se ha de entender dentro de estas perspectivas: la intención del escritor, las expectativas de los lectores y la función que asigna el cliente al encargar la traducción, que puede ser distinta a la del texto original. Si la función del original se mantiene, el texto traducido es un *texto dependiente*. En el caso contrario, estamos ante un *texto derivado*. En cuanto a las principales funciones que puede cumplir la traducción, Sager reconoce la de informar, ordenar, persuadir y evaluar.

4. *Estatus del texto meta en relación con el texto origen (relative status of source and target language)*

Es la interdependencia funcional entre el texto origen y el texto meta, que puede presentarse como una escala gradual que va desde la máxima a la mínima dependencia: *equal documents*, *parallel documents*, *dependent documents*, *derived documents* y *autonomous documents*.

5. *Noción de que se está leyendo un texto traducido (awareness)*

6. *Modelos textuales de traducción disponibles (translation text models)*

El traductor debe conocer las convenciones lingüísticas que determinan las expresiones idiomáticas, colocaciones y la estructura asociadas a tipos textuales, sobre todo, para la traducción de textos dependientes y derivados.

La utilidad de estos parámetros es indudable porque ayudan a concebir el proceso de traducción como una ruta en la que se presentan sucesivos cruces de caminos cuya señalización es cada una de las opciones que existen dentro de cada parámetro, y en la que hay que tomar decisiones constantemente. En algunos tipos de traducción como en la traducción técnica, el número de decisiones es mayor que para otros tipos de traducción. Con estos parámetros, el traductor puede resolver, guiado por algo más que la intuición, la tensión que se produce entre las condiciones que rodearon al texto origen y su audiencia por una parte, y la función del texto término y los posibles

lectores del mismo, por la otra. Igualmente, el traductor puede aprender a asociar determinadas estrategias con determinados tipos de traducción.

2.2.6. Paul Kussmaul: convención, tipos textuales y la codificación lingüística de los actos de habla

A la hora de hablar de tipos textuales, Kussmaul (1997) hace gala de una aproximación integradora que aúna aspectos cognitivos y pragmáticos de la traducción y encuentra una serie de nociones relacionadas: *convención*, *cultura*, *macroestructura* y *microestructura*.

Para ilustrar la importancia de estos aspectos en la traducción, Kussmaul (ibid: 67) cuenta como anécdota un malentendido ocasionado por la existencia de diferencias retóricas entre el discurso académico propio del mundo anglosajón y el del mundo germánico¹². Según Galtung (1985), son el resultado de la existencia de *prototipos culturales del discurso intelectual*. La moraleja de la anécdota puede ser que el desconocimiento de las convenciones textuales en la lengua origen y término puede llevar a grandes lagunas en la comunicación.

2.2.6.1. Convención, cultura y tipo textual

Kussmaul entiende las convenciones, no simplemente como lo arbitrario, sino como lo que permite que exista la conformidad y las expectativas en la comunicación (Lewis 1969: 78). En el nivel textual, la idea de convención se ha aplicado para designar las condiciones que han de cumplirse para el éxito de la fuerza ilocucionaria de un acto de habla (Wunderlich 1972: 11 y ss.). También Kussmaul destaca que, de los dos tipos de reglas que reconoce Searle (1969: 33), las *constitutivas* y las *regulativas*¹³, las más relevantes en el estudio de tipos textuales son las regulativas. Estas normas regulativas están determinadas por la cultura.

¹² Kussmaul explica cómo se interpretaron los comentarios de profesores alemanes y británicos en el turno de preguntas que siguió una conferencia pronunciada en una universidad británica. Para los asistentes británicos, el discurso académico de los alemanes fue algo ofensivo por contener críticas muy directas. Esta forma de emitir juicios críticos es la propia del discurso académico alemán, que difiere con respecto al anglosajón, que es más indirecto y sutil.

¹³ Las reglas *constitutivas* crean y definen el acto de habla y se pueden formular como *X counts as Y in the context of C*. Por otra parte, las reglas *regulativas* regulan la forma de actuar y se pueden formular como *If Y, do X*.

Para Kussmaul, el término *cultura* incluye mucho más que objetos e instituciones; incluye también principios de comportamiento (*behavioural knowledge*), un aspecto que a menudo pasa desapercibido en los estudios de traducción. Esta visión de la cultura es compartida por muchos académicos de la traducción (Reiss y Vermeer 1984) y de otras áreas afines como la antropología (Holland y Quinn 1987).

En cuanto al término *tipo textual*, Kussmaul reconoce la ambigüedad del mismo porque designa dos categorías identificadas por Reiss (1971) como *texttyp* y *textsorte*. (Véase apartado 2.2.2.1.). Para paliar esta ambigüedad, relaciona, al igual que Crystal y Davy (1969) y House (1977), el concepto de *tipo textual* con el de *situación*. Para estos eruditos, la situación, y por ende, el tipo textual quedan definidas por una serie de dimensiones como la individualidad, dialecto, tiempo, medio, participación (monólogo o diálogo), provincia, estatus, modalidad y actitud social¹⁴. Un cambio en una de estas dimensiones produce automáticamente otro tipo textual.

Asimismo, Kussmaul intenta demostrar cómo esas dimensiones situacionales se reflejan en las convenciones asociadas a cada tipo textual a un macronivel y a un micronivel.

2.2.6.2. Macroestructuras

Según Kussmaul, es difícil basar las macroestructuras en modelos lingüísticos. Algunos intentos se han centrado en la progresión de tema-remata (Gerzymisch-Arbogast 1986). Por otra parte, cuando la lingüística del texto trata las macroestructuras, examina la línea de pensamiento general de un texto y la secuencia de fragmentos de texto inherentes a un tipo textual. Y así, Göpferich (1995) ha visto cómo en textos científico-técnicos, la flexibilidad de las macroestructuras cambia según el tipo textual específico. En concreto, mientras menos técnico es un texto, más flexible es su macroestructura. A este respecto, el inglés y el alemán no difieren demasiado. Esto se puede explicar por el hecho de que la ciencia y la tecnología son campos internacionales. No obstante, pensamos que sería interesante confirmar esta hipótesis con otras lenguas que no tengan

¹⁴ House (1977) introdujo esta categoría para dar cuenta del grado de confianza o intimidad que existe entre los participantes, un factor que resulta esencial en la traducción porque determina el conocimiento compartido de los participantes y, por tanto, la información que se ha de hacer explícita en una traducción.

un origen germánico e investigar tipos textuales que estén más imbricados en culturas individuales.

Kusssmaul ilustra este tipo de investigación mencionando a Clyne (1981, 1987, 1991), uno de los representantes de la retórica contrastiva que compara el inglés y el alemán. Por ejemplo, Clyne examinó textos académicos y encontró que difieren en la cultura anglosajona y germánica. Los primeros se caracterizan por su linealidad, por el hecho de que los argumentos y el razonamiento lógico son más lineales. Los textos en alemán contienen más digresiones y tienden a mostrar una cierta erudición en notas a pie de página, citas y referencias al final del texto. Esta información adicional, al presentarla en textos en inglés, normalmente se incluye dentro del cuerpo del texto.

Con la finalidad de objetivar las convenciones retóricas propias del alemán y el inglés y explorar sus aplicaciones al campo de la traducción, Kusssmaul analiza las dimensiones situacionales implicadas en las convenciones macroestructurales. Tras comparar cómo difieren los textos académicos en inglés y en alemán en función de parámetros situacionales como la participación (monólogo o diálogo), la provincia, la actitud social, la individualidad, el dialecto, el tiempo, el medio, el estatus, la modalidad y la actitud social, Kusssmaul se plantea la posibilidad de que existan una serie de *macroestructuras culturales* que han de cumplirse para que el texto sea aceptado en la lengua término. Posteriormente asocia las convenciones macroestructurales a una serie de microestructuras, que también son convencionales.

2.2.6.3. Microestructuras

De las microestructuras que se pueden encontrar en textos académicos, Kusssmaul analiza la forma lingüística de tres tipos, los enunciados metacognitivos (*metacognitive utterances*), la actitud de relatividad y falsa modestia ante el conocimiento científico (*hedging*) y los actos de habla directivos. En su análisis está muy presente la teoría de los actos de habla de Searle (1969).

Las microestructuras que Kusssmaul denomina *metacognitive utterances* tienen como función asegurar la comprensión correcta de los textos (cfr. Göferich 1995:381 y ss.). En el corpus de textos académicos en que se fundamenta su análisis, las microestructuras que más abundan son las que anuncian (*announcements*) y las que

hacen referencia a algo ya mencionado (*back-references*). Además, estas se materializan de un modo distinto en alemán y en inglés. Y así, mientras que en inglés se utiliza el pronombre de primera persona del plural, que sugiere que tanto el autor como el lector se sienten muy implicados en la observación de un fenómeno y la resolución de un problema (ibid), el alemán muestra una preferencia por las formas impersonales. El reconocimiento de la relación entre formas lingüísticas y actos de habla en un texto puede ayudar a identificar tipos textuales, ya que, como indica Kussmaul, es posible que un mismo acto de habla se exprese de forma distinta en géneros distintos.

Kussmaul da buena prueba de esta afirmación cuando analiza las distintas formas lingüísticas asociadas a un tipo textual que contiene principalmente actos de habla directivos: las instrucciones. También estudia cómo estas formas se mantienen o cambian en distintos subtipos textuales, según compartan o no los mismos factores situacionales. Por ejemplo, las instrucciones que aparecen en un frasco de sales de baño contienen formas imperativas en inglés, y construcciones pasivas en alemán. Sin embargo, este imperativo que se emplea en inglés no se puede utilizar en otros textos con fuerza ilocucionaria distinta como las *regulations and rules*.

La explicación de esto la encontramos en el hecho de que un cambio en una dimensión situacional lleva aparejado un cambio en la expresión lingüística. En todos los ejemplos que presenta Kussmaul, se ve muy claro cómo, en traducción, es necesario respetar factores como el estatus y las reglas de cortesía (*politeness*) y la forma lingüística para que el acto de habla se desarrolle de una forma clara en la lengua término. Aquí nos gustaría hacer una aclaración. Kussmaul utiliza la unidad léxica *reglas de cortesía* como las normas de buena educación, y no en el sentido más usual en la lingüística, es decir, lo que es apropiado en una situación comunicativa.

Su enfoque es muy provechoso porque, aparte de presentar algunos tipos de malentendidos y malas traducciones ocasionados por una incorrecta codificación lingüística de los actos de habla, presenta resultados sacados de un corpus de textos en inglés y en alemán. Así, más que ver los errores, se ve qué es lo usual en cada lengua en la línea de la textología contrastiva postulada por Hartmann (1996).

Al traducir microestructuras y macroestructuras siempre aparece el dilema de si se deben preservar las estructuras que aparecen en la lengua origen o si el texto debe

adaptarse a las convenciones de la lengua meta, creando un texto que parezca “normal”. El consejo de Kussmaul es que en la traducción se debe sobreentender el acto de habla que aparece en el texto original, aunque utilizando expresiones idiomáticas de la lengua término. También recomienda tener en cuenta que existen indicadores de la fuerza ilocucionaria característicos de un tipo textual concreto porque lo ha establecido así la convención. Si se siguen las convenciones de la lengua término, se alcanzarán las *condiciones de éxito*.

Asimismo, la fuerza ilocucionaria es sensible a una serie de parámetros como el estatus, la actitud social, y la participación (monólogo, diálogo, etc.). Por este motivo, para que la fuerza ilocutiva del original no se pierda, el traductor debe cumplir las normas de cortesía (*politeness*) de la cultura término.

De entre estas dimensiones, la más relevante a la hora de resolver el dilema arriba mencionado es el *estatus*. Para resolver este dilema el traductor debe sopesar su estatus como traductor, el estatus del autor del texto original y el estatus de los lectores. Kussmaul (ibid: 81) lo expresa del siguiente modo: “As translators we will have to weigh the status of the source-text authors, our own status and the status of the readership carefully against each other when making these kinds of decision”. Como ejemplo, presenta el caso de un autor con reconocimiento en la comunidad científica que utilizaba un idiolecto que rompía las normas de cortesía y describe cómo el traductor de sus obras se podía permitir el lujo de mantener esa violación de la convención.

Como conclusión, podemos decir que Kussmaul reconoce que las convenciones difieren entre tipos de texto diferentes y entre tipos de texto similares pero que pertenecen a distintas culturas. Las convenciones están determinadas por factores situacionales, los cuales son esenciales para definir los tipos de texto y se manifiestan en las macroestructuras y en las microestructuras de los textos. Tanto las macroestructuras como las microestructuras están íntimamente relacionadas con las reglas que rigen los actos de habla, que también son algo propio de una comunidad de discurso. El análisis desde los estudios de traducción de estas convenciones puede resultar muy fructífero. Si este análisis se completa con las investigaciones llevadas a cabo en distintos campos como la textología contrastiva, la lingüística de corpus y la retórica contrastiva (Connor

1995, 1996; Connor y Kaplan 1987), una disciplina que, para nuestra sorpresa, Kussmaul no ha mencionado, los resultados pueden ser aún más fructíferos.

2.2.7. La aproximación pragmático-semiótica de Ian Mason y Basil Hatim

Hatim y Mason (1990, 1997) estudian la traducción desde una perspectiva pragmática y semiótica y enfatizan el papel del contexto socio-cultural y la ideología en la comunicación. Su análisis aplica los principios de la lingüística sistémica-funcional de Halliday (Halliday 1978; Halliday *et al* 1964), la teoría de los actos de habla de Austin (1962) y Searle (1969, 1979), el principio de cooperación de Grice (1975) y el método de análisis de la corriente denominada *Lingüística Crítica*¹⁵.

La palabra contexto designa dos realidades ya señaladas por Firth (1951). Una es el *contexto cultural* y otra es el *contexto situacional* (Cfr. Apartado 2.2.1.). Dentro del contexto cultural está el contexto situacional, que es el ambiente extratextual que moldea, y al mismo tiempo es moldeado por la expresión lingüística.

2.2.7.1. Contexto situacional y *registro*

La consideración del contexto en la traducción pasa por un análisis de registro del texto porque el TT ha de conseguir una equivalencia de registro. Hatim y Mason incluyen bajo el término *registro* no sólo la variedad con respecto al uso (Halliday, McIntosh y Stevens 1964: 87, cfr. Apartado 2.2.1.) sino también la variación lingüística según el hablante. Siguiendo a Halliday, Hatim y Mason (1990: 55) consideran que el estudio del registro pasa por la identificación de tres elementos: *campo*, *tenor* y *modo* del discurso.

[...] identifying the register membership of a text is an essential part of discourse processing; it involves the reader in a reconstruction of context through an analysis of what has taken place (field), who has participated (tenor), and what medium has been selected for relaying the message (mode).

Como ejemplo de esto, presentan un artículo sobre la salud mental desde el punto de vista sociomédico (*campo*), que está escrito por un investigador médico para

¹⁵ Esta corriente aplica el método de análisis de la lingüística sistémica al estudio de la ideología y lo denominan *linguistic criticism* (Fairclough 1992; Fowler 1991, 1996; Fowler *et al* 1979; Kress 1989).

una audiencia semiespecializada (tenor), y que ha sido escrito para ser leído, siguiendo las convenciones de los escritos académicos (modo).

Pese al papel destacado del registro en la traducción, Hatim (1997) y Hatim y Mason (1990) sostienen que no se puede hablar sólo de registro porque esto sería limitar el contexto situacional a su dimensión institucional-comunicativa, cuando realmente el contexto cuenta con tres dimensiones interdependientes:

- a) *la dimensión institucional-comunicativa*, que comprende aspectos del mensaje como la pertenencia a un determinado registro
- b) *la dimensión pragmática*, que quiere dar cuenta de la intencionalidad en el lenguaje y de la posibilidad que este nos brinda de poder hacer cosas con palabras
- c) *la dimensión semiótica*, que explica la intertextualidad y la pertenencia de los textos a un sistema de signos socio-culturales.

2.2.7.2. Las tres dimensiones del contexto

2.2.7.2.1. La dimensión institucional-comunicativa

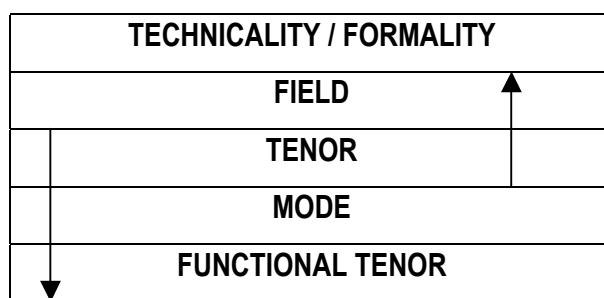
La dimensión institucional-comunicativa¹⁶ engloba los aspectos de la transacción entre emisor y receptor como hablantes de una lengua que utilizan un determinado dialecto. Como quiera que estas variables no son suficientes para explicar la comunicación, hay que ver las limitaciones de uso bajo las que los emisores y receptores se comunican, lo cual nos lleva a considerar el campo, el tenor o grado de formalidad y el modo de interacción.

De estos tres factores, Hatim (1997) defiende que el más importante es el tenor. En esto coincide con Kussmaul (1997), aunque este último lo denomina estatus. Otros términos para designar al tenor son *grado de formalidad*, *estilo* y *actitud*. El tenor es fundamental porque interactúa con el campo y con el modo. Al interactuar con el campo, el texto alcanza un determinado grado de tecnicidad o formalidad. Al interactuar con el modo, el tenor determina con qué finalidad se utiliza el lenguaje, por ejemplo, la

¹⁶ Hatim y Mason (1990) la denominan dimensión comunicativa.

de persuadir o la de informar. A esta última interacción, Hatim le da el nombre de *tenor funcional*. Por la adecuación del tenor con el modo y el campo, podemos decir que las transacciones institucionales-comunicativas adquieren un carácter interactivo. Coincidimos con Hatim en la importancia que tiene el tenor en la configuración de los textos. De hecho, en esta tesis se analizarán las diferencias cohesivas entre dos tipos de texto dentro de la oncología: textos destinados a profesionales de la medicina y textos destinados a pacientes. En estos últimos, el tenor incluye una actitud de objetividad y prudencia con respecto al contenido de los textos que evita la sensiblería y la posibilidad de crear falsas expectativas.

TABLA 2: El tenor en interacción con el campo y el modo. (Hatim 1997: 26)



2.2.7.2.2. La dimensión pragmática

En la dimensión pragmática, lo esencial es la *negociación del significado* entre el emisor y receptor y las intenciones y expectativas que se ponen en juego en la comunicación. Hatim y Mason (1990: 65) escriben al respecto:

The translator, in addition to being a competent processor of intentions in any SL text, must be in a position to make judgements about the likely effect of the translation on TL readers/hearers.

El análisis de la dimensión pragmática lleva aparejado el estudio de los actos de habla del texto, de su fuerza ilocucionaria y perlocucionaria y de cómo el efecto

acumulado de los actos de habla lleva a la percepción de un *acto textual*¹⁷. También conlleva el estudio de las *condiciones de éxito*, de cómo se cumplen y se violan las máximas de Grice y de las *implicaturas* y *presuposiciones* presentes en el texto.

Por otra parte, además de estudiar cómo un texto despliega su significado pragmático, hay que ver cómo los valores pragmáticos que hay en una secuencia de actos de habla ejercen una mutua interacción y se convierten en signos pertenecientes a un sistema general de valores apropiado a una cultura dada.

2.2.7.2.3. La dimensión semiótica

En los apartados anteriores, se ha visto cómo la comunicación se institucionaliza mediante el uso del lenguaje que hacen los hablantes. También se ha visto cómo para poder alcanzar determinados objetivos retóricos, para poder conseguir algo mediante las palabras, es necesaria una negociación del significado, una interacción entre el emisor y el receptor en el contexto mediante signos.

Ahora, en la dimensión semiótica, se produce una interacción a distintos niveles entre el emisor y el receptor con los textos, por una parte, y los textos con otros textos, por otra parte. Mediante esta interacción, los textos y los elementos contenidos en ellos adquieren el rango de signos o constructos semióticos que reflejan la cosmovisión de una determinada cultura y que contribuyen a la estabilización de las relaciones sociales.

Como resultado de la dimensión semiótica, los textos aumentan su significado denotativo y adquieren un significado connotativo de tipo *ideacional*, *textual* e *interpersonal*. Los textos adquieren *significados ideacionales* sobre procesos e instituciones sociales mediante el *campo*, *significados textuales* que nos indican la distancia física entre los participantes mediante el *modo*, y *significados interpersonales* que hacen referencia a relaciones de *poder* y *solidaridad* mediante el *tenor*.

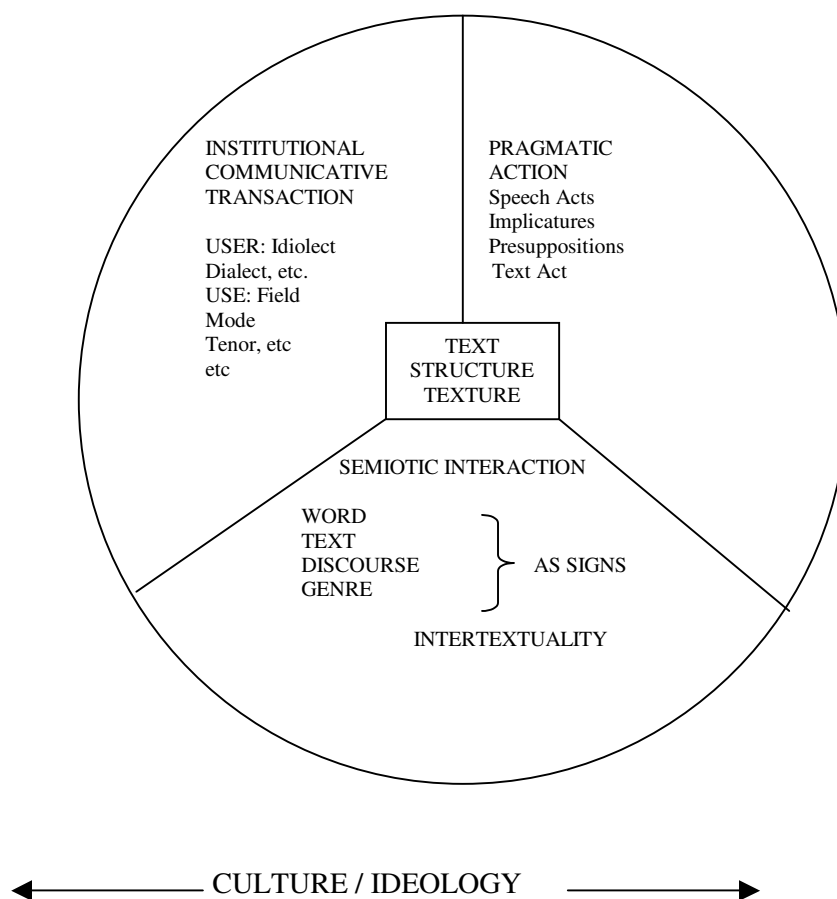
Asimismo, mediante la intertextualidad, vinculamos unos textos con otros que hemos encontrado con anterioridad en contextos parecidos y que son relevantes en nuestra cultura. La intertextualidad es la condición necesaria para la comprensión de los textos que supone que unos textos dependen de otros por el hecho de que son entidades

¹⁷ Horner (1975) define *acto textual* como la fuerza ilocucionaria predominante en una serie de actos de habla.

semióticas. Mediante la intertextualidad, se enlazan los enunciados con una ocasión social codificada en el lenguaje de un modo convencional (*género*), con una afirmación actitudinal (*discurso*) y con una función retórica (*tipo textual*). Estos tres elementos se encuentran en el núcleo de la dimensión semiótica.

A continuación presentamos un diagrama en el que se destaca cómo la dimensión semiótica subyace las otras dos dimensiones del contexto.

DIAGRAMA 4: Las tres dimensiones del contexto



(a) *La transferencia intersemiótica*

La existencia de la dimensión semiótica tiene implicaciones importantes en la traducción. El traductor, al transferir significados de un sistema cultural a otro, tiene que analizar por separado cómo las unidades léxicas, las oraciones y otras unidades

superiores a la oración codifican la dimensión institucional-comunicativa, pragmática y semiótica. Posteriormente, si el encargo de traducción así lo requiere, tiene que preservar las propiedades tanto semióticas como pragmáticas y comunicativas que los signos del TO despliegan. Con todo, el proceso de transferencia intersemiótica no está exento de restricciones genéricas, discursivas y textuales, que determinan cómo se ha de traducir un determinado signo.

(b) Géneros y discursos

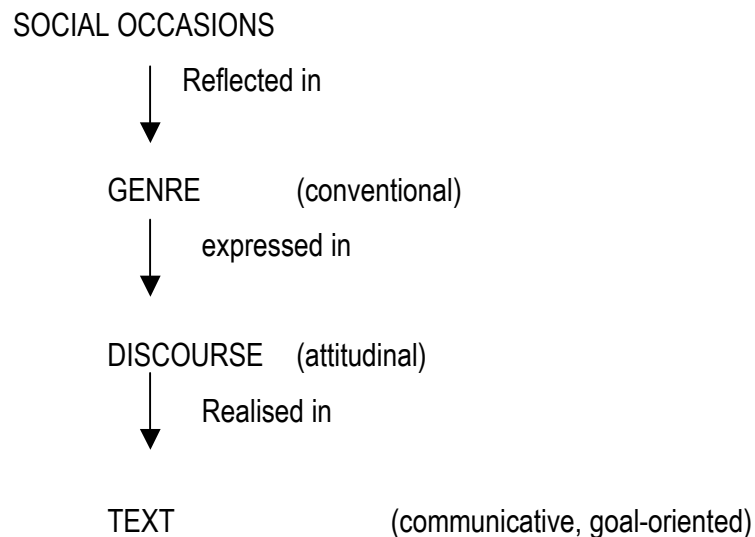
El carácter cultural de la dimensión semiótica hace que se incluyan en esta los géneros y discursos más relevantes en una sociedad. Hatim (1997) define el género como las formas textuales convencionales asociadas a una determinada ocasión social (por ejemplo, el editorial o la receta de cocina). En cuanto al discurso, es un modo de hablar y escribir que lleva a los participantes a adoptar una postura determinada ante ciertas áreas de la actividad sociocultural (por ejemplo, el discurso fundamentalista islámico o el discurso racista). Estas dos unidades están relacionadas en el sentido de que el género refleja situaciones sociales y el discurso refleja la actitud hacia esas situaciones sociales.

Los discursos dependen de fenómenos extralingüísticos porque tratan la ideología y las relaciones de poder, pero se codifican mediante el lenguaje, mediante determinados patrones léxicos, sintácticos y genéricos. Estos patrones lingüísticos consolidan la ideología y la cosmovisión imperante en un grupo social. De hecho, los discursos, al igual que los géneros, pueden llegar a ritualizarse como el discurso sexista o el fundamentalista.

Las últimas publicaciones de Hatim y Mason prestan una gran atención a cómo se enmascara la ideología detrás del lenguaje y cómo el lenguaje consolida determinadas ideologías e influye en nuestro comportamiento y pensamiento. Para descubrir esta ideología, Hatim y Mason (1997) aplica la metodología diseñada por la corriente denominada *Lingüística Crítica* y estudian cómo la gramática de la transitividad, el vocabulario, las transformaciones sintácticas, la modalidad, la deixis y la topicalización, entre otros, actualizan la función ideacional, interpersonal y textual del lenguaje.

En el siguiente diagrama queda claro cómo conciben Hatim y Mason la relación entre el texto, el género y el discurso:

DIAGRAMA 5: Relación jerárquica entre género, discurso y texto (Hatim 1990)



(c) Los tipos textuales

Hatim y Mason señalan que todos los intentos de elaborar una *tipología textual* aplicada a la traducción no han prosperado por no tener en cuenta la multifuncionalidad de los textos. Esta requiere que se tomen en cuenta diferentes perspectivas y variables. Por este motivo, tipologías que clasifican los textos con arreglo a un único criterio como el *campo del discurso* no resultan efectivas porque proponen categorías tan generales como la de *tipo de texto periodístico, religioso o científico*.

Como alternativa, Hatim y Mason proponen una tipología que considere distintos aspectos, pero que sea lo bastante flexible como para acomodarse a la diversidad real. En el modelo de análisis pragmático elaborado por estos autores, se exige que se contemplen tanto los contextos efectivos del uso lingüístico como las intenciones de los usuarios. Y esto es así porque, según estos autores, en cualquier punto concreto de una interacción es posible aislar un foco pragmático, que presumiblemente incluirá un conjunto de intenciones mutuamente relevantes y determinará la configuración de un determinado *tipo textual*. Su definición de tipo

textual es: “a conceptual framework which enables us to classify texts in terms of communicative intentions serving an overall rhetorical purpose” (ibid: 140).

Entre las intenciones que están al servicio del señalado propósito retórico global, identifican cuatro: la de llamar la atención de los lectores, anunciar un tema, expresar apoyo a un proyecto y justificarlo por argumentación.

Normalmente, es posible identificar en los textos más de una intención, fenómeno este que Hatim y Mason denominan *hibridación*. Sin embargo, por el hecho de que la focalización juega un papel esencial en la comunicación y la cognición, en todos los textos se puede percibir un propósito retórico principal y varios propósitos subsidiarios. Como se vio en el apartado 2.2.2.2., Werlich llama a este propósito retórico principal *foco contextual dominante*.

Hatim y Mason recogen la idea de propósito retórico y foco contextual con un solo término: *foco tipotextual*. El foco tipotextual representa los medios comunicativos, pragmáticos y semióticos en virtud de los cuales un texto se relaciona con su contexto y se define como ejemplar de un *tipo textual*.

Hatim y Mason (1990), adaptando la tipología de Werlich (1976), proponen tres tipos textuales básicos:

a) La exposición

Tipo textual en el que se presentan conceptos, objetos y acciones sin emitir una valoración sobre los mismos. Es posible reconocer tres tipos:

- Descripción: se centra en la observación de objetos en el espacio
- Narración: se centra en la presentación de acciones en el tiempo
- Exposición conceptual: se centra en el análisis de conceptos adoptando una postura de distanciamiento frente a los mismos.

b) La argumentación

Tipo textual en el que se valoran conceptos y creencias. Es posible reconocer dos tipos:

- Contraargumentación: en la que se presenta una tesis para ser rebatida.
- Argumentación íntegra: en la que se expone una tesis y se defiende con todos los argumentos posibles.

c) La instrucción

Tipo textual en el que el foco está en la formación de conductas futuras. Dentro de las instrucciones, las hay *con alternativa*, como la presente en textos publicitarios y en consejos al consumidor, y *sin alternativa*, como la presente en contratos y tratados internacionales

Las diferencias entre estos tipos textuales no son totalmente nítidas. De hecho, Hatim y Mason mencionan cómo Reiss (1976) incluye la *instrucción con alternativa* y la *argumentación* dentro de la categoría que ella denomina *textos operativos*.

Por último, Hatim (1997) reconoce que cada uno de estos tipos textuales lleva asociado una estructura textual, el plan de composición del texto. Por ejemplo, en una contraargumentación se puede detectar la siguiente estructura: *tesis, objeciones a esa tesis, confirmación y conclusión*. El reconocimiento de esta estructura en el texto se debe en cierto modo a que los hablantes las tienen interiorizadas y a que, según Werlich (1976) y Beaugrande (1981), los tipos textuales están asociados a estructuras cognitivas como los *marcos, guiones y escenarios*. Se corroborará esta opinión en los capítulos 11 y 12, donde se comparan marcos conceptuales identificados a partir de dos tipos textuales diferentes.

2.2.7.3. Tipos textuales y textura

Para que la estructura que define los tipos textuales sea efectiva y apropiada es necesario explotar los recursos lingüísticos que *tejen* el texto, la *textura*. Por este motivo, Hatim (1997: 76) concede una gran importancia a la *textura* de los textos, a la que define como el conjunto de aspectos organizativos y cohesivos del texto que contribuyen a que el texto constituya una unidad coherente. En la configuración de la textura de un texto, el elemento que juega un papel más importante es el tipo textual. En este planteamiento coincide con la hipótesis que se desarrolla en esta tesis de que existe una relación entre cohesión y tipología textual.

La textura, es decir, la realización lingüística de valores contextuales puede ser más o menos explícita de forma que puede variar de un tipo textual a otro y de unas lenguas a otras. De hecho, la lengua inglesa tiene una textura implícita en comparación con el árabe, a la que Hatim (ibid: 100) atribuye una textura explicativa. Estas

diferencias en cuanto a la textura pueden hacer que, a la hora de traducir, se planteen problemas en cuanto a la cohesión y la coherencia de los textos. Se explorará esta área en la sección 4.11, aunque en relación a textos pertenecientes al dominio de la biomedicina, a las lenguas inglesa y española y a las diferencias entre textos originales y textos traducidos.

2.3. Recapitulación

En los estudios de traducción, el término *tipo de texto* o *tipo textual* ha sido uno de los más populares junto con el de *equivalencia*. Esto se debe a que el proceso cognitivo de la traducción y las estrategias implicadas en el mismo están condicionadas por las características textuales del texto origen, por una parte, y por las características textuales que se esperan del texto traducido, por otra. Estas dos facetas, la tipología de textos aplicada al texto origen y la tipología de textos traducidos, son las que guían los principales estudios sobre tipología de textos en traducción.

Al revisar dichos estudios, hemos encontrado que, en cierto modo, utilizan la traducción como banco de pruebas de teorías lingüísticas surgidas de la interdisciplinariedad y plenamente consolidadas. Podríamos decir que la traducción se empapa de las corrientes en boga y evoluciona conforme a estas, con un ligero retraso temporal. Al igual que ocurre en la lingüística, las tendencias siguen el movimiento del péndulo. Se pasa de la primacía de la equivalencia a la de la función o *skopos*, de la primacía del texto origen a la primacía del texto término. Sólo desde la escuela de Leipzig con Neubert, Snell-Hornby y Kussmaul, se empiezan a buscar teorías integradoras que reconocen la importancia tanto del texto origen como del texto término y la posibilidad de aunar aspectos cognitivos y textuales.

A la primera tipología textual, elaborada por Reiss basándose en las funciones del lenguaje de Bühler, se le añadirán tipologías estrictas de base cognitiva que pretenden descubrir la estructura de los textos. Estas aproximaciones se nutren de las aportaciones de la lingüística textual alemana. Al aplicar las tipologías a textos concretos se percibe que no son categorías cerradas y que se debería aspirar a elaborar *prototipologías*, una idea que responde a la repercusión de la teoría de prototipos de Eleanor Rosch en la lingüística. Esta posición, iniciada por Neubert y Snell-Hornby,

supone el comienzo de una tendencia a la relativización, al eclecticismo y al funcionalismo. A esta relativización también contribuye la aparición, a raíz de la internacionalización de la ciencia, de nuevos tipos textuales, lo cual genera nuevos tipos de traducciones, y ocasionan la introducción de nuevos tipos de texto en la cultura término.

Otra gran influencia en los estudios sobre tipología textual la constituye la lingüística derivada de los postulados de Firth y Halliday, que concede una gran importancia al contexto, al lenguaje como acción y a la semiótica social. Estas ideas y las de Austin y Searle encuentran eco en la teoría del *skopos* (Vermeer y Reiss), en las aplicaciones de la teoría de los actos de habla a la traducción (Snell-Hornby, Kussmaul, Hatim y Mason) y en la teoría funcional de Nord. Esta última aúna el criterio de *lealtad* al texto origen con el de *finalidad* del texto meta.

En todas estas aproximaciones, el término *tipo de texto* o *tipo textual* se utiliza como un hiperónimo que designa distintas realidades asociadas a la variedad textual y a la variedad en la práctica de la traducción. También se utiliza para designar el propósito retórico de un texto. En el ámbito anglosajón, se han sistematizado estas diferencias de carácter textual con la introducción de los términos *registro* y *género*. La aplicación de estos conceptos a la traducción y el estudio de sus implicaciones semióticas e ideológicas vienen de la mano de Hatim y Mason.

La falta de consenso en la terminología sobre tipología textual quizá se deriva de la complejidad de elementos lingüísticos y extralingüísticos que se interrelacionan en la creación de textos. A este respecto, Beaugrande (1980: 197) escribe:

A text type is a distinctive configuration of relational dominances obtaining between or among elements of (1) the surface text; (2) the textual world; (3) stored knowledge patterns; and (4) a situation of occurrence.

En consecuencia, estimamos que cualquier estudio sobre tipología textual exige una aproximación interdisciplinar que considere los aspectos cognitivos, lingüísticos y sociales implicados en la producción y recepción de textos. Estos aspectos pueden explicar por qué situaciones comunicativas que son más frecuentes de lo normal en una

cultura determinada originan tipos textuales convencionales denominados géneros, de los que hablaremos en el capítulo siguiente.

3. La variación textual en la lingüística y la lengua para fines específicos

En el capítulo anterior, han aparecido una serie de términos que con frecuencia se confunden entre sí. En efecto, la distinción entre *registro*, *tipo textual* y *género* no ha sido sometida a consenso en el ámbito de la lingüística y la traducción. Esto se debe en parte a que son fenómenos que se superponen. Por ejemplo, el género epistolar y sus distintos subgéneros pueden actualizarse a través de distintos registros (la carta a los Reyes Magos publicada en un periódico con fines reivindicativos, la carta de amor a un/a desconocido/a, la carta comercial a un viejo amigo moroso, etc.), según el tema, los participantes en la comunicación y el modo en que se expresa el contenido. Asimismo, la colaboración interdisciplinar podría haber sido mayor, con lo que la traducción se hubiera beneficiado de investigaciones sobre tipología textual bastante consolidadas en áreas de la lingüística como la lengua para fines específicos o *LSP* (Swales 1990; Bhatia 1993).

La lengua para fines específicos tiene muchos puntos en común con la traducción y la interpretación. Tanto la lengua para fines específicos como la traducción enfatizan la idea de finalidad y de necesidades comunicativas. En la primera, estas necesidades comunicativas se pueden explicitar mediante un análisis de necesidades, mientras que en la segunda, es el encargo de traducción el que indica qué necesidades comunicativas pretende suplir el texto meta. Asimismo, en ambas disciplinas se trabaja con lenguas de especialidad o tecnolectos y hay un gran interés por las convenciones retóricas propias de una comunidad de discurso. Este interés ha cristalizado en investigaciones sobre el concepto de *género*, término preferido en lengua para fines específicos, frente al de *tipo textual*, término preferido en los estudios de traducción.

Antes de proponer una definición de los términos *registro*, *tipo textual* y *género*, resumimos las opiniones sobre un aspecto que subyace cualquier distinción entre tipos textuales y entre géneros. Es el de la concepción del texto como estructura o como organización.

3.1. El texto ¿estructura u organización?

Las principales perspectivas sobre la configuración del texto se pueden subsumir en dos tendencias: el texto como estructura o como organización. Descartamos de plano la idea de que el texto no tiene organización porque nuestra experiencia cotidiana y la tradición retórica que surgió a partir de la cultura griega parecen confirmar lo contrario.

3.1.1. El texto como estructura

La organización del texto permite una descripción estructural, al igual que la oración puede ser descrita estructuralmente. Esta idea es el fundamento de la gramática textual y la lingüística textual, que utilizaron la estructura de la oración para explicar la estructura del texto. Algunos trabajos fundamentados en esta aproximación son Van Dijk (1972), Grimes (1975), Pike y Pike (1977), Van Dijk y Kintsch (1978), Longacre (1983) y Graustein y Thiele (1987). En estos, la descripción estructural del texto suele ser prescriptiva y establece qué combinaciones de oraciones no son posibles, sobre todo, desde el punto de vista lógico.

El interés se encuentra en indagar cómo se representa el conocimiento en la memoria y cómo esta representación del conocimiento influye en la comprensión del discurso. El receptor de un mensaje lo comprende en función de unos esquemas cognitivos (*schemata*). Estos esquemas han sido estudiados en detalle por Van Dijk (1972, 1977, 1980).

Van Dijk (1977) desarrolló los conceptos de *macroestructuras*, *microestructuras* y *superestructuras*. Las *macroestructuras* representan la organización semántica global del discurso, el tema o el asunto del discurso. Las *microestructuras* denotan la estructura local de un discurso, es decir, la estructura de las oraciones y las relaciones de conexión y de coherencia entre ellas. El reconocimiento de las *macroestructuras* de un discurso se produce en función de unas reglas de proyección semántica (*macrorreglas*) que vinculan las proposiciones de las microestructuras textuales con las de las macroestructuras textuales. Por último, el significado contenido en las *macroestructuras* está relacionado sistemáticamente con una estructura formal global: las *superestructuras*. Las superestructuras constituyen la forma global del discurso que define la organización general del mismo y las relaciones jerárquicas de sus respectivos

fragmentos. Un ejemplo de superestructura lo hallamos en el esquema narrativo de los cuentos que comprende la introducción, la complicación, la resolución, la evaluación y la moraleja. Esta superestructura está vinculada a su macroestructura en el sentido de que, por ejemplo, la resolución de un cuento trata sobre la acción que concluye la narración. La investigación de la superestructura que ha realizado Van Dijk se ha centrado en narraciones, anuncios publicitarios y noticias periodísticas. Otro ejemplo de superestructura pero dentro del ámbito médico lo constituye la nota clínica, que cuenta con una sección de fundamentos, otra que presenta una observación clínica y por último, una sección de conclusiones.

En definitiva, en esta perspectiva se apuesta por descubrir la estructura del texto en función de unas estructuras cognitivas que el lector identifica en los textos para poder comprenderlos e interpretarlos. El lector percibe en el texto una estructura al confrontar el texto con las macroestructuras almacenadas en su memoria. El punto débil de esta aproximación es que, al conceder a los procesos cognitivos la primacía, el contexto y la interacción comunicativa quedan en un segundo plano.

3.1. 2. El texto como organización

Frente a esta concepción del texto, desde la lingüística sistémica se aboga por un mayor hincapié en el contexto y en la interacción. De aquí nace la idea de que el texto está organizado pero no constituye estructuras rígidas. Esta es la opinión compartida por dos tendencias dentro de la lingüística sistémica:

- El *análisis relacional de cláusulas* (*clause relational analysis*) de Winter y Hoey
- Los estudios sobre *género* (*genre analysis*)

3.1.2.1. El análisis relacional de cláusulas (Winter y Hoey)

La investigación sobre la organización del discurso realizada por Winter (1974, 1976) y posteriormente, por Hoey (Hoey 1979, 1983; Winter y Hoey 1986) pone el énfasis en el contexto y en cómo la gramática de las oraciones contribuye a su interpretación en contexto y a la articulación del *discurso*, término que estos autores utilizan en lugar del

demasiado explotado *texto*¹⁸. La interacción comunicativa cobra importancia (Hoey y Winter 1986):

[...] text analysis that fails to recognize the interactive nature of written discourse will be inadequate in some respects.

La visión de la organización textual de estos autores deriva de los siguientes postulados elaborados por Winter (*apud* Hoey 1983: 15 y ss.):

- 1) El discurso es el producto de relaciones semánticas entre oraciones y proposiciones. El discurso está organizado en parte de forma jerárquica. Esta organización no es inmutable.

[...] a discourse may be made up of clause relations which are themselves members of larger clause relations which are in turn members of an overall relation. (Hoey 1983: 32)

- 2) Hay un consenso entre los hablantes de una lengua sobre cómo se organizan los discursos y las secciones que se pueden encontrar en un discurso.

[...] people are capable of complaining of a play that it did not have 'a beginning, a middle or an end' or of a politician's answer that 'he went off the point'. (Hoey 1983: 33)

- 3) Las oraciones contiguas se enlazan entre ellas mediante repetición y anáfora, pero estos recursos no logran explicar la organización de los discursos.
- 4) Los discursos no están contruidos de oraciones individuales que se van yuxtaponiendo para construir párrafos. Los discursos tampoco son cadenas de párrafos.

¹⁸ Para referirse a una sección del discurso, Winter y Hoey utilizan la palabra *passage* en lugar de *texto*

El trabajo de Eugene Winter cubre dos facetas. Por una parte, pretende demostrar cómo la gramática y el significado de las oraciones sólo pueden ser entendido en su cotexto y en su contexto. Por otra parte, pretende identificar las relaciones *clausales* que aparecen en el texto y ver cómo están organizadas estas relaciones (*clause relational organisation*) en un fragmento.

A clause relation is the cognitive process, and the product of that process, whereby the reader interprets the meaning of a clause, sentence, or group of sentences in the context of one or more preceding clauses, sentences, or groups of sentences in the same discourse. It is also the cognitive process and the product of that process whereby the choices the writer makes from grammar, lexis, and intonation in the creation of a clause, sentence, or group of sentences are made in the context of the other clauses, sentences, or groups of sentences in the discourse. (Winter 1971)

Existen dos tipos de relaciones clausales. En las primeras, denominadas relaciones de comparación (*matching relations*), tanto el emisor como el receptor intentan procesar los fragmentos de información del texto como elementos que se pueden comparar. Es decir, intentan detectar las diferencias y las similitudes en la información presentada. El listado que Winter propone de estas relaciones es el siguiente: *contrast, compatibility, generalization-example, preview-detail* y *topic maintenance*.

En el segundo tipo de relaciones denominadas relaciones de secuencia lógica (*logical sequence relations*), los fragmentos de información del texto están conectados en una secuencia lógica. Es decir, se llama la atención sobre diferentes relaciones: *cause-consequence, condition-consequence, evaluation-basis, instrument-achievement* y *time sequence*.

Como resultado de estas relaciones, los fragmentos de texto se conciben como una totalidad, como algo más que la suma de sus partes. En esto concuerda con el principio de *gestalt* (véase sección 2.2.3.2.).

Estas relaciones se explicitan en el texto mediante recursos gramaticales y léxicos. Dentro de estos recursos léxicos, uno fundamental es la repetición, de la que hablaremos en el capítulo 4. Hoey (1979) muestra cómo algunos lexemas (*reason,*

difference, example, result, achieve, compare, contrast) indican la organización del discurso. Estos recursos lingüísticos que, sin ser conectores conjuntivos, mantienen al igual que estos el desarrollo argumental de los textos (*flow*)¹⁹ han recibido diferentes denominaciones: *discourse-organising words* (McCarthy 1991), *vocabulary 3* (Winter 1977 y 1978), *signalling words* (Jordan 1984), *anaphoric nouns* (Francis 1986: 15) o *procedural vocabulary* (Widdowson 1983, Robinson, 1988)²⁰. Definiremos pues el *vocabulario organizador del discurso* como las palabras o expresiones de un estatus intermedio entre las palabras plenas (*content words*) y las palabras forma (*function words*) que no sólo representan segmentos de texto, sino que también informan al lector sobre la estructura textual (*general - particular, cause- consequence, hypothetical - real*)²¹ que el autor ha elegido. En la parte de *Resultados* de esta tesis, se verá la función de este vocabulario en la cohesión del texto.

Sin embargo, estos tipos de organización basados en relaciones clausales no explican la organización global del discurso. Por eso, Winter postula que, si se combinan estos distintos tipos de relaciones, se consigue llegar al nivel discursivo. En este nivel, identifica una lista abierta de organizaciones discursivas. Algunas de las más frecuentes son: *problem-solution, general-particular, matching- contrast* y *matching-compatibility*.

El primero de estos patrones retóricos pasa del planteamiento de un problema a la propuesta de una solución y contiene los siguientes elementos, aunque está sujeto a variaciones: *Situation-Problem-Response-Result-Evaluation*. Cada uno de estos se presenta en el texto como una relación causal. Y así, el *problema* normalmente se presenta como una relación *causa-consecuencia*.

El patrón *problema-solución* juega un papel decisivo en la parte del artículo científico cuya escritura plantea más dificultad: la introducción (Zappen, 1983). Este investigador propone cinco elementos en la introducción de un artículo científico: *goal, current capacity, problem, solution* y *criteria of evaluation*. En una línea muy parecida se sitúa el modelo *Create a Research Space (CARS)* de Swales (1990), en el que se sostiene que la introducción ha de destacar la relevancia del área en que se enmarca la

¹⁹ Swales y Feak (1994: 21).

²⁰ McCarthy (1991: 76-78).

investigación y la existencia de una parcela (*niche*) que el estudio en cuestión pretende ocupar aportando una contribución original. Según Swales (1994: 175), la introducción del artículo científico normalmente presenta la siguiente estructura:

- Move 1 Establishing a research territory
 - a. by showing that the general research area is important, central, interesting, problematic, or relevant in some way (Optional)
 - b. by introducing and reviewing items of previous research in the area (Obligatory)
- Move 2 Establishing a niche (Obligatory)
 - a. by indicating a gap in the previous research, raising a question about it or extending previous knowledge in some way
- Move 3 Occupying the niche
 - a. by outlining purposes or stating the nature of the present research (Obligatory)
 - b. by announcing principal findings. (Optional)
 - c. by indicating the structure of the RP. (Optional)

La visión del texto de Hoey se sitúa en la línea de Winter. Para Hoey, la percepción del texto como una organización, en lugar de como una estructura, presupone que los emisores proporcionan a sus receptores una serie de patrones que, aunque no obligatorios, tienen una probabilidad de ocurrencia muy alta. Estos patrones del texto son formas de organización socialmente aceptadas y pueden ser descritas como relaciones semánticas entre oraciones y grupos de oraciones.

No hay un límite al número de patrones de organización y de hecho, un inventario cerrado de las mismas sería imposible, aunque las relaciones de las que se componen son limitadas. Los hablantes de una lengua están continuamente produciendo nuevos patrones organizativos a partir del listado limitado de relaciones disponibles.

Sobre cómo se construye esta organización, Hoey (1991) explica que la cohesión léxica es el fundamento de cualquier organización textual. La cohesión léxica es el recurso cohesivo más frecuente en todos los textos (ibid: 9) y pone de manifiesto la relación tan estrecha que existe entre léxico y el texto. Sobre este particular, Hoey (ibid: 217) afirma:

²¹ McCarthy (1994).

Every lexical selection affects or creates cohesive links that [...] help to organize the text; patterns of organization of a more conventional kind, such as problem-solution patterning, likewise only have reality in so far as they are made real by lexis. Conversely, relations between lexical items, for example, sets, collocations, are a function of their appearance in text. Furthermore, each textual selection constrains the lexical choices possible, and it is in the combination of the lexical and textual choices that writers or speakers make that their creativity is expressed.

De esta conexión tan estrecha entre léxico y texto, Hoey deduce que, al igual que podemos describir el léxico en función de la intersección entre el eje paradigmático y sintagmático, también podemos describir el texto en función de las relaciones horizontales que se establecen entre las oraciones y las opciones paradigmáticas abiertas en el texto (ibid: 219). La interrelación entre léxico y texto subyace la aplicación pragmática del Modelo Lexemático-Funcional de Martín Mingorance (1998) presente en Jiménez (1994). Jiménez defiende que cada lexema es un texto y que su definición lexicográfica puede indicar en qué tipo de texto puede aparecer.

3.1.2.2. Los estudios sobre género

La concepción del texto como organización también la defienden los estudiosos del género como Martin (1985), Swales (1985) y Ventola (1987), que sostienen que la organización depende tanto del contexto como de las características internas del texto. Es decir, cada género es el resultado de:

- una situación comunicativa convencional que impone una organización retórica bastante estricta. Un ejemplo de esta organización lo hallamos en la exigida en las normas de publicación de revistas biomédicas.
- una conjunción de elementos lingüísticos que tejen una organización flexible. Aquí hay que hablar de opcionalidad y probabilidad, en lugar de obligatoriedad.

La aplicación del concepto de *género* a los estudios de lingüística surge al constatar la capacidad que tienen los hablantes de una lengua de reconocer una serie de características y funciones que están prototípicamente presentes en algunos textos y

que están asociadas a situaciones sociales habituales en una determinada cultura, por ejemplo, el artículo experimental, las instrucciones de uso y las recetas de cocina.

Swales (1990: 58), desde el área del inglés para fines específicos, define el concepto de *género* del siguiente modo:

A class of communicative events, the members of which share some set of communicative purposes. The purposes are recognised by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constraints choice of content and style. [...] In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience. If all high probability expectations are realised, the exemplar will be viewed as prototypical by the parent discourse community.

Esta definición destaca los siguientes aspectos: función comunicativa, organización y prototipicidad, en el sentido de que cada texto es una instancia de un determinado género, que se considera central en una determinada comunidad de discurso.

Bhatia (1993) adopta esta definición e intenta identificar una relación entre la función comunicativa de un género determinado y su estructura cognitiva típica. Es decir, pretende reconocer las estrategias que el emisor utiliza para que su texto sea más efectivo. Según Bhatia (ibid: 19-20) existen unas regularidades en la organización de los géneros de carácter cognitivo. Estas regularidades reflejan las estrategias que los miembros de una comunidad de discurso o de una comunidad profesional utilizan en ese género con unas funciones comunicativas específicas. Esta organización cognitiva tiene un carácter colectivo, puesto que refleja el conocimiento social y convencional acumulado en un tipo de discurso o en una comunidad profesional.

La idea de función social y de secuencialidad ya estaban presentes en definiciones provenientes de la lingüística sistémica desarrollada en Australia, principalmente, Martin (1984) y Kress (1985). Estos autores hablan del principio de elección en el nivel léxico-gramatical y de cómo este tiene implicaciones ideológicas. También introducen la idea de estructura, en el sentido de que todos los géneros constan

de una serie de etapas. Según Martin (1984: 25), el género es “a staged goal-oriented purposeful activity in which speakers engage as members of our culture”.

En definitiva, un determinado género es principalmente el resultado de dos elementos. En primer lugar, es el resultado de una situación comunicativa convencional que impone una organización retórica formada por *etapas* o *movimientos* (Swales 1990) que responde a la función social que va a cumplir el género. Esta organización retórica puede ser bastante estricta, como la de los artículos experimentales médicos. Las normas de publicación de manuscritos enviados a publicaciones científicas, elaboradas por el *International Committee of Medical Journal Editors* imponen que el manuscrito presentado conste de los siguientes componentes²²:

- primera página
- resumen y palabras clave
- texto
- agradecimientos
- bibliografía
- tablas e ilustraciones

También imponen que las secciones del texto coincidan con el formato *IMRAD*, acrónimo para *Introduction, Methods, Results And Discussion* (Cfr. Swales 1990).

En segundo lugar, los géneros son el resultado de una conjunción de elementos lingüísticos que tejen una organización flexible. Aquí entran en juego la opcionalidad y la probabilidad, en lugar de la obligatoriedad.

Sería interesante investigar cómo se determinan esas etapas o segmentos textuales; especular sobre si han sido establecidas simplemente por convención o si existe una base cognitiva que las determine (Van Dijk 1980; Neubert, 1981). En consonancia con Bhatia (1993) y Paltridge (1994), pensamos que ambas perspectivas son complementarias.

²² Huth (1987) y Puerta y Mauri (1995).

3.1.2.3. Recapitulación

Haciendo balance de los dos grandes enfoques sobre el texto, el texto como estructura o como organización, opinamos que el segundo enfoque es más concreto porque le da una mayor importancia a la *forma* en que la superficie del discurso aporta pistas al lector o al oyente para que perciba de un modo exacto la organización del discurso. Asimismo, al reconocer que, aunque estas pistas sean limitadas, no lo son los patrones de organización del discurso, esta perspectiva da cabida a todas las instancias de género. De esta forma es posible juzgarlas en base a su adecuación respecto a las normas retóricas más aceptadas en una determinada comunidad de discurso.

Por otra parte, desde la lingüística del texto, la supremacía de factores cognitivos hace que no se le preste la debida atención a factores contextuales y a la interacción emisor-receptor. Por eso, no incluyen como objeto de estudio estructuras que contradicen nuestros esquemas mentales sobre el mundo, las cuales sí estarían incluidas en un enfoque *hallidayano*, como ocurrencias poco probables, pero seleccionadas con una clara función comunicativa.

De todas formas, como quiera que en el procesamiento del lenguaje se producen tanto procesos *bottom-up* como *top-down*, opinamos que estas perspectivas son complementarias. Es decir, la comprensión de los textos se alcanza gracias a unas macroestructuras o patrones de organización almacenados en la memoria que al confrontarlos con la superficie del discurso proporcionan un sentido al discurso. Al mismo tiempo, la superficie del discurso aporta la información necesaria para llegar a una organización del discurso que no siempre se amolda con patrones previos. Mediante procesos *bottom-up*, podemos adquirir nuevas formas de comunicación más efectivas y reconocer variantes de patrones previos almacenados en la memoria.

3.2. Registro, tipo textual y género

El concepto de *registro* intenta explicar el hecho de que utilicemos el lenguaje de un modo diferente en distintas situaciones (Halliday 1985). Intenta englobar las características lingüísticas de un texto que son el resultado de la interacción de tres factores situacionales: el *campo* (asunto), el *tenor* (relación social entre el emisor y el receptor) y el *modo* (forma en que se organiza el contenido que se está comunicando).

Según Trosborg (1997: 10), el registro impone restricciones en el nivel léxico y sintáctico, en contraposición al *género*, cuyas restricciones afectan la estructura discursiva del texto.

El término *tipo textual o tipo de texto* es muy habitual en los estudios de traducción como cajón de sastre el que se incluyen instancias de variación textual según el tema (legal, científico-técnico, literario), el tono (formal/informal), el medio (oral/escrito/escrito para ser leído), la función retórica (informar/evaluar/convencer) y la función en el contexto cultural (receta de cocina, editorial, noticia). Nuestra revisión del término *tipo de texto* en los estudios de traducción nos lleva a pensar que se puede entender en el sentido amplio que acabamos de mencionar o en un sentido restringido. Según este último, el concepto *tipo textual* designa el propósito retórico principal que domina un texto concreto. Según este propósito retórico dominante, los textos pueden ser expositivos, argumentativos y exhortativos. Este es el sentido que dan Werlich (1976) y Hatim y Mason (1990) al término *tipo textual*.

Aunque nos gustaría evitar la polisemia de los términos, en este trabajo de investigación, utilizamos *tipo textual* en el sentido amplio, que es el que se le da en los estudios de traducción.

Por último, el término *género* designa tipos textuales convencionales que se originan en situaciones comunicativas que son más frecuentes de lo normal en una cultura determinada para cumplir fines concretos. En este concepto se produce una conjunción entre lo cognitivo y lo social en el sentido de que se hacen convencionales aquellas organizaciones del texto que se acomodan más a las características de nuestra mente y que resultan más eficientes comunicativamente. Estas organizaciones se almacenan en nuestra memoria semántica y, al activarlas con frecuencia en la producción y recepción de textos, se convierten en parte de nuestro conocimiento procedimental. De esta forma se construye una comunidad de discurso y se imponen una serie de valores en los miembros de un círculo profesional o social (Bazerman y Paradis 1991). En este proceso de “domesticación”, que busca el consenso en aras de la eficacia, interactúan no sólo elementos lingüísticos, sino también tipográficos, epistemológicos y organizativos. Estos nos hacen percibir la diferencia entre distintos géneros.

3.3. La variación textual en el inglés con fines médicos

La necesidad de establecer una *tipología de textos* aplicada a la traducción es claramente acuciante en la traducción de textos pertenecientes al dominio de la biomedicina. En este dominio, la especialización del traductor exige la identificación de registros, géneros y tipos textuales dentro de la biomedicina, sin la cual, el texto traducido perdería su función comunicativa. Estos llevan asociados una serie de recursos morfosintácticos y tipográficos y una determinada organización retórica.

La comunidad médica distingue una serie de tipos textuales a los cuales se tienen que amoldar todos los trabajos científicos que quieran publicarse en una revista biomédica. La mayoría de los tipos de texto reconocidos en las ciencias médicas y de la salud se forjaron en lengua inglesa y, por el carácter internacional del saber científico, se han adaptado al resto de lenguas que producen literatura médica. Según Maher (1986), al convertirse el inglés en una *lingua franca* en la comunicación biomédica, que permite el acceso a la información, esta lengua es una especie de código entre los miembros del círculo profesional de la medicina.

A esta introducción de nuevos géneros en lenguas vernáculas ha contribuido el hecho de que los investigadores y facultativos del campo de la medicina lean los últimos avances de su especialidad en publicaciones internacionales con un elevado índice de impacto. Teniendo en cuenta que estas se escriben en inglés, sus lectores van asimilando poco a poco las convenciones retóricas de estos géneros con la aspiración de que algún día sus trabajos aparezcan en una de estas publicaciones. En consecuencia, al comparar revistas especializadas médicas en lengua inglesa y española, los géneros que aparecen en el índice de contenidos son prácticamente los mismos, al igual que su estructura global.

Numerosas publicaciones pertenecientes al área del Inglés con Fines Médicos han tratado los principales formatos que siguen los textos médicos. Por ejemplo, Adams Smith (1984) reconoce tres categorías de artículos que se publican en revistas médicas:

- *clinical case notes / short report*: este género breve consta de tres partes: *introduction*, *case report (methods)* y *comment*. Estas secciones siguen el orden de las notas que el médico toma sobre su paciente.
- *research papers (28-30)*: dentro de este género y según el tema, reconoce entre artículos basados en ensayos clínicos o basados en análisis epidemiológicos. Los primeros tienen una estructura más estricta (*Summary and conclusions, Introduction, Methods and Results, Discussion*). En los segundos hay una revisión de casos basados en historiales médicos y entrevistas, y su estructura es menos estricta.
- *editorials*: este género no cuenta con apartados, tiene una estructura muy libre y abunda en expresiones evaluativas.

Salager-Meyer (1994: 152) apunta que los géneros fundamentales en el ámbito de la medicina son *editorials*, *review articles*, *research papers* y *case reports*. Los dos primeros son más argumentativos, persuasivos y evaluativos, de ahí que abunden en ellos expresiones para articular matices de certeza (*hedging*). Los *research papers* y *case reports* son más informativos y descriptivos. Para más detalles sobre estos géneros también se puede consultar Salager-Meyer *et al.* (1989), Salager-Meyer (1991), Thomas y Hawes (1994), Mirador (1995) y Webber (1994).

Nwogu (1991) estudia un género al que denomina *journalistic reported version (JRV) of professional medical research reports*. Con este término designa las adaptaciones de artículos experimentales para profesionales de la medicina que aparecen en textos divulgativos: revistas de divulgación científica y periódicos. El autor reconoce los movimientos (*moves*) que suelen aparecer en ellos, los cuales son sensibles a condiciones pragmáticas tales como lector, función y medio de discurso. El autor de este género puede ser bien un científico, bien un periodista científico. La divulgación científica también será el objeto de estudio de Myers (1991) y Fernández Polo (1999).

Con posterioridad, Nwogu estudia la estructura y los movimientos del artículo experimental en su totalidad. Supone un avance pues con respecto a estudios anteriores que sólo han estudiado secciones de este género (Nwogu 1997: 120). A partir de un corpus de 30 textos publicados en cinco revistas médicas (*The Lancet*, *British Medical*

Journal, *The New England Journal of Medicine*, *The Journal of Clinical Investigation* y *The Journal of the American Association*) identifica una estructura esquemática con once movimientos. Cada movimiento es "un segmento textual compuesto de un conjunto de rasgos lingüísticos que otorgan uniformidad al segmento y señalan su contenido discursivo" (ibid: 122).

TABLA 3: Movimientos y funciones discursivas presentes en los artículos experimentales (Nwogu 1997: 125)

Move	Discourse function	Section
1	Presenting Background Information	Introduction
2	Reviewing related Research	
3	Presenting New Research	
4	Describing Data Collection Procedure	Methods
5	Describing Experimental Procedure	
6	Describing Data-Analysis Procedure	
7	Indicating Consistent Observations	Results
8	Indicating Non-Consistent Observations	
9	Highlighting Overall Research Outcome	Discussion
10	Explaining Specific Research Outcomes	
11	Stating Research Conclusions	

Pilegaard (1997: 159-160) reconoce tipos textuales que cambian según el grado de tecnicidad (textos más o menos técnicos), el campo (textos de medicina general o medicina especializada), el canal de comunicación (oral, escrito) y el grado de formalidad o el tenor. Con respecto al tenor, Pilegaard, siguiendo a Lankamp (1989: 21) reconoce cinco tipos textuales:

- *Textbooks*, que tienen una intención pedagógica
- *Journal articles*, que ponen en comunicación a los expertos
- *Popular medicine*, que acerca la medicina al público en general mediante el periodismo
- *Doctor / patient language*
- *Manuals*, que aportan al especialista conocimientos sobre los aspectos técnicos de la medicina

En cuanto al concepto de género, este autor opina que la elección de lexemas en la lengua término a menudo varía en función del género y elabora un listado con los géneros médicos tradicionales, cuyo significado explicaremos con posterioridad. Entre ellos existen diferencias en el nivel léxico, sintáctico, estilístico, pragmático y cognitivo.

- *Case report*
- *Disease review*
- *Treatment-focused report*
- *Research paper / Experimental report*
- *Dissertation*
- *Medical textbook*
- *Speech / printed version of speeches*

3.3.1. Géneros en la comunicación entre especialistas

Hemos creído conveniente contrastar las clasificaciones previas con la opinión de los especialistas²³ y las categorías que aparecen con más frecuencia en publicaciones médicas de mayor índice de impacto (entre otras, *New England Journal of Medicine*, *The Lancet*, *The Journal of the American Association*, *British Medical Journal*, *Proceedings of the Mayo Clinic*) y en *Medline*, la base de datos bibliográfica de la

²³ Los especialistas consultados durante la elaboración de esta tesis doctoral han sido el Dr. Nicolás Olea Serrano del Departamento de Radiología de la Universidad de Granada, el Dr. Pedro Ballesteros García del Servicio de Oncología del Hospital Universitario Virgen de las Nieves y la Dra. M^a Rosa Montes Ramírez del Departamento de Fisiología de la Universidad de Granada.

National Library of Medicine de Estados Unidos, una de las más utilizadas por los especialistas. En estos géneros la situación comunicativa y el tenor (especialista-especialista) son los mismos, aunque hay diferencias en cuanto a los objetivos, el contenido transmitido y los recursos lingüísticos.

Al imponerse el inglés como *lingua franca*, estos géneros están plenamente integrados en la comunidad médica española. Estos coinciden con los géneros que Puerta López-Cózar y Mauri (1995: 15-31) identifican en la revista médica en español con mayor índice de impacto, *Medicina Clínica*. Las distinciones entre estos géneros normalmente obedecen a los contenidos, los procedimientos utilizados, las características propias de la investigación y la estructura retórica. Entre paréntesis incluimos la nomenclatura de estos géneros en las revistas anteriormente mencionadas.

- a) Artículo original (*research article, journal article, original article, article, original investigation, original contribution*): es el género más frecuente, prestigioso y uniforme dentro de la literatura médica. Presenta el resultado de una investigación novedosa utilizando normalmente cuatro secciones retóricas, *Introduction, Methods, Results, Discussion*, que configuran una estructura denominada *IMRAD* en la página 88. Está precedido por un *abstract* o resumen. En los artículos en español, por ejemplo, en *Neoplasia*²⁴ estas secciones reciben los nombres de *Fundamentos, Pacientes y métodos, Resultados y Conclusiones*. Otras denominaciones para estas secciones son las de *Introducción, Material y métodos*²⁵ y *Discusión*.

Según el diseño investigador utilizado, los artículos originales pueden ser de tipo descriptivo o de tipo analítico. Los artículos originales de tipo descriptivo presentan unos resultados en la mayoría de los casos cuestionables. Dentro de los primeros, los más usuales son:

- El análisis de una serie de casos (*case-series article*): artículo en el que se describe una nueva enfermedad u observaciones que habían pasado previamente desapercibidas.

²⁴ Publicación bimestral del grupo editorial DOYMA, incluida en *Excerpta Medica*.

²⁵ En algunos artículos, esta sección reciben el nombre de *Sujetos y métodos*.

- El análisis epidemiológico descriptivo (*disease review*): artículo en el que se pretende obtener datos inferenciales mediante un análisis transversal de una muestra sesgada, aunque por su falta de rigor estadístico no lo consigue. En los estudios transversales, la causa y el efecto o desenlace de una enfermedad se miden al mismo tiempo, a diferencia de los estudios longitudinales.
- Ensayo clínico no controlado (*randomized controlled clinical trial*): artículo en el que se describe un experimento para determinar la eficacia de una intervención preventiva, terapéutica o rehabilitadora. A pesar de que la asignación del tratamiento de prueba se realiza al azar, al no haber un grupo de control o testigo, dicho estudio carece de objetividad.

Entre los estudios analíticos, destaca el ensayo clínico controlado (*randomized controlled clinical trial*), en el que se describe un experimento para determinar la eficacia de una intervención preventiva, terapéutica o rehabilitadora. A diferencia del ensayo clínico no controlado, hay un grupo de prueba y un grupo de control. La asignación a cada grupo se realiza al azar y se presenta una comparación estadística entre los resultados obtenidos en ambas poblaciones.

b) Artículo de revisión (*review article*)

Consiste en la revisión o análisis de un determinado tema por un experto desde dos perspectivas: su experiencia propia y la resultante del estudio de la bibliografía publicada sobre el particular. En la primera perspectiva (*case-series article*) hay un estudio de casos en el que los sujetos han sido tratados directamente por los investigadores o la organización a la que pertenecen. En la segunda perspectiva (*case-review article, review of reported cases, subject review*) se presenta un estudio de casos en los que los datos han sido extraídos de informes publicados y no de la experiencia de los autores del artículo. El artículo de revisión tiene una estructura poco estricta que suele ser la de Introducción, Bases metodológicas y Aplicaciones clínicas. En inglés, su estructura normalmente es: *Purpose and design, Results* y *Discussion*. Teniendo en cuenta que las revisiones van dirigidas no sólo a especialistas, sino también a médicos de familia y estudiantes, su lenguaje es más general y sencillo que el utilizado en un artículo original.

c) Editorial (*editorial*)

El editorial es un artículo breve cuyo objetivo es expresar una opinión o interpretar hechos u opiniones ajenas. Tiene una estructura muy libre y está muy marcado desde el punto de vista actitudinal con evaluaciones, recomendaciones y consejos, lo que lo hace más complejo sintáctica y lingüísticamente que otros géneros. Los editoriales discuten una única enfermedad o un tratamiento para un grupo de enfermedades, por lo que son una buena ilustración del pensamiento y la práctica médica contemporánea. Por su brevedad, no suelen incluir subtítulos, aunque sí incluyen citas bibliográficas.

d) Correspondencia o Carta al Director (*letter to the editor, correspondence, letters*)

En la sección *Cartas al director* de las revistas médicas especializadas, se presentan con brevedad y concreción, objeciones o comentarios relativos a artículos publicados recientemente en misma, y observaciones o experiencias que, por sus características, puedan ser resumidas en un texto breve. Esta sección es idónea para describir los efectos de un tratamiento o resultados preliminares de un estudio.

Al título de la carta sigue la fórmula "Sr. Director:" o "Al Director:" (*Sir / To the editor/ Editor*). No cuenta con subtítulos pero el texto puede ir acompañado de una figura y una tabla. Suele llevar bibliografía de diez citas como máximo. Con mucha frecuencia las cartas van seguidas de la réplica del autor del artículo o el director de la revista, precedido por la fórmula "Réplica del autor" (*Dr. X replies: / Author's reply / In reply / In response*).

e) Nota clínica (*clinical note, case report*): Es la comunicación de una o más observaciones clínicas (diez o menos casos) que aporta una novedad al conocimiento de la fisiopatología de un proceso, la descripción de nuevos procesos técnicos, etc. Este género es principalmente descriptivo y expositivo. Las revistas de especialidad son más receptivas a este tipo de artículos, dado que las notas clínicas presentan casos que, por sus peculiaridades o rareza, difícilmente se hallan expuestos en los libros de texto. Una nota clínica consta de las siguientes partes: Introducción / Fundamentos, Observación clínica / Descripción del caso (*case report / case description, case summary*) y Discusión (*comment / discussion*).

- f) Diagnóstico y Tratamiento (*treatment-focused report*): artículo que describe nuevos procedimientos diagnósticos y terapéuticos.
- g) Artículo especial: artículo que versa sobre temas que, sin ser estrictamente médicos, son de interés. Algunos de estos temas son la gestión de bibliografía, la calidad de vida, la docencia y la redacción de artículos científicos.
- h) Conferencia clinicopatológica (*consensus-development-conference*): publicación de una conferencia en la que un médico expone su opinión de experto y su diagnóstico sobre un caso clínico problemático del que sólo conoce datos que se le han ofrecido con varias semanas de antelación. En la conferencia, el experto expone las conclusiones a las que ha llegado mediante el análisis y un razonamiento hipotético-deductivo. Después se abre un debate con la audiencia y con el patólogo encargado del caso.
- i) Conferencia clínica (*clinical conference*): publicación de una disertación sobre una materia sobre la que el ponente es un experto y en la que únicamente interviene él.
- j) La imagen de la semana: conjunto formado por una o varias imágenes referentes al problema clínico de un paciente y un pie en cuyo texto se explica el contenido y el mensaje de la imagen.
- k) Crítica de libros / reseña de libros (*Book, software, or audiovisual-product review*)
Las reseñas de libros suelen comenzar con una breve introducción para presentar el libro. Le sigue una descripción formal de su contenido, es decir, se menciona el número de capítulos y la estructura general del mismo (bloques, bibliografía, etc.), y se comenta la calidad tanto del texto como de las ilustraciones. A continuación, se establece cuál es la principal aportación del libro. Por último, se comentan los problemas o errores que se hayan detectado, y se hace un comentario final, en el que se suele recomendar el libro a futuros lectores.

Además de estos géneros, en las publicaciones en lengua inglesa también hemos encontrado las referencias a congresos celebrados (*meeting report*) y otros géneros menos especializados como *commentary, news and views*.

3.3.2. Propuesta simplificada de tipos textuales médicos

Las clasificaciones mencionadas son tan exhaustivas que hacen difícil el estudio de las variaciones cohesivas con respecto a tipos textuales. De ahí que en este estudio se adopte el término *tipo textual* en lugar del más restrictivo *género*.

Por otra parte, teniendo en cuenta que la cohesión cambia sobre todo respecto a las dificultades comunicativas que impone el contexto situacional y el canal de comunicación y con respecto a la relación y al conocimiento que comparten los participantes en la comunicación, proponemos una adaptación de la taxonomía de Lévy-Leblond (*apud* Gutiérrez Rodilla 1998: 20) que se puede aplicar a los textos sobre oncología:

TABLA 4: Taxonomía textual según el tenor y el canal de comunicación basada en Lévy-Leblond 1996

TENOR	CANAL DE COMUNICACIÓN	
	Escrito	Oral
Intercambio especializado (oncólogo – oncólogo) (oncólogo – profesional de la medicina) (oncólogo – estudiante de medicina)	Publicaciones especializadas (artículos experimentales, notas clínicas, etc.) Libros de texto	↑ Conferencias, coloquios, sesiones clínicas, seminarios ↓
Intercambio público (oncólogo / médico / investigador – público con cultura científica) (periodista científico – público general) (oncólogo / médico – público general/ paciente)	Publicaciones de divulgación semiespecializada Public. de divulgación general Folletos de salud pública Sitios web sobre cáncer	Programas de divulgación (radio, televisión...)
Intercambio “familiar” Entre profesionales de la salud y/o investigadores	Diarios de laboratorio, cartas, correo electrónico...	Discusiones informales “de pasillo”, telefónicas

La fila *Intercambio especializado* recoge las tipologías textuales del lenguaje científico destinado a la comunicación entre médicos especialistas en oncología, otros profesionales de la medicina y futuros médicos.

Bajo *Intercambio público*, se citan tipos correspondientes a aquellas situaciones en que se habla de asuntos científicos, no entre profesionales sino cuando el receptor del mensaje es el público en general. Estamos hablando de la divulgación científica. El término divulgación designa "todo intento de comunicación de los descubrimientos generados en el ámbito de la ciencia a un público no experto" (Fernández Polo 1999: 86). Dentro de esta, Myers y Nwogu (*apud* Fernández Polo 1999: 86), reconocen dos tipos de divulgación que obedecen a diferencias en el nivel de formación de los lectores previstos. Por una parte, está el tipo de divulgación que aparece en diarios o semanarios de interés general, y por otra, el tipo de divulgación "sofisticada" de revistas tales como *Scientific American* o *New Scientist*. Fernández Polo deja la puerta abierta a más tipos de divulgación cuando sostiene que los distintos tipos de divulgación obedecen a dos criterios: el medio empleado para la misma (escrito, imagen, sonido y objeto) y las características del divulgador (periodista o investigador). Por eso, en esta tesis se incluirán dentro de la divulgación recursos sobre salud y medicina disponibles en sitios web dirigidos al público general, entre ellos los pacientes de cáncer, tanto si la información se gestiona desde la sede de alguna sociedad médica como si se publica en los sitios de periódicos y revistas.

Por último, se presenta la variante *familiar* del lenguaje científico, es decir, la comunicación entre científicos en una situación comunicativa ajena al estrecho espacio de la comunicación seria, de las grandes reuniones y de sesiones puramente técnicas.

4. Revisión del concepto de *cohesión*

Uno de los parámetros que determinan los segmentos en los que se puede fragmentar un determinado tipo textual es la cohesión (Hasan 1978). En este capítulo se citan los estudios principales sobre cohesión, sobre todo cohesión léxica e indagamos sobre si existe una correlación entre tipo textual y cohesión.

Tras mencionar muy brevemente la diferencia entre los términos *coherencia* y *cohesión*, repasaremos algunos de los modelos de análisis más conocidos en el ámbito anglosajón: Halliday y Hasan (1976), Hasan (1984), Winter (1974, 1979), Hoey (1983, 1991, 1995) y Morris y Hirst (1991). Posteriormente, comentaremos algunas de las aplicaciones de los estudios sobre cohesión, por una parte, para una mejor comprensión de la organización del texto y de los tipos de texto, y por otra, para el reconocimiento automático de la cohesión. También se verá qué dice la *Teoría Sentido Texto* sobre la cohesión.

4.1. Coherencia y cohesión

En torno a los años setenta, en el Reino Unido surge una gramática del texto (Quirk, Greenbaum, Leech and Svartvik, 1972) en la que adquiere importancia el concepto de *cohesión*. Las especulaciones sobre la cohesión cristalizaron en Halliday y Hasan (1976) y se extendieron fuera del Reino Unido. Y así, desde la lingüística textual alemana, se empezó a hablar de *coherencia* y *cohesión* como dos criterios fundamentales de textualidad (Beaugrande y Dressler 1981).

En esencia, la cohesión es la forma en que determinados elementos léxicos o gramaticales de una oración pueden conectar esa oración a elementos anteriores o posteriores dentro de un texto (Hoey 1991: 11). La cohesión se presenta de un modo explícito en el texto, es una propiedad objetiva del texto. Por el contrario, la coherencia es subjetiva porque es la unidad de significado que percibe el receptor de un mensaje al relacionar el texto con la realidad extralingüística. Ambos están relacionados en el sentido de que la cohesión ayuda a que el lector perciba la coherencia del texto.

4.2. Halliday y Hasan (1976)

Halliday y Hasan (1976) presenta una clasificación de los medios disponibles en inglés para crear cohesión en los textos. Para Halliday y Hasan, la organización del texto (*texture*) se construye a partir de relaciones entre elementos del texto tanto gramaticales como léxicas. A estas relaciones las denominan *enlaces cohesivos* (*cohesive ties*).

Los cinco tipos de enlaces cohesivos que identifican son la *conjunción*, *referencia*, *sustitución*, *elipsis* y *cohesión léxica*.

4.2.1. La conjunción

En esta categoría, se incluyen las conjunciones y otros *conectores* textuales que marcan de forma explícita relaciones semánticas (consecuencia, adición, oposición, etc.) entre oraciones o entre unidades superiores.

- (2) In the present study, loss of Rb protein expression was not a statistically significant prognostic variable as an individual factor in the cohort examined. **However**, our data suggested that [...]

(Cfr. Apéndice II: texto 1, pág. 2)

4.2.2. La referencia

Este recurso cohesivo no marca relaciones semánticas sino que es una relación semántica que se establece cuando elementos, principalmente pronominales y demostrativos, indican que la identidad de la entidad o acción a la que hacen referencia puede recuperarse del cotexto próximo. Hay dos tipos de referencia, anafórica y catafórica. En los ejemplos, los elementos relacionados mediante referencia aparecen en negrita.

Referencia anafórica:

- (3) **Inactivation of the retinoblastoma (Rb) gene** has been documented in various types of cancers, including lung cancer. **It** has also been associated with aggressive tumor behavior and poor clinical outcome in specific types of cancer.

(Cfr. Apéndice II: texto 1, pág. 1)

Referencia catafórica:

- (4) Survival was analyzed for the 89 patients who met **the following criteria**: (1) survived for more than 3 months after surgery; (2) did not die of causes other than lung cancer within 5 years after surgery; and (3) were followed for more than 3 years after surgery (for patients who remained alive).

(ibid, pág.2)

4.2.3. La sustitución

Relación gramatical por la que un elemento reemplaza a un término precedente cuyo referente no es la misma entidad, sino una entidad a la que el mismo término puede designar.

- (5) **A tougher substance** is needed. This **one** is too soft

(Kennedy y Bolitho 1984: 132)

4.2.4. La elipsis

Este recurso se emplea cuando se omite una frase o idea que está explícitamente expresada en la oración previa porque no se considera necesaria para la comprensión de la segunda oración.

- (6) The exploitation of energy will guarantee a more pleasant way of life. But for how long? (**will the exploitation of energy guarantee a more pleasant way of life?**)

(ibid)

4.2.5. La cohesión léxica

Bajo el nombre de cohesión léxica se incluyen una variedad de relaciones semánticas que pueden existir entre lexemas. La cohesión léxica es el tipo de enlace cohesivo más frecuente e importante, como se constata en el análisis realizado por Halliday y Hasan. Este punto de vista es compartido por Hoey (ibid: 10):

Lexical cohesion is the only type of cohesion that regularly forms multiple relationships (though occasionally reference does so too). If this is taken into account, lexical cohesion becomes the dominant mode of creating texture.

Halliday y Hasan identifican dos formas de cohesión léxica: la *reiteración* y la *colocación*. Esta terminología se modificará en Hasan (1984).

4.2.5.1. La reiteración

La *reiteración* es la forma en que un lexema puede evocar o traer a la memoria el sentido de un lexema anterior. Esta se consigue mediante la repetición literal de esa palabra o mediante la utilización de un sinónimo, sinónimo parcial, hipónimo o hiperónimo.

- (7) **Investigadores** de Baltimore acaban de usar en ratones una vacuna contra este tipo de **tumor** y la experiencia ha resultado exitosa. El estudio, publicado en la revista *Cancer Research*, revela que los **científicos** utilizaron una proteína, llamada E7, que juega un papel fundamental en la progresión del **cáncer** de cuello de útero.

(Cfr. Apéndice II, texto 2)

4.2.5.2. La colocación

La *colocación* aparece definida de un modo tan ambiguo que incluye relaciones tan variadas como las existentes entre *boy—girl*; *wet—dry*; *ill—doctor*; *try—succeed*.

Esta ambigüedad ha hecho que, en estudios posteriores (Hoey 1991), la colocación se entienda en el sentido que le da Firth (1957): "the company a word keeps". Este autor entiende la colocación como la relación que un lexema establece con otros lexemas que aparecen con una probabilidad significativa en su cotexto inmediato. Normalmente, desde la lingüística de corpus (Sinclair 1991), este cotexto abarca cuatro palabras a ambos lados del lexema en cuestión.

A diferencia de otras combinaciones habituales de palabras, según Tercedor (1999: 89), en las colocaciones participan dos elementos, de los cuales, la base o palabra clave mantiene su significado y goza de plena autonomía, de ahí que admitan cambios de orden y variación sintáctica. En la idea de autonomía de sus componentes también coincide Montero Martínez (2000) cuando afirma que el significado de las colocaciones puede inferirse de las partes.

Para ilustrar el concepto de colocación, mencionaremos una serie de ejemplos en inglés y en castellano. Stubbs (1996) analiza un corpus de textos periodísticos publicados en el Reino Unido durante el gobierno del partido conservador y destaca la frecuencia con que la palabra *nation* “coocurre”²⁶ con otros lexemas como *biggest*, *European*, *favoured*, *industrialized*, *largest*, *trading* y *wealthiest*. En este ejemplo, se destaca el papel que puede jugar la repetición de determinados patrones lingüísticos en la consolidación de una determinada ideología (Fowler 1996; Hatim 1997).

El ejemplo correspondiente al castellano ha sido sacado de un corpus de textos periodísticos sobre educación publicados en *El Mundo* y *El País* en los años 1995 y 1996 (López Rodríguez 1997). En este corpus, la palabra *enseñanza* aparece con mucha frecuencia en compañía de las palabras *pública*, *obligatoria*, *calidad*, *secundaria*, *primaria* y *centros*. Estos resultados fueron obtenidos con el programa informático *Wordsmith Tools*²⁷.

TABLA 5: Colocaciones de la palabra *enseñanza*²⁸

N	WORD	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	ENSEÑANZA	129	0	0	0	0	0	0	0	129	0	0	0	0	0
2	PÚBLICA	19	0	19	0	0	0	0	0	0	19	0	0	0	0
3	OBLIGATORIA	17	1	16	1	0	0	0	0	0	6	10	0	0	0
4	CALIDAD	17	13	4	1	0	8	4	0	0	0	2	0	0	2
5	SECUNDARIA	14	0	14	0	0	0	0	0	0	13	0	1	0	0
6	PRIMARIA	12	0	12	0	0	0	0	0	0	10	1	0	0	1
7	CENTROS	9	8	1	0	1	3	4	0	0	0	0	0	0	1
8	FEDERACIÓN	8	8	0	2	0	0	6	0	0	0	0	0	0	0
9	GENERAL	8	6	2	5	0	0	1	0	0	0	0	2	0	0
10	ESO	8	3	5	0	1	0	0	2	0	0	0	5	0	0
11	ENSEÑANZA	7	0	0	0	0	0	0	0	7	0	0	0	0	0
12	CCOO	6	1	5	0	0	1	0	0	0	0	4	1	0	0
13	SINDICATOS	5	4	1	1	1	1	1	0	0	0	0	0	0	1
14	TRABAJADORES	5	4	1	0	0	4	0	0	0	0	0	0	0	1
15	CASTELLANO	5	0	5	0	0	0	0	0	0	0	3	0	0	2

²⁶ El verbo *coocurrir* se utiliza con frecuencia en la lingüística de corpus y es un calco del inglés *co-occur*

²⁷ Programa creado por Mike Scott [<http://www.liv.ac.uk/~ms2928/homepage.html>] de la *University of Liverpool* y distribuido por Oxford University Press.

²⁸ Se incluyen sólo las palabras plenas en su contexto.

4.3. Hasan (1984)

En 1984, Hasan publica un estudio en el que explora la relación entre coherencia y cohesión en la literatura infantil e intenta medir la coherencia sobre la base de la comprensión de los relatos por parte de los lectores. Hasan (1984) también reformula las categorías de cohesión léxica propuestas en 1976 y las organiza en dos categorías: *general e inmediata (instantial)*.

TABLA 6: Categorías de cohesión léxica

COHESIÓN LÉXICA GENERAL	COHESIÓN LÉXICA INMEDIATA
Repetición: <i>leave, leaving, left</i>	Equivalencia: <i>the <u>sailor</u> was their <u>daddy</u></i>
Sinonimia: <i>leave, depart</i>	Designación (<i>naming</i>): <i>the dog was called <u>Toto</u></i>
Antonimia: <i>leave, arrive</i>	Apariencia (<i>semblance</i>): <i>the <u>deck</u> was like a <u>pool</u></i>
Hiponimia: <i>travel, leave</i>	
Meronimia: <i>hand, finger</i>	

Hasan (1984: 202)

Después de analizar los enlaces cohesivos en relatos infantiles, Hasan descubre que no existe una relación entre el número de enlaces cohesivos y la coherencia que encuentran los lectores. Por eso, Hasan sostiene que la función de los enlaces cohesivos es formar una especie de *cadena* (*chains*) y que estas interactúan unas con otras. A esta interacción la denomina *armonía cohesiva*. Son las cadenas y no los enlaces cohesivos aislados los que crean cohesión.

Una cadena cohesiva se forma cuando un elemento cohesivo refiere hacia un elemento anterior, que a su vez también está relacionado cohesivamente con otro elemento previo. Existen dos tipos de cadenas: la de *identidad* y de *similitud*.

La cadena de identidad está compuesta de cadenas cohesivas que comparten los mismos referentes. En inglés es necesario hacer referencia mediante pronombres, reiteraciones o elementos equivalentes al tema del que se está tratando en sucesivas oraciones.

La cadena de similitud se da cuando distintas entidades están implicadas en procesos o descripciones paralelas y, por tanto, comparten relaciones cohesivas con palabras de significado parecido. Por ejemplo, si en una descripción tres entidades son descritas con el adjetivo *little*, se formará una cadena de similitud formada por las distintas instancias de *little*.

Como ilustración, presentamos un fragmento de uno de los cuentos estudiados por Hasan (Hoey 1991: 15).

- (8) **1** Once upon a time there was a little girl
 2 and she went out for a walk
 3 and she saw a lovely little teddy bear
 4 and so she took it home
 5 and when she got it home she washed it
 6 and she had the teddy bear for many many weeks and years.

En este tipo de texto, las cadenas son muy obvias. Es posible reconocer tres cadenas de identidad y dos de similitud.

Cadenas de identidad

1 girl **2** she **3** she **4** she **5** she, she **6** she

3 teddy bear **4** it **5** it, it **6** teddy bear

4 home **5** home

Cadenas de similitud

2 went out **5** got ... home

4 took **6** had

La conclusión a la que llega Hasan es que la presencia de cadenas aisladas no garantiza la coherencia del texto, pero sí la interacción entre cadenas. Esta conclusión será crucial en la metodología de Hoey, en la que los recursos cohesivos adquieren sentido sólo cuando están interrelacionados.

4.4. Winter (1974, 1979, 1986)

La investigación de Winter gravita en torno a dos ejes: la forma en que la gramática de las oraciones contribuye a su interpretación en contexto y la función de los enlaces cohesivos en la organización del texto.

La principal función de los enlaces cohesivos es la *repetición*, un recurso que llama la atención sobre determinada información del texto. La repetición puede adquirir distintas formas:

- a) La *sustitución*, que es similar a la *referencia* de Halliday y Hasan (1976)
- b) La *eliminación (deletion)*, que corresponde a la *elipsis*
- c) y la *repetición léxica*, que coincide con la *cohesión léxica*.

De todas formas, Winter (1979: 101) presta atención a la repetición que tiene un carácter sistemático. Es decir, su énfasis está en la repetición significativa de uno o varios de los elementos constituyentes de una *cláusula* en la estructura de otra. Esta repetición proporciona a cada cláusula una serie de elementos constantes que permiten reconocer cuál es la información nueva y qué relevancia tiene esta en el cotexto.

La repetición también hace necesarios una serie de cambios en la estructura de la nueva cláusula. Winter, siguiendo a Dixon (1965), Quirk *et al.* (1972), denomina a estos cambios *reemplazos (replacements)*. El reemplazo no se ha de entender como la sustitución de una palabra por otra sino la presentación de nueva información en un contexto que ya ha sido previamente presentado.

A continuación, se incluye un ejemplo tomado de la publicación *New Statesman* y la representación de cómo se producen estos reemplazos en Hoey (1983), que destaca qué elementos son constantes y cuáles variables.

- (9) Pressures built up on all sides: **his** father, a ‘moderately successful plumbing contractor’ (said Time) **demande**d performance. **His** mother, who left her husband in Florida and moved to Austin to be near her son, **demande**d love. Whitman could provide neither.

TABLA 7: Reemplazos y elementos constantes del ejemplo (9)

	His His	father mother	demanded demanded	performace love
REPETICIÓN REEMPLAZO	His	*****	demanded	*****
CONSTANTE VARIABLE	His	parent which sex	demanded	response by him what kind of response

Mediante la repetición y el reemplazo, se consigue la interpretación de pares de oraciones en el texto y Winter consigue explicar cómo la gramática y la cohesión interactúan en los textos. A partir de ahí, Winter especula sobre la relación entre las oraciones y la organización del texto.

En conjunto, el trabajo realizado por Winter contribuye a una mejor comprensión del papel que juega la cohesión. Los puntos clave de su aproximación aparecen resumidos en Hoey (1991: 20) y tendrán una gran repercusión en este discípulo de Winter.

- La cohesión ha de ser interpretada en el contexto de las oraciones en las que ocurre.
- Para estudiar la cohesión hay que ver cómo la repetición forma agrupaciones entre pares de oraciones.
- Es más importante ver la función de repetición de la cohesión que las clasificaciones de tipos de cohesión.
- La repetición tiene valor informativo y proporciona un marco para interpretar los elementos que han cambiado en cada nueva oración.
- Las relaciones de repetición que se establecen entre oraciones son normalmente múltiples y entre oraciones no contiguas.

4.5. Metodología de Hoey (1991, 1995)

Hoey (1983, 1991, 1995) continúa en la línea de Winter, Halliday y Hasan y propone una metodología de análisis textual basada en el estudio de la cohesión, sobre todo, de la cohesión léxica. Su metodología hace posible detectar la sistematicidad de las relaciones cohesivas de un texto y representar gráficamente cómo los rasgos cohesivos se combinan para organizar los textos y crear continuidad en los mismos. El objetivo último es demostrar que la coherencia, la cohesión y la organización (*texture*) del texto se crean principalmente por las relaciones léxicas entre las palabras y entre las oraciones²⁹.

A esta metodología se la conoce como *patterns of lexis* y ha destacado por su capacidad de evaluar la red compleja de relaciones léxicas que caracterizan los textos naturales.

En su revisión a la taxonomía de enlaces cohesivos de Halliday y Hasan, Hoey (1991: 7) destaca que la reiteración léxica y la colocación son principalmente formas de relación léxica, y sólo de un modo secundario, marcadores de relaciones textuales. En esto se diferencian de la conjunción, la referencia, la sustitución y la elipsis. Las relaciones léxicas creadas por reiteración léxica y la colocación contribuyen a la creación y a la organización o *textura* del texto, a la vez que este proporciona el contexto para la creación y la interpretación de relaciones léxicas.

La mayoría de estos enlaces cohesivos, a excepción de la conjunción, se basan en la repetición porque permiten que se mencione de nuevo lo que ya se ha mencionado. La repetición tiene como función mostrar las relaciones léxicas entre las palabras y las oraciones. Por este motivo, Hoey, al igual que Winter, estima necesario remodelar la clasificación de recursos cohesivos de Halliday y Hasan para dar cuenta de aquellos que se sustentan en la repetición.

²⁹ A la cohesión que se establece entre palabras, Sinclair (1993) la denomina *point-to-point cohesion*. Este tipo de cohesión tiene una naturaleza léxica, en contraposición con la cohesión entre oraciones o complejos de cláusulas, que, según Sinclair, tiene una naturaleza textual o estructural.

4.5.1. Cohesión y tipos de repetición

A continuación presentamos los recursos cohesivos que Hoey considera esenciales en el establecimiento de las relaciones léxicas de un texto. Las categorías propuestas se basan en la repetición.

4.5.1.1. Repetición léxica simple

Se produce cuando un lexema que ya ha aparecido en el texto se repite tal cual o con la única alteración de un morfema gramatical que no modifica las relaciones sintagmáticas. Hoey sólo aplica esta repetición a palabras plenas (*open-set lexical items*).

Presentamos tres fragmentos de oraciones que contienen las siguientes instancias de repetición léxica simple: *authors—authors*; *prognostic significance—prognostic significance* y *study—studies*.

- (10) [...] and the **authors** of this article have reported previously on the **prognostic significance** of both of them.
- (11) In the present **study**, the **authors** evaluated the **prognostic significance** of [...]
- (12) Larger prospective **studies** with Stage I NSCLC patients are necessary to confirm the current findings.

(Cfr. Apéndice II: texto 1, pág. 1)

4.5.1.2. Repetición léxica compleja

Ocurre cuando se repiten dos lexemas que comparten un morfema léxico, aunque son formalmente diferentes, o que son formalmente idénticos pero no tienen la misma función sintáctica. Por ejemplo, el sustantivo *humans* y el adjetivo *human*. Dos elementos forman una repetición de este tipo si se pueden parafrasear en el contexto en que aparecen de modo que la paráfrasis de uno incluya al otro. Y así, *human* es igual que *of humans*. Veamos ahora un ejemplo tomado de nuestro corpus.

- (13) [...] and p53 protein expression in the same **tumor** specimens, and [...]
- (14) [...] those with Rb-/ras+ **tumors** survived for a significantly shorter period

(ibid)

4.5.1.3. Paráfrasis simple

Es el tipo de repetición en la que una unidad léxica sustituye a otra en contexto sin que se produzca una pérdida o ganancia en especificidad y sin que se produzca ningún cambio de significado. Hasan denominaba a este tipo de repetición *sinonimia*. La paráfrasis simple puede ser *parcial*, si la sustitución sólo funciona en una dirección o *mutua*.

- (15) Aunque el 40% de los varones mayores de 50 años tiene **cáncer** de próstata, sólo el 8% de todos los **tumores** acabará produciendo un problema al **paciente**. Por eso, es necesario encontrar métodos y fórmulas que concreten mucho más el qué hacer con los **enfermos**.

(Cfr. Apéndice II, texto 3)

4.5.1.4. Paráfrasis compleja

Se da cuando una unidad léxica del texto se puede definir en función de otra unidad léxica del texto con la que no comparte ningún morfema y cuando se cumple una de las siguientes situaciones:

- Cuando entre las dos unidades léxicas existe una relación de antonimia. El término de antonimia engloba un gran número de casos que Hoey no ha sabido sistematizar. Según Lyons (1977), los distintos tipos de oposiciones léxicas relacionadas con la antonimia entran en alguna de estas categorías:

(a) *antónimos graduales: hot—cold*

(b) *antónimos binarios: dead—alive; same—different*

(c) *términos inversos (converses): lend—borrow; parent—child*

(d) *términos pertenecientes a sistemas de incompatibilidad múltiple (open-ended systems of multiple incompatibility):*

*Secretary of State—Minister—Minister of State—Parliamentary
Under-Secretary of State*

January—February—March ...

Esta falta de sistematicidad le lleva a eliminar la antonimia de su clasificación de recursos cohesivos (Hoey 1995), aunque añadiendo una categoría, la pertenencia a un conjunto cerrado de palabras plenas.

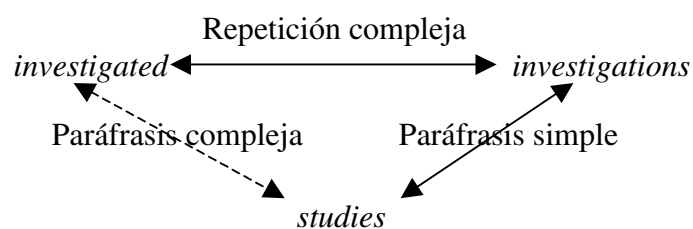
- Cuando el lexema que se analiza es una repetición compleja de otro elemento y, al mismo tiempo, una paráfrasis simple de un tercer elemento. En este caso se establece una paráfrasis compleja entre el segundo y el tercer elemento. Veamos un caso de paráfrasis compleja en el siguiente texto

(16) For this reason, induction chemotherapy is being **investigated** in patients expected to have a poor prognosis after standard surgery and radiotherapy. [...] Clinical **investigations** have reported an improvement of survival in stage III NSCLC after induction chemotherapy by using different combinations of cytotoxic drugs. [...] The combination of mitomycin, vindesine or vinblastine and cisplatin has produced encouraging results in several **studies** [...]

(Cfr. Apéndice II, texto 4)

A continuación presentamos un triángulo que ilustra cómo se genera una paráfrasis compleja a raíz de las relaciones de repetición que se establecen entre un elemento y otros dos.

DIAGRAMA 6: Triángulo de conexiones



En el diagrama se ve cómo, si una palabra (*investigations*) está en relación con otras dos (*investigated* y *studies*), se crea también un *enlace putativo* entre *investigated* y *studies*. El nombre que Hoey asigna a esta relación es el de *link triangle* o triángulo de conexiones.

- Otra situación en la que se habla de paráfrasis compleja es cuando uno de los elementos del triángulo de enlaces que funciona de mediador de relaciones no aparece y el lector infiere este elemento. Este tipo de paráfrasis se da, por ejemplo, entre los lexemas *investigadores* y *estudio*. El elemento implícito que sirve de mediador es la palabra *investigación*.

(17) **Investigadores** de Baltimore acaban de usar en ratones una vacuna contra este tipo de tumor y la experiencia ha resultado exitosa. El **estudio**, publicado en la revista [...]

(Cfr. Apéndice II, texto 2)

4.5.1.5. Repetición superordinada, hiponímica y correferencial

Se produce cuando el significado de un término está incluido en el significado del otro y sus referentes son idénticos. En esto se diferencia de los casos mencionados con anterioridad, en los que los lexemas implicados contenían la misma información.

Por ejemplo, la relación entre *genetic* y *biologic* o *cancer* y *lung carcinoma* (ibid, texto 1), o la repetición correferencial que se produce al mencionar *tumor maligno más frecuente en el hombre* como referencia al *cáncer de próstata* (texto 3). Creemos interesante resaltar que para automatizar este último tipo de repetición es necesaria información enciclopédica.

4.5.1.6. Otras formas de repetición que no son léxicas.

Bajo esta categoría, Hoey incluye la *sustitución* y la *elipsis*. La sustitución es una forma de repetición de elementos textuales que dependen absolutamente de otros elementos para ser interpretados semánticamente porque no tienen significado proposicional por sí solos. Es decir, en la sustitución se emplean *palabras forma* que hacen referencia no sólo a un elemento léxico del texto sino también a una realidad externa al texto. Hoey utiliza el término sustitución en un sentido más amplio que Halliday y Hasan, quienes denominaron a este tipo de repetición *reference*. La sustitución recurre a pronombres personales, pronombres y determinantes demostrativos y proformas como *one*, *do*, *so*, *(an)other* y *similar*.

Hoey también incluye como forma de repetición la elipsis. Sin embargo, en su método de análisis, Hoey presta atención principalmente a las formas de repetición léxica, que son más fáciles de reconocer por medios computacionales.

Para que quede más claro qué categoría asignar a cada tipo de repetición, Hoey (ibid: 58-60 y 69) elabora unos diagramas con preguntas (Véase Apéndice III).

En 1995, Hoey elimina de esta lista de elementos cohesivos la paráfrasis compleja (en la que se incluía la antonimia), la repetición superordinada e hiponímica y la elipsis, y añade una categoría relacionada con la colocación, la pertenencia a un conjunto ordenado de palabras plenas como *March—April*.

4.5.2. Conexiones, vínculos y redes

A cada repetición de un elemento léxico de una oración en otra oración, ya sea contigua o no, Hoey la denomina *link*. De ahora en adelante, utilizaremos el término *conexión* como traducción para *link*. Se supone que, al repetir elementos léxicos en otra oración (*conexión*), estamos estableciendo una relación entre los mismos y entre las oraciones implicadas.

Al proponer la palabra *conexión*, Hoey está marcando una diferencia con respecto al término *cohesive tie* de Halliday y Hasan. El término *conexión* es más restringido porque designa las relaciones cohesivas léxicas y no-léxicas entre los elementos de dos oraciones pero sin incluir las conjunciones, colocaciones y algunos tipos de referencia que el término *cohesive tie* abarca. Además, la palabra *conexión* no parece indicar la dirección que siguen las relaciones.

A continuación presentamos las dos primeras oraciones del capítulo introductorio de un libro de texto para universitarios sobre filosofía política tomado de Hoey (1991: 246)³⁰. Hemos marcado las conexiones entre elementos léxicos mediante flechas.

³⁰ En el Apéndice IV aparecen las dieciséis primeras oraciones del texto.

What is attempted in the following volume is to present to the reader a series of actual excerpts from the writings of the greatest political theorists of the past; selected and arranged so as to show the mutual coherence of various parts of an author's thought and his historical relation to his predecessors or successors; and accompanied by introductory notes and intervening comments designed to assist the understanding of the meaning and importance of the doctrine quoted. The book does not purport to be a history of political theory, with quotations interspersed to illustrate the history.

Estas conexiones adquieren importancia en la organización de los textos en tanto en cuanto aparecen en conjunción con otras conexiones (*clustering of links*). Por tanto, no interesan las conexiones aisladas entre oraciones sino los haces de conexiones que llegan a tejer una especie de cuerda entre dos oraciones.

Por este motivo, Hoey introduce un nuevo término para referirse a la relación estrecha entre dos oraciones que comparten más conexiones de lo que es normal para un determinado tipo textual. Este término es el de *bond*, que hemos traducido como *vínculo*, porque el significado de este último lexema es el de una unión significativa entre dos objetos³¹. Con el fin de delimitar aún más los términos *link* y *bond*, Hoey (ibid: 91) afirma: "Lexical items form *links*, and sentences sharing three or more links form *bonds*".

En textos que no son ni narrativos ni especializados, Hoey establece que el número mínimo de conexiones para establecer un *vínculo* es de tres. Este umbral cambia de unos textos a otros y no se puede establecer *a priori*. Y así, en algunos textos científicos hay tantas conexiones entre oraciones que es necesario fijar un umbral de seis oraciones. Es decir, en esos textos sólo se computan como oraciones vinculadas aquellas que comparten un mínimo de seis conexiones.

El conjunto de vínculos que se interrelacionan en un texto (*net*) juegan un papel esencial en la organización del texto en el sentido de que crean su textura (véase p. 76). En esta *red*, un lector es capaz de calificar dos oraciones entre las que existe un vínculo como portadoras de proposiciones mutuamente relevantes. También es capaz, previa alteración de algunos elementos cohesivos, como la sustitución de una proforma por su

³¹ Fernández Polo (1999: 194) traduce *links* como "lazos" y *bonds* como "uniones".

referente, de entender esas oraciones como coherentes y como si estuvieran contiguas en el texto.

4.5.3. Cómputo y representación gráfica de las conexiones y vínculos

En los apartados anteriores hemos visto qué tipos de conexiones se pueden establecer entre los elementos léxicos de un texto y cómo un determinado número de conexiones pueden formar un vínculo entre oraciones. Ahora, nos disponemos a presentar cómo se pueden computar estas conexiones y vínculos de una forma sistemática. En aras de la claridad, hemos presentado las indicaciones metodológicas de Hoey en forma de etapas.

(1) Identificación de repeticiones y elaboración de una matriz de repetición

En primer lugar, se reconocen los casos de repetición léxica. Para hacerlo de una forma consistente, Hoey propone que se computen las conexiones siempre como si fueran anafóricas. Es decir, al leer la segunda oración de un texto, se empiezan a establecer las conexiones con la primera oración.

También apunta que las conexiones se establecen no sólo con elementos de la oración precedente, sino con *todas* las unidades léxicas de las que son una repetición en otra oración. A continuación se identifican los casos de paráfrasis, sustitución y elipsis.

Para evitar errores en el cómputo, las repeticiones se representan en una *matriz de repetición*. Esta matriz permite registrar el número y el tipo de conexiones que se establecen entre elementos de forma que se puedan visualizar las conexiones de cada oración con otras, las *cadena léxicas* del texto. Para comprender lo que es una matriz de repetición, sería conveniente consultar las tablas 1 y 2 del Apéndice V.

En la Tabla 1 del Apéndice V aparece la primera columna ampliada de la matriz de repetición correspondiente a las primeras oraciones del texto sobre filosofía política mencionado con anterioridad (véase Apéndice IV). En esta se perciben las conexiones basadas en repetición léxica de la primera oración con las quince siguientes, de ahí que sólo aparezcan palabras plenas.

La Tabla 2 representa la matriz de repetición completa e incluye además de los casos de repetición léxica, los de elipsis y referencia deíctica. En la matriz aparecen filas y columnas encabezadas por números que corresponden a cada una de las oraciones del

texto. En las filas, se presentan las conexiones de cada oración con las que le anteceden; en las columnas, aparecen las conexiones con oraciones posteriores. En cada casilla, hay pares de palabras entre los que se establece un tipo de conexión. La primera palabra del par pertenece a la oración precedente, indicado por el número de columna, mientras que la segunda pertenece a la oración con el número de la fila.

Así por ejemplo, la primera casilla de la tabla, representa las conexiones entre la primera oración (número de columna) y la segunda oración (número de fila). Los seis pares de palabras indican que hay seis conexiones entre la primera y la segunda oración. En el par (*volume-book*), que es un ejemplo de paráfrasis simple parcial, *volume* pertenece a la primera oración (primera columna) y *book*, a la segunda oración (segunda fila).

(2) Ajuste de las conexiones computadas

Siguiendo el principio de que un elemento sólo puede formar una conexión por oración³², en el caso de que un elemento repita dos elementos de una oración anterior o de que una oración contenga dos elementos que repiten un elemento anterior, se registra sólo una conexión

En estos casos hay que plantearse qué conexión incluir en el análisis. Hoey establece que existen conexiones que tienen más peso y que deben primar en el análisis. Las presentamos en orden decreciente.

- Repetición léxica simple
- Repetición léxica compleja
- Paráfrasis simple mutua
- Paráfrasis simple parcial
- Paráfrasis compleja por antonimia
- Otros tipos de paráfrasis compleja
- Sustitución
- Correferencia
- Elipsis

Hoey también elimina las conexiones internas de las oraciones y no las computa en el análisis. No estamos de acuerdo con esta decisión porque dejaría al margen casos de cohesión importantes como la colocación y la relación tan frecuente en textos

³² Esto no impide que un elemento establezca conexiones con más de una oración.

divulgativos entre un término incomprensible para un lector lego y el lexema de la lengua general que le aclara el significado de ese término (véase capítulo 13).

(3) *Elaboración de una tabla de conexiones*

Si eliminamos las unidades léxicas de cada casilla y ponemos en su lugar el número de conexiones correspondiente a cada par de oraciones, es fácil identificar a golpe de vista qué oraciones tienen más conexiones y cuáles no tienen ninguna conexión. Según Hoey, las primeras constituyen las oraciones centrales del texto y las segundas, las oraciones marginales. En la siguiente tabla, se ve por ejemplo cómo la décima oración tiene dos conexiones con la primera oración y ninguna, con la segunda. Los números entre corchetes indican el número de conexiones incluyendo los casos dudosos.

TABLA 8: Tabla de conexiones del texto del Apéndice IV

(1)																
(2)	6	(2)														
(3)	2	1	(3)													
(4)	5	1	2	(4)												
(5)	1	∅	1	∅	(5)											
(6)	3	1	2	1	1	(6)										
(7)	4	∅	3	2	2	2	(7)									
(8)	5	1	2	4	∅	1	1	(8)								
(9)	1	1	∅	2	∅	∅	∅	∅	(9)							
(10)	2	∅	∅	2	∅	∅	1	2	1	(10)						
(11)	1[3]	∅	3	1	1	2	2	1	∅	∅	(11)					
(12)	3[4]	∅	1	3	∅	1	2	1[2]	∅	2	1	(12)				
(13)	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	1	(13)
(14)	2[3]	∅	1[2]	2[3]	∅	[1]	2[3]	[1]	∅	1	[1]	3	∅	(14)		
(15)	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	2	(15)
(16)	1	∅	∅	1	∅	∅	1	1	∅	1	1	2	∅	3	3	

(4) *Decisión sobre el número de conexiones necesarias para establecer un vínculo*

Como expusimos con anterioridad, de estas conexiones, las únicas relevantes son las que son más frecuentes de lo normal y aportan una organización al texto.

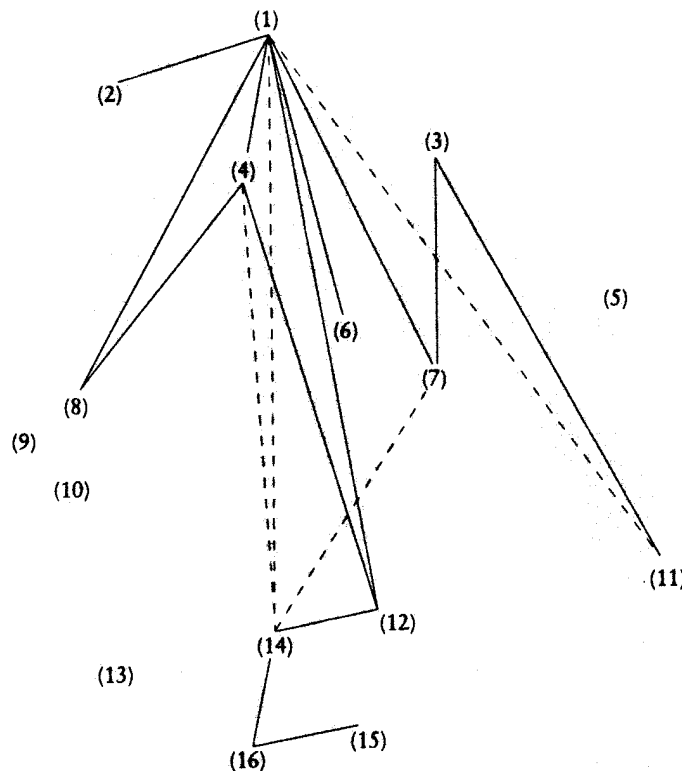
Observando la tabla de conexiones se establece qué umbral es apropiado para determinar un vínculo. Para los textos no-narrativos, Hoey lo establece en tres conexiones. La red de vínculos de un texto se puede representar gráficamente.

(5) *Representación de una red de vínculos: el diagrama topológico*

A partir de la tabla de conexiones se puede representar gráficamente la red de vínculos del texto. Hoey denomina a esta representación *diagrama topológico*. En un diagrama topológico, cada oración está representada por un número y cada trazo que la une con otra oración es un vínculo. Los trazos discontinuos indican que el vínculo no es claro. Sólo hay vínculos entre oraciones que tienen un número de conexiones por encima del umbral establecido.

El siguiente diagrama representa la red de *oraciones vinculadas* del texto objeto de análisis, que son las que aportan al texto coherencia. Esta red está tejida principalmente por patrones de repetición léxica, como vimos en la matriz de repetición.

DIAGRAMA 7: Diagrama topológico del texto del Apéndice IV



Los resultados de este diagrama se pueden representar como una tabla parecida a la del apartado (3), sólo que ahora estamos computando los vínculos entre oraciones. En esta, se le asignan a cada oración dos coordenadas. La primera se refiere al número de vínculos con las oraciones precedentes, y la segunda, al número de oraciones que le siguen. Cuando aparecen dos grupos de coordenadas, estamos ante vínculos formados por conexiones dudosas. Sumando las dos coordenadas, se obtiene el número total de vínculos de cada oración con las oraciones de este fragmento. Por ejemplo, la cuarta oración está vinculada con tres oraciones, una precedente y dos que le siguen.

TABLA 9: Tabla de vínculos del texto del Apéndice IV

1	(-, 6) (-, 8)	9	(0, 0)
2	(1, 0)	10	(0, 0)
3	(0, 2)	11	(1, 0) (2, 0)
4	(1, 2)	12	(2, 1)
5	(0, 0)	13	(0, 0)
6	(1, 0)	14	(1, 1) (4, 1)
7	(2, 0)	15	(0, 1)
8	(2, 0)	16	(2, 0)

Hoey (1995) presenta un formato de tablas de vínculos, adecuado para representar textos más extensos. En el eje de ordenadas aparecen todas las oraciones, y en el de abscisa, las oraciones precedentes (a la izquierda) y las oraciones que siguen (a la derecha). El umbral para formar un vínculo se estableció en tres conexiones. Especificamos entre paréntesis el número de conexiones igual o superior a tres. El texto del Apéndice IV quedaría del siguiente modo:

TABLA 10: Tabla de vínculos (adaptación de Hoey 1995: 77)

	1	2(6) 4(5) 6(3) 7(4) 8(5) 11? 12(3) 14?
1(6)	2	
	3	7(3) 11(3)
1(5)	4	8(4) 12(3) 14?
	5	
1(3)	6	
1(4) 3(3)	7	14?
1(5) 4(4)	8	
	9	
	10	
1? 3(3)	11	
1(3) 4(3)	12	14(3)
	13	
1? 4? 7? 12(3)	14	16(3)
	15	16(3)
14(3) 15(3)	16	

Antes de pasar a ver las aplicaciones de estos diagramas, se harán tres puntualizaciones. En primer lugar, Hoey ha diseñado esta metodología de análisis de conexiones, vínculos y redes de vínculos sobre la base de textos no-narrativos. En segundo lugar, cuando se produce una *coocurrencia* de repeticiones entre oraciones, estas oraciones quedan relacionadas más allá de las conexiones cohesivas individuales que existen entre las palabras. Por último, las conexiones y los vínculos se dan tanto entre oraciones contiguas como entre oraciones separadas y son el resultado de los patrones léxicos propios de cada texto.

4.6. Aplicaciones de la metodología de Hoey

Hoey (1991) apunta algunas de las aplicaciones que puede tener la identificación de *redes de vínculos* en la mejor comprensión de los textos. Antes de nada, es necesario analizar los diagramas topológicos correspondientes a un texto y la tabla con las coordenadas que indican el número de vínculos con oraciones previas y posteriores.

4.6.1. Fragmentación de textos y didáctica de las lenguas

Por una parte, se pueden identificar las oraciones centrales y las oraciones marginales de un texto. Las oraciones centrales son aquellas que contienen un número de vínculos

con otras oraciones mayor de lo que se considera normal para un tipo textual determinado. A raíz de esta identificación y tras suprimir las oraciones marginales, se pueden obtener resúmenes coherentes e inteligibles del texto original. Esto es posible porque, como dijimos en las páginas 116-117, si se yuxtaponen dos oraciones vinculadas, aunque estas no sean contiguas en el texto analizado, el lector las reconoce como mutuamente relevantes y coherentes entre sí.

Por otra parte, estas redes léxicas y la dirección que siguen los vínculos permiten identificar oraciones que inician o finalizan las distintas secciones en las que se divide un texto (*topic-opening, topic-closing sentences*).

El interés por la fragmentación de textos en función de la cohesión léxica ha sido continuado por algunos discípulos de Hoey como Berber Sardinha (1994a, 1994b, 1997a) y Renouf y Collier (1995), como veremos en el apartado 4.8.1. Berber Sardinha también ha explorado la distancia que existe entre las oraciones vinculadas (*bonding distance*). Esta distancia puede ser de muchas oraciones, e incluso de capítulos como indican Phillips (1985) y Hoey (1991). La opinión de Berber Sardinha (1995) al respecto es que la vinculación es más frecuente entre oraciones no contiguas y que, probablemente, esta distancia sea sensible al tipo textual.

Otra aplicación de la metodología de Hoey la encontramos en la didáctica de las lenguas y en los aspectos cognitivos relacionados con la lectura y la escritura. Sobre la hipótesis de que la comprensión lectora se basa en los vínculos oracionales que establece el lector, Hoey incide en la necesidad de explotar en el aula de idiomas, entre otras, una forma de lectura no-lineal que ayude al lector a descubrir los vínculos entre oraciones y a plasmarlos por escrito. Según este autor, la explotación de esta técnica posiblemente aumenta la comprensión, ya que las oraciones más vinculadas son las más relevantes para la interpretación de otras. Además, sostiene que esta técnica despierta la creatividad. En España, Díez Prados (1999) y Suau Jiménez (1998) han seguido esa línea de investigación.

4.6.2. Vinculación e intertextualidad (Hoey 1995)

Hoey (1995, 1996) explora la relación que existe entre el fenómeno de vinculación y la intertextualidad. La intertextualidad es un fenómeno que explica el hecho de que

podamos relacionar unos textos con otros y de que existan distintos *géneros*. Estos géneros no sólo sirven de signos de identidad de una determinada comunidad de discurso, sino que también le aportan vitalidad. Esta intuición de una posible relación entre *vinculación* y *género* ya está presente en Hoey (1991: 190).

Antes de adentrarse en el fenómeno de la intertextualidad, Hoey (1995) busca una base psicológica a la capacidad de identificar en los textos patrones léxicos y vínculos entre oraciones. Estos vínculos han sido identificados en textos en lengua inglesa, portuguesa, española, francesa y alemana. Su razonamiento está guiado por las siguientes preguntas:

- a) ¿El lector tiene realmente consciencia de los vínculos entre oraciones? ¿Son estos vínculos un constructo que sólo puede descubrir el analista?
- b) Si el lector realmente construye estos vínculos ¿qué nos dice este fenómeno sobre el modo en que procesamos el lenguaje en nuestra mente?

4.6.2.1. La percepción de vínculos por parte del lector

Aunque no se ha realizado un estudio experimental que demuestre que los lectores son conscientes de los vínculos que se establecen entre oraciones, Hoey aporta como argumento el hecho de que se hayan podido reconstruir fragmentos de la Biblia y de otros textos literarios antiguos a partir de fragmentos separados de texto. Los estudiosos de estos textos han sido capaces de establecer conexiones entre oraciones que no son contiguas e incluso que pertenecen a partes distintas del texto.

Otro posible indicio de la percepción de vínculos por parte del lector lo podemos encontrar en el ejercicio que Eugene Winter diseñó en 1970 para comprobar que los hablantes de inglés comparten una intuición sobre la construcción del discurso. Este ejercicio aparece descrito en Hoey (1983: 4-7). En el mismo, Winter presentaba a sus alumnos una serie de diez oraciones pertenecientes a un discurso, que habían sido desordenadas y les pedía que reconstruyeran el texto original en 20 minutos. La mayoría de los alumnos fueron capaces de reconstruir de una forma bastante aproximada el texto original. Hoey (1983: 5) afirma que los alumnos estaban utilizando un conocimiento instintivo que Winter (1976) llama *consenso*. Este consenso se basa en la habilidad de

establecer conexiones entre oraciones mediante distintos recursos como la repetición, la anáfora y la catáfora.

Por último, es probable que los lectores establezcan vínculos puesto que, al leer detenidamente un texto, volvemos varias veces a oraciones previas e intentamos relacionarlas con oraciones anteriores y posteriores. Este tipo de lectura no es lineal y permite que el lector detecte qué oraciones son relevantes para la interpretación de otras. Curiosamente, las oraciones relevantes de un texto son las que tienen más contenido en común, las que tienen más vínculos entre ellas.

4.6.2.2. Creación de vínculos y procesamiento del lenguaje

Para intentar responder cómo se crean los vínculos, conjugaremos la opinión de Hoey (1983, 1991, 1995) y la de estudiosos del procesamiento de la información y del lenguaje (Baddeley 1992; Bell 1991; Givón 1998a, 1998b; Faber 1998a).

Como ya se señaló en el apartado sobre el modelo cognitivo de Bell (1991), los procesos mentales para interpretar un texto no son lineales sino que son interactivos y suponen un movimiento tanto de avance como de retroceso a través del texto.

Cuando empezamos a leer un texto, nuestra memoria a corto plazo nos permite concentrar nuestra atención en un máximo de tres oraciones. En la memoria a corto plazo, hay un módulo denominado memoria de trabajo (*working memory*) que va a permitir realizar las operaciones mentales necesarias para procesar cada oración (Baddeley 1992).

Para analizar de forma simplificada los procesos mentales que pueden explicar el fenómeno de vinculación, consideraremos las primeras líneas del quinto texto del Apéndice II.

(18) ONCOLOGÍA

El triunfo del taxol

El fármaco pasa a primera línea en el tratamiento del cáncer de ovario

Ha dejado de ser un segundón. El taxol (paclitaxel en su nombre genérico), el avance quimioterápico frente al cáncer más importante de la última década, ha

adquirido hoy la categoría de producto oncológico de primera línea. Con la publicación en el número de este jueves del *New England Journal of Medicine* de un trabajo en el que se demuestra la superioridad de una combinación de taxol y cisplatino sobre la quimioterapia estándar del cáncer de ovario, los oncólogos tendrán que pensar en él desde el principio cuando se enfrenten a un tumor maligno de la gónada.

En el trabajo, realizado con el patrocinio de los Institutos Nacionales de la Salud de EEUU, se comprobó el resultado de tratar inicialmente a 386 mujeres con cáncer del ovario con dos tipos de quimioterapia diferentes. La mitad de ellas - elegidas de forma aleatoria- recibió cisplatino y ciclofosfamida y la otra mitad cisplatino y taxol. Los resultados al cabo de cuatro años de seguimiento clínico fueron significativamente mejores -tanto en supervivencia como en número de recidivas- para el grupo que recibió inicialmente el taxol que para las féminas tratadas de forma estándar.

Al empezar a leer el texto, el lector se concentra en el título. Con el fin de procesarlo, su memoria de trabajo conecta con la memoria semántica, que, como dijimos en el apartado 2.1., es un compartimento permanente de la memoria a largo plazo. La memoria semántica dispone de un repositorio con las palabras más frecuentes, otro con las estructuras sintácticas más frecuentes y, según Hoey, de un archivo donde están almacenadas las concordancias más frecuentes.

La memoria de trabajo busca en la memoria semántica las palabras y unidades poliléxicas del título con el fin de llegar a una interpretación preliminar del mismo. Si la búsqueda es infructuosa, entran en acción los mecanismos de interpretación descritos por Bell: el mecanismo de búsqueda léxica y el analizador sintáctico. Estos mecanismos son también activados por la memoria de trabajo, que es la que controla la interacción entre la memoria semántica y la memoria episódica.

Una vez que la memoria de trabajo ha llegado a una primera interpretación del título, las palabras plenas quedan almacenadas en la memoria a corto plazo en forma de líneas de concordancia en cuyo nodo central puede aparecer el término *taxol* o *cáncer de ovario*.

(19)

Oncología. El triunfo del **taxol**. El fármaco pasa a primera línea en el tratamiento del cáncer ovárico.
tratamiento del **cáncer ovárico**.

Según Hoey, estas líneas de concordancia se activarán cada vez que aparezca en el texto el lexema que hace de nodo. Puede haber más de una línea de concordancia por oración, según las palabras plenas que se consideren como nodo.

Conforme el lector avanza en la lectura del texto, se van almacenando en la memoria a corto plazo más líneas de concordancia. Hoey (1995: 79) lo explica del siguiente modo:

[...] every time we encounter a sentence we sort the words in it into a set of on-going concordances for those words, thereby creating a multiple record of the sentence which is then capable of being accessed as information through those records.

Aún se desconoce cuántas palabras se almacenan en cada línea de concordancia. Los nodos de estas concordancias son las palabras que la memoria episódica a corto plazo ha reconocido como instancias de otras palabras del texto. Por ejemplo, cuando el lector ha acabado el primer párrafo, es probable que haya almacenado las siguientes concordancias, extraídas con el programa *Microconcord* de *Oxford University Press*.

(20)

1 ONCOLOGIA El triunfo del **taxol** El fármaco pasa a primera línea
2 ico Ha dejado de ser un segundón. El **taxol** (paclitaxel en su nombre genérico),
3 a superioridad de una combinación de **taxol** y cisplatino sobre la quimioterapia

1 primera línea en el tratamiento del **cáncer** ovárico Ha dejado de ser un segundó
2 el avance quimioterápico frente al **cáncer** más importante de la última década,
3 sobre la quimioterapia estándar del **cáncer** de ovario, los oncólogos tendrán
4 primera línea en el tratamiento del **cáncer** ovárico Ha dejado de ser un segundó

1 línea en el tratamiento del cáncer **ovárico** Ha dejado de ser un segundón. El
2 quimioterapia estándar del cáncer de **ovario**, los oncólogos tendrán que pensar

1 n de taxol y cisplatino sobre la **quimioterapia** estándar del cáncer de ovario,

Posteriormente, cuando el lector se encuentra de nuevo uno de los lexemas que hacen de nodo, por ejemplo, el lexema *taxol* del segundo párrafo, su memoria de trabajo lo busca en el archivo de concordancias, activando así las concordancias de otras

palabras de su radio colocacional: *cáncer*, *ovario* y *quimioterapia*. Se han señalado con un subrayado. Como resultado de esta operación, el lector reconoce las conexiones entre lexemas similares a *taxol* y entre los lexemas que aparecen en su cotexto, por ejemplo, *oncología* y *quimioterapia*. Si las conexiones entre los elementos de dos oraciones son más frecuentes de lo habitual, el lector establecerá un vínculo entre estas dos oraciones.

(21)

ONCOLOGIA El triunfo del **taxol** El fármaco pasa a primera línea en el
 ico Ha dejado de ser un segundón. El **taxol** (paclitaxel en su nombre genérico), e
 a superioridad de una combinación de **taxol** y cisplatino sobre la quimioterapia

primera línea en el tratamiento del **cáncer** de ovárico Ha dejado de ser un segund
 sobre la quimioterapia estándar del **cáncer** de ovario, los oncólogos tendrán que
 el avance quimioterápico frente al **cáncer** más importante de la última década,

línea en el tratamiento del cáncer **ovárico** Ha dejado de ser un segundón. El
quimioterapia estándar del cáncer de **ovario**, los oncólogos tendrán que pensar en
 n de taxol y cisplatino sobre la **quimioterapia** estándar del cáncer de ovario, lo

El proceso que se acaba de describir no es el único que tiene lugar en la creación de conexiones entre oraciones. También son necesarias una serie de operaciones léxicas, sintácticas y discursivas. De estas operaciones, podemos mencionar la identificación de relaciones de sinonimia, o de correferencia, la habilidad para hallar el referente de un pronombre o de expandir una expresión léxica utilizando el conocimiento enciclopédico.

De todas formas, Hoey hace hincapié en los procesos que pueden explicar que, en los textos, las conexiones se establezcan entre las oraciones y no simplemente entre las palabras. Por este motivo, Hoey sostiene que el lenguaje se almacena casi sin procesar en forma de concordancias y que continuamente nuevas ocurrencias se añaden a estas conforme leemos y hablamos. Cuando el sistema que guarda en nuestra memoria las concordancias indica que una concordancia tiene más conexiones de lo normal, nuestra mente la interpreta en función de las líneas de concordancia con las que está relacionada. En consecuencia, este sistema permite que se emparejen nuevos enunciados con antiguos enunciados y que se creen conjuntos de concordancias interconectadas gracias a los procesos léxicos, sintácticos y discursivos arriba

mencionados. Estos procesos hacen que exista un mayor número de paralelismos entre las oraciones y permite que se emparejen unas con otras.

Esta hipótesis sobre el fenómeno de vinculación tiene como ventaja que aporta una explicación a otros fenómenos como la adquisición del lenguaje, la colocación, las citas, las bromas y la intertextualidad. También parece estar conforme con la idea de que aprendemos las palabras en contexto y que somos capaces de guardar en la memoria a largo plazo frases y oraciones que se han activado con frecuencia en la memoria a corto plazo. De esta forma, podemos pronunciarlas de una forma automatizada que reduce el esfuerzo a realizar por la memoria de trabajo.

Sin embargo, son perceptibles algunas inconsistencias en su explicación. En primer lugar, nos ha sorprendido la primacía que Hoey otorga a las líneas de concordancia, a pesar de que no haya incluido la colocación como recurso cohesivo.

Por otra parte, si Hoey pretende dar una explicación cognitiva a los procesos de vinculación, debería haber comentado algo sobre la teoría de marcos de Fillmore (1977) y haber citado algunos modelos del lexicón mental como el de Aitchison (1987, 1994) o el de Martín Mingorance (1998). Estos modelos explican cómo las palabras almacenadas en nuestro lexicón están relacionadas entre sí formando una red y cómo al utilizar una palabra se activan sus conexiones con otras palabras que aparecen con frecuencia en el mismo contexto o que tienen un significado o un significante parecido. Estas conexiones podrían ser la base de las líneas de concordancia a las que hace alusión Hoey. Asimismo opinamos que la investigación de Hoey podría haberse enriquecido con una mención a los sistemas de la memoria humana. De esta forma, podría haber encontrado una base neurológica al fenómeno de la vinculación.

4.6.2.3. Intertextualidad

Hoey sostiene que, si los procesos descritos en el apartado anterior se producen en la realidad, constantemente estamos almacenando líneas de concordancia independientemente de si estas proceden del mismo texto o no. En consecuencia, los vínculos que establecen los hablantes no son sólo entre oraciones del mismo texto, sino también entre textos producidos en fechas similares o extraídos de las mismas fuentes. Incluso, Hoey sostiene que ha encontrado vínculos entre textos que no tienen una fuente

común inmediata. Todas estas hipótesis están respaldadas por una argumentación coherente, por un estudio experimental previo (Abraham 1993) y por tres textos sobre astronomía.

Para apoyar la idea de intertextualidad entre textos extraídos de una misma fuente, Hoey argumenta que, cuando escuchamos o leemos un texto y posteriormente decidimos hablar o escribir sobre el tema del que trata, utilizamos léxico proveniente de dicho texto porque tenemos almacenadas en nuestra mente líneas de concordancia provenientes del mismo. De esta forma, establecemos vínculos no sólo con el texto que hemos procesado con anterioridad, sino con las palabras que lo codifican. De ahí que se puedan encontrar vínculos entre el texto de un escritor y los textos que consultó antes de escribirlo.

En cuanto a la vinculación entre textos que no tienen una fuente común inmediata o cuyas fuentes se basan en un texto muy remoto, Hoey ha constatado su existencia en diversos textos, aunque, claro está, la vinculación es menor. Su opinión al respecto queda clara en la siguiente cita:

If we indeed bond the sentences/utterances we write or speak with those we have received and are storing our mental concordances, then, we would expect the authors of the secondary sources to write sentences that bond with those in the primary sources and in turn would expect the authors of the newspaper articles to write sentences that bond with their secondary sources.

(Hoey 1995: 90)

Ante la posible objeción de que no es extraordinario que se utilicen las mismas palabras para hablar de un mismo tema, Hoey arguye que sí es llamativo que las oraciones que están vinculadas sean coherentes y pertinentes entre ellas.

Haciendo balance, la metodología de Hoey puede ser un instrumento valioso para encontrar una relación entre el fenómeno de la intertextualidad y los procesos mentales por los que almacenamos los enunciados con los que interactuamos al comunicarnos. El hecho de que podemos vincular unos textos con otros es una evidencia indirecta de la vinculación constante que hacemos entre nuestros enunciados y los que hemos escuchado con anterioridad.

Una consecuencia de este almacenamiento de colocaciones y expresiones previas en la memoria a largo plazo es que, al interactuar con los textos de una comunidad de discurso, inconscientemente vamos adoptando las convenciones y el sistema de valores de la misma y, en consecuencia, construimos y hacemos que se consolide esa comunidad de discurso (Bazerman y Paradis 1991). Hoey comparte esta idea cuando afirma (ibid: 90):

If one bonds with what one has heard and read, and if everyone else in the field is doing the same, then certain expressions are going to become typical of that field and to distinguish the well-read from the novice.

Hasta aquí se han repasado algunas aplicaciones interesantes de la metodología de Hoey. Se ha hecho hincapié sobre todo en la última, puesto que fundamenta la intertextualidad y la existencia de diversos géneros en las relaciones cohesivas.

4.7. Morris (1988), Morris y Hirst (1991)

Morris y Hirst sostienen que el contenido de un texto se crea a partir de *cadena léxica* que tratan sobre el mismo tema. A diferencia de los autores que hemos citado, su perspectiva se encuadra dentro de la lingüística computacional. En consecuencia, dan menos importancia a la cohesión como fenómeno lingüístico interoracional y conceden la primacía al contenido semántico del texto.

En Morris y Hirst (1991: 22-23) se definen las cadenas léxicas como “sequences of related words (...) spanning a topical unit of the text”. Estas cadenas de palabras que tratan sobre el mismo tema pues, están formadas por lexemas semánticamente relacionados en un texto. Por ejemplo, si un texto contiene las palabras *cáncer* y *melanoma* estas pertenecerían a una misma cadena léxica porque *melanoma* es un tipo de *cáncer*.

Estos autores crean un algoritmo para la identificación de estas cadenas, aunque realizan el análisis de forma manual, buscando las palabras plenas susceptibles de formar parte de una cadena y la relación entre ellas en un tesoro, en concreto, el *Roget's Thesaurus*. Se asigna a cada lexema un número que lo asocia a una categoría

del tesoro. El tesoro permite que se computen no sólo las relaciones entre palabras idénticas, sino también entre palabras que pertenecen a la misma categoría semántica o entre palabras pertenecientes a una categoría distinta pero que están directamente relacionadas. Por ejemplo, estos autores incluyen dentro de la misma cadena cohesiva las formas *cow*, *cows* y *sheep*, pero no computan la relación indirecta que *cow* y *wool* establecen mediante *sheep*³³.

Asimismo, también especulan sobre si se debe incluir dentro de la misma cadena cohesiva lexemas que están distantes en el texto y deciden establecer una distancia máxima de tres oraciones. Si la distancia es mayor, se computa lo que Morris y Hirst denominan *chain return*.

Así pues las cadenas léxicas son conjuntos de palabras de un texto que están relacionadas por vínculos léxicos y por proximidad. El *Roget's Thesaurus* (en adelante, *Roget*) sirve de base de datos para establecer vínculos léxicos o relaciones que, según Morris y Hirst, pueden ser de varios tipos. Las relaciones de *identidad* (*identity*) son las que se establecen entre lemas o inflexiones de la misma palabra (*aberrations-aberration*). Las relaciones *categoriales* (*category*) vinculan miembros que pertenecen a una misma categoría semántica del tesoro, por ejemplo, la relación entre *aberration* y *abnormalities*, que se establece mediante la categoría ABNORMALITY, que sirve de hiperónimo de ambas. Las relaciones de *grupo* (*group*) se establecen entre palabras pertenecientes a categorías vecinas dentro de un mismo *grupo de categorías*, teniendo en cuenta que las entradas del tesoro aparecen asociadas en grupos de categorías jerárquicamente organizados. Esta relación es la que existiría entre dos lexemas pertenecientes al grupo **DEFICIENCY: prognosis**, asociado a la categoría PATHOLOGY, y *disease*, asociado a la categoría DISEASE.

Por otra parte, las palabras contenidas en una categoría del *Roget* a veces hacen referencia a otras categorías. Por eso, se establece una relación entre dos palabras siempre que la entrada de una de ellas haga alusión a la entrada que contiene la otra palabra. A esta relación, Morris y Hirst la denominan *one-point relation*. Por último, Morris y Hirst proponen otra categoría en la que las palabras en cuestión contienen

³³ Estos autores denominan la distancia entre los miembros de distintas categorías transitividad y fijan la transitividad en un vínculo. Por eso excluyen la relación de dos vínculos entre *cow* y *wool*.

entradas que apuntan hacia una tercera categoría. Este último tipo de vínculo genera relaciones muy difusas que pueden distorsionar las cadenas léxicas ya identificadas.

Finalmente, esta aproximación postula que cada cadena cohesiva representa una porción de la estructura cohesiva de un documento. Por este motivo, se afirma que la organización de las cadenas léxicas de un documento refleja en cierto modo la estructura discursiva de ese documento.

Si bien es verdad que los creadores de la noción de cadena léxica no la aplicaron al análisis automatizado de textos, sus investigaciones han servido de base a una serie de estudios cuyo objetivo es la identificación automática de la cohesión con la ayuda de un tesoro. La utilización de un tesoro tiene la ventaja de permitir un análisis más completo puesto que, además de la cohesión entre palabras idénticas (repetición léxica), el ordenador también es capaz de reconocer la que existe entre palabras semánticamente relacionadas.

4.8. Aplicaciones informáticas para el reconocimiento automático de la cohesión

En la década de los noventa, una serie de autores han estudiado el modo de reconocer la cohesión de forma automática. Las diversas aproximaciones siguen principalmente dos tendencias, la de Hoey y la de Morris y Hirst. La primera se deriva de los estudios de Hoey y, ante la dificultad inherente al estudio de la cohesión, limita la cohesión a la repetición de formas léxicas. La segunda tendencia además trata de reconocer la cohesión derivada de relaciones conceptuales. Seleccionamos a continuación aquellas que han servido de base a programas informáticos que, en la mayoría de los casos, están aún en periodo de prueba.

4.8.1. Aproximaciones basadas en la repetición léxica

Entre las aplicaciones de la metodología de Hoey (1991), Renouf y Collier (1995) computan las conexiones y vínculos dentro de textos periodísticos expositivos con la intención de generar resúmenes a partir de los mismos. En este estudio sólo se reconoce la repetición léxica simple (*parliament – parliament*), limitación presente en el

programa informático derivado, *Automatic Abridgement System*³⁴, cuya versión comercial está preparándose en la Universidad de Liverpool.

En esta misma universidad, Berber Sardinha (1997a) ha estudiado cómo contribuye la cohesión a la organización del texto, en concreto, cómo la cohesión puede ayudar a identificar las secciones en las que se divide un texto. Para ello, diseña un programa de *segmentación textual*³⁵, con el que analiza un corpus de 300 textos pertenecientes a tres géneros: el artículo experimental, el informe comercial (*business report*) y la entrada de enciclopedia. Los resultados obtenidos confirman que las oraciones que delimitan las secciones tipográficas de un texto son las más relevantes desde el punto de vista de la cohesión léxica. Berber Sardinha sostiene que estas oraciones son las que están vinculadas con más oraciones, y no las que tienen un mayor número de conexiones con una determinada oración. Para justificar el hecho de que el programa sólo reconoce instancias de repetición léxica simple (*important – important*) y compleja (*important –importance*), Berber Sardinha remite a Benbrahim (1996), donde se estima que, de todos los recursos cohesivos identificables en un texto, entre un 92 y un 97% se fundamentan en estos dos tipos de repetición.

La investigación de Mohamed Benbrahim y Khurshid Ahmad se encuentra en la línea divisoria que separa los programas citados y los que reconocen algo más que la repetición del mismo lexema en un texto. Benbrahim y Ahmad (1994) aplica la metodología de Hoey (1991) con miras a la producción automática de resúmenes y bancos terminológicos. Este estudio cristaliza en el programa de análisis textual *Telepattan*, que identifica las conexiones y vínculos de un texto y las representa gráficamente³⁶. En Benbrahim (1996) se ofrece un análisis detallado del número y tipo de conexiones que aparecen en cinco textos académicos en inglés y se extraen las conclusiones expuestas en el párrafo anterior. El cómputo de las instancias de paráfrasis simple y compleja se realiza en función de las categorías del tesoro del proyecto *Wordnet* (Miller 1991).

³⁴ Comunicación personal con Antoinette Renouf (ajrenouf@liv.ac.uk).

³⁵ El programa se denomina *Words* aunque, por falta de financiación, su autor aún no ha procedido a su distribución.

³⁶ Mohamed Benbrahim (mohamed@ceanet.po.my) está programando en C++ una versión que se distribuirá de forma gratuita entre investigadores.

Telepattan ha servido de prototipo de dos programas de gestión de terminología desarrollados también por la Universidad de Surrey: *Sebriba* y *System Quirk*. *Sebriba*³⁷ genera automáticamente un corpus buscando en la red aquellos textos cuyo contenido se ajuste mejor a la palabra clave de búsqueda e identifica los términos del corpus resultante. *System Quirk* permite elaborar córpora textuales, aunque a partir de las páginas de Internet determinadas por el usuario. Asimismo es de utilidad en la gestión de córpora textuales, ya que dispone de una tipología textual que clasifica los textos en función de siete atributos: *texto*, *autor*, *publicación*, *lengua*, *dominio*, *derechos de autor* y *textos relacionados*. Por último, *System Quirk* también extrae términos. Por este motivo, la investigación que sustenta este proyecto se centra en la actualidad en el estudio de unidades fraseológicas y la elaboración de ontologías específicas para un determinado campo del saber³⁸.

4.8.2. Aproximaciones que incluyen relaciones conceptuales

Dado que la cohesión también se construye a partir de relaciones semánticas, el concepto de *cadena léxica* ideado por Morris y Hirst (1988, 1991) está presente en una serie de estudios que pretenden la identificación de la cohesión mediante un ordenador dotado de tesoro (Green 1997; Ellman 1998, 1999; Ellman, Klincke y Tait 1998).

Green (1997)³⁹ utiliza las nociones de *cadena léxica* y *similitud semántica* en la creación automática de vínculos de hipertexto entre las palabras de un artículo periodístico o entre distintos textos periodísticos. La base de datos léxica de la que se sirve es *Wordnet*.

Ellman (1997, 1998, 1999) crea un generador de cadenas léxicas para optimizar la recuperación de información en Internet mediante búsquedas basadas no sólo en la forma de las palabras, sino también en su significado, de ahí que incorpore un tesoro. Este sistema, que su autor denomina *Hesperus*, analiza las cadenas léxicas de cualquier

³⁷ Los autores de *Sebriba* son Steve Collingham (S.Collingham@surrey.ac.uk) y K. Ahmad (k.ahmad.mcs@surrey.ac.uk).

³⁸ En el sitio de Internet de *System Quirk* [<http://www.mcs.surrey.ac.uk/SystemQ>] también hay una aplicación que elabora en poco tiempo un resumen de cualquier texto que se introduzca en el recuadro destinado a tal efecto.

³⁹ Documento de Internet disponible en <http://sjgreen.mpce.mq.edu.au/~sjgreen/Documents/sjgthesis.ps.gz>.

texto y otorga un perfil genérico del documento (*Generic Document Profile*) que permite comparar la similitud semántica entre ese texto y otros. A partir de textos prototípicos (*exemplars*) de un determinado tema, el sistema elabora el perfil correspondiente y activa un motor de búsqueda que encontrará textos cuyo perfil indique una similitud conceptual.

Al igual que Morris y Hirst, Ellman utiliza el *Roget's Thesaurus* en lugar de *Wordnet* (Miller 1991). Justifica esta elección aduciendo que el *Roget* computa un mayor número de relaciones y relaciones más simples. Además, su estructura es más sencilla y equilibrada, 1024 categorías, en contraposición a las 91595 de *WordNet 1.5*, lo cual permite una puesta en marcha más eficaz.

Hesperus extrae el conocimiento léxico de una versión electrónica que Ellman confecciona a partir de dos versiones del *Roget*⁴⁰. Este tesoro contiene bastante terminología médica, en parte porque su autor, Peter Mark Roget, era un médico que pretendía aumentar su capacidad expresiva⁴¹.

Aunque se constatará esto en el análisis, veamos ahora qué encontramos en el índice bajo la palabra *cancer*. Esta palabra nos remite a *carcinosis* 651 n, es decir, a la categoría ILL - HEALTH - DISEASE, que comprende una serie de sustantivos (n). Bajo *carcinosis* aparecen los siguientes lexemas:

(22)

Epithelioma, cancer; neoplasm, growth; tumour, indolent t., benign t.; malignant tumour, cancerous growth; melanoma, black cancer.

[*Roget's Thesaurus of English Words and Phrases* (1982), Harmondsworth: Penguin books: 251]

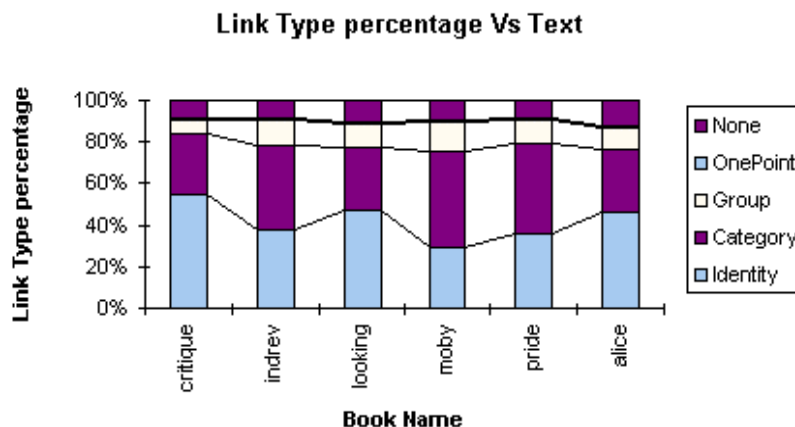
Teniendo en cuenta que se mencionan otras enfermedades en este apartado y que los síntomas y el tratamiento constituyen apartados por sí mismos, parece que el *Roget* puede ser útil en la detección automática de relaciones conceptuales en textos sobre oncología, a pesar de ser un tesoro de la lengua general. De hecho, con esta base de conocimiento léxico, *Hesperus* halla la relación entre *carcinoma* y *melanoma*.

⁴⁰ La primera, publicada en 1911, consta de 56000 palabras y ha sido utilizada por el proyecto Gutenberg. La segunda cuenta con 250000 palabras.

⁴¹ La primera edición, publicada en Londres en 1852 se llamaba *Thesaurus of English Words and Phrases classified and arranged so as to facilitate the expression of ideas and assist in literary composition*.

Por otra parte, la etiqueta que *Hesperus* concede a los vínculos es la de ID (*identity*), CAT (*category*), GRP (*group*) y ONE (*one-point*). De estos tipos de vínculos, en contra de lo esperado, la relación ID es la más frecuente. En Ellman (1998) se afirma que el 60-80% de las cadenas léxicas se deben a la repetición de términos. A diferencia de Morris y Hirst (1991), el programa de Ellman sólo computa como relación ID la existente entre formas idénticas o entre sustantivos que cambian sólo respecto al número. Quedan por tanto fuera de esta categoría las inflexiones de un mismo verbo (*includes-include*), que cuenta como CAT. Siguen en importancia a la relación ID, las denominadas CAT y GROUP. La categoría ONE es muy inusual. Al reconocer la preponderancia de relaciones basadas en repetición léxica, Ellman coincide con lo defendido en Benbrahim (1996) y Berber Sardinha (1997a). Veamos en la siguiente tabla el porcentaje que corresponde a cada tipo de vínculo en seis libros analizados por Ellman, que son los que aparecen en el eje horizontal:

DIAGRAMA 8: Porcentaje de los principales tipos de conexión en seis libros (Ellman, 1999)⁴²



Critique:	<i>The Critique Of Pure Reason</i> de Immanuel Kant
Indrev:	<i>Lectures on The Industrial Revolution in England</i> de Arnold Toynbee
Looking:	<i>Through The Looking Glass</i> de Lewis Carroll
Moby:	<i>Moby Dick</i> de Herman Melville
Pride:	<i>Pride And Prejudice</i> de Jane Austen
Alice:	<i>Alice's Adventures In Wonderland</i> de Lewis Carroll

⁴² Disponible en <http://osiris.sunderland.ac.uk/~cs0jel/web1.gif>.

Aunque describiremos los datos que proporciona *Hesperus* sobre los textos en el apartado de metodología, por el momento, anticipamos que vincula palabras separadas por un máximo de 500 palabras. También enumeramos los pasos que sigue *Hesperus* para elaborar el *perfil genérico del documento*, que, en adelante, denominaremos *perfil*.

1. Lematización y anotación del texto.
2. Supresión de las palabras forma (*stopwords*).
3. Búsqueda de cada lexema en la versión electrónica del tesaurus.
4. Creación de cadenas léxicas.
5. Elaboración del *perfil*.

De todos los programas que hemos revisado, creemos que *Hesperus* es el que mejor se ajusta al estudio de la cohesión en textos extensos, de ahí que este sea el programa utilizado en nuestro análisis. *Hesperus* tiene la ventaja de disponer de un tesaurus, lo cual permite reconocer relaciones semánticas entre palabras. Además, proporciona datos relevantes sobre los textos, tales como el número de palabras plenas, el tipo de relación cohesiva y el porcentaje de activación de los conceptos en el texto.

Aunque *Hesperus* aún no está disponible, nuestra investigación ha sido posible gracias a la desinteresada ayuda de Jeremy Ellman, quien sometió parte de nuestro corpus de textos sobre oncología al programa *Hesperus* y nos envió los resultados como fichero comprimido.

4.9. La cohesión desde la *Teoría Sentido Texto*

4.9.1. Consideraciones generales

La Teoría Sentido Texto (*Meaning Text Theory*), propuesta en Mel'cuk y Zholkovskij (1970, 1988), es un modelo lingüístico que pretende relacionar los textos con sus representaciones semánticas y presentar estas relaciones de una forma rigurosa, ya que, al fin y al cabo, en la lengua las relaciones léxicas están institucionalizadas.

La TST ha sido aplicada principalmente en lexicografía, en concreto, en el *Explanatory Combinatorial Dictionary* (Mel'cuk *et al.* 1984), que es un diccionario formal cuyo objetivo es especificar el universo sintáctico, semántico y léxico de las unidades léxicas. Asimismo, según Saint-Dizier y Viegas (1995: 27), se ha tenido muy en cuenta este modelo en la lingüística computacional. Desde la traducción, Faber (1999a) y Tercedor (1999: 149-159) también destacan la validez de este modelo para la descripción léxica y terminográfica en dos lenguas, estableciendo como *interlingua* las funciones léxicas propuestas dentro de la TST.

Las 53 funciones léxicas propuestas inicialmente⁴³ son un intento de sistematizar las relaciones semánticas y sintácticas que se dan en los textos. Las funciones léxicas, de ahora en adelante FL, son relaciones léxicas institucionalizadas según la definición de Wanner (1996: 1):

A lexical relation is institutionalized if it holds between two lexical units L1 and L2 and L2 has the following characteristics: if L1 is chosen to express a particular meaning M, its choice is predetermined by the relation of M to L2 to such an extent that in case M and L2 are given, the choice of L1 is a language-specific automatism.

Como ilustración, Wanner cita la relación existente entre *aircraft – crew*, *sheep – flock*, *influence – exert* y *attention – pay*. Esta relación se manifiesta de un modo sistemático en la comunicación entre los miembros de una comunidad de discurso.

La representación formal de las FL es similar a la de las funciones matemáticas. Mel'cuk (1996: 39-40) define *función léxica* (f) como la función o correspondencia que asocia una determinada expresión léxica (X), que se denomina *argumento* o *palabra clave*⁴⁴, con un conjunto ($\{Y_i\}$) de expresiones léxicas—*valor* de f— que expresan, al depender de X, un significado específico asociado con f:

⁴³ A estas se irán sumando algunas más. Mel'cuk (1996) propone 64 funciones léxicas simples y sugiere una serie de funciones léxicas complejas derivadas de estas.

⁴⁴ Otros términos para denominar al *argumento* (*palabra clave*) y al *valor* son los de *base* y *colocador* (Hausmann 1985: 119; Irsula 1991: 160; Tercedor 1999: 90). Opinamos que estos dos últimos hacen alusión a relaciones sintagmáticas (fraseología). Debido a que nuestro estudio va a estudiar no sólo estas sino también las relaciones paradigmáticas, creemos oportuno utilizar la nomenclatura de *palabra clave* y *valor*.

$$(23) f(X) = \{Y_i\}$$

$$f(\textit{palabra clave}) = \{\textit{valor}\}$$

$$\textit{Anti}(\textit{malignant}) = \textit{benign}$$

En este ejemplo, la función **Anti** (*Antónimo*) relaciona el lexema *malignant* (palabra clave) con un término de significado opuesto, el lexema *malignant* (valor).

Según Mel'cuk (1996: 41), las expresiones léxicas que constituyen el valor de una determinada función léxica pueden ser de cinco tipos: lexema, frase hecha, ejemplo libre⁴⁵ (*free phrase*), base de palabra compuesta y morfema derivativo.

TABLA 11: Tipos de Funciones léxicas (Mel'cuk 1996: 41)

Type	LF example
a lexeme	Magn (smoker) = heavy
a phraseme	Magn (to smoke) = like a chimney Magn (to rain) = cats and dogs
a free phrase	Magn (to smoke) = without stopping lighting up a cigarette from the other, ...
a compounding stem	Magn (smoker) = chain- [chain-smoker]
a derivational means	S ₁ (to smoke) = ~ -er [smoker]

Como este modelo ya ha sido aplicado a la representación del conocimiento terminológico sobre oncología por Tercedor (1999: 194), en lugar de incluir un listado de FL referentes a la lengua general, se incluirá el repertorio de funciones terminológicas que esta autora elabora a partir de las FL de Mel'cuk.

A ₀	adjective derived from keyword A ₀ (<i>tumor/tumour</i>) = <i>tumoral</i>
A ₁ , A ₂	typical adjective for numbered participant A ₁ (<i>tumor/tumour</i>) = <i>malignant</i> A ₂ (<i>tumor/tumour</i>) = <i>benign</i>

⁴⁵ Proponemos esta traducción para *free phrase* siguiendo a Rodríguez Reina (2000: 326), que define el *ejemplo libre* como un "tipo de frases o combinaciones de palabras que no se someten a las leyes de frecuencia de aparición ni a reglas de colocación".

Able ₁ , Able ₂	ability of numbered participant Able ₁ (<i>tumour</i>) = <i>tumorous</i>
Adv ₀	adverb from keyword Adv ₀ (<i>cancer</i>) = <i>oncogenically (growth)</i>
Adv ₁ , Adv ₂	adverb from numbered participant Adv ₁ (<i>treatment</i>) = <i>innovative</i>
Anti	antonym (exact or near) Anti (<i>malignant</i>) = <i>benign</i>
Bon	standard praise for keyword Bon (<i>growth</i>) = <i>normal</i>
Caus	cause Caus ₁ (<i>cancer</i>) = <i>smoking</i>
Centr	center of Centr (<i>cell</i>) = <i>nucleus</i>
Cont	continue Cont (<i>grow</i>) = <i>keep, keep on</i>
Contr	non-antonymic contrast Contr (<i>apoptosis</i>) = <i>necrosis</i>
Conv _{ijk}	conversive (opposite where participants switch roles) Conv (<i>attack</i>) = <i>suffer an attack</i>
Culm	culmination of Culm ₁ (<i>cancer</i>) = <i>invasion</i> Culm ₂ (<i>cancer</i>) = <i>metastasis</i>
Degrad	degradation of (dimension can be specified: color, movement, etc) Degrad (<i>cancer: evolution</i>) = <i>stages</i>
Epit	standard epithet (representing a part of keyword) Epit (<i>cancer</i>) = <i>advanced</i>
Excess	excessive functioning of (dimension can be specified as in Degrad) Excess (<i>cancer</i>) = <i>metastasis</i>

Fact _{0,1,2}	verb meaning “the realization of” with the keyword as grammatical subject and the participants as objects Fact ₁ (<i>diagnosis</i>) = <i>confirm</i>
Figur	metaphor Figur (<i>cancer</i>) = <i>war</i>
Fin	stopping of Fin (<i>cancer</i>) = <i>cure</i> Fin (<i>tumor</i>) = <i>surgery</i>
Func _{0,1,2}	verb which takes the keyword as subject of first participant, second participant, etc. Func ₁ (<i>diagnosis</i>) = <i>includes CT scans</i>
Gener	generic word Gener (<i>MR</i>) = <i>diagnostic test</i>
Germ	the core of Germ (<i>problem</i>) = <i>crux</i>
Imper	the command associated with Imper (<i>care</i>) = <i>Watch out!</i>
Involv	verb meaning non-participant involvement Involv (<i>scant</i>) = <i>fill</i>
Incep	the beginning of Incep (<i>cure</i>) = <i>treat</i>
Labor _{ij}	verb which takes the participants ij as subject and object and the keyword as secondary object Labor ₁₂ (<i>esteem</i>) = <i>hold</i> (i.e., holds y in esteem)
LabReal _{ij}	verb meaning “the realization of” with the first two participants as subject and object and the keyword as a secondary object (a combination of Labor and Real) LabReal ₁₂ (<i>mind</i>) = <i>bring to</i> (x brings y to mind)
Liqu	the elimination of Liq (<i>cancer</i>) = <i>eradicate</i>
Loc _{in}	preposition for “in”
Loc _{ab}	preposition for “from”
Loc _{ad}	preposition for “to”

	Loc _{in} (<i>cancer</i>) = <i>in</i>
Magn	intensity (can specify dimension of quantity or duration) Magn ₁ (<i>treatment: time</i>) = <i>duration</i> Magn ₁ (<i>treatment: amount</i>) = <i>dose</i>
Manif	is manifest in, with the keyword as subject Manif (<i>small-cell lung cancer</i>) = <i>lung</i>
Minus	less of Minus (<i>process</i>) = <i>revert</i>
Mult	regular aggregate of Mult (<i>cell</i>) = <i>tissue</i>
Nocer	to harm, injure, or impair Nocer (<i>tissue</i>) = <i>affect</i>
Obstr	to function with difficulty Obstr (<i>organs</i>) = <i>malfunction</i>
Oper _{1,2,3}	verb which takes numbered participant as subject and keyword as object Oper ₁ (<i>cancer</i>) = <i>develop</i> Oper ₂ (<i>cancer</i>) = <i>suffer from</i>
Perm	permit or allow Perm (<i>treatment</i>) = <i>tolerate</i>
Plus	more of Plus (<i>cancer</i>) = <i>metastasis</i>
Pos _{1,2,3}	positive attributes of numbered participants Pos ₁ (<i>cancer</i>) = <i>curable</i>
Pred	copula for nouns and adjectives Pred (<i>cancer</i>) = <i>to develop</i>
Propt	preposition for “because of” Propt (<i>cancer</i>) = <i>as a result of</i>
Prox	to be on the verge of Prox (<i>tumor</i>) = <i>precancerous</i>

Qual _{1,2,3}	Highly probable qualities of numbered participants Qual ₁ (<i>cancer</i>) = <i>mortal or fatal</i> Qual ₂ (<i>cancer</i>) = <i>curable</i> Qual ₃ (<i>cancer</i>) = <i>preventive</i> Qual ₄ (<i>cancer</i>) = <i>detectable</i>
Real _{1,2,3}	verb meaning to realize with keyword as object and numbered participants as subject Real ₁ (<i>ambition</i>) = <i>realize</i>
S ₀	noun for keyword S ₀ (<i>immune</i>) = <i>immunity</i>
S _{1,2,3}	typical noun for numbered participant S ₁ (<i>cancer</i>) = <i>victim</i> S ₂ (<i>cancer</i>) = <i>patient</i>

Según Tercedor, las FL denominadas con la letra **S** y un subíndice son útiles para detallar los instrumentos y pruebas utilizados en el diagnóstico y el tratamiento del cáncer, la localización del tumor, el medio y el modo, junto con el resultado. Estas funciones pueden combinarse entre sí para una descripción exhaustiva de los conceptos.

S _{instr}	typical instrument S _{instr} (<i>cancer of the stomach</i>) = <i>endoscope</i>
S _{loc}	typical location S _{loc} (<i>endocrinal cancer</i>) = <i>pancreas</i>
S _{med}	typical means S _{means} (<i>drug</i>) = <i>oral</i> ⁴⁶
S _{mode}	typical mode S _{mode1} (<i>cancer</i>) Ligu ₁ = <i>chemotherapy</i>
S _{res}	typical result S _{res} (<i>apoptosis</i>) = <i>cell death</i>
Sing	one instance of Sing (<i>treatment</i>) = <i>course</i>

⁴⁶ Tercedor utiliza esta función para designar factores de riesgo, que opinamos que están relacionados con S_{cause}.

	Sing (<i>tratamiento</i>) = <i>pauta</i>
Son	to emit typical sound Son (<i>difficult breathing</i>) = <i>whistling</i>
Sympt	to be a physical symptom of Sympt (<i>small-cell lung cancer</i>) = <i>coughing up blood</i>
Syn	synonym and near synonym Syn (<i>cancer</i>) = <i>malignant neoplasm</i>
V ₀	verb for keyword V ₀ (<i>metastasis</i>) = <i>metastasize</i>
Ver	true, correct, or proper Ver (<i>treatment</i>) = <i>effective</i>

En 6.6. *Funciones léxicas en la generación de textos cohesivos*, clasificaremos estas funciones de acuerdo con diversas aproximaciones dentro de la TST y con la forma en que se crea cohesión.

4.9.2. Aproximaciones a la cohesión

En la TST hemos encontrado una gran precisión en la descripción semántica de los lexemas y algunas alusiones a la cohesión (Alonso *et al.* 1995: 355-365; Lee 1996: 299-306; Mel'cuk 1996: 95-96) pertinentes para nuestra investigación, a pesar de que no logren explicar el fenómeno cohesivo en textos completos. Sus ejemplos sólo implican una relación semántica unidireccional o bidireccional entre dos oraciones pertenecientes al mismo párrafo, cuando, como se ha visto en este capítulo, las relaciones semánticas normalmente son multidimensionales y trascienden más allá del párrafo.

Mel'cuk alude a cómo las FL pueden ser útiles en la selección de expresiones capaces de establecer relaciones anafóricas tales que garanticen la cohesión del texto y la secuencia oracional sin incurrir en repeticiones tediosas. La base de todas estas relaciones se fundamenta en el conocimiento léxico. El ejemplo de cohesión que presenta es el siguiente:

(24) Sales *increased* slightly in Quebec and Ontario. Modest *gains* were also reported in British Columbia.

En este fragmento, en lugar de repetir elementos léxicos de la primera oración y decir *Sales also increased in British Columbia*, se utiliza el valor de la función léxica S_2 ⁴⁷ asociada al lexema *increase*. Formalmente, S_2 (increase) = gain_N [the amount by which X increased]. La aplicación de esta FL genera un texto más variado y elegante desde el punto de vista retórico.

Alonso Ramos *et al.* (1995) también estudian la referencia anafórica como fenómeno cohesivo y cómo las FL paradigmáticas pueden formalizar fenómenos léxico-semánticos implicados en la generación de textos. Al hablar de cohesión, la sitúan en el plano paradigmático y en un contexto textual amplio que abarca las oraciones anteriores y posteriores que contienen elementos léxicos en relación semántica con el lexema en cuestión.

En su análisis de la anáfora como recurso cohesivo léxico, estos autores establecen tres criterios que ha de cumplir toda expresión que se considere anáfora de otra:

1. Tener la misma referencia que su antecedente
2. Ser una palabra plena, con lo que se restringe el campo de estudio a la cohesión léxica
3. Compartir un vínculo léxico-semántico con el antecedente. De ahí que no se considere como anáfora vínculos de naturaleza extralingüística como el existente entre *the Prime Minister of France* y el nombre de la persona que ocupa ese cargo.

Para poder delimitar el vínculo léxico-semántico arriba mencionado, esquematizan en forma de marco la representación conceptual subyacente a los lexemas analizados. En este marco se sistematizan los roles semánticos de los argumentos que aparecen en cada oración. Para sistematizar aún más estas relaciones léxico-semánticas

⁴⁷ Typical noun for numbered participant.

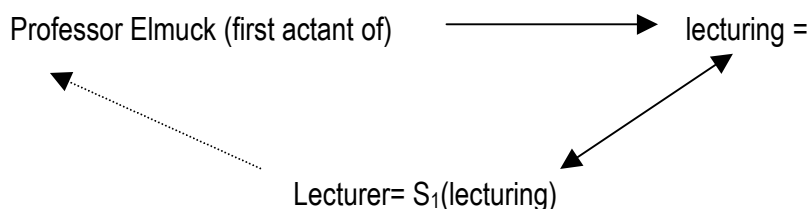
y darle un alcance más allá de la oración, estos autores aplican las funciones paradigmáticas de la TST al estudio de relaciones anafóricas. En concreto, utilizan las funciones léxicas **Gener**, **Syn**, **Conv**, **S₀**, **S_{instr}**, **S_{med}**, **S_{res}**, **S_{loc}** y **S_{mod}**, **Contr** y **Anti**, tanto en su forma simple (**Gener**) como en su forma compuesta (**Gener S₀**).

Por otra parte, según estos autores, la referencia entre el antecedente y la anáfora se puede establecer directamente (*direct lexical coreference*) o de un modo indirecto (*indirect lexical coreference*). En el último caso, la referencia se establece a través de otra expresión que está relacionada paradigmáticamente con el antecedente. Como ejemplo podemos mencionar la relación referencial indirecta entre *Professor Elmuck* y *lecturer* en el siguiente texto:

(25) *Professor Elmuck* was *lecturing* on Lexical Functions to third year students. The *lecturer* was interesting.

Esta relación indirecta entre *Professor Elmuck* y *lecturer* se establece mediante el lexema *lecturing*, cuyo primer actante o argumento es *lecturer* [**S₁** (lecture) = *lecturer*]. Formalmente, hemos presentado esta relación correferencial indirecta mediante líneas discontinuas, mientras que las relaciones semánticas las hemos marcado con líneas continuas.

DIAGRAMA 9: Referencia anafórica directa e indirecta



En cierto modo, esta forma de referencia es similar a la que Hoey (1991: 64-65) denomina paráfrasis compleja mediante triángulo de conexiones (*link triangle*).

En conclusión, si bien los criterios de definición de la referencia anafórica, especialmente el que la restringe a la que se establece entre palabras plenas, y el

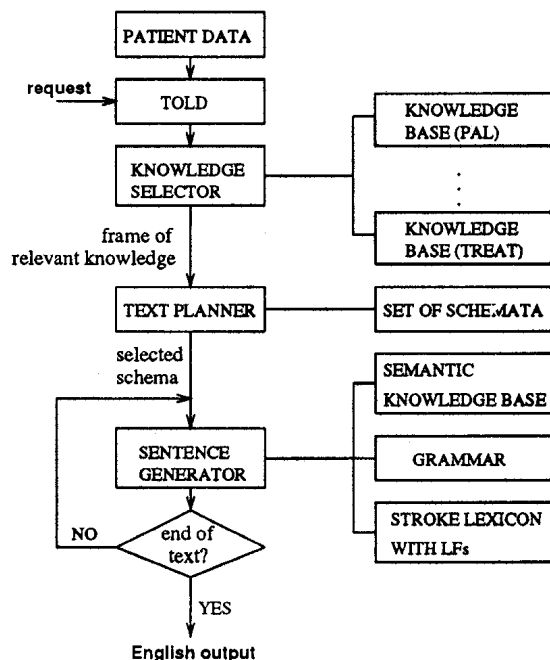
concepto de marco propuesto por Alonso Ramos *et al.* nos pueden ser de utilidad para nuestro estudio, creemos conveniente aumentar la taxonomía de FL implicadas en la cohesión. También consideramos necesario incluir las relaciones sintagmáticas, dado que la repetición de las mismas tiene un valor cohesivo y aplicar todos estos planteamientos a textos completos y a un corpus extenso de textos.

Otros estudiosos de la cohesión dentro de la TST como Lee y Evens (1996) aplican el concepto de cohesión léxica a la creación de un sistema experto capaz de generar párrafos explicativos sobre el diagnóstico y tratamiento de pacientes afectados de apoplejía. Su análisis de la cohesión es el más completo dentro de la TST y su alusión a los trabajos de McKeown (1985), Morris y Hirst (1991) y Myers (1991) demuestra un mayor conocimiento de dicho fenómeno.

Al igual que nosotros, Lee y Evens reconocen dos tipos de cohesión léxica, aunque los denominan con la nomenclatura de Halliday y Hasan (1976): reiteración y colocación, en lugar de acuñar un término en línea con la TST, algo así como relaciones cohesivas paradigmáticas y sintagmáticas.

Este generador de textos cuenta con: a) un selector de conocimiento que consulta una base de datos y elige el marco más apropiado; b) un planificador textual que selecciona entre un conjunto de superestructuras (*schemata*); c) un generador de oraciones conectado a tres repositorios, uno con información semántica, otro con información gramatical y el tercero, un lexicón en el que se especifican las FL en las que pueden participar los lexemas del mismo. Los componentes de este generador no difieren demasiado con respecto a los propuestos en el modelo cognitivo de la traducción de Bell (1991).

DIAGRAMA 10: La organización del generador (Lee y Evens 1996: 301).



Estos tres elementos se corresponden con los tres pasos que sigue el generador a la hora de producir textos. El primer paso consiste en la selección de uno de los marcos contenidos en la base de datos conceptual. Como ilustración del tipo de marcos que contiene esta base de datos, veamos el correspondiente a la aspirina (ibid: 301):

(26) Drug frame (

(generic	: Aspirin)
(brand_name	: Bufferin, Excedrin, Easprin)
(class	: Anti-Platelet)
(avg_oral_dose	: 325_mg)
(min_oral_dose	: 325_mg)
(indication	: cerebral_embolism_of_cardiac_origin)
(contraindication	: peptic_ulcer_disease)
(risk	: NIL)
(efficacy	: NIL)
(chemistry	: NIL))

En un segundo paso, se procede a la planificación de los textos. Un planificador textual selecciona el esquema textual o *superestructura* más apropiada para la situación

comunicativa que se presenta. En el caso de la receta de un medicamento, la superestructura que guía el mensaje del facultativo es básicamente la siguiente: el nombre del medicamento, la posología, la vía (oral por defecto), y las indicaciones y contraindicaciones.

Por último, el generador presta atención al nivel oracional y a la elección de lexemas para transmitir el significado pretendido y para mantener la cohesión del texto pero sin incurrir en una excesiva repetición. En esta etapa son importantes dos facetas: la generación de expresiones referenciales y la generación de colocaciones apropiadas.

En la primera faceta, las FL relacionan los lexemas y determinan cómo se articulan y cómo se pueden intercambiar estos según la información que se presente sea nueva o no, lo cual determina el artículo que debe acompañar a cada lexema. Así por ejemplo, el hecho de que *aspirin* esté en relación con *anticoagulant* mediante la función léxica **Gener**, hace que la expresión *this anticoagulant*, donde *anticoagulant* es un término genérico de *aspirin*, pueda hacer referencia a *aspirin*. A este tipo de referencia Lee y Evans lo denominan *sustitución superordinada*:

(27) *Aspirin* was prescribed for Mrs. Washington. The dosage was 325 mg., the average oral dose. *This anticoagulant* is indicated in cerebral embolism of cardiac origin.

En la búsqueda de colocaciones apropiadas para un término, el generador busca aquellas que hacen de *valor* de la función léxica que se pretenda expresar, una información que tiene almacenada. En uno de los ejemplos que presentan Lee y Evans, el generador tiene que seleccionar la expresión adecuada para expresar la idea de descenso de la temperatura corporal. Dado que, entre la información almacenada por el generador figura que el lexema *temperature* se relaciona mediante la función **CausPredMinus** con la expresión *[to]reduce*, es posible producir la oración:

(28)

Dilation of the blood vessels in the skin reduces the temperature of the body

CausPred (*temperature*) = *[to] reduce*

En definitiva, esta última aproximación pone de manifiesto las aplicaciones de la TST a la lingüística computacional. En concreto, las funciones léxicas pueden servir de fundamento a la especificación de sublenguajes, es decir, de "subsistemas de la lengua restringidos para hacer referencia a un determinado campo del saber" (Kittredge y Lehrberger 1982). Como afirma Frawley (1988: 344), los análisis textuales de sublenguajes científicos que han llegado a buen término tienen una base léxica.

4.9.3. Implicaciones

En vista de todas estas aproximaciones al fenómeno cohesivo, parece plausible sistematizar las relaciones semánticas entre los términos que crean cohesión mediante funciones léxicas abstractas. No obstante, creemos que la TST deja un poco de lado la faceta textual de la cohesión, el hecho de que la cohesión se crea entre palabras que conforman un texto y no entre palabras aisladas. A raíz de esto, sus representantes circunscriben el análisis a fragmentos de texto y no se despegan demasiado de las especulaciones sobre la cohesión de Halliday y Hasan (1976). Los estudios que más se aproximan a nuestra propuesta son los de Alonso Ramos et al. (1995), Lee y Evens (1996) y Escalier y Fournier (1997). En este último estudio, se propone una representación del significado que tenga en cuenta las diferentes realizaciones léxicas y sintácticas de este en una lengua determinada y cómo estas quedan relacionadas mediante la paráfrasis. Intuimos que estas expresiones con un significado muy parecido se van repitiendo a lo largo de todo un texto mediante determinadas funciones léxicas paradigmáticas. Se estudiará esto en el capítulo 10.

Creemos pues posible aplicar la rigurosidad de descripción semántica de la TST al estudio de la cohesión en textos no muy extensos (análisis intensivo de la cohesión). En este tipo de análisis, se combinarán las funciones léxicas con una adaptación del modelo de Hoey, que concede gran relevancia a criterios formales y textuales. En el capítulo 7, *Esquemas de representación para el análisis intensivo de la cohesión*, se aplicará la TST a la clasificación de los recursos cohesivos léxicos de los textos y se interpretarán estos recursos a la luz de las cadenas cohesivas del texto analizado en el capítulo 10.

4.10. Cohesión y tipo textual

Hasta aquí se han repasado distintos estudios que apuntan hacia la doble naturaleza léxica y textual de la cohesión. Es ahora oportuno prestar atención a las líneas de investigación que han explorado la relación que puede existir entre la cohesión, principalmente, la cohesión léxica y los tipos textuales.

4.10.1. Cohesión referencial y género

Uno de los primeros estudios que pretende encontrar una relación entre *recursos cohesivos* y *género* es el de Frawley y Smith (1989).

Desde el presupuesto de que un texto es un texto porque sus constituyentes presentan coherencia entre ellos, Frawley y Smith (1989) quieren determinar si la coherencia y la forma de establecer relaciones cohesivas basadas en la referencia cambia de unos géneros a otros.

Basándose en la premisa de que la referencia es la forma cohesiva más utilizada⁴⁸, estos autores se concentran en el estudio de la distribución y la función cohesiva de los pronombres, que establecen principalmente conexiones endofóricas y exofóricas en los textos. Su hipótesis es que la forma en que se distribuyen estos pronombres es indicativa del tipo de discurso en el que se establecen esas conexiones. Frawley y Smith utilizan un corpus de 16 000 palabras extraído del *Brown Corpus* y que contiene instancias de los siguientes discursos, que Frawley y Smith denominan erróneamente géneros: periodismo, ficción, religión y ciencia.

Aunque reconocemos el carácter pionero de su estudio y el valor de haber utilizado un corpus informatizado, hemos encontrado algunos fallos metodológicos. En primer lugar, los autores no aclaran qué recursos establecen la coherencia y qué pronombres concretos son los que se están teniendo presentes. Tampoco definen qué entienden por *género*, de forma que llegan a confundir este concepto con el de *discurso*, que implica dominios como la ciencia y la religión.

En un primer momento, estudian la distribución de pronombres nominativos de tercera persona del singular, que, según Frawley y Smith, son indicio de cohesión

⁴⁸ En contraposición a esta postura, Hoey (1991) sostiene que el enlace cohesivo más frecuente es la cohesión léxica.

endofórica o intratextual en un texto concreto. Los resultados correspondientes a los cuatro tipos de discurso mencionados indican claras variaciones. Por ejemplo, se constata una parquedad de pronombres en los textos científicos, en comparación con los otros tipos de discurso. Esta falta de referencias pronominales ha sido señalada por Ellegard (1978) y Gopnik (1972) y pretende evitar la ambigüedad. Otra explicación de este fenómeno se encuentra en que, en los textos científicos, la cohesión endofórica se establece mediante repetición léxica o la mención de palabras léxicamente relacionadas.

En cuanto a la utilización de pronombres de referencia exofórica, que, según Halliday y Hasan, son los de primera y segunda persona, también se perciben diferencias. Y así, los pronombres exofóricos también son mucho menos frecuentes en los textos científicos.

A grandes rasgos, sus conclusiones son las siguientes. El género de ficción es el que más utiliza la endófora y la exófora, al igual que los textos religiosos. Los textos periodísticos recurren con frecuencia a la endófora y con poca frecuencia a la exófora. Los textos científicos hacen un uso escaso de la endófora y la exófora. A partir de estos resultados, se puede inferir que distintos *tipos textuales* llevan asociados distintas configuraciones de la *cohesión referencial* mediante pronombres, que tiene una naturaleza textual.

4.10.2. Cohesión léxica y conocimiento científico en textos especializados y divulgativos

Los patrones cohesivos que se dan en los artículos científicos especializados difieren con respecto a los que se dan en los artículos científicos de divulgación. Myers (1991) encuentra una explicación a esto en el hecho de que los lectores de dichos textos no poseen los mismos conocimientos sobre el mundo científico. Este conocimiento es el que permite al lector unir una oración a la siguiente, relacionar referencias a la misma entidad y distinguir referencias a distintas entidades.

Los lectores de textos científicos necesitan conocer las relaciones léxicas que existen entre los lexemas del texto para ver la cohesión implícita del mismo. Por el contrario, los lectores de artículos de divulgación tienen que reconocer las relaciones cohesivas del texto, que están marcadas de una forma explícita, para inferir relaciones

léxicas que desconocen debido a que son legos en la materia de la que trata el artículo. Mediante estas relaciones cohesivas son capaces de unir los campos semánticos pertenecientes a un dominio especializado al dominio de la vida diaria.

Al leer textos científicos especializados, lectores no expertos como estudiantes, profesores de idiomas, traductores y periodistas tienen problemas a la hora de entender la terminología y a la hora de conectar una oración con la siguiente. Esto es así porque no es suficiente saber el sentido de los términos y la función de los *conectores*. También es necesario conocer las relaciones léxicas existentes dentro del texto, que no son explícitas en los artículos especializados.

En el caso de los textos de divulgación, el lector no necesita aportar este conocimiento léxico al texto porque la cohesión explícita en el mismo le ayuda a inferir los conocimientos sobre la materia que no posee. Estos recursos cohesivos, que son mucho más variados en los textos de divulgación, construyen una especie de puente entre el vocabulario cotidiano de los lectores y la terminología científica que aparece en el texto. Mientras que, en los textos especializados, el recurso más empleado es la cohesión léxica, en los textos divulgativos, además de este recurso, se emplean con frecuencia pronombres, conjunciones y elipsis.

Para poder entender la cohesión de un texto, es necesario tener una serie de conocimientos (Brady y Berwic 1983; Lehnert y Ringle 1982). Con la intención de especular sobre el conocimiento que el lector lego requiere para entender textos especializados, Myers se adentra en dos dominios cuya principal preocupación es la representación del conocimiento: el procesamiento del lenguaje natural y la sociología de la ciencia.

Para los primeros, el conocimiento se entiende de un modo cognitivo, como algo que puede ser representado en una serie de proposiciones (Brady y Berwick 1983; Lehnert y Ringle 1982). El conocimiento se encuentra en los enunciados y es posible representarlo en un gráfico como predicados lógicos y redes semánticas. Para la sociología de la ciencia (Latour 1987; Bazerman 1988), el conocimiento es algo que poseemos los seres humanos para poder encontrar sentido a nuestra interacción diaria con otros seres humanos y con objetos, con nuestro entorno. Desde la sociología de la

ciencia, el conocimiento está enfocado a la toma de decisiones y a la acción; es algo eminentemente práctico.

Myers defiende que ambos tipos de conocimiento son necesarios para poder detectar la cohesión de un texto. Por una parte, el lector ha de comprender la estructura semántica de un texto y, por otra parte, debe entender qué función juega el texto en la comunidad de discurso y en la construcción del conocimiento científico.

Por ejemplo, en un artículo de genética, el lector tiene que tener dos tipos de conocimiento. En primer lugar, tiene que saber qué moléculas componen el ADN. En segundo lugar, tiene que saber qué hacen los expertos con esas moléculas, los procesos, técnicas e instrumentos implicados en la investigación. También tiene que estar familiarizado con el género del artículo experimental y los recursos retóricos y movimientos (Swales 1981, 1990, 1994) que se emplean en este género. Este último conocimiento sobre el género tiene implicaciones en la cohesión porque permite que los lectores establezcan relaciones entre elementos aparentemente inconexos. Un lector consciente de que, en la introducción de un artículo experimental, se indica lo inapropiado de aproximaciones anteriores y lo que pretende aportar ese artículo al campo de la genética, tendrá más posibilidades de interpretar correctamente las relaciones entre palabras.

En su estudio, Myers compara los recursos cohesivos empleados en la introducción a algunos artículos científicos con los que aparecen en textos divulgativos. El corpus incluye distintos tipos textuales: artículos científicos, noticias científicas, artículos de divulgación y secciones de libros de texto.

4.10.2.1. Textos especializados: requisitos epistemológicos y cohesión

El conocimiento especializado que ha de tener el lector lego para identificar las entidades que aparecen en los textos e inferir conocimiento de los mismos comprende básicamente los siguientes aspectos:

(a) Reconocer casos de homonimia y polisemia

El lector lego necesita saber con qué sentido se utilizan los términos especializados y qué sentido adquieren en el contexto de la ciencia palabras de la lengua

general. Así, el lector lego necesita ampliar el marco de algunos verbos o nombres de la lengua general y aprender qué argumentos llevan en ese dominio.

Por ejemplo, en un texto especializado, el verbo *to infect* puede tener el siguiente marco según el primer argumento tenga un rol semántico de *fuerza* o de *agente* (Dik 1989).

<p>(29) <i>one organism infects another organism</i> <i>a researcher can infect one organism with another organism</i></p>	<p>(<i>organism</i>) _{Fuerza} (<i>researcher</i>) _{Agente}</p>
---	---

Este marco no se encuentra en la lengua general común, en la que el primer argumento del verbo *to infect* tiene el rol semántico de *Fuerza*.

También se encuentran casos de unidades léxicas repetidas con la misma forma aunque se refieren a distintas entidades o a distintos aspectos de la misma entidad. Por ejemplo, en artículos sobre genética, el ADN aparece con distintos sentidos: *unidad, molécula con dos filamentos, sustancia de la que se toman cantidades o calificador de entidades*.

(b) *Asociar palabras que tienen el mismo o un significado parecido: sinonimia, hiponimia y antonimia*

El lector no especializado tiene problemas a la hora reconocer la misma entidad bajo distintas formas. No obstante, esto no es demasiado común en los artículos especializados, en los que se utiliza la repetición léxica en lugar de la sustitución por sinónimos, para evitar la falta de precisión. Cuando hay una sustitución de una palabra por lo que parece un sinónimo, los expertos infieren que tiene que haber una diferencia en su significado porque, de lo contrario, se hubiera recurrido a la misma palabra.

Además de sinónimos (*paráfrasis simple*), la cohesión léxica también depende de la utilización de antónimos (*paráfrasis compleja*) y términos pertenecientes a un conjunto ordenado de palabras plenas (*sistemas de incompatibilidad múltiple*). Para entender estas oposiciones, el lector necesita tener conocimiento experto puesto que las

pistas sintácticas y estructurales en artículos especializados no son tan explícitas como para permitir reconocer esos contrastes.

Las relaciones de hiperonimia e hiponimia son más frecuentes en los textos de divulgación que en los especializados. Si aparecen en los últimos, normalmente indican un cambio de énfasis y no se hacen explícitas mediante la cohesión.

(c) Determinar relaciones de dominancia en los sintagmas

El lector ha de conocer distintas colocaciones y saber cuál es el nodo en esa colocación para reconocer las relaciones de dominancia que se dan en los extensos sintagmas nominales típicos de los textos científicos. Este conocimiento es compartido por los expertos en la materia.

Además de la extensión de los sintagmas, otro problema es que algunos de los lexemas se pueden interpretar como sustantivos o como adjetivos. En contraposición con los textos especializados, los textos de divulgación evitan esta confusión convirtiendo estos sintagmas en oraciones de relativo y en sintagmas preposicionales. Para poder detectar estas relaciones de dominancia es necesario una labor terminográfica seria que establezca todas las relaciones posibles entre los términos y las relaciones que se activan cada vez que se activa un término.

(d) Definir segmentos discursivos en los que existen funciones comunicativas distintas

Este conocimiento va más allá del conocimiento de relaciones léxicas y supone reconocer unidades estructurales superiores a la oración. Supone la capacidad de reconocer la intención del autor en las distintas secciones de un artículo. Por ejemplo, un lector comprenderá mejor la introducción de un artículo si sabe que el contenido de esta se organiza en función de tres funciones comunicativas:

Move 1	Establishing a research territory
Move 2	Establishing a niche (Obligatory)
Move 3	Occupying the niche

(véase *Modelo de Swales* en 3.1.2.1.)

El desconocimiento de esta estructura probablemente impide que el lector lego entienda la función social del artículo y que encaje las oraciones dentro de un *esquema* que codifique los actos de habla que se suceden en el texto.

4.10.2.2. Textos de divulgación: recursos cohesivos y adquisición de conocimiento

La importancia de los textos no especializados para el lector lego radica en que le permiten adquirir conocimiento de experto mediante recursos cohesivos explícitos. Estos recursos ayudan al lector a reconocer las relaciones léxicas entre los términos y establecen un puente entre dos registros, el del especialista y el del aficionado.

En los textos de divulgación, el lector hace inferencias a partir de la cohesión que encuentra en el texto. Como Halliday y Hasan (1976: 299) defienden:

It is the continuity provided by cohesion that enables the reader or listener to supply all the missing pieces, all the components of the picture which are not present in the text but are necessary to its interpretation

Uno de los recursos cohesivos más frecuentes es la utilización de términos especializados definidos en función de términos de la lengua común o de términos semiespecializados. Por ejemplo, Myers (ibid: 17) menciona la siguiente definición:

(30) chromosomes ... are chemical threads that carry the instructions needed to run every *cell*.

Después de estas definiciones, se repite el término especializado para que el lector aumente su conocimiento. También aparecen casos en los que hay un término especializado y, unas líneas más abajo, aparece un término superordinado que ayuda al lector a comprender el significado del término especializado.

Las metáforas son también un recurso frecuente en los textos divulgativos, como se percibe en este ejemplo:

- (31) Instructions for the protein repertoire in both prokaryotes and eukaryotes are **written in the four letter alphabet** of DNA.

También los lectores de textos divulgativos pueden percibir relaciones entre oraciones viendo el uso explícito de conjunciones. En los artículos especializados estas relaciones no suelen ser explícitas.

4.10.2.3. Recapitulación

Myers aporta a los estudios sobre cohesión una interesante reflexión sobre la interrelación entre conocimiento y cohesión, y entre registro y cohesión. Los distintos tipos de conocimiento que requieren los lectores de artículos científicos y los lectores de artículos de divulgación determinan que se utilicen distintos recursos cohesivos.

Mientras que los lectores de artículos especializados tienen que conocer las relaciones léxicas entre términos para detectar la cohesión implícita propia de estos textos, los lectores de textos de divulgación tienen que reconocer las relaciones cohesivas, que son explícitas en estos tipos textuales, para inferir las relaciones léxicas entre términos.

Esta diferencia puede explicar los distintos recursos cohesivos utilizados en los dos tipos textuales. En los artículos especializados, la cohesión se crea principalmente mediante repetición léxica. En los textos de divulgación, esta también crea cohesión, pero se utilizan otros recursos como la referencia, la sinonimia y la conjunción.

Teniendo en cuenta que en los artículos científicos no se utilizan apenas pronombres o sinónimos como recursos cohesivos, estos textos resultan de difícil lectura para los lectores no especialistas. Por el contrario, los textos de divulgación contienen más recursos que hacen explícita la cohesión y, en consecuencia, es más fácil que el lector detecte las relaciones entre las oraciones y haga inferencias sobre el significado de términos desconocidos.

De todas formas, según Myers, la lectura de textos no especializados no es suficiente. Es necesaria la lectura de distintos géneros que muestren distintas situaciones sociales como artículos de divulgación a distintos niveles, artículos de revisión, informes sobre el desarrollo de la investigación y libros de texto, entre otros. Es decir, la

lectura de artículos de divulgación no puede servir de sustituto a la lectura de textos especializados porque, como dice Myers (ibid: 23),

Learning the language of molecular genetics, like learning any language, involves learning collocations and idioms. The reader (or translator) needs to know which words and phrases count as technical terms (like *gene expressions* or *translation*), which are the most common collocations in the field (such as coding *for* a protein of DNA *strand*), and which are drawn from general usage (such as *directs* or *instructions*). Scientific language is by its nature formulaic: a relatively limited set of verbs goes with a relatively limited set of nouns. Popularizations explore the relations between this register and other registers. But they do not necessarily show institutionalization of expressions.

4.10.3. Cohesión y dificultad comunicativa

Tanskanen (1995) establece que la cohesión léxica es sensible a un aspecto extralingüístico esencial en la configuración de un determinado género: las dificultades comunicativas que impone el contexto situacional. En su estudio, se identifican distintas categorías de cohesión léxica y se estudian su frecuencia y densidad con el fin de descubrir los recursos cohesivos que se adaptan mejor a situaciones en las que la comunicación es más difícil o más fácil. Para ello, investiga cómo se establece la cohesión léxica en tres géneros que responden a una situación comunicativa diferente: la conversación cara a cara, el discurso preparado y la prosa académica.

Su clasificación de los artificios que crean cohesión léxica aúna las taxonomías de Halliday y Hasan (1976) y Hasan (1984). Así, reconoce dos grandes grupos, la reiteración y la colocación y, dentro del primero, identifica unas subcategorías que son bastantes parecidas a las de Hasan (1984).

Las subcategorías de reiteración que Hasan denominaba repetición, sinonimia, antonimia, hiponimia y meronimia se denominan ahora repetición, equivalencia, contraste, especificación (*public transport–underground*) y coespecificación (*buses – underground*) y generalización, respectivamente. La diferencia lleva consigo cambiar una terminología propia de la semántica léxica por una más enfocada a las relaciones que son específicas al discurso.

En cuanto a la colocación, esta comprende principalmente dos categorías:

- a) Pertenencia a un conjunto ordenado de lexemas como los días de la semana o los números: *Monday – Thursday*.
- b) Implicación, en la que la asociación entre elementos no es sistemática sino que se basa en la tendencia que tienen a aparecer en el mismo contexto: *Winter – cold; armies – war*.

Con respecto a la colocación, Tanskanen sostiene que es una categoría difícil de identificar en los análisis de la cohesión léxica porque no implica una relación tan sistemática como la que se da en la reiteración (Halliday y Hasan 1976: 288; Hoey 1991: 83-84). También sostiene que la colocación es una relación más local que la reiteración. Si estas afirmaciones de Tanskanen fueran ciertas, no hubiera sido posible crear programas informáticos que computan colocaciones como *Microconcord* o *Wordsmith Tools* ni tampoco hubiera sido posible el estudio de Phillips (1985) sobre cómo la colocación crea cohesión entre capítulos de libros.

En su análisis no hemos encontrado detalles sobre el corpus estudiado, lo cual nos hace poner en entredicho los resultados de este estudio y destacar la necesidad de replicar investigaciones de este tipo.

En primer lugar, Tanskanen analiza la frecuencia de casos de reiteración y colocación en los tres géneros y detecta que la reiteración es mucho más frecuente que la colocación. Un mayor uso de la reiteración en comparación con la colocación se da en aquellos géneros en los que la comunicación es más difícil como en los discursos preparados y en la conversación cara a cara entre más de dos participantes. En situaciones en las que las demandas comunicativas son más moderadas, como en la prosa académica, las diferencias entre el recurso más frecuente (la reiteración) y el que le sigue en frecuencia (la colocación) se reducen.

El orden de frecuencia de las subcategorías para los tres géneros es bastante similar para los tres géneros. Y así, la repetición es el recurso cohesivo más frecuente, seguido de la implicación, la equivalencia, la generalización, la especificación y la co-especificación con prácticamente la misma frecuencia. El contraste y el conjunto

ordenado aparecen como los recursos menos frecuentes. De aquí se deduce que el orden de frecuencia de las subcategorías no indica una diferencia entre los géneros.

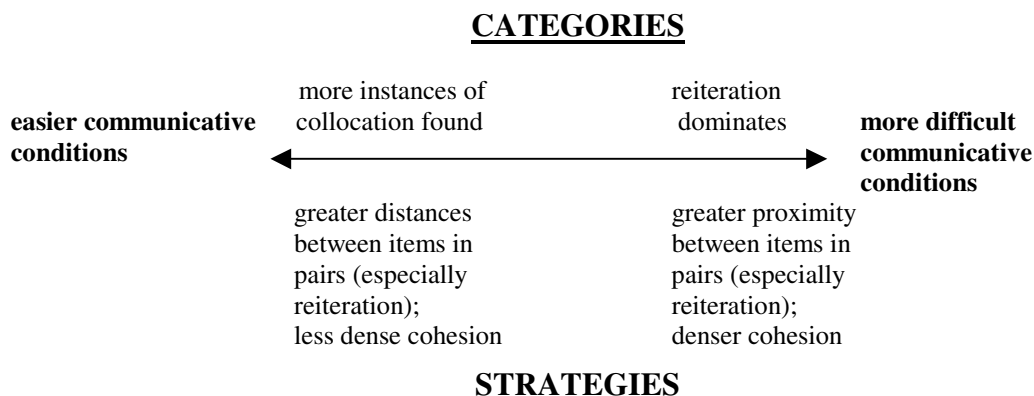
En segundo lugar, Tanskanen observa la variación de la densidad de recursos cohesivos y especula sobre su valor como indicio de situaciones comunicativas difíciles. La densidad de recursos cohesivos depende de las frecuencias normalizadas de cohesión y colocación⁴⁹ y de la distancia, en oraciones, entre los elementos del par.

El discurso preparado es el más denso en cuanto a reiteración, seguido por la prosa académica. La conversación es el género con menor densidad de reiteración. En cuanto a la distancia, los pares entre los que hay una relación reiterativa están separados por más oraciones que los que están relacionados mediante colocación. La distancia entre pares reiterativos es mayor en las conversaciones, sobre todo, cuando hay unos cuantos interlocutores. Es algo menor en la prosa académica y bastante menor en el discurso preparado, en el que los elementos del par cohesivo se encuentran en la misma o en una oración adyacente. También se encuentran relacionados en la misma oración o en una oración adyacente los pares relacionados por colocación, lo cual hace que la densidad de colocación normalmente se mantenga alta.

En esencia, el discurso preparado es el más denso en cuanto a la cohesión léxica, ya que hay más pares cohesivos que en los otros géneros y las distancias entre los elementos de los pares son menores. Tanskanen deduce que, si la densidad de recursos cohesivos es alta, sobre todo cuando son de naturaleza *reiterativa*, nos hallamos ante un género que cobra existencia en condiciones comunicativas difíciles.

⁴⁹ La frecuencia normalizada corresponde al número de *pares* en relación cohesiva por cada mil palabras.

DIAGRAMA 11: Cohesión y dificultad comunicativa (Tanskanen 1995: 537)



Tanskanen presenta sus conclusiones en un diagrama, en el que se indica que la reiteración se utiliza de un modo más frecuente que la colocación, y domina principalmente en situaciones en las que la comunicación es más difícil. La colocación se utiliza en un mayor grado en aquellas situaciones en las que la comunicación es más fácil. Asimismo, las estrategias de utilización de la cohesión léxica también cambian según las condiciones comunicativas, al menos en lo que se refiere a la densidad cohesiva: cuanto más difícil resulta la comunicación, mayor es la proximidad entre los elementos entre los que se establece la cohesión, y por tanto, mayor la densidad cohesiva.

Los recursos empleados en situaciones más difíciles están orientadas al receptor, mientras que los que se utilizan en condiciones en las que la comunicación es más fácil están orientadas al productor del mensaje.

Esta aproximación parece demostrar que la cohesión léxica se ajusta a las demandas comunicativas planteadas en determinadas situaciones y, en cierto modo, apoya la hipótesis de que puede existir una correlación entre cohesión y tipos textuales.

4.11. Aproximaciones a la cohesión desde la traducción

La redacción de textos cohesivos facilita la transmisión de la función comunicativa de los mismos y su comprensión por parte del lector. En consecuencia, este fenómeno lingüístico es crucial para el traductor, que detecta la conexión entre los elementos del texto origen y reformula esta unidad conceptual en el texto meta mediante recursos cohesivos.

Los estudios de traducción, apenas han explorado la construcción de la cohesión en los textos, con la excepción de los trabajos de Blum-Kulka (1986), Lotfipour-Saedi (1997) y Fernández Polo (1999). En todo caso, la investigación de la cohesión no ha puesto el énfasis en la cohesión léxica y no se ha llevado a cabo ni en textos extensos completos ni mediante métodos informáticos.

Aunque no se centren en la cohesión léxica, estos ponen de manifiesto que la traducción normalmente lleva consigo cambios en la cohesión. El estudio más citado es el de Blum-Kulka (1986: 19), donde se propone la llamada hipótesis de la explicitación, que sostiene que los textos traducidos cuentan con una mayor redundancia que lleva aparejada un mayor grado de explicitación cohesiva.

Blum-Kulka (ibid: 23) define la coherencia como una relación semántica implícita entre las partes del texto que se hace explícita gracias a la interpretación que hace el lector. La cohesión es una relación semántica explícita entre partes de un texto, expresada por marcadores lingüísticos.

Asimismo, añade que la traducción puede llevar consigo cambios en la cohesión en dos sentidos (Blum-Kulka 1986: 18). Por una parte, los textos traducidos son más explícitos o menos explícitos que el original. Por otra, el potencial semántico explícito o implícito del texto origen a veces varía en el texto meta.

Algunos de estos cambios se deben a la gramática, a motivos estilísticos, es decir, al hecho de que diferentes lenguas imponen diferentes normas que rigen la utilización de recursos cohesivos y, por último, al proceso traductor en sí, en el que afloran las normas retóricas correspondientes a la lengua origen y término. Lotfipour-Saedi (1997) comenta que la cohesión léxica contribuye a la textura del texto. Lo aplica a un fragmento literario de sólo 276 palabras en el que Lotfipour-Saedi sólo analiza *una* cadena léxica. Este autor parece desconocer los estudios sobre cohesión de Hoey,

Morris y Hirst, como indica el hecho de que conciba la *colocación* en el sentido que le dan Halliday y Hasan. Este sentido dista del sentido más aceptado en lingüística, que es el que le da Firth, es decir, la colocación es la relación que un lexema establece con otros lexemas que aparecen con una probabilidad significativa en su cotexto inmediato.

De todas formas, propone unas etapas que permiten evaluar si una traducción mantiene el valor discursivo pretendido por el TO:

- a) determinar las principales cadenas léxicas centrales al tema del TO
- b) describir las relaciones semánticas entre los nodos de las cadenas
- c) determinar la distancia en palabras entre los nodos
- d) determinar hasta qué punto los miembros de la cadena están subsumidos en la jerarquía textual
- e) evaluar la naturaleza de la equivalencia traductológica comparando el texto origen y meta según estos rasgos textuales

Por último, presentamos un trabajo que está plenamente relacionado con la hipótesis formulada en esta tesis. Fernández Polo (1995, 1999) detecta cómo al traducir textos se producen cambios en la cohesión y coherencia de los mismos a raíz de la adecuación a las normas retóricas asociadas a una determinada lengua y a un determinado tipo textual. El interés de esta propuesta merece una sección aparte.

4.11.1. Cambios cohesivos en la traducción de textos de divulgación científica

Fernández Polo (1995) parte del presupuesto de Myers (1991) de que en textos especializados se utilizan diferentes elementos cohesivos en comparación con sus versiones de divulgación y sostiene que estas mismas diferencias se encuentran al confrontar textos divulgativos en inglés con sus traducciones al español. Los textos profesionales muestran una preferencia especial por la repetición, mientras que en la divulgación esta es menos habitual.

En los textos divulgativos, la conexión entre los conceptos es mucho más explícita, dado el limitado conocimiento de los lectores, y esto se ve reflejado en una serie de elementos cohesivos: la referencia, la cohesión léxica, la conjunción y el tema,

principalmente. Estos elementos son necesarios para mantener la efectividad comunicativa, y por tanto, hay que prestarle bastante atención a la hora de traducir.

En este trabajo se indagan las causas de la variación con respecto a la cohesión que se produce como resultado de la traducción de textos divulgativos del inglés al español. Para ello, se analiza un reducidísimo corpus formado por dos artículos de la revista *Scientific American* y la versión traducida al castellano que publica *Investigación y Ciencia*. Fernández Polo (1999) amplía el número de artículos de estas publicaciones hasta conseguir un corpus de aproximadamente 65 000 palabras formado por cinco originales en inglés y sus correspondientes traducciones, y cinco artículos originales en castellano.

Las conclusiones de ambos trabajos apuntan hacia distinciones cohesivas que obedecen, entre otros motivos, a la estructuración retórica propia de las lenguas inglesa y española. Asimismo, explican el papel de la repetición en los textos originales ingleses con el fin de detectar posibles pérdidas de coherencia en las versiones traducidas. En esencia, la repetición léxica presenta un determinado referente mediante diversas expresiones que pueden ser sinónimas o estar relacionadas mediante hiponimia. Con estas se establece una red conceptual en torno a la cual se articula el tema. Es decir, se presenta un determinado referente desde distintos ángulos con la esperanza de que al menos una de las formulaciones permita al lector acceder a la idea que se le desea transmitir.

La referencia a entidades previamente mencionadas en los artículos del *Scientific American* se actualiza mediante la repetición léxica, incluso en casos en los que la distancia entre la segunda mención y el antecedente es mínima, como se ve en el ejemplo que aporta Fernández Polo (1995: 258).

- (32) Since then, the sophisticated satellite has been making the first comprehensive astronomical detections of **gamma rays, the most energetic form of electromagnetic radiation**. When viewed **at gamma ray energies**, the universe appears unfamiliarly dynamic and capricious.

No ocurre lo mismo en la versión española, en la que se evita la repetición léxica y en su lugar aparecen formas de referencia menos explícitas cognitivamente: elipsis, pronombres, sinónimos y términos superordinados.

(33) Desde entonces, este satélite de 16 toneladas de peso ha venido realizando rastreos exhaustivos de **los rayos gamma celestes, la radiación del extremo más energético del espectro**. A **esas frecuencias** aparece un universo voluble y dinámico.

En este sentido, la versión española se aproxima más a las convenciones de los textos divulgativos ingleses, que utilizan pronombres y sinónimos en lugar de repetición léxica. En la traducción al español (33), los sinónimos para "rayos gamma celestes", "el espectro" y "esas frecuencias", no dejan muy claro cuál es su referente, ya que, según Polo, este no está explícito en el cotexto anterior.

Para explicar este fenómeno, Fernández Polo remite a la norma retórica que considera la repetición léxica como un signo claro de un estilo pobre (Montaño-Harmon 1991: 420-421; Aznar Más *et al.* 1989: 30). El cumplimiento de esta convención hace que el lector español tenga más problemas a la hora de asignar el referente a un pronombre, establecer la conexión entre dos sinónimos o entre un hiperónimo y su hipónimo, y ver si dos formas similares hacen referencia a la misma entidad. Estas operaciones mentales requieren un mayor conocimiento del tema y un mayor esfuerzo por parte de la audiencia española.

Fernández Polo también presta atención a dos recursos cohesivos que Hoey (1991) no incluye en su análisis: la progresión temática y la conjunción. Mientras que en el texto original la posición temática incluye el rema de la oración precedente, en las traducciones aparecen segmentos de información conocida en posición de rema, con lo cual, se destruye la cohesión. Según este autor, estas diferencias responden a que cada género lleva asociados unos patrones temáticos y estos están determinados en gran medida por la audiencia. Hay géneros cuyos patrones temáticos contribuyen más a la comprensión por parte del lector que otros. Por este motivo, los traductores han de ser conscientes del papel de la progresión temática en el establecimiento de la cohesión.

En cuanto al uso de conjunciones y conectores, estos son más explícitos en los artículos del *Scientific American* que en *Investigación y Ciencia*, en cuyos artículos se suprimen algunos de los conectores presentes en el texto original. Esto parece contradecir la hipótesis de la explicitación de Blum-Kulka (1986: 19), es decir, el hecho de que la traducción supone un mayor grado de explicitación cohesiva. Una respuesta a este fenómeno se puede encontrar en la retórica de contrastes (Connor y Kaplan 1987; Hinds 1987; Mauranen 1993; Tyler 1992), que sostiene que hay diferencias en el grado de explicitación entre lenguas y culturas. El inglés es muy explícito o, en palabras de Hinds (1987) y Hatim (1997) es *writer-responsible*. El español es mucho menos explícito y se sitúa entre el inglés y las denominadas *reader-responsible languages* como el japonés y el coreano. Algunas investigaciones como las de Saint-John, Montaña-Harmon y Reppen y Grabbe (*apud* Fernández Polo 1999: 179) sugieren que los hispanohablantes consideran los textos que siguen la retórica anglosajona como excesivamente simples, aburridos e incluso infantiles; y a la inversa, para un anglosajón, la retórica del español le resulta confusa y difícil de comprender. En consecuencia, la supresión de conectores no es más que una adaptación del texto original a la estructura retórica de la lengua española.

A pesar de lo importante que resulta en traducción respetar las convenciones retóricas de las lenguas de partida y de llegada, una aplicación irreflexiva de las mismas puede conllevar modificaciones en la cohesión textual. Estas modificaciones requieren un conocimiento experto mayor por parte del lector español, que puede pensar que distintas formas léxicas corresponden a distintas realidades cuando, en realidad, corresponden a la misma entidad.

En consecuencia, Fernández Polo concluye que es necesario realizar estudios contrastivos sobre cómo se amolda la cohesión a distintos géneros en las lenguas inglesa y española. También es necesario encontrar el equilibrio entre las convenciones retóricas de una lengua y de un género determinado y los recursos cohesivos asociados a esa lengua y a ese género.

5. La lingüística de corpus al servicio de la traducción y la terminología

5.1. Introducción

La metodología denominada *lingüística de corpus* ha tenido y sigue teniendo una importante repercusión en el estudio del lenguaje⁵⁰ y en el desarrollo de programas informáticos para el procesamiento del lenguaje natural⁵¹ (Butler 1992; McEnery y Wilson 1996: 87-145). Podemos encontrar una explicación de esto en el hecho de que, desde sus comienzos, esta metodología ha tenido la pretensión empiricista de estudiar el funcionamiento real de la lengua mediante el estudio de ingentes cantidades de texto producidos por los hablantes de una lengua. Asimismo, la importancia económica que reportan los corpóra informatizados a los grandes editores de diccionarios y a las industrias de la lengua⁵² pueden explicar este vigor. Otros factores que también han propiciado el desarrollo de corpóra informatizados han sido la creciente capacidad de los nuevos ordenadores y una mayor accesibilidad a textos en soporte electrónico. Esta se ha logrado gracias a la comunicación por Internet y a instituciones que se dedican a recopilar corpóra reutilizables como ACL (Association for Computational Linguistics), ECI (European Corpus Initiative), LDC (Linguistic Data Consortium), ICAME (International Computer Archive of Modern and Medieval English), ACL/DCI (Association for Computational Linguistics Data Collection Initiative) o ELRA (European Language Resources Association).

De todas formas, antes de proseguir, es preciso aclarar qué se entiende por *corpus*, ya que este término se suele confundir con otros relacionados como *archivo*, *colección* o *textoteca*. Por otra parte, en los estudios de traducción, el término *corpus* designa un conjunto muy limitado de textos en soporte papel del que difícilmente se

⁵⁰ Algunas de las disciplinas que se pueden beneficiar de esta metodología son la semántica, lexicología y lexicografía, fonética, gramática, pragmática y análisis del discurso, sociolingüística, estilística, lingüística diacrónica, didáctica de las lenguas, psicolingüística, estudios culturales y traducción.

⁵¹ En este sentido, se han desarrollado infinidad de aplicaciones para el procesamiento del lenguaje y del habla que han sido aprovechadas en lexicografía monolingüe y bilingüe, en el análisis y etiquetaje sintáctico automático de textos, en la traducción automática y en sistemas de gran utilidad para el traductor como los bancos terminológicos y las memorias de traducción.

⁵² El País, jueves 2 de abril de 1998.

pueden extraer conclusiones generalizables. A continuación presentamos una definición que puede aclarar el sentido del término en el ámbito de la lingüística de corpus (McEnery y Wilson 1996: 24).

So a corpus in modern linguistics, in contrast to being simply any body of text, might more accurately be described as a finite-sized body of machine-readable text, sampled in order to be maximally representative of the language variety under consideration. (McEnery y Wilson 1996: 24)

A la idea de tamaño limitado, formato electrónico, muestreo y representatividad, McEnery y Wilson añaden que todo corpus debe constituir una referencia estándar a la lengua que pretende representar. Para que el corpus constituya una referencia reutilizable (Hockey y Walker 1993: 235), ciertas organizaciones como TEI (Text Encoding Initiative), en el pasado, y EAGLES (Expert Advisory Group on Language Engineering Standards), en la actualidad, se esfuerzan por elaborar normas de estandarización. Otros autores como Leech y Fligelstone (1992: 116) también inciden en que los corpóra suelen ser representativos de una determinada lengua o tipo textual y añaden que son recopilados con un claro objetivo en mente.

Teniendo en cuenta que una revisión de los avances alcanzados por la lingüística de corpus constituiría de por sí una tesis, en este apartado se destacarán los aspectos de la lingüística de corpus relacionados con la traducción y las aplicaciones de esta metodología a la traducción. En Biber *et al.* (1998), Kennedy (1998), Pérez Guerra (1998) y en los archivos de la lista de distribución *Corpora* [<http://www.hd.uib.no>] se pueden consultar algunas de las referencias bibliográficas, publicaciones periódicas, listas de distribución, organizaciones y direcciones de Internet más relevantes.

5.2. Evolución de la lingüística de corpus

La utilización de corpóra en el estudio del lenguaje entronca, por una parte, con la tradición anglosajona iniciada por Firth (1935, 1957a, 1957b) y continuada principalmente por sus discípulos, Halliday y Sinclair, y por otra parte, con la labor de estructuralistas norteamericanos como C.C. Fries, H. Kuxera y W.N. Francis (*apud* Stubbs 1996: 22-50). La primera influencia preconiza que el significado de una palabra

depende de su coocurrencia con otras palabras⁵³ y vislumbra lo que será una concepción contextual del significado léxico (Cruse 1986; Pustejovski 1995; Wanner 1996). La segunda se enmarca en el auge que los métodos empíricos y estadísticos experimentaron en la década de los cincuenta.

Estos planteamientos cristalizaron en la creación del corpus *Brown* de inglés americano (Francis y Kuxera 1982), al que siguieron otros como el corpus *London-Lund* o el corpus *LOB (Lund-Oslo-Bergen)*. En la década de los setenta, la investigación en traducción automática y lingüística computacional fue impulsada en Canadá y Europa Occidental, en las que se aspiraba a una política multicultural. La demanda de traducciones de documentación legal, administrativa, técnica y científica fue la que hizo que se construyeran los primeros corpórea bilingües, que contenían textos en una lengua junto con su traducción a otra. Un buen ejemplo lo constituye el *Canadian Hansard*, en el que se transcriben en inglés y en francés las reuniones del parlamento canadiense de mediados de los setenta (Brown *et al.* 1993). El millón de palabras que lo conforman fue utilizado en traducción automática.

Mientras se elaboraban los primeros corpórea bilingües y multilingües en soporte electrónico, se produjo un importante esfuerzo de compilación de diccionarios en inglés, que ya se había convertido en *lingua franca*. Este interés sería el motor del primer proyecto lexicográfico basado en corpus: la elaboración del diccionario *Collins COBUILD* (1987), auspiciado por la editorial HarperCollins y la Universidad de Birmingham. Este diccionario se fundamentaba en un corpus de 20 millones de palabras llamado *Birmingham corpus*, que se convirtió en un instrumento capaz de aportar para cada palabra no sólo información semántica sino también información sobre su combinatoria, su sintaxis más típica y sus patrones estructurales y colocacionales. Al mismo tiempo, este corpus, que ha seguido creciendo hasta alcanzar 329 millones de palabras en julio de 1998 y que ahora se denomina *Bank of English*⁵⁴, sentó la base para

⁵³ Esto queda magistralmente expresado en la famosa cita de Firth (1957): “You shall know a word by the company it keeps”. Esta cita también vislumbra lo que se denominará *ecología del lenguaje* (Walker 1991), es decir, la tendencia a relacionar los textos con un contexto que sirva de marco de referencia.

⁵⁴ El servicio *CollinsDirect* [http://titania.cobuild.collins.co.uk/direct_info.html] hace accesibles concordancias y colocaciones extraídas a partir de 50 millones de palabras de este corpus, que contiene textos en inglés británico principalmente, aunque también norteamericano (9 millones), tanto orales como escritos.

ingentes córpora monolingües posteriores como el *British National Corpus (BNC)*, el *International Corpus of English (ICE)* o el *Survey of English Usage (SEU)*.

También en la década de los ochenta, proliferaron córpora especializados destinados a sistemas de traducción automática de sublenguajes (McKeown 1992) y a la didáctica del inglés. En esta última categoría, entraría un corpus científico-técnico desarrollado en China por Yang (1985) o el *Guangzhou Petroleum English Corpus* destinado a la industria petrolífera (Zhu 1985).

La década de los noventa va a suponer un avance significativo en cuestiones como el diseño del corpus, que intenta reflejar las necesidades del usuario, la recopilación, la anotación y tratamiento informático de los textos, la accesibilidad, la estandarización y la evaluación⁵⁵ de resultados.

Al igual que ocurriera en los cincuenta, se ha producido un renacer del interés por métodos empíricos y estadísticos, en cierto modo motivado por la existencia de ordenadores mucho más potentes. Junto con esto, las relaciones políticas, comerciales y científicas de ámbito internacional, que demandan más que nunca la edición de documentos en más de una lengua, y la necesidad de aprender idiomas suponen un empuje para los diccionarios monolingües y, por ende, para los córpora bilingües. Y así, la década de los noventa asiste a un aumento del interés por córpora bilingües y multilingües por parte de investigadores en traducción automática (Brown *et al.* 1993)⁵⁶, lexicografía bilingüe (Klavans y Tzoukermann 1990), didáctica de lenguas extranjeras (Barlow 1994, 1996) y en terminología (Budin y Wright 1997; Cabré, 1997). Estos córpora, llamados también *córpora paralelos* (véase 5.3.), pueden proporcionar información documentada sobre cómo textos escritos en una lengua se pueden trasladar a otra de acuerdo con una serie de factores contextuales como el estilo, registro,

⁵⁵ Este tema fue el hilo conductor del congreso internacional que *ELRA (European Language Resources Association)* celebró en Granada en 1998 con el título: *First International Conference on Language Resources and Evaluation*.

⁵⁶ A finales de los ochenta y principios de los noventa, la traducción automática se ha relanzado de nuevo, una vez asumidas sus limitaciones y la necesidad de restringir sus objetivos (Teubert 1996: 242). El resultado ha sido la creación de sistemas tanto de traducción automática como de traducción humana asistida por ordenador de gran eficacia. Entre los principales sistemas de traducción automática, podemos citar Systran, Logos, Eurotra y Globalink Translation Systems (GTS). En cuanto a los programas de traducción asistida, los más conocidos son IBM Translation Manager 2, Star Transit, Trados Translator's Workbench y Déjà Vu.

dominio, etc. La elaboración de estos corpórea se ve beneficiada también por la experiencia adquirida en la creación de corpórea monolingües y por la posibilidad de edición en CD-Rom.

Barlow (1996: 49) menciona una serie de proyectos relacionados con la elaboración de corpórea bilingües y multilingües. Entre los primeros, se encuentran *Intersect* (francés-inglés), *Contragram*, el *Canadian Hansard Corpus* y un corpus paralelo inglés-noruego recopilado en la Universidad de Oslo. En cuanto a los corpórea multilingües, Barlow hace mención al corpus *Aarhus* de derecho contractual (danés-inglés-francés), al corpus *Triptic* (inglés-francés-neerlandés) y a una serie de proyectos europeos como *Lingua*, *Multext* y *Multex-East*⁵⁷.

Hasta aquí, la mayoría de los corpórea mencionados han sido concebidos para representar la lengua general⁵⁸. Sin embargo, la utilización del inglés como *lingua franca* en los dominios de especialidad hace que estos corpórea no suplan las necesidades de sus usuarios. Surge entonces una acuciante necesidad de crear corpórea con fines específicos o corpórea de sublenguajes que, sin ser tan extensos como los de la lengua general, puedan facilitar la comunicación en inglés entre expertos y la traducción especializada de o hacia el inglés. En este sentido, Gledhill (1996: 109) afirma que la mayor parte de la investigación en lingüística de corpus ha estado orientada hacia la elaboración de corpórea generales representativos y sólo recientemente se ha prestado atención a variedades especializadas del inglés.

En consecuencia, en la actualidad se están desarrollando numerosos corpórea monolingües y multilingües especializados y también proyectos dentro del sector de la ingeniería lingüística⁵⁹. En estos corpórea y proyectos están implicados investigadores de

⁵⁷ Para más información sobre estos corpus se puede visitar la página sobre textos paralelos elaborada por este investigador [<http://www.ruf.rice.edu/~barlow/para.html>].

⁵⁸ A este tipo de corpórea el Observatorio Español de Industrias de la Lengua [<http://cervexp.cervantes.es/oeil/oeilin2.html>] lo denomina *corpus de la lengua general con fines generales*.

⁵⁹ Dentro del Programa de Aplicaciones Telemáticas de la Comisión Europea hay abundantes proyectos en marcha como MABLE (*Multilingual Authoring of Business Letters*), AVENTINUS (*Advanced Information System for Multinational Drug Enforcement*), MultiMeteo (*Multilingual production of weather forecasts*) o ARISE (*Automatic Railway Information Systems for Europe*). Para obtener más información sobre los mismos se puede consultar el folleto *Language Engineering: Progress and Prospects*, editado por el equipo LINGLINK de la empresa Anite Systems [<http://www.anite-systems.lu>].

áreas como la lingüística computacional, la traducción automática, el inglés para fines específicos y la terminografía. En la siguiente tabla presentamos algunos ejemplos de corpórea especializados:

TABLA 12: Ejemplos de corpórea monolingües y multilingües especializados⁶⁰

	NOMBRE	LENGUA (S)	DOMINIO DE ESPECIALIDAD
MONOLINGÜES	Cranfield collection	Inglés	Aerodinámica
	RAT (University of Reading Academic Text) corpus	Inglés	Artículos experimentales y tesis doctorales de diferentes dominios
	IBM corpus	Inglés	Informática (sublenguaje de los manuales de IBM)
	Medlars collection	Inglés	Medicina
	PSC (Pharmaceutical Sciences Corpus)	Inglés	Medicina y Farmacología aplicado a la oncología
	Medicor	Inglés americano	Medicina
	Corpus Textual del Español Periodístico	Español	Textos periodísticos
	LAN (Micro Focus SA)	Español	Técnica
MULTILINGÜES	LEJES	Español	Derecho (textos académicos)
	Danish-German-Spanish Corpus of Biotechnology	Danés-alemán-español	Biotecnología
	TEST corpus	Inglés-italiano	Debates parlamentarios
	CRATER (Corpus Resources and Terminology Extraction)	Francés-Inglés-Español	Telecomunicaciones
Corpus textual plurilingüe especializado	Catalán-español-inglés-francés-alemán	Economía, derecho, medio ambiente, medicina e informática	

Sin menoscabar la calidad y utilidad de los multilingües, estimamos que estos corpórea no dan cuenta de la variedad textual que se produce en estos dominios de especialidad. Siendo esto así, queda mermada la enorme potencialidad que brindan los corpórea en el estudio de los distintos géneros (Carne 1996: 135). Asimismo, el hecho de que contienen una importante proporción de textos traducidos hace que las colocaciones

⁶⁰ Para un estudio exhaustivo remito al ya citado Observatorio Español de Industrias de la lengua, a los mensajes de la lista de distribución *CORPORA*, de la que se puede obtener información en <http://hd.uib.no>, y a la bibliografía presentada en 5.1.

y convenciones retóricas asociadas a los tipos textuales de una determinada lengua aparezcan distorsionados por la influencia del texto origen. Por este motivo, creemos conveniente que estos *córpore* multilingües tengan dos componentes: un corpus de *textos paralelos* y otro de *textos comparables*.

En este proceso se han acuñado principalmente dos términos para designar los *córpore* multilingües: *corpus paralelo* y *corpus comparable*.

5.3. *Córpore* paralelos y *córpore* comparables

En la literatura de la lingüística de corpus, a la hora de referirse a los *córpore* multilingües aparecen los términos *corpus paralelo* y *corpus comparable*. Aunque no existe un consenso en el uso de estos términos, en la actualidad, su significado más extendido es el que presentamos a continuación.

Un *corpus paralelo* es aquel que presenta el mismo texto en más de una lengua, es decir, un texto y su traducción a una o más lenguas (Mc Enery 1994: 312; 1996: 58). Según Barlow, estos textos normalmente aparecen alineados para facilitar la búsqueda de equivalencias de traducción. Por ejemplo, el *Canadian Hansard corpus* y *Crater*, un corpus paralelo con textos en francés, inglés y español del dominio de especialidad de las telecomunicaciones.

En cuanto a los *córpore* comparables (Peters *et al.* 1996: 69), son un conjunto de textos en más de una lengua que, sin ser traducciones, por coincidir en el tema, motivación situacional y función comunicativa, proporcionan una excelente base para la comparación de dos o más lenguas. La mayoría de estos *córpore* representan un determinado sublenguaje. En cierto modo, estos son un conjunto de pequeñas colecciones de *córpore* monolingües que han sido recopilados siguiendo criterios muy parecidos. Como ejemplo podemos mencionar el corpus *Aarhus*, que contiene textos originales del subdominio especializado del derecho contractual en danés, inglés y francés. Algunos de los investigadores pioneros en la utilización de estos tipos de corpus como Hartmann (1996)⁶¹, Laffling (1992) y Zanettin⁶² (1994, 1996) los han

⁶¹ Para evitar la ambigüedad, Hartmann denomina *bitextos* a los textos en el que uno es el texto origen y el otro, el texto término.

denominado *córpora* paralelos, siguiendo la terminología más usual en los estudios de traducción. En esta última disciplina, los textos paralelos son textos lingüísticamente independientes aunque funcionalmente equivalentes resultantes de una situación comunicativa idéntica o muy similar.

Los *córpora* paralelos fueron los primeros en aparecer. De hecho, hasta hace poco, la mayoría de los *córpora* multilingües contenían textos originales y su traducción, sobre todo en Europa, donde todas las lenguas oficiales de la Unión Europea gozan de un estatus igual. Hartmann (1980) fue uno de los primeros en hablar de textos paralelos, un término con el que incluía tanto los textos resultantes de una traducción como los que, sin ser traducciones, eran funcionalmente similares. En 1996, Hartmann distingue entre cuatro tipos de textos paralelos y los presenta en la tabla que presentamos abajo.

TABLA 13: Tipos de textos paralelos

	Approach	Parallel text type	Text corpus	
Metaphrase	Translation Text comparison	Bitext	'aligned', e.g. Canadian Hansard	'sampled' multilingual, e.g. MULTEXT
	Crosscultural Text adaptation	Twinned text	'normalised', e.g. ISSCO	
	Contrastive Text typology	Paired text	'domain-specific' e.g. corpus by Göpferich (1995)	
Paraphrase	Intralingual Text typology	Intertext	'register-diversified', e.g. BNC 'region-specific', e.g. ICE 'period-specific', e.g. Helsinki	

Teubert (1996: 245) también distingue entre distintos tipos de *córpora* paralelos. En primer lugar, aquellos que incluyen textos escritos en una lengua A y su traducción a una lengua B (y C...). Otro tipo lo forman textos que tienen una igual proporción de textos originales en lengua A y B y sus respectivas traducciones. Por último, Teubert

⁶² En concreto, Zanettin utiliza el término *textos globalmente paralelos* y los define así: "collections of texts which are not translations, but rather texts which share a wide range of features, to serve as the basis for comparison and then for contrastive analysis between languages". (Zanettin, 1996:101).

habla de *córpore* paralelos compuestos exclusivamente de las traducciones a las lenguas A, B y C, de textos escritos inicialmente en una lengua Z.

Por otra parte, los *córpore* paralelos han sido criticados sobre la base de que no presentan el uso real del lenguaje natural sino textos influenciados por el texto origen, lo cual hace que no quede representada la lengua término en su forma más genuina (Hartmann 1994). Estos *córpore* son útiles en tanto que presentan ejemplos de equivalencias de traducción y de cómo el traductor resuelve problemas traductológicos, lo cual puede aportar datos relevantes sobre el proceso de la traducción.

Por eso, se empezaron a elaborar *córpore* bilingües/multilingües comparables. Sobre los criterios que garantizan la comparabilidad de los textos, Corazzari y Picchi (1994) consideran que es necesario que exista un vocabulario común. En consecuencia, proponen una combinación de criterios externos como dominio, tema y periodo temporal, y criterios internos, como la selección de textos que contengan una serie de palabras clave extraídas del vocabulario especializado de un sublenguaje. De ahí que, normalmente, los textos comparables contengan textos pertenecientes a lenguajes restringidos.

En cuanto a la conveniencia de utilizar uno u otro tipo de corpus en la traducción de textos, estimamos que ambos son complementarios y necesarios para llegar a comprender el proceso y el producto de la traducción⁶³; y en esto coincidimos con Tercedor (1999) y Teubert (1996). En consecuencia, en nuestra investigación los datos emanarán de dos repositorios: por una parte, un corpus comparable de textos sobre cáncer de pulmón compuesto por una proporción similar de tipos textuales para el inglés y el español; y por la otra, un corpus paralelo formado por textos en inglés sobre el mismo tema y su traducción al español.

5.4. Aplicaciones de la lingüística de corpus a la traducción

Siguiendo la tendencia a la interdisciplinariedad de la que se alimenta la traducción, numerosos investigadores en el área de los estudios de traducción han defendido la

⁶³ No obstante en la traducción de términos culturales, resulta más apropiada la utilización de corpus comparables, ya que “es este tipo de corpus el que puede ayudar al traductor a reflejar en la lengua término no sólo el significado proposicional del texto sino también su valor semiótico dentro del contexto más amplio de la ideología y de la cosmovisión de una sociedad” (López Rodríguez 1998).

utilización de corpórea informatizados (Baker *et al.* 1993, 2000; Barlow 1996; Peters y Picchi 1997; Hartmann 1994; López Rodríguez 1998; Muñoz Martín y Sánchez Trigo 1994; Teubert 1996; Zanettin 1996).

Baker *et al.* (1993) busca un punto de encuentro entre los partidarios de estudios descriptivos dentro de la traducción y las propuestas de la lingüística de corpus de extraer inferencias sobre el funcionamiento de la lengua basadas en datos reales. De hecho, Baker sostiene que las técnicas estadísticas y la metodología puestas en funcionamiento por Sinclair (1991) pueden contribuir al paso de estudios prescriptivos a descriptivos. De todas formas, Baker no quiere limitar esta metodología al estudio de la lengua origen y la lengua término, sino que pretende que se profundice en el conocimiento del proceso de la traducción mediante la búsqueda de tendencias en la traducción ya sean de carácter universal o establecidas por una norma (Toury 1980), el análisis de los distintos borradores de una traducción y la especulación sobre cuál es la unidad de traducción. Baker (2000) compara diversos rasgos lingüísticos tal y como se presentan en algunos tipos textuales del *BNC* y los compara con un corpus de traducciones hacia el inglés que contiene tipos textuales parecidos.

Michael Barlow, uno de los principales especialistas en corpórea paralelos⁶⁴, también señala la utilidad de los corpórea informatizados en la traducción:

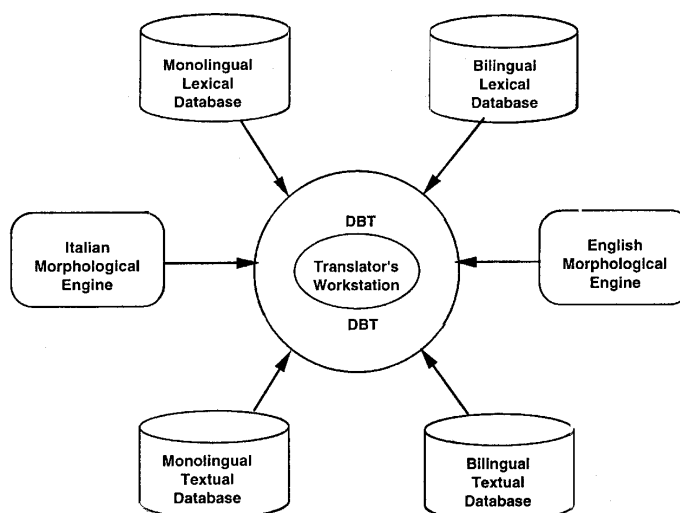
By looking at multiple instances of the translation of a word or structure, the idiosyncratic influences will be overpowered by the consistent and meaningful regularities. The accumulation of motivated translation choices revealed by a concordance program allow the general patterns to be perceived. Hence generalisations emerge from the aggregation of large numbers of individual instances". (Barlow 1996: 50)

También menciona algunas de las áreas de dificultad dentro de la traducción que pueden esclarecerse con la utilización de corpórea bilingües alineados: la fraseología, la polisemia, la organización del discurso y las estructuras retóricas (ibid: 52-53).

⁶⁴ De hecho, Barlow ha creado una página web llamada *Parallel Corpora* [<http://www.ruf.rice.edu/~barlow/para.html>] y un programa informático que crea concordancias multilingües y facilita la búsqueda de equivalencias de traducción.

Peters y Picchi (1997: 247) también reconocen lo provechosos que resultan avances producidos en lingüística de corpus y en lexicografía computacional como los léxicos electrónicos y córpora de referencia tanto para los profesionales de la traducción como para los académicos en el campo de los estudios de traducción. Los primeros necesitan encontrar equivalentes de traducción con mucha rapidez. Los segundos se plantean posibles realizaciones de un concepto en distintas lenguas al mismo tiempo que consideran una serie de factores como el uso, estilo, registro o dominio. Sus esfuerzos han cristalizado en una *estación de trabajo* de traducción humana asistida por ordenador para las lenguas italiana e inglesa desarrollada en el Instituto de Lingüística Computacional en Pisa. Esta consta de dos bases de datos (BD) léxicas y dos BD textuales, tanto monolingües como bilingües y dos repositorios con información morfológica sobre el italiano y el inglés. Un motor de búsqueda permite encontrar equivalentes, colocaciones, frases hechas e información morfológica, sintáctica y semántica a partir de los diccionarios y el corpus paralelo y comparable que sustentan el sistema.

DIAGRAMA 12: Prototipo de estación de trabajo para la traducción humana asistida por ordenador. (Peters y Picchi 1997: 270)



La traducción automática y la traducción asistida por ordenador también han encontrado un apoyo en la utilización de córpora. Laffling (1991, 1992), que Peters y

Picchi consideran uno de los primeros en defender la utilización de *córpora* paralelos, ha empleado un corpus de textos comparables en la construcción del lexicón de transferencia de un sistema de traducción automática, aunque él los denomina textos paralelos. Según Laffling, estos textos facilitan la traducción porque proporcionan al ordenador una base de conocimiento parecido al repertorio lingüístico del que dispone el traductor. También proporcionan información sobre las colocaciones más típicas en determinados tipos textuales y permiten la identificación de contextos equivalentes desde el punto de vista léxico que van a facilitar la traducción.

Por último, no sólo la teoría y la práctica de la traducción se han abonado con las conclusiones derivadas del estudio de *córpora* comparables. Algunos profesores de traducción (Muñoz Martín y Sánchez Trigo 1994; Zanettin 1994, 1996) han utilizado *córpora* comparables elaborados a partir de textos periodísticos en dos idiomas en la formación de futuros traductores. Según Zanettin (1996: 100), los textos periodísticos “proporcionan instancias de uso de la lengua actual en una gran variedad de tipos textuales y ofrecen información valiosa sobre la cultura y la sociedad”. Este autor ha ensamblado los textos en una base de datos textual de consulta para los alumnos de la Escuela de Traductores e Intérpretes de Bolonia.

5.4.1. La utilización de *córpora* en la traducción biomédica

La ventaja de utilizar una metodología de corpus en la traducción biomédica reside en que estos contienen textos que proporcionan información que los diccionarios bilingües no contienen. Desde que Tomaszczyk (1989: 180) observara que los diccionarios especializados no son adecuados para la traducción científica y técnica, son muchas las voces que han destacado las limitaciones de estos diccionarios, cuanto más si son bilingües (Faber 1998b, Tercedor 1999, Williams 1996). En efecto, estos suelen presentar equivalentes de traducción y omiten información fundamental en el nivel oracional y supraoracional como los términos y unidades fraseológicas más apropiados en determinados registros y géneros, información gramatical, estilística y retórica.

Por este motivo, parece obvia la necesidad de otras fuentes de información. Aparte de la consulta a especialistas, Williams (1996: 296) sostiene que los textos paralelos—en nuestra terminología, textos comparables—son los que aportan una

información más fidedigna al traductor. De ahí que sostenga que un método combinado de diccionarios especializados y, al menos, tres textos paralelos pueden satisfacer el 85-90% de las necesidades del traductor. Siendo esto así, si a la hora de traducir el traductor dispusiera, no de tres textos, sino de un conjunto de textos tanto paralelos como comparables en soporte electrónico, su labor se agilizaría significativamente.

No obstante, han sido escasos los corpórea biomédicos compilados hasta el presente y se han realizado únicamente en una lengua. Entre estos podemos mencionar una serie de corpórea biomédicos en inglés:

- a) El corpus Medlars, que consta de 1 033 *abstracts*, por lo cual sólo se puede considerar representativo de este tipo textual
- b) El *PSC (Pharmaceutical Sciences Corpus)*, que cuenta con 150 artículos experimentales sobre cáncer publicados en veinte revistas médicas y farmacéuticas entre 1990 y 1993.
- c) El Corpus *Medicor* de textos médicos en inglés norteamericano actual, recopilado en la Universidad de Helsinki⁶⁵. Incluye los siguientes tipos textuales: artículos experimentales, muestras de libros de texto, artículos de divulgación, editoriales de revistas especializadas, manuales para profesionales, folletos informativos / guías (*popular guidebook*).

También podemos mencionar algunas fuentes que pueden proporcionar un acopio de textos biomédicos en soporte electrónico. Una de ellas es el *BNC*, que, por su gran extensión, contiene una importante representación de textos biomédicos, etiquetados y listos para cualquier análisis. Un vistazo a la base de datos bibliográfica de este corpus puede facilitar la selección de textos de este tipo⁶⁶. Así, el *BNC* incluye

⁶⁵ Su compiladora, Minna Vihla, ha elaborado este corpus con el objetivo de estudiar la modalidad en textos médicos. Por este motivo, ha primado la representatividad en cuanto a tipos textuales incluidos sobre la cantidad, 397 311 palabras.

⁶⁶ Esta idea fue sugerida en la lista de distribución *Corpora* por David Lee (Departamento de Lingüística de la Universidad de Lancaster) el 23 de octubre de 1998 como respuesta a una pregunta sobre la existencia de corpórea de textos científicos.

bajo las categorías *ciencia aplicada* y *ciencia pura*, textos pertenecientes a las siguientes publicaciones periódicas:

- *Journal of Gastroenterology* (713 164 palabras)
- *The Lancet* (135 850)
- *British Medical Journal* (449 961)

Asimismo, muchas revistas médicas ponen a disposición de sus suscriptores una versión electrónica⁶⁷ y hay numerosos sitios sobre medicina en Internet, lo cual facilitará cada vez más la elaboración de corpóra biomédicos comparables y paralelos.

Siempre que el análisis sea el adecuado, estos corpóra informatizados pueden corroborar o corregir los resultados obtenidos en estudios previos basados en reducidos corpóra textuales médicos y completar la literatura sobre los patrones léxicos, colocacionales y retóricos presentes en textos médicos en inglés y en otras lenguas. De estos estudios previos, nos gustaría destacar el trabajo en terminología realizado por Thomas y Hawes (1994), que analiza los complementos del verbo, y las investigaciones en el análisis del discurso de artículos experimentales médicos realizados por Adams-Smith (1984), Busch-Lauer (1995), Kretzenbacher (1990), Myers (1990, 1992), Nwogu (1991, 1997) y Salager-Meyer (1989, 1994). Ya revisamos algunos de estos estudios en 3.3.

En este área también podemos situar la excelente labor de Gledhill (1996, 1997), compilador del anteriormente citado *PSC (Pharmaceutical Sciences Corpus)*. Tanto el diseño del corpus como su extensión, 500 000 palabras, hacen que se pueda considerar representativo de este subdominio. Su objetivo es el de detectar elementos que son especialmente significativos en las secciones retóricas de artículos experimentales⁶⁸ y representar las variaciones de la fraseología propia en cuanto a sus elementos gramaticales a lo largo de estas secciones.

Resulta difícil citar investigaciones similares que se enmarquen en los estudios de traducción, a no ser las de Gledhill (1996, 1997), Tercedor (1999), García de

⁶⁷ Existe un exhaustivo listado de publicaciones médicas en soporte electrónico disponible en www.gen.emory.edu/MEDWEB/medweb.html.

⁶⁸ Es decir, título, resumen, introducción, métodos, resultados y discusión.

Quesada (tesis doctoral en curso), López Rodríguez (2000) y Pérez Hernández (2000), que giran en torno al subdominio de la oncología. Tercedor estudia las variaciones en la fraseología de tipo léxico en inglés y español correspondientes a los registros más usuales en este subdominio.

Como punto final a este breve repaso de cómo la lingüística de corpus puede ser un valioso instrumento en los estudios de traducción, se hará una aclaración. Si bien esta metodología empírica y cuantitativa ha supuesto una superación de la lingüística basada exclusivamente en la introspección y la intuición en favor de una lingüística basada en datos auténticos (Biber y Finegan 1991), no obstante, no se debería considerar la lingüística de corpus como una panacea. Los estudios lingüísticos, tanto los basados como los guiados por los datos extraídos de un corpus, presentan algunos puntos débiles como su dependencia de la labor de abstracción por parte del lingüista, no exenta de subjetivismo, y el hecho de que la calidad del estudio depende en gran medida de la calidad y amplitud del corpus (Márquez 1998: 157). Estas limitaciones son aún mayores cuando el objeto de estudio, la cohesión, es un fenómeno que se manifiesta a lo largo de todo un texto y tiene un alcance intraoracional e interoracional.

**SEGUNDA PARTE:
METODOLOGÍA**

6. Fundamentos metodológicos para el estudio de la cohesión

6.1. Introducción

La metodología seguida en esta tesis modifica diversas aproximaciones que han resultado muy fructíferas en la comprensión del fenómeno de la cohesión. Pretende también aportar una sistematización que ayude al traductor a reconocer los recursos cohesivos presentes en las lenguas naturales y a utilizarlos en la redacción de textos.

De todas las investigaciones del capítulo anterior, las más decisivas para nuestro análisis son las de Hoey (1991, 1995), la Teoría Sentido Texto, Morris y Hirst (1991) y Ellman (1998). Las dos primeras nos servirán para analizar textos breves de forma detallada (*análisis intensivo*), mientras que las otras dos orientarán el estudio cuantitativo de textos extensos (*análisis extensivo*).

Antes de detallar la metodología a seguir, expondremos nuestra interpretación del fenómeno de la cohesión en cinco postulados que encabezan los apartados que siguen.

6.2. La cohesión como fenómeno interoracional e intraoracional

La cohesión tiene una doble naturaleza. Por una parte, es un fenómeno *interoracional*, ya que vincula entre sí las oraciones de un texto mediante recursos gramaticales y léxicos. La cohesión interoracional hace que las oraciones estén integradas en unidades textuales mayores: *macroestructuras* (van Dijk 1977b: 152)⁶⁹, *elementos discursivos* (Paltridge 1997: 90)⁷⁰, *actos de habla* (Searle 1979) y *movimientos* (Swales 1981, 1990). Hoey adopta esta concepción con la finalidad de delimitar el objeto de estudio. Y así, prescinde de los vínculos intraoracionales—los existentes entre palabras pertenecientes a una misma oración—porque pertenecen al dominio proposicional y no recurren a la repetición sistemática, que es el principal recurso sobre el que se sustenta la cohesión. De todos modos, la cohesión logra unir algo más que oraciones concretas: activa marcos conceptuales a través de las *palabras clave* del texto, independientemente de las oraciones en las que aparecen. Es decir, la cohesión trasciende la división en oraciones.

⁶⁹ "Sequences [of a text] which somehow 'belong together'".

⁷⁰ "Divisions in a text in terms of the way in which their parts function to convey various types or categories of information"

Por otra parte, la cohesión es un fenómeno *intraoracional* que hace que diversos conceptos queden asociados en una proposición y que determinados elementos que suelen coocurrir juntos queden relacionados semánticamente. Desde esta perspectiva, la colocación es un recurso cohesivo intraoracional que cuando se repite a lo largo del texto, también crea vínculos interoracionales.

Seguiremos la primera concepción en el análisis intensivo, con el fin de analizar las oraciones más informativas, y prescindiremos de la división en oraciones para el análisis extensivo. Esto se debe en parte a que la unidad de traducción no es la oración sino el enunciado y a que al establecer límites oracionales se hace difícil un estudio exhaustivo. Por eso, en el análisis extensivo el objeto de estudio serán cadenas léxicas que unen unidades léxicas⁷¹ independientemente de la oración a la que pertenecen.

6.3. La cohesión léxica como principal recurso cohesivo

De las cinco formas de cohesión identificadas por Halliday y Hasan (conjunción, referencia, sustitución, elipsis y cohesión léxica), al igual que estos autores y que Hoey, hemos destacado la *cohesión léxica*⁷² como el principal recurso cohesivo y el que más contribuye a la organización de los textos.

En consecuencia, nos hemos circunscrito al estudio de la cohesión léxica, lo cual supone profundizar en el modo en que se articula la terminología en los textos especializados, uno de los principales escollos a los que se enfrenta el traductor. El análisis de la cohesión léxica redundará pues en el conocimiento de la terminología y la textura.

Los intentos de establecer categorías de cohesión léxica se han atenido a la división que hicieron Halliday y Hasan (1976) entre *reiteración* y *colocación*⁷³, y se han centrado principalmente en el estudio de la primera. Resulta más fácil reconocer los elementos que se repiten en un texto que los que coocurren. Además, la formulación que Halliday y Hasan hicieron de la colocación es algo difusa e incluye relaciones tan

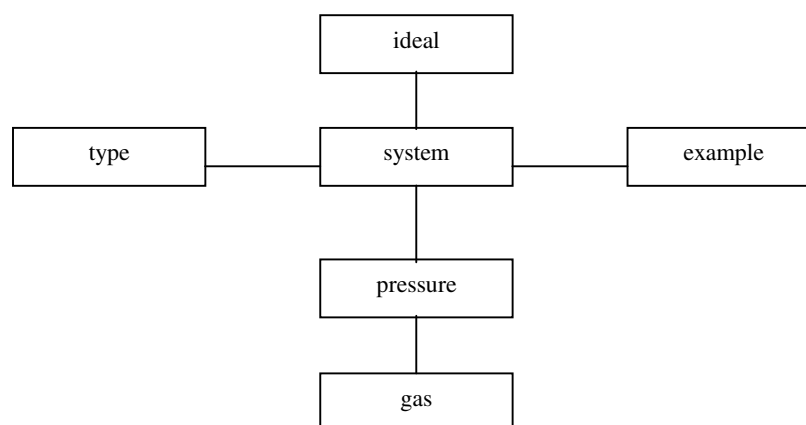
⁷¹ Dado que la competencia general y especializada del hablante están integradas de forma que, fuera de contexto comunicativo, no se puede discernir si una unidad léxica es una *palabra* o un *término*, siguiendo a Cabré (1999: 102), llamaremos a las unidades dotadas de referencia con la denominación neutra *unidades léxicas*.

⁷² Recordamos que es la que vincula palabras plenas en función de las relaciones semánticas que pueden existir entre ellas.

variadas como las existentes entre *boy—girl*, *wet—dry*, *ill—doctor*, *try—succeed*, las cuales se incluyen en gran parte de la literatura dentro de la relación de antonimia (Lyons 1968, 1977). Asimismo, su propuesta no coincide con la de Firth, que es la más extendida en lingüística, de forma que Halliday y Hasan no consideran como colocación la relación entre *sour—milk*, *commit—offense*, *dream—come true*. Como se vio en 4.2.1.2., Firth entiende la colocación como la relación que un lexema establece con otros lexemas que aparecen con una probabilidad significativa en su cotexto inmediato.

La colocación como recurso cohesivo léxico sólo ha recibido atención por parte de Phillips (1985)⁷⁴. Este autor analiza el papel de la colocación en la organización de libros de texto científicos siguiendo una perspectiva estadística. De su análisis extrae redes colocacionales correspondientes a cada capítulo, y, comparándolas, identifica cómo se van repitiendo a lo largo de capítulos separados a veces por muchas páginas. Esta consistente repetición de determinados patrones colocacionales le lleva a vincular aquellos fragmentos de texto que comparten dichos patrones. A continuación presentamos una de las redes colocacionales que se configuran en diversos capítulos de un libro de texto sobre termodinámica a partir de términos que suelen aparecer en el mismo cotexto.

DIAGRAMA 13: Redes colocacionales (Phillips 1985: 131-132)



⁷³ Entre otros, podemos mencionar los realizados por Hoey (1991) y Tanskanen (1995).

⁷⁴ La falta de estudios sobre la colocación y su relación con la cohesión también ha sido destacada por Michael Hoey (Comunicación personal).

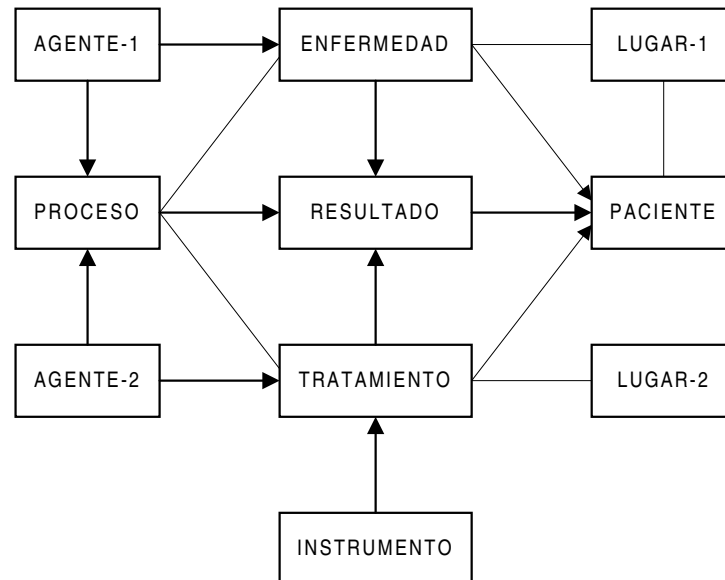
Este vacío de estudios sobre la colocación como recurso cohesivo nos ha llevado a prestar atención no sólo a los recursos basados en la repetición, sino también al uso consistente de colocaciones en textos sobre oncología. Esta innovación se completa con la propuesta de un nuevo marco de análisis acorde con las numerosas investigaciones posteriores a 1976. Este marco fundamenta la cohesión en la repetición de elementos léxicos relacionados paradigmáticamente o sintagmáticamente.

6.4. Cohesión léxica y activación de marcos

La cohesión léxica activa un universo léxico de relaciones paradigmáticas y sintagmáticas. Este universo léxico se podría representar en forma de marcos conceptuales, como proponen Fillmore y Atkins (1998). A partir de la premisa de que las palabras no se relacionan directamente sino en virtud de sus vínculos con marcos, de los que derivan su significado, Fillmore y Atkins (*apud* Faber y Mairal 1999: 74) han elaborado un inventario de dominios semánticos dentro del proyecto *FrameNet*. Según sus autores, los marcos no se han extraído meramente de la intuición del hablante, sino de córpora textuales.

Estos marcos o representaciones mentales de experiencias y creencias que se pueden verbalizar con lexemas son los que nos permiten interpretar una sucesión de oraciones como un todo. Esto es así porque la cohesión hace que se sucedan en el texto algunos de los lexemas que componen un determinado marco. Esta sucesión repetitiva llama la atención sobre la información relevante de los textos de forma que, ante la infinidad de palabras que se activan al leer en un texto, el lector sea capaz de discernir cuáles son fundamentales y cómo interpretar el texto. Este potencial que tiene la cohesión de destacar la información relevante es el fundamento de los programas informáticos citados en 4.8.

En el caso del subdominio de la oncología, Faber (1997, 1999b) ha propuesto un marco prototípico denominado *EVENTO MÉDICO* en el que se pueden integrar la mayoría de las unidades léxicas que aparecen en textos sobre oncología.

DIAGRAMA 14: Elementos del *EVENTO ONCOLÓGICO*

Dentro de este marco se pueden activar unidades léxicas relacionadas paradigmáticamente o sintagmáticamente en función de su pertenencia a uno de los campos conceptuales dentro de la oncología. Faber (1999b: 99) ha reconocido los siguientes:

1. SÍNTOMAS
2. PROCEDIMIENTOS QUIRÚRGICOS
3. FÁRMACOS
4. EFECTOS SECUNDARIOS
5. FACTORES DE RIESGO
6. PRUEBAS DIAGNÓSTICAS
7. PARTES DEL CUERPO
8. TUMORES
9. TRATAMIENTOS
10. INSTRUMENTOS MÉDICOS

Por ejemplo, dentro del campo TRATAMIENTOS, los términos *radioterapia* y *quimioterapia* están relacionados paradigmáticamente mediante una relación de cohiponimia. *Radioterapia* y *láser* establecen la relación paradigmática TRATAMIENTO-INSTRUMENTO MÉDICO. Asimismo, el término *radioterapia* está sintagmáticamente relacionado con otros como *curar*, *recibir*, *tratar* o *aplicar*.

La reiteración de estas unidades léxicas y de estas relaciones en las oraciones que componen un texto es lo que le confiere cohesión al texto. Es decir, la cohesión se fundamenta en el uso reiterado de unidades léxicas que están relacionadas, bien en el plano paradigmático o en el plano sintagmático. Estas van tejiendo vínculos dentro de las oraciones y entre oraciones.

Por todo esto, en lugar de adoptar la distinción de Halliday y Hasan entre reiteración y colocación, nosotros haremos referencia a la REPETICIÓN EXACTA, REPETICIÓN PARADIGMÁTICA, REPETICIÓN SINTAGMÁTICA y REPETICIÓN MIXTA.

6.5. Cohesión y repetición en el eje paradigmático y sintagmático

La cohesión léxica se fundamenta en algo tan simple como la repetición de elementos léxicos (Winter 1979; Hasan 1984; Hoey 1991, 1995) en el eje paradigmático o sintagmático. Al revisar la bibliografía, no hemos encontrado referencias a la cohesión basada en relaciones paradigmáticas o sintagmáticas. De ahí que propongamos una clasificación que distinga entre cohesión en el eje paradigmático o sintagmático, al mismo tiempo que se consideran criterios formales, textuales y léxicos.

La adopción de la dicotomía saussureana entre lo paradigmático y lo sintagmático contribuye a delimitar claramente la cohesión, un fenómeno en el que entran en juego básicamente dos tipos de relaciones⁷⁵:

- a) las relaciones de similitud, contraste y sustitución que se pueden establecer entre lexemas en contextos específicos mediante asociación mental (*relaciones paradigmáticas*)
- b) las relaciones de co-ocurrencia entre lexemas que aparecen en el mismo sintagma u oración (*relaciones sintagmáticas*).

En cierto modo, esta distinción no es muy diferente de la existente entre la *reiteración* y la *colocación* (Halliday y Hasan 1976). Tampoco existe gran diferencia entre la cohesión basada en relaciones paradigmáticas y la *cohesión léxica general* de Hasan (1984)⁷⁶, bajo la que se incluyen, junto a la repetición, relaciones semánticas pertenecientes al eje paradigmático: sinonimia, antonimia, hiponimia y meronimia (véase p. 106).

Desde el punto de vista formal, hemos clasificado todos los casos de repetición según el grado de similitud formal de los elementos repetidos. Distinguimos tres casos:

- REPETICIÓN EXACTA: enlaza formas idénticas desde el punto de vista morfofonológico y sintáctico de una misma unidad léxica. Hacemos la distinción entre repetición exacta y simple porque determinadas formas de un mismo paradigma pueden tener distintos cotextos y funciones.

(34) *biopsia* - *biopsia*

(35) *treatment* – *treatment options*

En esta categoría también se incluirán sustantivos en posición de modificadores nominales. Por ejemplo, el sustantivo *treatment* en *treatment options* sería una instancia de repetición exacta, porque en lengua inglesa, la función del adjetivo puede ser realizada por un sustantivo antepuesto a otro sustantivo. En este caso, la función gramatical no es distinta (Hoey 1991: 55).

- REPETICIÓN SIMPLE: enlaza a) formas idénticas (*word-type*) de una misma unidad léxica con distinta función sintáctica o b) variantes morfológicas (*word-tokens*) de una misma unidad léxica. En lengua inglesa, también es la que se produce entre un lexema que pertenece a más de una categoría gramatical como resultado de un proceso de formación de palabras llamado *conversion* o *zero derivation*.

⁷⁵ Las definiciones que siguen son una adaptación de Wanner (1996: 3, 12)

⁷⁶ Los estudios de cohesión hasta Hoey (1991) suelen recurrir a la dicotomía, en consonancia con lo que ocurre en otras áreas de la lingüística. Buena prueba de ello son las dicotomías *reiteración /colocación* (Halliday y Hasan 1976) o *cohesión léxica general – cohesión léxica inmediata* (Hasan 1984).

- (36) lung cancer – lung cancers⁷⁷
- (37) treat – treated⁷⁸
- (38) release (Verbo) – release (Sustantivo)⁷⁹

- REPETICIÓN VARIADA: enlaza unidades léxicas diferentes desde el punto de vista formal pero relacionadas semánticamente. Aquí es posible distinguir dos categorías según la variación se fundamente en un cambio léxico o sintáctico:

REPETICIÓN VARIADA LÉXICA:

- (39) cancer – carcimoma.
- (40) anomalías – alteraciones

REPETICIÓN VARIADA SINTÁCTICA:

- (41) investigar – investigaciones⁸⁰
- (42) cáncer de pulmón – cáncer pulmonar

Desde el punto de vista textual, las repeticiones exacta y simple marcan la cohesión de forma explícita, mientras que la repetición variada puede ser explícita, semiexplícita o implícita según cómo se presente la relación entre lexemas en el cotexto. Esta distinción es perceptible al comparar textos especializados y textos divulgativos. En los primeros, la repetición variada tiene lugar de una forma implícita, es decir, no existen en el texto ningunos indicadores de la relación existente entre las unidades léxicas que se repiten. En este texto sobre oncología, es el conocimiento especializado del lector el que le permite reconocer la cohesión entre el término superordinado *chromosome 9 aberrations* con respecto a *deletion of chromosome 9p*.

- REPETICIÓN VARIADA IMPLÍCITA

(43)

- 1 Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary *Chromosome 9 Aberrations*
- 2 BACKGROUND. *Deletion of chromosome 9p* has been reported in numerous tumor types.

(Cfr. Apéndice II: texto 7, pág.1)

⁷⁷ Repetición léxica simple en Hoey (1991).

⁷⁸ Repetición léxica simple en Hoey (1991).

⁷⁹ Repetición léxica compleja en Hoey (1991).

⁸⁰ Repetición léxica compleja en Hoey (1991).

Por el contrario, en los textos divulgativos, la repetición variada se presenta de forma semi-explícita o explícita en el texto:

- REPETICIÓN VARIADA SEMI-EXPLÍCITA: En esta noticia médica publicada en el diario *El Mundo* (25-1-1996), el título y el subtítulo presentan próximos los lexemas *prostáticas* y *glándula*, anticipando la relación de sinonimia entre *próstata* y *glándula*. Esta se da a entender al lector, repitiendo tres veces el mismo cotexto: *cáncer* de*. Aunque quizás la repetición variada se deba más a motivos estilísticos—no repetir demasiado la palabra *próstata*, que didácticos.

(44)

Fórmulas prostáticas

Sigue la polémica sobre la terapia del cáncer de la glándula

[...] se aconseja qué es lo que hay que hacer con los cánceres de próstata, [...]

[...] en otros tantos enfermos afectados de cáncer de la glándula [...]

[...] los autores del trabajo concluyen que ante un cáncer de próstata [...]

(Cfr. Apéndice II: texto 3)

- REPETICIÓN VARIADA EXPLÍCITA

En este párrafo de una página *web* sobre el cáncer⁸¹ se hace explícita la relación entre dos unidades léxicas susceptibles de crear cohesión mediante la locución *determinados ... como* o el paréntesis. Hay también un caso de repetición semi-explícita en el que se relaciona *causas externas* con *factores externos* (repetición del lema *extern**), y estos, a su vez con *factores ambientales* (repetición del lexema *factores* y utilización del determinante *otros*).

(45)

El cáncer puede ser causado por *causas externas* al cuerpo, o por causas internas.

Los *factores externos* que pueden causar el cáncer incluyen el estar expuesto a **determinados productos químicos como** el *benceno* y el *asbesto (amianto)*. Los pintores, los fabricantes de neumáticos, los destiladores y los fabricantes de zapatos, están expuestos a menudo al *benceno*. Los trabajadores de minas, aislamiento y astilleros están expuestos a menudo al *asbesto*. Otros *factores ambientales* que causan cáncer incluyen [...]

(Cfr. Apéndice II: texto 6)

⁸¹ Texto extraído de la página web de información sanitaria *EcoMedic*, patrocinada por Sarenet, Pulso 100 y Sanitas [<http://ecomedic.com/em/indice.htm>].

Al hablar de repetición, pues, no nos hemos limitado a la reiteración exacta de la representación grafofonológica (REPETICIÓN EXACTA) de una unidad léxica, sino que también hemos incluido otras instancias de repetición. La nomenclatura de estas instancias incluye los términos *paradigmática* o *sintagmática*, en consonancia con nuestra distinción entre lo paradigmático y lo sintagmático:

- a) REPETICIÓN PARADIGMÁTICA SIMPLE: reiteración de variantes morfológicas (*word-tokens*) de una misma unidad léxica (*word-type*) sin que se produzca un cambio en la relación sintagmática.

(46) child – children

- b) REPETICIÓN PARADIGMÁTICA SINTÁCTICA: reiteración del mismo significado encarnado en unidades léxicas diferentes desde el punto de vista sintáctico.

(47) water (*sustantivo*) – water (*verbo*): SIMPLE

(48) buy (*verbo*) – purchase (*sustantivo*): VARIADA LÉXICA

Como se percibe en el ejemplo, las unidades léxicas no siempre están morfológicamente relacionadas.

- c) REPETICIÓN PARADIGMÁTICA SEMÁNTICA: reiteración de palabras relacionadas semánticamente por compartir el mismo significado básico o por aparecer relacionadas en el contexto extralingüístico, en cuyo caso, la relación semántica es más difusa. La relación entre estas palabras puede ser, entre otras, de hiponimia, sinonimia, oposición, meronimia o circunstancia. Bajo esta categoría, surgen variantes, de entre las que destacamos:

REPETICIÓN PARADIGMÁTICA HIPONÍMICA:

(49) fármaco – taxol

REPETICIÓN PARADIGMÁTICA SINONÍMICA:

(50) letal – mortífero

REPETICIÓN PARADIGMÁTICA OPOSITIVA:

(51) maligno – benigno

REPETICIÓN PARADIGMÁTICA MERONÍMICA:

(52) lung – lobe

REPETICIÓN PARADIGMÁTICA ASOCIATIVA CIRCUNSTANCIAL (INSTRUMENTO):

(53) resección – láser

- d) REPETICIÓN SINTAGMÁTICA: reiteración de combinaciones de lexemas relacionados sintagmáticamente con una frecuencia significativa en nuestro corpus o en el texto analizado. Tercedor (1999) clasifica las combinaciones de lexemas en una escala que va desde unidades con vinculación fija en la cadena sintagmática (unidades lexicalizadas)⁸² hasta la combinación sintagmática de lexemas que suelen aparecer próximos entre sí aunque con una cierta libertad (colocaciones)⁸³, pasando por una categoría intermedia denominada unidad fraseológica no lexicalizada. Es decir, las unidades fraseológicas no lexicalizadas tienen mayor fijación que las colocaciones y menos que las unidades lexicalizadas. De todas formas, a estas dos categorías, Tercedor las denomina unidades fraseológicas (UF) siguiendo a Meyer y Mackintosh (1996: 3), dado que el límite entre unas y otras no está claramente fijado. Corpas (2000: 484) define las *unidades fraseológicas* como "combinaciones estables formadas por al menos dos palabras y [...] se caracterizan por la alta frecuencia de aparición en la lengua y de coaparición de sus elementos integrantes, así como por la institucionalización, estabilidad, la idiomatización y la variación potencial que dichas unidades presentan en diverso grado". La expresión "diverso grado" incide de nuevo en que no existe una división tajante entre unidades fraseológicas y colocaciones, de ahí que en esta tesis utilicemos UF para referirnos a los dos. Bajo *repetición sintagmática*, sólo aparecerán casos en los que la repetición de

⁸² Para Tercedor (1999: 297, 299), las unidades lexicalizadas son unidades poliléxicas sin unión gráfica que transmiten un concepto y no permiten alteración en el orden de los elementos que las componen ni tampoco en los constituyentes de las mismas.

⁸³ Véase definición en p. 104.

UF sea exacta, simple o variada sintáctica y la relación entre ellas sea la de sinonimia. Analizando distintos ejemplos surgen una serie de categorías:

REPETICIÓN SINTAGMÁTICA EXACTA :

(54) family_history_of_cancer – family_history_of_cancer

REPETICIÓN SINTAGMÁTICA SIMPLE :

(55) prueba_citológica—pruebas_citológicas

REPETICIÓN SINTAGMÁTICA VARIADA SINTÁCTICA:

(56) metástasis_cerebrales—metástasis_en_el_cerebro

Cuando hay repetición variada léxica, las unidades fraseológicas—basadas en relaciones sintagmáticas—pueden establecer entre ellas una relación paradigmática, como veremos en el apartado siguiente.

e) REPETICIÓN MIXTA: reiteración variada léxica de unidades tanto monoléxicas como poliléxicas en la que entran en juego tanto relaciones sintagmáticas como paradigmáticas. En esta categoría mixta, vamos a distinguir entre:

- REPETICIÓN MIXTA DE UNIDADES FRASEOLÓGICAS⁸⁴: vincula dos unidades fraseológicas relacionadas semánticamente mediante hiponimia, oposición, sinonimia, etc.

(57) adenocarcinoma_of_the_lung – lung cancer (REPETICIÓN SINTAGMÁTICA HIPONÍMICA DE UNIDADES FRASEOLÓGICAS)

- REPETICIÓN MIXTA DE UNIDAD FRASEOLÓGICA Y LEXEMA: es la que existe entre una unidad fraseológica y un lexema que están relacionados semánticamente mediante hiponimia, sinonimia, meronimia, acronimia, etc.

(58) digestive_system – stomach (REPETICIÓN MIXTA MERONÍMICA DE UNIDAD FRASEOLÓGICA Y LEXEMA)

(59) positron_emission_tomography – PET (REPETICIÓN MIXTA ACRONÍMICA DE UNIDAD FRASEOLÓGICA Y LEXEMA)

6.6. Funciones léxicas en la generación de textos cohesivos

Es posible presentar las relaciones paradigmáticas y sintagmáticas de un texto en forma de funciones léxicas abstractas cuya sucesión a lo largo del texto configura su textura.

Las funciones léxicas propuestas dentro de la Teoría Sentido Texto (TST), que describimos en 4.9., aportan una importante sistematización de relaciones en torno a criterios léxicos, aplicable al estudio de la cohesión.

6.6.1. Propuesta de funciones léxicas cohesivas

6.6.1.1. Repetición exacta

En la TST no hemos encontrado ninguna función léxica para designar la relación entre el mismo lexema a lo largo de un texto. Dada la importancia de la repetición del mismo lexema en la creación de textura, hemos propuesto la función **Aequ** (*aequalis*), que indica la igualdad entre dos lexemas desde el punto de vista formal y semántico. En el caso de la repetición exacta, **Aequ** también indica igualdad sintáctica.

Aequ equal term
Aequ (*tumor*) = *tumor*

6.6.1.2. Repetición paradigmática

En la repetición paradigmática, la cohesión se establece mediante FL que vinculan significados que aparecen asociados con regularidad a la palabra clave en función de su pertenencia a un sistema de relaciones semánticas⁸⁵. La repetición paradigmática por lo general se sustenta en relaciones semánticas. En consecuencia, la mayoría de las propuestas de relaciones paradigmáticas incluyen algunas de las relaciones de sentido formuladas por Lyons (1968, 1977). Este es el caso de las relaciones paradigmáticas propuestas por el Modelo del Léxico Generativo (Pustejovsky 1995): sinonimia, antonimia, hiponimia y herencia, meronimia, implicatura y presuposición. Wanner (1996) por su parte señala la hiponimia, meronimia, sinonimia y oposición.

Además de las cuatro últimas relaciones semánticas, incluiremos otras más difusas bajo el nombre de *repetición paradigmática asociativa*. También incluiremos la

⁸⁴ No especificamos que la repetición es variada porque se infiere del significado de los términos *hiponimia, sinonimia, meronimia*.

relación entre las variantes morfológicas de una misma palabra y entre unidades léxicas que son semánticamente equivalentes pero sintácticamente diferentes (repetición paradigmática sintáctica).

6.6.1.2.1. REPETICIÓN PARADIGMÁTICA SIMPLE

La FL sobre la que se sustenta la repetición paradigmática simple es **Aequ** (*aequalis*) que indica la repetición en el mismo texto de dos lexemas iguales desde el punto de vista formal y semántico.

Aequ equal term
 Aequ (*disease*) = *diseases*

Como bajo la repetición simple se incluyen formas idénticas de una misma unidad léxica con distinta función sintáctica, **Aequ** también relacionaría el sustantivo *water* con el verbo *water*, aunque como ha habido un cambio sintáctico, **Aequ** aparece combinada con **V₀** (véase apartado siguiente).

6.6.1.2.2. REPETICIÓN PARADIGMÁTICA SINTÁCTICA

Este tipo de reiteración pone en juego una serie de FL entre lexemas con el mismo significado pero que no pertenecen a la misma categoría sintáctica. Dado que estas FL acarrearán un cambio de categoría sintáctica, Mel'cuk (1996: 52) las denomina *syntactic derivatives*, Wanner (1996), *syntactic paradigmatic relations* y Alonso Ramos y Tutin (1996), *syntactic derivations*. Con esta FL nos referimos a lexemas que comparten la misma raíz.

V₀ verb for keyword
 V₀ (*cautery*) = cauterize
 AequV₀ (*water*) = *water*
 S₀ noun for keyword
 S₀ (*tolerate*) = tolerance

 A₀ Adjective derived from keyword
 A₀ (*clinic*) = clinical

⁸⁵ Esta definición es una adaptación de la propuesta por Steele y Meyer (1990: 42) para las funciones léxicas paradigmáticas.

Adv₀ Adverb from keyword
Adv₀ (*oral*) = orally

Dentro de la repetición paradigmática sintáctica podemos incluir dos tipos de FL que hemos incluido en la sección 6.6.1.2.3. bajo la repetición paradigmática asociativa, en concreto, las funciones **Able**, **S_{1,2,3}** y **A_{1,2}**. Cuando estas dos últimas asocian dos lexemas de igual raíz, tendremos un caso de repetición paradigmática sintáctica calificativa o actancial, según corresponda.

Able_{1,2} ability of numbered participant
Able₁ (*regression*) = regressive
Able₂ (*detect*) = detectable
REPETICIÓN PARADIGMÁTICA SINTÁCTICA CALIFICATIVA

S_{1,2,3} typical noun for numbered participant
S₁ (*anesthetize*) = anesthetist
REPETICIÓN PARADIGMÁTICA SINTÁCTICA ACTANCIAL

A₁, A₂ typical adjective for numbered participant
A₁ (*to cure*) = curative
A₂ (*to cure*) = curable
REPETICIÓN PARADIGMÁTICA SINTÁCTICA ACTANCIAL

6.6.1.2.3. REPETICIÓN PARADIGMÁTICA SEMÁNTICA

En este apartado se incluyen FL que establecen relaciones jerárquicas y no jerárquicas entre términos que comparten el significado básico o que aparecen relacionados en el contexto extralingüístico. Aunque las funciones agrupadas bajo la categoría de meronimia aparecen en Alonso Ramos y Tutin (1996: 157) como sintagmáticas, nosotros las consideraremos paradigmáticas siempre y cuando la palabra clave y el valor no coocuran.

a) REPETICIÓN PARADIGMÁTICA SINÓNÍMICA

En este tipo de recurso cohesivo, los lexemas relacionados tienen el mismo significado pero, a diferencia de la repetición paradigmática sintáctica, pertenecen a la misma categoría sintáctica y no tienen la misma raíz. Somos conscientes de que la sinonimia absoluta no existe y que suele haber diferencias

entre los sinónimos con respecto a los siguientes parámetros: dialecto, registro, situación, valor axiológico y radio colocacional (Palmer *apud* Saeed 1997: 65-66).

Syn synonym and near synonym
Syn (*doctor*) = physician

Como existen diferentes grados de sinonimia, Mel'cuk (1996: 47) propone tres funciones léxicas relacionadas con **Syn**:

Syn \supset including synonym (with more specific meaning)
Syn \supset (*to fire upon*) = *to shell*
Syn \subset included synonym (with less specific meaning)
Syn \subset (*to modify*) = *to change*
Syn \cap intersecting synonym
Syn \cap (*to escape*) = *to elude, to avoid*

La sinonimia parcial puede adoptar también las funciones **Figur** y **Gener**.

Figur metaphor
Figure (*cancer*) = beast⁸⁶

Gener generic word
Gener (*pain*) = feeling

En relación con la sinonimia, introducimos una FL no reconocida en la TST que establece un vínculo entre una unidad fraseológica y su acrónimo. La repetición es paradigmática si hay una gran fijación de los elementos de la UF, de forma que apenas se reconoce una relación sintagmática entre ellos. Cuando esto no ocurre, estamos ante un caso de repetición mixta de UF y lexema (véase p. 216).

Acro acronym
Acro (*Small cell lung cancer*) = *SCLC* REPETICIÓN PARADIGMÁTICA
ACRONÍMICA

⁸⁶ Metáfora usual entre los pacientes de cáncer basada en la Biblia que hemos identificado en numerosos correos electrónicos de la lista de distribución LUNG-ONC [<http://www.medinfo.org/listserv.html>].

b) REPETICIÓN PARADIGMÁTICA HIPONÍMICA

La función que asocia un término hiponímico con su hiperónimo es **Gener**.

Gener generic word
Gener (*Edatrexate*) = drug

c) REPETICIÓN PARADIGMÁTICA MERONÍMICA

A pesar de que la relación entre el concepto comprensivo y el concepto partitivo o relación *parte – todo* no está siempre institucionalizada (Wanner 1996: 5), es posible identificar siete funciones léxicas cohesivas basadas en la relación de meronimia.

- Relación *miembro – conjunto*:
 - Mult regular aggregate of
Mult (*cell*) = tissue
- Relación *conjunto social – personal*:
 - Equip the staff /crew of
Equip (*hospital*) = staff
- Relación *organización – máximo representante*
 - Cap the head of
Cap (*hospital*) = director
- Relación *todo – parte uniforme*
 - Sing one instance /unit of
Sing (*tissue*) = cell
- Relación *todo – culminación /centro /núcleo*
 - Culm culmination of
Culm (*cancer*) = metastasis
 - Centr center of
Centr (*thorax*) = breastbone
 - Germ the core of /the origin of
Germ (*cell*) = nucleus

d) REPETICIÓN PARADIGMÁTICA OPOSITIVA

Activa tres FL que se encuentran muy en la línea de la distinción de Lyons (1968) entre antonimia, conversión y complementariedad. La denominación por parte de la TST es la de antonimia, conversión y contraste.

Anti	antonym (exact or near) Anti (<i>malignant</i>) = benign
Conv _{ijk}	conversive (opposite where participants switch roles) Conv (<i>cure</i>) = heal ⁸⁷
Contr	non-antonymic contrast Contr (<i>auricle</i>) = ventricle

e) REPETICIÓN PARADIGMÁTICA ASOCIATIVA⁸⁸

Esta categoría se aplica para designar relaciones semánticas más difusas que las de sinonimia, hiponimia, oposición y meronimia. La repetición se fundamenta en el hecho de que los referentes de los lexemas aparecen relacionados tanto en el contexto extralingüístico como lingüístico, lo cual permite una *asociación* mental entre dichos lexemas que crea cohesión.

- CALIFICATIVA: Vincula a la palabra clave con adjetivos que califican a los actantes⁸⁹ potenciales en la situación designada por la misma. En el caso de que estos adjetivos coocuran con el sustantivo, nos encontramos ante un caso de repetición sintagmática. Los subíndices indican si el valor es el primer, segundo o tercer actante o argumento de la situación a la que hace alusión la palabra clave.

Able _{1,2}	ability of numbered participant Able ₁ (<i>cisplatin</i>) = curative Able ₂ (<i>resect</i>) = resecable
---------------------	---

⁸⁷ Nos referimos a la acepción '*to become healthy again*' del verbo *to heal*, presente en la oración "The patient healed".

⁸⁸ La nomenclatura de las categorías que hemos incluido procede de Alonso Ramos y Tutin (1996: 153).

⁸⁹ En la TST, el término actantes es similar al de argumentos (Dik 1978) y al de participantes (Halliday 1985). Mel'čuk y Wanner (1996: 214) definen los actantes semánticos de un lexema como los argumentos del predicado que constituyen el significado de ese lexema. Alonso Ramos y Tutin (1996: 166) emplean el término *participantes* para referirse a la situación designada por la palabra clave y por tanto, sostienen que un lexema no tiene participantes, sino actantes.

Qual_{1,2,3} Highly probable qualities of numbered participants

Qual₁(*cancer*) = invasive

Qual₂(*cancer*) = fatal

Qual₃(*cancer*) = curable

- ACTANCIAL: Bajo esta categoría se agruparán una serie de funciones para formar nombres, adjetivos y adverbios que hagan referencia a los actantes de la situación que la palabra clave designa. De nuevo, los subíndices indican si el valor es el primer, segundo o tercer actante o argumento de la situación a la que hace alusión la palabra clave.

S_{1,2,3} typical noun for numbered participant

S₁ (*examine*) = doctor

S₂ (*cancer*) = patient

A_{1, A2} typical adjective for numbered participant

A₁ (*to cure*) = effective

A₂ (*to cure*) = early stage (tumor)

- CIRCUNSTANCIAL: Este tipo de repetición cataliza una serie de FL que relacionan la palabra clave con un sustantivo que designa las circunstancias que acompañan a la palabra clave. En el dominio de la oncología esta función es sumamente productiva en la generación de cohesión ya que pone en contacto algunos de los elementos participantes en el EVENTO ONCOLÓGICO (Faber 1997, 1999b): instrumento (INSTRUMENTO), lugar (LUGAR), medio (AGENTE-2), modo (TRATAMIENTO) y resultado (RESULTADO). El TST reconoce estos elementos en cinco funciones circunstanciales (**S_{inst}**, **S_{loc}**, **S_{med}**, **S_{mod}** y **S_{res}**), que relacionan un verbo (palabra clave) con el sustantivo correspondiente. A estos elementos circunstanciales, pensamos que se pueden añadir el de causa (**S_{cause}**) y síntoma (**S_{Sympt}**).

En nuestro corpus hemos encontrado muchos ejemplos en los que la palabra clave es un sustantivo, en lugar de un verbo, que se relaciona con otro sustantivo. Para evitar una multiplicidad excesiva de funciones, utilizaremos la misma nomenclatura para estos casos.

S_{instr}	typical instrument S_{instr} (<i>diagnosis</i>) = bronchoscope (Instrumento para diagnosticar el cáncer de pulmón) [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (INSTRUMENTO)]
S_{loc}	typical location S_{loc} (<i>hypernephroma</i>) = kidney (Localización del tumor) [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (LUGAR)]
S_{med}	typical means S_{med} (<i>chemotherapy</i>) = cisplatin [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (MEDIO)]
S_{mode}	typical mode S_{mode1} (<i>cancer</i>) $Liqu_1$ = chemotherapy [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (MODO)]
S_{res}	typical result S_{res} (<i>treatment</i>) Pos= survival [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (RESULTADO)]
S_{cause}^{90}	typical cause (risk factor) S_{cause} (<i>asbestos</i>) = cancer [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (CAUSA)]
S_{effect}	typical effect of S_{effect} (<i>chemotherapy</i>) = ulcerations [REPETICIÓN PARADIGMÁTICA CIRCUNSTANCIAL (EFECTO)]
S_{Sympt}	to be a physical symptom of S_{Sympt} (<i>lung cancer</i>) = dyspnea [REPETICIÓN MIXTA CIRCUNSTANCIAL (SÍNTOMA)]

En este apartado también se podrían incluir otras relaciones entre sustantivos más difusas como las de contenedor-contenido, material-producto y origen-destino. De ahí que hayamos propuesto las siguientes funciones: $S_{Contain}$, S_{Mater} , S_{Germ} , S_{Destin} .

$S_{Contain}$	container of the keyword $S_{Contain}$ (<i>urine</i>) = bladder
---------------	--

⁹⁰ Esta función es diferente de la función sintagmática **Caus**, que asocia un sustantivo con el verbo causativo del que es argumento.

S_{Mater}	material/substance from which the keyword is made S_{Mater} (<i>platinol</i>) = cisplatin <i>Platinol</i> es el nombre comercial de una sustancia química que se llama <i>cisplatin</i> (cisplatino).
S_{Germ}	origin of the keyword S_{Germ} (<i>lymphocyte</i>) = bone marrow
S_{Destin}	destination to which the keyword is going. S_{Destin} (<i>lung cancer</i>) = lymph nodes

- VERBAL : Asocia verbos entre los que existe una relación semántica como la de causa-efecto, acción-resultado, etc.

Effect	a verb meaning ‘to be the effect of ‘ Effect (<i>to exercise</i>) = to sweat
Result	a verb meaning ‘to be the expected result of ‘ Result (<i>to fast</i>) = [to] lose weight

Aunque Alonso Ramos y Tutin (1996: 155) incluyen esta última como relación sintagmática aspectual, nosotros creemos que es una relación asociativa que se produce entre dos verbos. De hecho, es una relación similar a la que existe entre términos conversos, aunque la diferencia reside en que la función **Result** relaciona verbos en los que el sujeto gramatical es el mismo. No obstante, como veremos en 6.6.1.3. bajo el apartado c) *Con verbos de realización*, cuando esta función se combina con otra para definir la coocurrencia de un verbo y un sustantivo, nos encontramos ante una función sintagmática con verbo aspectual.

6.6.1.3. Repetición sintagmática

Con este nombre hacemos referencia a la cohesión que se fundamenta en la repetición en distintas oraciones de relaciones sintagmáticas, esto es, de las relaciones que se derivan del comportamiento colocacional de la palabra clave dentro de una red semántica⁹¹. En nuestro análisis de la repetición sintagmática, no vamos a computar la relación entre una palabra y las que se encuentran en su radio colocacional como

⁹¹ Adaptación de Steele y Meyer (1990: 42).

cohesiva, a menos que sea frecuente en el corpus o recurrente en un determinado texto, lo cual crea vínculos entre las oraciones del texto.

Al igual que ocurre con las relaciones paradigmáticas, es posible sistematizar las funciones léxicas que participan en este tipo de cohesión de acuerdo con la *Teoría Sentido Texto*. En esta ocasión la clasificación de las FL obedece a la categoría sintáctica del *valor* de la función léxica: sustantivo, adjetivo, adverbio o verbo (Alonso Ramos *et al.* 1995: 354-355; Alonso Ramos y Tutin 1996: 154-157; Mel'cuk 1996: 56-72). Hemos suprimido la categoría que estos autores denominan *funciones léxicas preposicionales* porque nuestro principal interés está en la cohesión léxica y en aquellos valores que son palabras plenas, y porque el estudio de textos sobre cáncer requiere especialmente una aproximación terminológica, en la que prima el vocabulario especializado. De este, se concederá una especial relevancia a los sustantivos y los verbos y, claro está, a los lexemas que constituyen su radio colocacional.

6.6.1.3.1. FUNCIONES LÉXICAS NOMINALES

Entre las funciones cuyo valor es un sustantivo, Alonso Ramos y Tutin (1996) incluyen una serie de funciones que se detallaron en la categoría repetición paradigmática semántica: **Gener**, **Figur**, **Sing**, **Cap**, **Culm**, **Mult**, **Centr** y **Germ**. Creemos que en el caso de que tanto la palabra clave como el valor aparezcan en el mismo cotexto, estas funciones establecen una relación sintagmática entre estos. El hecho de que puedan comportarse tanto paradigmáticamente (si la palabra clave y el valor comparten un componente semántico) como sintagmáticamente (si la palabra clave y el valor forman una frase) ha motivado que se consideren *funciones mixtas*. Este término es el que utilizan Ramos *et al.* (1995: 355) para designar las funciones con un doble comportamiento⁹², algo muy habitual en la función léxica **Centr**.

Centr	center of
	Centr (<i>life</i>) = the prime [of life]

⁹² Estos autores las ilustran con los siguientes ejemplos: **Mult** (ship) = fleet; **Cap** (school) = principal.

6.6.1.3.2. FUNCIONES LÉXICAS ADJETIVALES /ADVERBIALES

Las funciones sintagmáticas adjetivales /adverbiales son las que relacionan en la cadena sintagmática a las siguientes categorías gramaticales:

Adjetivo + Sustantivo: *deep regret*

Verbo + Adverbio: *to complain bitterly*

Adverbio + Adjetivo: *highly competent*

Entre estas podemos mencionar cuatro funciones que se pueden considerar afines por ser positivas desde el punto de vista axiológico:

Ver true, correct, or proper (as it should be)
Ver (*dose*) = appropriate

Bon standard praise for the keyword coming from the speaker
Bon (*growth*) = normal

Pos_i positive attributes of numbered participants coming from the agent of the keyword
Pos₁ (*operation*) = successful

Magn intensity (can specify dimension of quantity or duration)
Magn (*malignant mesothelioma*) = advanced

Nótese cómo en el subdominio del cáncer, el valor de la función **Magn** suele ser negativo, al contrario de lo que ocurre en la mayoría de los dominios en los que la abundancia y la intensidad son valoradas positivamente. Otras funciones relacionadas con la idea de cantidad son las de **Plus** y **Minus**, que aparecen con frecuencia combinadas con la función **Pred** y con funciones que informan sobre el aspecto verbal.

Plus more of
IncepPredPlus (*temperature*) = rising

Minus less of
IncepPredMinus (*pressure*) = decreasing

Es posible combinar las funciones **Magn**, **Ver**, **Bon** y **Pos_i** con la función **Anti** (antonym) para obtener FL complejas⁹³.

AntiMagn (*temperature*) = low

AntiPos₂ (*results*) = negative

La función **Bon** también se puede combinar con **Minus**, con lo que genera una función compleja denominada **Pejor**.

Pejor (*lifestyle*) = deteriorating

Por último, mencionaremos la función **Epit**, que relaciona un sustantivo con un adjetivo sin que se especifique si el adjetivo supone un juicio de valor positivo o negativo. En el caso de la lengua española, creemos necesario especificar con una flecha (→) si el adjetivo antecede (1) o sigue (2) al sustantivo.

Epit standard epithet (representing a part of keyword)

Epit (*cure*) = new

Epit₁ (→ *cura*) = nueva

Epit₂ (*etapa* →) = avanzada

6.6.1.3.3. FUNCIONES LÉXICAS VERBALES

Las funciones léxicas verbales relacionan el sustantivo que actúa de palabra clave con un valor que pertenece a la categoría verbal.

a) CON VERBOS COPULATIVOS.

Mel'cuk reconoce dos funciones, **Copul** y **Pred**, que conectan sustantivos con verbos copulativos. La función **Copul** relaciona un sustantivo con un verbo cuyo significado es 'ser'. En cuanto a **Pred**, relaciona un sustantivo o un adjetivo L con un verbo cuyo significado es 'ser un L', con lo que el significado del verbo cubre el significado de la combinación sintáctica de la palabra clave y su cópula.

Copul verb with the meaning of a copula

Copul (*example*) = to serve as an [example]

Pred verb meaning 'to be an L'

Pred (*smoker*) = smoke

⁹³ Alonso Ramos *et al.* (1995: 355) las definen como "combinations which cannot be decomposed [...], that is, roughly LFs for which we cannot deduce the composition from the parts".

Nuestra opinión respecto a la función **Pred** es que sólo se debería considerar sintagmática cuando la palabra clave y el valor coocurren, algo que no sucede con mucha frecuencia. Opinamos que **Pred** es una función mixta que, en la mayoría de los casos, se comporta paradigmáticamente. De hecho, su significado es el inverso del de la función paradigmática **S₁**, que relacionaba un verbo con el sustantivo correspondiente al agente de la acción verbal (véase p. 199).

b) CON VERBOS DE APOYO

Los verbos de apoyo son aquellos que en una función léxica relacionan un sustantivo que funciona como palabra clave con sus actantes. Esta denominación hace referencia al hecho de que estos verbos están desprovistos de significado en presencia de la palabra clave. Por este motivo, Mel'cuk denomina a estos verbos semiauxiliares. Asimismo, en los ejemplos de la TST revisados, es perceptible un cierto grado de fijación entre la palabra clave y el valor, lo cual marca la diferencia con respecto a los verbos del apartado siguiente. Las funciones léxicas asociadas a los verbos de apoyo, **Oper_i**, **Func_i** y **Labor_{ij}**, se diferencian entre sí por la función sintáctica de la palabra clave (L) en relación con el verbo, que puede ser la de objeto principal (*Deep-Syntactic Actant II*), sujeto (*DsyntA II*) u objeto secundario (*DsyntA III*)⁹⁴.

TABLA 14: Definición de los verbos de apoyo (Mel'cuk 1996: 64).

LF	Deep-syntactic-role with respect to the LF value		
	DsyntA I	DSyntA II	DSyntA III /IV
Oper _{0/1/2}	Dummy /Ist /IInd DSyntA of L	L	—
Func _{0/1/2}	L	none /Ist / IInd DsyntA of L	—
Labor _{12/21}	Ist /IInd DsyntA of L	IInd /Ist DsyntA of L	L

⁹⁴ En la Gramática Funcional (Dik 1978, 1989), estas funciones sintácticas son las que desempeña el segundo argumento (objeto principal), el primer argumento (sujeto) y el tercer argumento (objeto secundario).

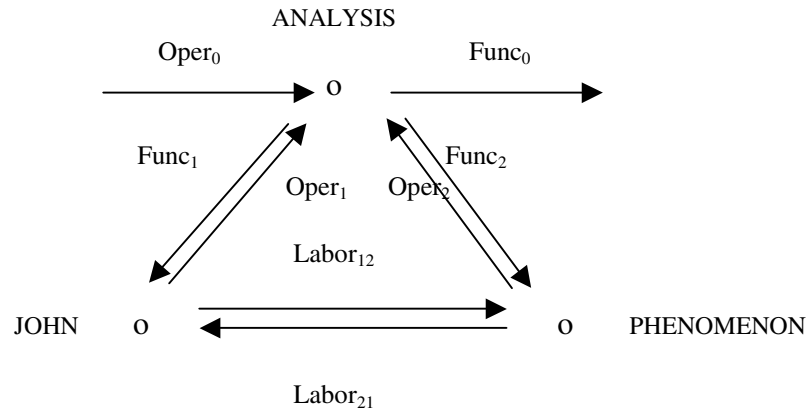
Según este cuadro, en la función **Oper**, la palabra clave desempeña el papel de objeto (segundo argumento) del verbo que hace de valor. En **Func**, la palabra clave hace de sujeto (primer argumento), mientras que en **Labor**, hace de objeto secundario (tercer argumento) del verbo. La distinción entre el subíndice de las funciones indica la perspectiva desde la que se considera la acción en relación a la estructura subyacente de la oración, algo que Mel'cuk alude con el nombre de *estructura comunicativa de la oración*. Así, el número 1 indica que el verbo es el correspondiente a un sujeto agente, mientras que el subíndice 2 indica un sujeto paciente. En el caso de que el verbo fuera impersonal o intransitivo, el subíndice sería 0. Veamos algunos ejemplos

- Oper_i verb which takes numbered participant as subject and keyword as object.
 Oper₁ (*support*) = to lend [support to someone]
 Oper₂ (*support*) = to receive [support from someone]
- Func_i verb which takes the keyword as subject and the nth participants as objects.
 Func₀ (*silence*) = reigns
 Func₁ (*proposal*) = comes, stems [from N].
- Labor_{ij} verb which takes the participants i, j as subject and object, an the keyword as secondary object.
 Labor₁₂ (*esteem*) = hold [x holds y in esteem]

A continuación hemos incluido un ejemplo tomado de Mel'cuk (1996: 63) en el que el lexema *analysis* aparece con los verbos de apoyo que más frecuentemente lo acompañan. Las combinaciones léxicas que aparecen en el ejemplo son aplicables a muchos de los sustantivos propios de la epistemología de la ciencia⁹⁵. Para mejorar la comprensión de los ejemplos, hemos incluido un diagrama (ibid) en el que se percibe la relación entre los actantes del verbo.

⁹⁵ Estos sustantivos se incluyen dentro de lo que hemos llamado "palabras organizadoras del discurso".

DIAGRAMA 15: Verbos de apoyo que se combinan con la palabra clave *analysis*. Relaciones sintácticas en la estructura profunda.



- Oper₁ (*analysis*) = [to] carry out
[John carries out the analysis of the phenomenon]
- Oper₂ (*analysis*) = [to] undergo
[The phenomenon underwent (careful) analysis by John]
- Func₁ (*analysis*) = is due
[The analysis of this phenomenon is due to John]
- Func₂ (*analysis*) = covers, concerns
[John's analysis concerns this phenomenon]
- Labor₁₂ (*analysis*)= [to] submit
[John submits this phenomenon to a (careful) analysis]
- Labor₂₁ (*analysis*)= ---
- Func₀ (*analysis*) = takes place, occurs
[John's analysis of the phenomenon is taking place]
- Oper₀ (*analysis*) = [one] sees
[One sees an analysis of the phenomenon by John]

c) CON VERBOS DE REALIZACIÓN (*FULFILL*, HACER)

Estas funciones se comportan sintácticamente de una forma similar a las que hemos denominado *funciones léxicas con verbos apoyo*. Sin embargo, en lugar de repetir el significado de la palabra clave como ocurría con los verbos de apoyo, aportan a la palabra clave el significado LLEVAR A CABO, CUMPLIR, O SATISFACER. Además, la fijación entre el verbo y la palabra clave es menor que en el caso de los verbos de apoyo. Por otra parte, el significado de la palabra clave normalmente hace referencia a la condición previa o requisito necesario para que se desarrolle la acción verbal..

Real_i verb meaning ‘to fulfil’ which takes the nth participants as subject and the keyword as object.
Real₂ (*hypothesis*) = [to] confirm ARTICLE [hypothesis]

Fact_i verb meaning ‘to fulfil’ which takes the keyword as subject and the nth participants as objects.
Fact₀ (*chemotherapy*) = [to] work [well]
Fact₀ (*hope*) = [to] come true

Labreal_{ij} verb meaning ‘to fulfil’ which takes the participants i, j as subject and object, and the keyword as secondary object.
Labreal₁₂ (*Navelbine*) = [to] combine *drug* (i.e. gemzar) with [navelbine]

d) CON VERBOS ASPECTUALES

Con esta categoría englobamos los verbos llamados *phrasal verbs* y *aspectual verbs* en el TST. Consideramos que el término *phrasal verb* no es apropiado porque induce a confusión con verbos con preposición y/o adverbio como *give in* o *look forward to* y que lo fundamental de estos verbos es que informan sobre el aspecto verbal. Por esto, y por el hecho de que aparecen en combinación con otras funciones léxicas verbales, apostamos por el término *funciones léxicas con verbo aspectual*.

Para aclarar el significado de cada una de estas FL aspectuales, vamos a utilizar la nomenclatura correspondiente al sistema aspectual español (Fente Gómez 1971) porque permite ordenar estas funciones de acuerdo con un

gradiente temporal que va desde que la acción está a punto de comenzar hasta que la acción ha terminado:

Prox	be on the verge of (aspecto ingresivo) ProxEpit (<i>tumor</i>) = precancerous
Incep	the beginning of (aspecto incoativo) IncepOper ₂ (<i>chemotherapy</i>) = to begin [chemotherapy of <i>drug</i>].
Cont	continue (aspecto durativo) ContOper ₂ (<i>attention</i>) = to hold [attention]
Perf	perfective (aspecto perfectivo) ⁹⁶ PerfFunc ₀ (<i>tumor</i>) = to be dead
Result + LF	the result of (aspecto resultativo) ResultOper ₂ (<i>lung carcinoma</i>) = to develop

Dentro de los tipos de repetición paradigmática, en concreto, dentro de la repetición asociativa verbal (p. 200), vimos que la función Result es mixta. Esta función se comporta sintagmáticamente en combinación con otra función léxica verbal. De esta forma, hemos encontrado una solución intermedia entre la propuesta de Mel'cuk (1996: 55), que la considera paradigmática, y la de Alonso Ramos y Tutin (1996: 155), quienes apuestan por una interpretación sintagmática de la misma.

e) CON VERBOS CAUSATIVOS

Con esta denominación se incluirán tres funciones que relacionan sustantivos y verbos o verbos entre sí, de forma que el valor de la función tiene un significado causativo.

Caus	cause CausFunc ₁ (<i>hope</i>) = to raise hope in someone
Perm	permit or allow Perm (<i>die</i>) = let

⁹⁶ A esta función Mel'cuk (1996: 64) la denomina **Fin** y la define como 'to cease'.

Liqu the elimination of
Liqu (*tumor*) = resect

A continuación hemos subdividido un conjunto de funciones que Mel'cuk (ibid: 69-72) agrupa aduciendo que aluden a una situación nueva relacionada con la palabra clave. Para tal subdivisión, hemos adoptado la propuesta de Alonso Ramos y Tutin (ibid: 155-156). Nos ha movido la gran diferencia que existe semánticamente entre estas funciones y el hecho de que se hayan incluido funciones aspectuales como **Prox** y **Stop**, similar a la función **Fin**.

f) CON VERBOS CUYO SIGNIFICADO ES NEGATIVO

Alonso Ramos y Tutin (ibid: 155) clasifican bajo una misma categoría una serie de verbos que indican un empeoramiento de la situación en que la se encuadra la palabra clave.

Degr degradation of (dimension can be specified: e.g. color, movement, etc.)
S₁Degrad (*chromosome* →) = chromosome aberrations

Obstr to function with difficulty
S₁Obstr (*bowel* →) = bowel obstruction

Excess excessive functioning of (dimension can be specified, as in Degrad)
Excess (*cancer cells*) = [to] proliferate

Nocer to harm, injure, or impair
NocerReal₁ (*blood supply*) = [to] cut ART [the blood support of the tumor]

g) CON VERBOS QUE INDICAN UNA SITUACIÓN NUEVA

Involv verb meaning non-participant involvement
Involv (*CT scan*) = [to] show the size, shape, and location of a tumor.

Manif is manifest in, with the entry as subject
Manif (*non-small cell lung cancer*) = [to] affect the lungs

Sympt to be a physical symptom of
CausExcess^{size}Sympt₂₁ (*lung cancer*) = [to] cause ART [N's] neck and face to be swollen

La función **Sympt**, a diferencia de **S_{Sympt}**, hace que un verbo y un sustantivo aparezcan en el mismo contexto. En cuanto a los subíndices que acompañan a estas dos funciones, hacen referencia a la parte del cuerpo implicada (1), el dueño de esa parte del cuerpo o el sujeto paciente de una sensación o estado físico (2) y la sensación o estado físico en sí (3).

h) CON VERBOS DE SIGNIFICADO MISCELÁNEO

Prepar to get ready for the appropriate use
PreparReal₁ (*surgical instruments*) = [to] sterilize

Son to emit a typical sound
A₁[ObstrSon (*breath*)] S_{Sympt} (*lung cancer*) = wheezy breathing

6.6.1.4. Repetición mixta de unidades fraseológicas

Dado que este tipo de repetición vincula dos unidades fraseológicas relacionadas semánticamente mediante hiponimia, oposición, sinonimia, etc., las FL implicadas son las mismas que utilizamos en el apartado de repetición paradigmática:

TABLA 15: Propuesta de funciones léxicas paradigmáticas

Aeq	Qual _{1,2,3}
V ₀	S _{1,2,3}
S ₀	A ₁ , A ₂
A ₀	S _{instr}
Adv ₀	S _{loc}
Syn	S _{med}
Figur	S _{mode}
Gener	S _{res}
Mult	S _{cause}
Equip	S _{effect}
Cap	S _{Sympt}
Sing	S _{Contain}
Culm	S _{Mater}
Centr	S _{Germ}
Germ	S _{Destin}
Anti	Effect
Conv _{ijk}	Result
Contr	

Mencionamos un ejemplo de este tipo de repetición, en el que se vinculan las UF *undifferentiated small cell lung cancer* con *small cell lung cancer*. La cohesión se produce al combinar la relación paradigmática mediante **Gener** con la relación sintagmática mediante **AntiMagn** y **Manif**. No obstante, como la UF tiene un alto grado de lexicalización, la relación sintagmática pasa bastante desapercibida.

(60) Gener (*undifferentiated small cell lung carcinoma*) = *small cell lung cancer*

AntiMagn (*cell*) = *small*

Manif (*small cell* → *carcinoma*) = *lung*

6.6.1.5. Repetición mixta de unidad fraseológica y lexema

Con esta categoría nos referimos a la utilización de una unidad fraseológica y de un lexema que está relacionado con la misma semánticamente mediante hiponimia, sinonimia, meronimia, acronimia, etc. y que se puede sistematizar con las FL del apartado anterior. En cierto modo, esta categoría es similar a la anterior, a excepción de que la relación no se establece entre dos unidades fraseológicas.

Dentro de la relación de sinonimia vamos a incluir la relación entre una unidad fraseológica y su acrónimo, que es un caso de repetición mixta cuando la relación paradigmática marcha al mismo nivel que la relación sintagmática, que es bastante obvia. En el ejemplo, la relación sintagmática se establece mediante S_{med} .

S_{med} typical means (syntagmatic)

S_{med} (→ *tomography*) = *positron emission*

Acro acronym (paradigmatic)

Acro (*Positron emission tomography*) = *PET*

Como conclusión, nuestro intento de estudiar la cohesión de un modo exhaustivo que contemple su complejidad semántica y textual nos ha llevado a proponer nuevas categorías de repetición. Entre estas, destacamos la repetición sintagmática y mixta, que dan cuenta de la importancia de la fraseología en los textos científicos.

Estas categorías se incluirán en matrices de repetición en las que la tipografía resulta bastante informativa. Se incluirán en los capítulos 7 y 10.

7. Esquemas de representación para el análisis intensivo de la cohesión

A la hora de plasmar gráficamente la cohesión, hemos necesitado más de una forma de representación dado que la extensión de los textos sobre cáncer y la exhaustividad de nuestro análisis dejan bastante limitada la capacidad de las *matrices de repetición*, presentadas en 4.5.

El primer tipo de representación utilizado en esta tesis es una adaptación de dichas matrices que permite un análisis detallado. Se aplicará esta representación al análisis de textos breves, ya que la segmentación en oraciones permite detectar empírica y sistemáticamente todos los recursos cohesivos desplegados en el texto. En efecto, los textos breves nos brindan la oportunidad de concebir la cohesión como un fenómeno *interoracional* que se amolda a algunas exigencias *intraoracionales*, como, por ejemplo, la de minimizar la repetición del mismo lexema⁹⁷ en la misma oración o en el cotexto cercano⁹⁸. Esta doble faceta interoracional e intraoracional de la cohesión permite tomar la oración como punto de referencia para analizar exhaustivamente textos breves. Este tipo de análisis *interoracional* susceptible de ser representado mediante matrices de repetición es el que hemos llamado *análisis intensivo*.

El otro modelo de representación del que nos vamos a servir se basa en el principio de *lexical chain* (Morris y Hirst 1991) y la aplicación del mismo al reconocimiento automático de la cohesión por parte de Ellman (véase 4.8.2.). Estimamos que este puede ser idóneo para el análisis de textos más extensos y de un corpus textual, como parece indicar el hecho de que este modelo haya sido aplicado al procesamiento automático de textos.

7.1. Modificaciones necesarias a las matrices de repetición de Hoey (1991)

La representación gráfica para el análisis intensivo de la cohesión es un desarrollo de la propuesta por Hoey y permite un reflejo detallado de las relaciones semánticas entre términos que crean cohesión. Para esto han sido necesarias la modificación de las categorías propuestas por Hoey, la inclusión de la repetición sintagmática y la descripción de las relaciones semánticas entre términos.

⁹⁷ Un lexema es el significado que se actualiza mediante diversas formas léxicas asociadas a la misma raíz.

A pesar de haber modificado las categorías de repetición de Hoey, hemos intentado mantener la idea de que la repetición es un elemento fundamental en la construcción de la cohesión, de ahí que hayamos utilizado también la palabra *repetición*. En la siguiente tabla, exponemos la posible correspondencia entre nuestra propuesta y la de Hoey. A pesar de definir las, Hoey (1991) excluye del análisis la repetición superordinada e hiponímica y la paráfrasis compleja porque considera que no contribuyen demasiado a la cohesión del texto.

TABLA 16: Modificación de las categorías de cohesión propuestas por Hoey (1991)

HOEY (1991)	NUESTRA PROPUESTA
Repetición léxica simple	Repetición exacta: <i>cancer – cancer</i> Repetición simple: <i>treat – treated</i>
Repetición léxica compleja	Repetición simple: <i>release (verb) – release (noun)</i> Repetición variada paradigmática sintáctica
Paráfrasis simple	Repetición paradigmática semántica (sinonímica)
Paráfrasis compleja por antonimia ⁹⁹	Repetición paradigmática opositiva (antonímica)
Paráfrasis compleja ¹⁰⁰	Repetición paradigmática asociativa
Correferencia	-----
Repetición superordinada e hiponímica	Repetición paradigmática semántica (hiponímica)
-----	Repetición paradigmática semántica (meronímica)
-----	Repetición paradigmática opositiva (conversión y complementariedad)
-----	Repetición paradigmática asociativa
-----	Repetición sintagmática (exacta, simple y variada)
-----	Repetición mixta de unidades fraseológicas
-----	Repetición mixta de unidad fraseológica y lexema

En esta tabla, las casillas superiores contienen los recursos que, según Hoey, aportan al texto un mayor grado de cohesión. Estos son más explícitos y más reconocibles de forma automática, por ejemplo, mediante listas de frecuencia lematizadas que indiquen los lexemas con mayor peso específico en los textos. Sin embargo, la presencia consistente de los demás recursos en redes semánticas puede resultar tan cohesiva como la repetición exacta de lexemas.

⁹⁸ Con esto queda patente que el nivel intraoracional sí tiene importancia en la cohesión.

⁹⁹ Esta categoría desaparece en Hoey (1995).

Por otra parte, creemos que cualquier análisis sistemático de la cohesión debe incluir la repetición sintagmática y la que hemos denominado repetición mixta. Pensamos que las matrices de repetición deben reflejar la cohesión entre oraciones vinculadas no sólo por lexemas sino también por unidades fraseológicas. De hecho, nuestro análisis pone de manifiesto que en el lenguaje científico, la cohesión se crea principalmente por la reiteración de unidades terminológicas construidas a partir de más de una palabra, es decir, por repetición sintagmática y mixta.

De ahí que, a diferencia de Hoey, propongamos incluir en las casillas de las matrices no sólo lexemas, sino también unidades fraseológicas, cuyos componentes están separados por guiones bajos (Cáncer_de_pulmón).

En cuanto a la representación de la cohesión mediante funciones léxicas, se identificarán las FL que más contribuyen a la cohesión. En algunos casos se verá cómo al determinar la función léxica que vincula dos unidades léxicas¹⁰¹, hay que distinguir entre si se está tomando un punto de vista lexicográfico-terminológico o textual. Desde la primera perspectiva, se considera el papel de la FL que relaciona dos lexemas en el sistema de la lengua, que normalmente va a ser el mismo que se actualiza en los textos concretos (perspectiva textual).

Por la complejidad que conlleva la identificación de relaciones sintagmáticas y de FL, ha sido conveniente elaborar para cada texto dos tablas con el formato de una matriz de repetición. La primera tabla representa las relaciones cohesivas más obvias en el texto y se denominará *matriz de repetición de las relaciones cohesivas primarias*. En la segunda, llamada *matriz de repetición de las relaciones cohesivas secundarias*, se identifican una serie de relaciones que, aunque más difusas, crean cohesión. Desdoblaremos la primera tabla en dos, una tabla A y una tabla B, con el objetivo de visualizar en papel el análisis detallado que se llevará a cabo en el capítulo 10. La tabla de relaciones más difusas sirve meramente para que el lector se haga una idea de cómo recursos cohesivos que no son tan efectivos ni obvios contribuyen también significativamente a crear cohesión, si se activan frecuentemente.

¹⁰⁰ El carácter más difuso de esta categoría hace que Hoey (1995) prescindiera de ella.

7.1.1. Matriz de repetición de relaciones cohesivas primarias (Tabla A)

En este apartado se perfila el formato de las casillas de una tabla de tipo A, extraída de la matriz de repetición del texto analizado en el capítulo 10. En concreto, presentamos las casillas (2, 1), (3, 1) y (3, 2), donde la casilla (2, 1) objetiva la cohesión de la segunda oración con respecto a la primera¹⁰². Dentro de cada casilla, hemos otorgado un número a cada instancia de cohesión para facilitar la referencia. Al lado de cada número aparece, o bien una unidad léxica aislada, en el caso de que ambas oraciones contengan exactamente el mismo lexema (riesgo), o bien un par de unidades léxicas separadas por una barra oblicua (*Cáncer_pulmonar/ carcinoma_pulmonar*). En este último caso, la cursiva señala aquellos elementos que, por no ser iguales, crean repetición léxica variada.

TABLA 17: Fragmento de TABLA tipo A

	(1)		
2	1. Alteraciones_del_cromosoma_9 / Alteraciones_en_el_cromosoma_9 2. Cáncer_de_pulmón/ cáncer_ pulmonar 3. Cáncer_de_pulmón/ Tumor	(2)	
3	4. Alteraciones_del_cr9/ Anomalías_del_cr9 5. Cáncer_de_pulmón/ Carcinoma_pulmonar	12. Alteraciones_en_el_cr9/ anomalías_del_cr9 13. Riesgo 14. Investigadores/ investigaciones 15. Cáncer_pulmonar / carcinoma_pulmonar 16. Marcadores / predictor 17. Tumor / carcinoma_pulmonar	(3)

Así pues, la casilla (2, 1) contiene tres instancias de cohesión léxica, es decir, indica que la segunda oración tiene tres vínculos cohesivos con la primera oración. El primer y el segundo caso son ejemplos de repetición variada sintáctica mixta en la que se produce también la repetición exacta, por una parte, de los lexemas *alteraciones* y *cromosoma*, y por otra parte, del lexema *cáncer*:

¹⁰¹ Nos estamos refiriendo a lo que en inglés se conoce como *word-type*, es decir, a “las variantes morfológicas de una misma entrada del léxico” (Márquez 1998: 85).

(61) *alteraciones del cromosoma 9 / alteraciones en el cromosoma 9*: REPETICIÓN SINTÁCTICA MIXTA (SINONÍMICA)

(62) *cáncer de pulmón / cáncer pulmonar*: REPETICIÓN SINTÁCTICA MIXTA (SINONÍMICA)

La repetición mixta aúna una relación cohesiva paradigmática (sinonimia) con una sintagmática entre una alteración y su localización o entre el cáncer y su localización. La relación sintagmática se consigue mediante las funciones léxicas **Degrad_s**, **Loc_{POSS}**, **Loc_{IN}**, y **Loc_A**.

Los términos relacionados en un par están separados por una barra oblicua, de forma que el primero de ellos pertenece a la oración que aparece antes en el texto. Así, por ejemplo, en el par *alteraciones del cromosoma 9 / alteraciones en el cromosoma 9* de la casilla (2, 1), el primer término corresponde a la primera oración y el segundo, a la segunda.

Nótese también cómo la cursiva indica repetición variada, en contraposición a la grafía normal, con la que se marcará la repetición exacta. Asimismo, con la representación tipográfica de estos lexemas en negro indicaremos que comparten una relación de sinonimia, entendiendo esta en el sentido laxo que defendimos en el apartado 6.6.1.2. El color azul, que aparece en el tercer caso de cohesión de la tabla, indica una relación de hiponimia entre los términos.

En definitiva, la tabla A presenta todas las instancias de cohesión del texto y, mediante la tipografía, aporta información básica, que se va a completar en la tabla B.

7.1.2. Matriz de repetición de relaciones cohesivas primarias (Tabla B)

La tabla B objetiva el tipo de repetición y las funciones léxicas que comparten las unidades léxicas y terminológicas identificadas en la tabla A. En esta, la casilla (2, 1) sería algo así:

¹⁰² Recordamos que Hoey consideraba la cohesión como una relación con oraciones precedentes. Por eso, la referencia a las casillas mediante coordenadas sigue este formato: (oración posterior, oración precedente). Esta presentación nos permitirá aludir a las casillas de matrices de repetición que no encajan en formato A4.

TABLA 18: Fragmento de TABLA tipo B

	(1)	
2	1. Mixta de UF (v sint) / Syn 2. Mixta de UF (v sint) / Syn 3. Mixta de UF y L (v léx) / Gener <i>Syn (cáncer de pulmón) = este tipo de tumor</i>	(2)

Esta casilla indica que los lexemas del primer par, se relacionan mediante repetición mixta y que son dos unidades fraseológicas. La abreviatura *v sint* nos dice que la repetición es variada y se fundamenta en diferencias sintácticas. Si miramos la información que la tabla A aporta sobre el ejemplo 1, “Alteraciones *del* cromosoma_9 / Alteraciones *en el* cromosoma_9”, la cursiva indica que la variación de los términos lo es con respecto a la preposición utilizada.

Después de la mención al tipo de repetición, la tabla de tipo B especifica la función léxica que vincula esas dos UF desde el punto de vista lexicográfico y textual, que en el ejemplo 1 es la función **Syn**. No obstante, a veces vamos a asociar a un mismo ejemplo dos funciones léxicas. Cuando esto ocurre, la primera función es la que comparten los lexemas desde el punto de vista lexicográfico o en una ontología, es decir, sin tener en cuenta el contexto. En cuanto a la segunda función, es la que se actualiza en el texto en función del cotexto que acompaña al término en cuestión. Este cotexto hace que el valor de la FL desde el punto de vista lexicográfico no coincida con el valor que adquiere en el texto. Una buena ilustración de esto lo constituye el tercer caso de cohesión que se presenta en la casilla (2,1)

(63)

(2, 1)

3. *Cáncer de pulmón / tumor*

Mixta de UF y L (v léx) / Gener

Syn \supset (*cáncer de pulmón*) = este tipo de tumor

En este ejemplo, el término genérico asociado a *cáncer de pulmón* (primera oración) es *tumor* (segunda oración), por lo que podemos decir que desde el punto de vista lexicográfico estos dos términos se relacionan mediante **Gener**. Sin embargo, el cotexto precedente de la palabra *tumor* (*este tipo de*) hace que en este texto *cáncer de pulmón* (palabra clave) y *este tipo de tumor* (valor) queden asociados mediante sinonimia (**Syn**). Es decir, en el texto, la función **Gener** se comporta como la función **Syn** e indica que estos dos términos tienen la misma referencia. Cabe destacar que la expresión que hace de palabra clave de la FL es la que aparece en la oración precedente.

Intuimos que tanto la lengua inglesa como la española se sirven de un repertorio de expresiones que, al acompañar a los términos, contribuyen a tales cambios de comportamiento de la FL¹⁰³. En consecuencia, siempre que ocurra esto haremos explícito el cotexto del valor de la función en la tabla B.

Una vez que se han identificado los tipos de repetición que se dan en un texto, las FL que se ponen en juego y las cadenas cohesivas que se establecen, es posible completar aún más el análisis representando los posibles *marcos conceptuales* que sustentan la producción y la comprensión de textos cohesivos. El hecho de que los marcos correspondientes a cada texto se perfilarán a partir de los datos lingüísticos extraídos de las matrices de repetición o las cadenas léxicas les confiere una gran validez y utilidad para la traducción. Esto es así porque presupone que el marco conceptual para el texto original debe ser muy similar al correspondiente al texto traducido.

Como conclusión, nuestro intento de estudiar la cohesión de un modo exhaustivo que contemple su complejidad semántica y textual, nos ha llevado a proponer nuevas categorías de repetición. Entre estas, destacamos la repetición sintagmática y mixta, que dan cuenta de la importancia de la fraseología en los textos científicos. Además de incluir estas nuevas categorías, hemos modificado las matrices de repetición de Hoey en el siguiente sentido. En primer lugar, nuestras matrices se desdoblaron en dos, una tabla A y una tabla B, para que así puedan especificar las funciones léxicas del texto. En segundo

lugar, hemos hecho que la tipografía de los lexemas de las casillas resulte informativa. Las matrices de repetición en un análisis intensivo son útiles para identificar las oraciones con más vínculos, que, según Hoey (1991), son las que contienen el contenido fundamental de un texto.

7.2. Esquemas de representación complementarios

La información que aportan las matrices de representación se ha extraído de las listas de frecuencia elaboradas por el programa informático de análisis léxico *Wordsmith Tools*. Este programa además de proporcionar una tabla con datos estadísticos para cada texto como, por ejemplo, la media de palabras por oración, nos va a ayudar a detectar qué lexemas son más frecuentes, qué unidades fraseológicas constituyen y qué distribución tienen determinadas palabras a lo largo del texto.

Las listas de frecuencia para cada texto analizado de un modo intensivo van a ser las siguientes: una lista con lexemas individuales y listas con las combinaciones de dos, tres y cuatro palabras del texto. A este último tipo de listas se les denomina en lingüística de corpus *n-grams*.

Presentamos tanto los datos estadísticos como las tablas de frecuencia de un texto sobre posibles mutaciones genéticas asociadas al cáncer de pulmón. Se trata del *abstract* de un artículo experimental publicado en la revista *Cancer* de John Wiley & Sons (véase capítulo 10). Los *abstracts* son el género de mayor circulación entre la comunidad de discurso médica y constituyen una unidad coherente. En consecuencia, creemos que es posible analizar los *abstracts* independientemente del resto del artículo experimental. En el capítulo 10, se analizará la adaptación al español en forma de noticia publicada por la revista *Jano Medicina y Humanidades*.

(64)

1 Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary Chromosome 9 Aberrations

2 #ABSTRACT# BACKGROUND. Deletion of chromosome 9p has been reported in numerous tumor types.

¹⁰³ Se presentará el repertorio correspondiente a la lengua inglesa en 13.1. Expresiones puente entre el lenguaje del especialista y el del lector lego.

3 The authors demonstrated in an earlier study that spontaneous chromosome aberrations on chromosome 9 in peripheral blood lymphocytes (PBLs) were a significant risk predictor for lung carcinoma.

4 METHODS. The current study evaluated the relationship between self-reported family history of cancer and spontaneous chromosome aberrations in the PBLs of 97 previously untreated lung carcinoma patients.

5 The authors' hypothesis was that individuals exhibiting specific chromosome aberrations might have inherited genetic instability and thus be more likely to report a family history of cancer.

6 For each individual, a personal interview was conducted to construct a detailed family history, and 100 metaphases from PBLs were analyzed for spontaneous aberrations by G-banding.

7 RESULTS. The patients reported having 829 first-degree relatives, including 74 (8.9%) with cancer.

8 A significantly elevated odds ratio (OR) of 2.7 was noted in smokers for having chromosome 9 aberrations and a first-degree relative with cancer.

9 When the family history of cancer was dichotomized into lung carcinoma or other cancers, the OR associated with chromosomal aberrations was 8.5 for lung carcinoma but only 2.3 for other cancers.

10 In addition to chromosome 9 aberrations, other spontaneous chromosome aberrations and family history of cancer were also evaluated, but no associations were found.

11 There were no associations between age, gender, ethnicity, or smoking status and the chromosome 9 aberration profile.

12 CONCLUSIONS. The findings of this study suggest that chromosome 9 aberrations may be a marker of cancer susceptibility and may be associated with familial aggregation of cancer.

(Cfr. Apéndice II, texto 7)

7.2.1. Datos estadísticos sobre el texto/ corpus

A la hora de estudiar la cohesión, puede ser de utilidad conocer la proporción de formas distintas (*type*) por cada 100 ocurrencias léxicas (*token*), es decir, la *type-token ratio*. Si se comparan corpórea amplios de igual extensión, una *type-token* elevada puede indicar un mayor número de formas distintas y una mayor complejidad conceptual. A efectos de cohesión, puede ser el indicio de una mayor repetición variada, lo que supone un esfuerzo cognitivo por parte del lector y un despliegue mayor de recursos cohesivos. De todas formas, este índice no es demasiado informativo para corpórea pequeños, en los que adquiere un valor muy elevado.

Otros datos de la tabla son el número de oraciones (*sentences*) y el número de palabras por oración (*sent. length*), que siempre es mayor para la lengua española en relación con la lengua inglesa. Un número elevado de letras en las palabra puede indicar la

presencia de términos médicos de origen grecolatino, ya que estos suelen ser más extensos que los vocablos anglosajones.

TABLA 19: Datos estadísticos sobre el *abstract* en inglés

Text File	CHROMOS.TXT	3-letter words	37
Bytes	1.919	4-letter words	34
Tokens	285	5-letter words	14
Types	122	6-letter words	22
Type/Token Ratio	42.81	7-letter words	21
Standardised Type/Token		8-letter words	18
Ave. Word Length	5.53	9-letter words	17
Sentences	15	10-letter words	22
Sent.length	19.00	11-letter words	21
sd. Sent. Length	10.62	12-letter words	7
Paragraphs	1	13-letter words	2
Para. length	285.00	14-letter words	1
1-letter words	31		
2-letter words	38		

7.2.2. Listas de frecuencia

A continuación presentamos listas de frecuencia con los lexemas individuales y las combinaciones de dos, tres y cuatro palabras que aparecen más de dos veces. En la primera columna (*N*), se nos indica el rango de cada lexema con respecto a su frecuencia.

Bajo *word* aparece la forma léxica más frecuente del lexema en cuestión. El resto de formas léxicas asociadas a esa misma raíz están en la casilla *lemmas*. De los dos sentidos del término *lema*¹⁰⁴, se optará por el más amplio, es decir, lema es el conjunto de formas derivadas de la misma raíz independientemente de su categoría gramatical. Y así, las formas *findings* y *found* estarían agrupadas bajo el mismo lema, que etiquetaremos con la forma más frecuente del texto/corpus en mayúscula: FINDINGS.

Las tablas también indican la frecuencia absoluta (*Fr.*) y relativa (%) de un determinado lema. Nótese cómo después de la vigésimo cuarta unidad léxica, hemos sumado las frecuencias relativas para calcular la representación de los lexemas que

¹⁰⁴ El sentido más restringido de *lema* es el conjunto de formas flexivas asociadas a un mismo lexema. Desde esta perspectiva, *water** sería el lema que agrupa las formas {*water, waters*} del lexema *water*. {*Water, watered, watering, waters*} están asociadas al lexema verbal *water** (Cfr. *Corpora List*, 3/11/1999).

participan en relaciones de repetición exacta, simple y variada sintáctica. Obviamente, son aquellas palabras plenas lematizadas cuya frecuencia relativa es superior a 1. La suma, 34,7%, es mucho menor que el porcentaje de repetición léxica simple y compleja (variada sintáctica en nuestra terminología) obtenido por Benbrahim (1996), que era del 92-97%, o Ellman (1998), entre un 60 y un 80%. La única explicación para esta divergencia es que estos autores han computado como repetición léxica, la repetición entre palabras forma (*function words*). Además, ellos hablan de número de conexiones y nosotros, de número de lexemas que participan en esas conexiones, un número a todas luces diferente. Creemos que nuestro porcentaje es bastante más objetivo porque no caben errores en el cómputo de conexiones y porque no hemos fijado la oración como medida de la cohesión.

TABLA 20: Frecuencia de lexemas aislados

N	Word	Frq.	%	Lemmas
1	Chromosome	12	4,21	chromosomal(1)
2	Aberations	11	3,86	aberration(1)
3	Cancer	11	3,86	cancers(2)
4	Family	7	2,46	familial(1)
5	History	6	2,11	
6	Carcinoma	5	1,75	
7	Lung	5	1,75	
8	Associated	4	1,40	associations(2)
9	Spontaneous	4	1,40	
10	Patients	3	1,05	
11	PBLs	3	1,05	
12	Reported	3	1,05	report(1)
13	Study	3	1,05	
14	Authors	2	0,70	
15	Evaluated	2	0,70	
16	Findings	2	0,70	found(1)
17	First-degree	2	0,70	
18	Having	2	0,70	
19	Individual	2	0,70	individuals(1)
20	Lymphocyte	2	0,70	lymphocytes(1)
21	May	2	0,70	
22	Relative	2	0,70	relatives(1)
23	Significant	2	0,70	significantly(1)
24	Smokers	2	0,70	smoking(1)
	Repet. exacta, simple y variada sintáctica	99	34,7	
25	Abstract	1	0,35	
26	Addition	1	0,35	
27	Age	1	0,35	
28	Aggregation	1	0,35	
29	Analyzed	1	0,35	
30	Background	1	0,35	
31	Blood	1	0,35	
32	Conclusions	1	0,35	
33	Conducted	1	0,35	
34	Construct	1	0,35	
35	Current	1	0,35	
36	Deletion	1	0,35	
37	Demonstrated	1	0,35	
38	Detailed	1	0,35	
39	Dichotomized	1	0,35	
40	Earlier	1	0,35	
41	Elevated	1	0,35	
42	Ethnicity	1	0,35	
43	Exhibiting	1	0,35	
44	G-banding	1	0,35	
45	Gender	1	0,35	
46	Genetic	1	0,35	
47	Hypothesis	1	0,35	
48	Including	1	0,35	
49	Inherited	1	0,35	
50	Instability	1	0,35	
51	Interview	1	0,35	
52	Likely	1	0,35	
53	Marker	1	0,35	
54	Metaphases	1	0,35	
55	Methods	1	0,35	
56	Might	1	0,35	
57	Noted	1	0,35	
58	Numerous	1	0,35	
59	Odds	1	0,35	
60	Peripheral	1	0,35	
61	Personal	1	0,35	
62	Predictor	1	0,35	
63	Previously	1	0,35	
64	Primary	1	0,35	
65	Profile	1	0,35	
66	Ratio	1	0,35	

67	Relationship	1	0,35	
68	Results	1	0,35	
69	Risk	1	0,35	
70	Self-reported	1	0,35	
71	Specific	1	0,35	
72	Status	1	0,35	

73	Suggest	1	0,35	
74	Susceptibility	1	0,35	
75	Tumor	1	0,35	
76	Types	1	0,35	
77	Untreated	1	0,35	

TABLA 21: Combinaciones de dos palabras

N	Word	Freq.	%
1	Of cancer	7	2,45
2	Chromosome #	6	2,10
3	Family history	6	2,10
4	History of	5	1,75
5	Lung carcinoma	5	1,75
6	Chromosome aberrations	4	1,40
7	Spontaneous chromosome	3	1,05
8	A family	2	0,70
9	Aberrations and	2	0,70
10	Associated with	2	0,70

11	Cancer and	2	0,70
12	Carcinoma patients	2	0,70
13	For lung	2	0,70
14	May be	2	0,70
15	No associations	2	0,70
16	Of #	2	0,70
17	Other cancers	2	0,70
18	PBLs were	2	0,70
19	The authors	2	0,70
20	With cancer	2	0,70

TABLA 22: Combinaciones de tres palabras

N	Word	Freq.	%
1	Family history of	5	1,75
2	History of cancer	5	1,75
3	Chromosome # aberrations	4	1,40
4	Spontaneous chromosome aberrations	3	1,05
5	A family history	2	0,70
6	For lung carcinoma	2	0,70
7	Lung carcinoma patients	2	0,70
8	Of cancer and	2	0,70

TABLA 23:
Combinaciones de cuatro palabras

N	Word	Freq.	%
1	Family history of cancer	5	1,75
2	A family history of	2	0,70
3	History of cancer and	2	0,70

7.2.3. Matriz de repetición

Aunque el texto sólo cuenta con 12 oraciones¹⁰⁵, resulta bastante difícil representar la matriz de repetición en soporte papel, como se constata al representar sólo la primera columna de la matriz de repetición (Tabla 24). En esta hay un boceto de las conexiones entre la primera oración y las demás.

Según esta matriz de repetición, el título (primera oración) es muy informativo, ya que está vinculado con todas las oraciones. La relación se establece con respecto a las grandes áreas conceptuales del eje horizontal (MÉTODO CIENTÍFICO, ENFERMEDAD,

¹⁰⁵ En la tabla estadística aparecían 15 oraciones pero hemos considerado los encabezados (*Abstract*, *Background*) como parte de la oración que preceden.

PACIENTE, FACTORES DE RIESGO, CÉLULA, INSTRUMENTO, CARACTERÍSTICAS, DIAGNÓSTICO, TRATAMIENTO SÍNTOMAS, GENÉTICA). Todas estas áreas configuran un marco que podríamos denominar BIOLOGÍA MOLECULAR DEL CÁNCER DE PULMÓN (CITOGENÉTICA).

[TABLA 24](#): Primera columna de la matriz de repetición del texto 6 (desarrollo de Hoey 1991)

En el capítulo 10, se verá cómo es posible exportar estos datos a diagramas de sectores. Con estos se consigue una forma de representación empírica de *marcos conceptuales*.

De todas formas, por la extensión del texto, es preciso representar la matriz de repetición en forma de pares de coordenadas. En la siguiente tabla, se presenta cómo la primera oración está vinculada con el resto de las oraciones mediante 35 vínculos cohesivos, y la segunda, mediante 19. En el apéndice VI se incluyen las matrices para el resto de oraciones, en las que se llega a identificar un total de 153 vínculos. Cada instancia de cohesión está representada por un par de unidades léxicas pertenecientes a dos oraciones distintas (*par cohesivo*) o por un solo par en el caso de la repetición exacta.

TABLA 25: Cohesión con respecto a la primera oración (tabla A)

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR		
(2, 1)	1. Lung carcinoma / tumor types 2. Chromosome 9 aberrations / Deletion of chromosome 9p	(7, 1)	17. Patients 18. Family / Relatives 19. Cancer
(3, 1)	3. Chromosome 9 aberrations / Spontaneous chromosome aberrations 4. Lymphocyte / Peripheral blood lymphocytes (PLBs) 5. Lung carcinoma	(8, 1)	20. Cancer/ smokers 21. Chromosome 9 aberrations 22. Family / relative 23. Cancer
(4, 1)	6. Family history of cancer 7. Chromosome 9 aberrations / Spontaneous chromosome aberrations 8. Lymphocyte / PLBs 9. Lung carcinoma patients	(9, 1)	24. Lung carcinoma (2) 25. Family history of cancer 26. Cancer (2) 27. Chromosome 9 aberrations / chromosomal aberrations
(5, 1)	10. Patients / individuals 11. Chromosome 9 aberrations / chromosome aberrations 12. Family history of cancer	(10, 1)	28. Chromosome 9 aberrations 29. Chromosome 9 aberrations / spontaneous chromosome aberrations 30. Family history of cancer
(6, 1)	13. Patients / individual 14. Family history of cancer / family history 15. Lymphocyte / PLBs 16. Chromosome 9 aberrations / spontaneous aberrations	(11, 1)	31. Cancer / smoking 32. Chromosome 9 aberrations / chromosome 9 aberration
		(12, 1)	33. Chromosome 9 aberrations 34. Cancer / smoking 35. Family history of cancer / familial aggregation of cancer

TABLA 26: Cohesión con respecto a la segunda oración (tabla A)

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 2)	Véase (2, 1)
(3, 2)	36. Deletion of chromosome 9p / spontaneous chromosome aberrations 37. Tumor types /lung carcinoma
(4, 2)	38. Deletion of chromosome 9p / spontaneous chromosome aberrations 39. reported / self-reported 40. Tumor types/lung carcinoma
(5, 2)	41. Deletion of chromosome 9p / chromosome aberrations 42. report / reported 43. Tumor types/ lung carcinoma
(6, 2)	44. Deletion of chromosome 9p / spontaneous aberrations

(7, 2)	45. Tumor types / cancer
(8, 2)	46. Deletion of chromosome 9p / chromosome 9 aberrations 47. Tumor types / cancer
(9, 2)	48. Tumor types / cancer 49. Deletion of chromosome 9p / chromosomal aberrations
(10, 2)	50. Deletion of chromosome 9p / chromosome 9 aberrations 51. tumor types / cancer
(11, 2)	52. Deletion of chromosome 9p / chromosome 9 aberration (profile)
(12, 2)	53. Deletion of chromosome 9p / chromosome 9 aberrations 54. tumor types / cancer

Por último, para poder computar todos estos enlaces, proponemos dos tablas en las que aparecen las conexiones en las que participa cada oración y el número de oraciones con las que cada oración está vinculada.

TABLA 27: Cómputo de las conexiones en las que participa cada oración

Oración precedente con la que está vinculada (Nº de conexiones)											Nº Orac	Oración posterior con la que está vinculada (Nº de conexiones)
1	2	3	4	5	6	7	8	9	10	11	1	
2											2	
3	2										3	
4	3	3									4	
3	3	3	4								5	
4	1	1	3	3							6	
3	1	1	2	3	2						7	
4	2	2	2	2	2	1					8	
4	2	2	2	2	2	1	3				9	
3	2	2	3	2	2	1	2	3			10	
2	1	2	2	2	1	1	2	3	3		11	

3	2	2	3	2	2	1	2	4	3	3	12
---	---	---	---	---	---	---	---	---	---	---	----

TABLA 28: Tabla de las conexiones en las que participa cada oración

Oración	Cantidad de oraciones con las que se vincula (precedentes, posteriores)
1	(0, 11)
2	(1, 10)
3	(2, 9)
4	(3, 8)
5	(4, 7)
6	(5, 6)
7	(6, 5)
8	(7, 4)
9	(8, 3)
10	(9, 2)
11	(10, 1)
12	(11, 0)

En vista de la tabla, se infiere que el texto está muy cohesionado ya que cada una de las oraciones está vinculada con once oraciones. Por su carácter sintético, su concisión y su brevedad, los *abstracts* son géneros con un alto grado de cohesión. Las oraciones que incluye son aquellas que contienen la información principal y las palabras clave del texto. Como ya se dijo, estas son las que tienen mayores vínculos cohesivos.

8. Esquemas de representación para el análisis extensivo de la cohesión

El hecho de que la mayoría de los textos tengan una longitud incompatible con el análisis intensivo o con la metodología de la lingüística de corpus nos ha impulsado a diseñar una dinámica de trabajo diferente para el *análisis extensivo de la cohesión*. Se contará con el soporte de cuatro elementos:

- (1) El programa de análisis léxico *Wordsmith Tools*
- (2) El generador de cadenas léxicas *Hesperus* (véase pp. 134-138)
- (3) La hoja de cálculo *Microsoft Excel*
- (4) El análisis del investigador, que aporta su conocimiento de la lengua y lo complementa con la información que proporciona el experto en oncología.

8.1. Descripción de los resultados obtenidos con el programa informático *Hesperus*

Para cada texto que sirve de *input* a *Hesperus*, este programa confecciona un conjunto de documentos en HTML con los siguientes elementos:

- a) perspectiva general (*Overview*)
- b) perfil genérico del documento (*Generic document profile*)
- c) versión del texto completo donde se destacan las distintas cadenas léxicas reconocidas
- d) representación de cada cadena con todos los lexemas que la componen y el tipo de relación entre ellos (*Chains*).

Para explicar en qué consiste cada uno, hemos seleccionado el *output* a partir de un texto para profesionales de la salud sobre el cáncer no microcítico de pulmón (QDT2) (Cfr. Apéndice II, texto 10).

8.1.1. Perspectiva general

En la *perspectiva general*, aparecen todos los textos analizados y se les atribuye un número con dos decimales que mide la similitud de los mismos con respecto al texto

que *Hesperus* ha seleccionado como prototipo, el texto QDT1 (Cfr. Apéndice II, texto9). Para los textos QDT1 y QDT2, aparecería en pantalla lo siguiente¹⁰⁶:

TABLA 29: Perspectiva general correspondiente a los textos QDT1 Y QDT2.

Case: 35
[QDT1.txt](#) 1.00 (505 words, 10536 total)
[\(Chains,Profile\)](#)

Case: 34
[QDT2.txt](#) 0.20 (2470 words, 46217 total)
[\(Chains,Profile\)](#)

El número 0.20 que sigue a [QDT2.txt](#) indica el grado de similitud de ese texto con respecto a QDT1. A este texto prototípico, se le atribuye un valor de 1.00, ya que es similar a sí mismo. El número 2470 representa el número de palabras plenas del texto, porque, como vimos en la p. 136, *Hesperus* elimina en el análisis las *palabras forma*, ya que no son susceptibles de crear cohesión léxica.

En el mismo paréntesis, después del 2470, un número (*raw score*) representa la suma de todos los valores que *Hesperus* otorga a las categorías conceptuales del *Roget* activadas. Para visualizarlas sólo hay que hacer click en el perfil del documento (*profile*). Este valor no indica por sí sólo la calidad cohesiva del texto dado que, a mayor longitud del texto, se obtiene una puntuación mayor. Por este motivo, se relacionará este valor con el número total de palabras. En el perfil del documento aparece desglosado este valor.

Desde el perfil también podemos visualizar la versión del texto ([QDT1.txt](#), [QDT2.txt](#)) que contiene hipervínculos hacia las cadenas léxicas ([Chains](#)).

8.1.2. Perfil del documento

El *perfil del documento* incluye las categorías conceptuales del tesoro que se activan gracias a la cohesión que construyen las palabras plenas del texto. Como este perfil es demasiado extenso, fragmentamos el perfil completo y sólo incluimos los conceptos

¹⁰⁶ La perspectiva general completa aparece en el Apéndice VIIa.

más representativos y los menos representativos. Editamos el perfil completo en el apéndice VIIIb.

TABLA 30: Fragmento del perfil del texto QDT2.

NON-SMALL CELL LUNG CANCER

Document Profile: No file name (default), 196 Features, Total 46217 (0)

therapy_658_4268_n	Percent: 21.07,	Value: 9737
sick-person_651_4202_n	Percent: 7.02,	Value: 3246
show_522_3381_v	Percent: 5.69,	Value: 2628
minuteness_196_1271_n	Percent: 3.46,	Value: 1600
surgery_658_4267_n	Percent: 3.02,	Value: 1396
medical_658_4274_a	Percent: 2.72,	Value: 1255
experiment_461_2959_n	Percent: 2.64,	Value: 1220
cancer_651_4195_n	Percent: 2.56,	Value: 1183
respiratory-disease_651_4192_n	Percent: 1.90,	Value: 880

[...]

modality_7_44_n	Percent: 0.11,	Value: 50
lasting_113_699_a	Percent: 0.11,	Value: 50
cardiovascular-disease_651_4193_n	Percent: 0.11,	Value: 50
unprosperous_731_4742_a	Percent: 0.11,	Value: 50

Total Percent: 97.47

Antes del listado de conceptos, se especifica la cantidad de categorías conceptuales que subyacen las 2470 palabras plenas del texto. A esto hace alusión la referencia *196 Features*. No obstante, de estas 196 categorías, sólo aparecen en el perfil aquellas que contribuyen más de un 0,11 % a la cohesión del texto. Le sigue el ya mencionado *raw score* del documento (46217).

A continuación, cada categoría lleva una referencia a la jerarquía del *Roget's Thesaurus* (_658_4268) y a la categoría gramatical (nombre, verbo, adjetivo o adverbio) asociada a la misma. Le sigue un porcentaje que se obtiene de dividir el *raw score* del texto entre el valor asignado a esa categoría conceptual. Por ejemplo, el concepto principal del texto, THERAPY, pertenece a la categoría 658 (IMPROVEMENT), se actualiza mediante un sustantivo y representa un 21,07% de los conceptos activados en el texto. Confirmaremos este porcentaje cuando veamos que el concepto TREATMENT es más relevante en el texto analizado.

8.1.3. Versión del texto con hipervínculos hacia las cadenas léxicas

Hesperus también confecciona una versión del texto en la que las palabras pertenecientes a la misma cadena léxica aparecen con el mismo color (véase Apéndice VIIc). Presentamos el primer párrafo del texto en cuestión, donde las palabras que introducen cinco cadenas léxicas están subrayadas, en color azul y llevan un superíndice: *small*, *lung-cancer*, *cure*, *produce* y *partial*. Estos lexemas inician las cadenas que se hilan, respectivamente, a partir de los conceptos: MINUTENESS, PATHOLOGY OF LUNG CANCER, TREATMENT, CAUSE y COMPOSITION.

TABLA 31: Versión HTML del texto QDT2.

```
----- NON- small3 cell lung-cancer1 -----
***** GENERAL INFORMATION
Non- small cell lung-cancer (NSCLC) is a heterogeneous aggregate of at
least three distinct histologies of lung-cancer including epidermoid or
squamous carcinoma, adenocarcinoma, and large cell carcinoma. These
histologies are often classified together because, when localized, all have
the potential for cure0 with surgical resection. Systemic chemotherapy can
produce2 objective partial13 responses and palliation of symptoms for short
durations. Local control can be achieved with radiation in a large number
of patients with unresectable disease, but cure is seen only in a small
minority of patients.
```

Como vemos en el texto, el primer elemento de cada cadena lleva un superíndice con un número cuyo color indica que todas las palabras con tipografía en ese color van a encajar en esa cadena. Y así, en la primera línea, el término *lung cancer* va acompañado de un número en verde azulado. Todos los lexemas del texto con este color [*lung cancer* (2), *carcinoma* (2), *symptoms*, *patients* (2) y *disease*] están relacionados con *lung cancer* (primera línea) y forman una cadena, a la que se puede acceder mediante un vínculo de hipertexto. Según la tabla 31, pertenecen a esta cadena los lexemas *lung cancer*, que aparece en tres ocasiones, *carcinoma* y *patients*, que se repiten en dos ocasiones, *symptoms* y *disease*.

La tipografía de los lexemas “encadenados” es informativa del tipo de relación semántica que se da, de acuerdo con las siguientes convenciones:

- **identity** (negrita): relación entre formas idénticas o entre sustantivos que cambian sólo respecto al número. Ejemplo: *patient — patients*.
- category (subrayado): relación entre lexemas que pertenecen a una misma categoría del tesoro *Roget*. Ejemplo: *diagnosis – prognosis*, asociados a la categoría pathology 651 4203 n.
- *group* (cursiva): relación entre lexemas pertenecientes a categorías vecinas dentro de un mismo *grupo de categorías* del tesoro. Ejemplo: *lung cancer – carcinoma*, relacionados respectivamente a las categorías respiratory- disease 651 4192 n y cancer 651 4195 n. Ambas están incluidas en el *grupo* disease 651 n.

Según esta tipografía, en el fragmento de cadena que comentamos hay seis relaciones de identidad (ID), cuatro de grupo (GRP) y tres categoriales (CAT). Según Ellman (1998), el 60-80% de las cadenas léxicas se deben a la relación ID, que se basa en la repetición léxica.

8.1.4. Representación de cadenas léxicas (*Chains*)

El primer lexema de cada cadena nos remite a una página HTML que representa las cadenas léxicas individuales. Esta *representación* especifica las palabras en esa cadena, el tipo de vínculo léxico entre ellas (ID, CAT, GRP, ONE) y las categorías del tesoro mediante las que se establecen los vínculos. Si nos situamos con el ratón encima de [lung-cancer](#)¹, que inicia la cadena 1, una de las 14 cadenas del texto, encontraríamos 276 lexemas correspondientes a 23 formas léxicas distintas (véase Apéndice VIId). Presentamos una sección de esta cadena:

TABLA 32: Fragmento de una de las 14 cadenas léxicas del texto QDT2.

Chain 1

lung-cancer, lung-cancer, lung-cancer, carcinoma, carcinoma, symptoms, patients, disease, patients, diagnosis, patients, disease, patients, prognosis, patients, disease, patients, lung-cancer, patients, disease, patients, diagnosis, symptoms, patients, patients, confined, symptoms, patients, prognosis, symptoms, patients, lung-cancer, patients, patients, patients, patient, lung-cancer, lung-cancer, critical, cases, lung-cancer, carcinoma, cancers, cancers, carcinoma, cancer, lesion, cancer, patients, cancer, patients, lung-

cancer, lung-cancer, carcinoma, carcinoma, carcinoma, carcinoma, critical, patients, critical, disease, patients, patients, patients, lung-cancer, lung-cancer, cancer, cancer, patient, lesion, lesions, lesions, lesion, cancer, bronchial, carcinoma, bronchial, lung-cancer, patients, cases, patient, carcinoma, lung-cancer, cancers, patients, disease, patients, disease, disease, cancers, lung-cancer, lung-cancer, lung-cancer, patient, lung-cancer, patients, disease, lung-cancer, lung-cancer, carcinoma, cancers, patients, patients, cancers, lung-cancer, patients, lung-cancer, patient, condition, patient, critical, patients, lung-cancer, patients, cancer, patients, patients, patients, patients, patients, inoperable, patients, disease, patients, inoperable, patients, inoperable, patients, patients, critical, patients, patients, smokers, lung-cancer, patients, cancers, cancers, cancers, cancers, patients, cancers, cancers, patients, patients, patients, lung-cancer, patients, lung-cancer, patient, condition, patient, critical, inoperable, patients, disease, patients, patients, inoperable, patients, disease, disease, critical, patients, patients, patients, carcinoma, patients, disease, patients, lung-cancer, patients, lung-cancer, patients, patients, patients, patients, disease, patient, patients, disease, patients, patients, patients, patients, patients, cancer, patients, patients, patients, lung-cancer, patients, patients, disease, cases, disease, cases, patients, patients, cases, patients, lung-cancer, patients, lung-cancer, patients, patients, patients, disease, patients, patient, patients, disease, cough, chest-pain, patient, patients, patients, disease, patients, inoperable, patients, patient, patients, disease, patients, lung-cancer, toxic, patients, patients, patients, lung-cancer, lung-cancer, patients, inoperable, disease, lung-cancer, patients, patients, lesions, bronchial, pain, cases, lesions, critical, patient, patient, lesion, toxic, patients, lesions, lung-cancer, patients, lung-cancer, patients, lesion, disease, toxic, patients, patients, patients, patients, patients, patients, carcinoma, patients, cancers, lesion, cancer, patients, lesion, patients, patients, disease, patients, patients, lesions, patients,

:lung-cancer :Wd 4, Sent 1, Para 1, ???-> 0, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 11, Sent 1, Para 3, ID-> 4, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 25, Sent 1, Para 3, ID-> 11, Value 0
[respiratory-disease 651 4192 n](#)
:carcinoma :Wd 31, Sent 1, Para 3, GRP-> 25, Value 0
[cancer 651 4195 n](#)
:carcinoma :Wd 36, Sent 1, Para 3, ID-> 31, Value 0
[cancer 651 4195 n](#)
:symptoms :Wd 65, Sent 1, Para 3, GRP-> 36, Value 0
[illness 651 4186 n](#)
:patients :Wd 81, Sent 1, Para 3, GRP-> 65, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 84, Sent 1, Para 3, GRP-> 81, Value 0
[disease 651 4187 n](#)
:patients :Wd 95, Sent 1, Para 4, ID-> 81, Value 0
[sick-person 651 4202 n](#)
:diagnosis :Wd 97, Sent 1, Para 4, GRP-> 95, Value 0
[pathology 651 4203 n](#)
:patients :Wd 98, Sent 1, Para 4, ID-> 95, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 112, Sent 1, Para 4, ID-> 84, Value 0
[disease 651 4187 n](#)
:patients :Wd 120, Sent 1, Para 4, ID-> 98, Value 0
[sick-person 651 4202 n](#)
:prognosis :Wd 139, Sent 1, Para 4, CAT-> 97, Value 0
[pathology 651 4203 n](#)

Se especifica con qué lexema previo está vinculado cada elemento de la cadena, que es el precedido por la flecha (->). Por ejemplo, la undécima palabra del texto (Wd 11), *lung-cancer*, está vinculada mediante la relación de identidad con la cuarta palabra, *lung cancer* (ID-> 4). Las referencias al número de oración y de párrafo no son exactas, un aspecto que el autor del programa piensa modificar¹⁰⁷. Después de esto, encontramos la entrada del tesoro bajo la que aparece, acompañada del número que se le atribuye en el tesoro y de una letra (n) que hace mención a la categoría gramatical: [respiratory-disease 651 4192 n.](#)

Por último, aclararemos que las cadenas identificadas representan el 70 % de los conceptos presentes en el *perfil*. Es decir, además de estas cadenas, habría otras que o bien son menos importantes o bien se establecen entre lexemas separados por más de 500 palabras. Estas contribuyen al restante 30 % de la cohesión del texto.

8.2. Limitaciones de los resultados aportados por *Hesperus*

El *output* de *Hesperus* no es suficiente para plasmar las relaciones cohesivas del texto. Como acabamos de ver, no se explicitan todas las cadenas y no se computan las relaciones entre lexemas separados por más de 500 palabras. Además, a diferencia del lector humano, un ordenador no detecta la interacción entre la maquetación del texto, por una parte, y el género textual, la deixis, la estructura discursiva y la estructura informativa (*tema - rema; información principal - información secundaria*) del texto, por otra. Esto se constata cuando el ordenador pasa por alto que las palabras en negrita son más importantes desde el punto de vista cohesivo que las que no lo están. Si esta interacción ayuda a inferir más rápidamente las palabras clave sobre las que gira la cohesión, podemos decir que el ordenador carece de esta fuente de información sobre la cohesión del texto.

En relación con el tesoro, nos gustaría señalar que, por ser de la lengua general, no contiene todos los términos propios de la oncología, muchos de los cuales son UF que *Hesperus*, salvo en contadas ocasiones, no tiene la capacidad de reconocer. Y así,

¹⁰⁷ Comunicación personal.

vemos cómo se identifica la UF *lung cancer* pero no otras muy frecuentes en oncología como, por ejemplo, *small cell lung cancer* o *magnetic resonance imaging*.

Por otra parte, los fenómenos de polisemia, homonimia, ambigüedad y connotación presentes en el lenguaje dificultan la búsqueda de casos de cohesión por parte de un ordenador. Aunque en el lenguaje científico estos fenómenos se producen con menos frecuencia, seguimos con la limitación de utilizar un glosario de la lengua general donde los términos no están estructurados de acuerdo con el conocimiento sobre oncología. En muchas ocasiones *Hesperus* no capta el contexto y el significado que adquieren estos términos en los textos sobre cáncer. Por ejemplo, en la cadena 13 del texto QDT1, *nodes* queda asociado al concepto COMPOSICIÓN, que es la idea imprecisa que subyace a las distintas acepciones del término *node*, perteneciente a diferentes campos del saber: botánica, anatomía, astronomía, física, etc. En textos sobre oncología, designa unos componentes del sistema linfático, los ganglios linfáticos, a los que emigran con bastante facilidad las células cancerosas de un tumor cercano.

node n.

- 1 Bot. a the part of a plant stem from which one or more leaves emerge. b a knob on a root or branch.
- 2 Anat. a natural swelling or bulge in an organ or part of the body.
- 3 Astron. either of two points at which a planet's orbit intersects the plane of the ecliptic or the celestial equator.
- 4 Physics a point of minimum disturbance in a standing wave system.
- 5 Electr. a point of zero current or voltage.
- 6 Math. a a point at which a curve intersects itself. b a vertex in a graph.
- 7 a component in a computer network.

(*Concise Oxford Dictionary*)

Tampoco percibe los solapamientos entre cadenas que se derivan de la realidad extralingüística, que en el ámbito de la ciencia está en continuo cambio. *Hesperus* no señala unidades léxicas susceptibles de aparecer en dos cadenas diferentes. Sin embargo, este fenómeno de doble pertenencia es frecuente dado que las palabras no tienen límites perfectamente definidos. De hecho, cuando para cada texto hagamos nuestra propuesta definitiva de cadenas léxicas, marcaremos con vínculos de hipertexto aquellas palabras que pertenecen a más de una cadena. Por ejemplo, el término *clinical trials* pertenece al mismo tiempo a las cadenas léxicas *TREATMENT* y *RESEARCH*.

Para que un programa informático detecte la cohesión en un texto especializado de un modo eficaz, sería preciso adaptar su tesoro al subdominio al que pertenece y actualizarlo cada cierto tiempo. Los diferentes sentidos de los lexemas de la lengua general quedarían restringidos a los sentidos que adquieren en ese campo del saber. Además, en el caso de lexemas polisémicos, el concepto asociado se determinaría en función de las palabras de su radio colocacional. Esta tarea de desambiguación, que nos hará modificar las cadenas léxicas, requiere la ayuda de un programa de análisis léxico que elabore concordancias y listas de frecuencia. En estas últimas, queda patente la repetición de la misma forma léxica, por lo cual, reflejan muy bien el papel de la repetición en la creación de la cohesión. Se empleará para este fin el programa *Wordsmith Tools*.

Para ilustrar estas limitaciones, mencionaremos un ejemplo tomado de una de las cadenas reconocidas en el texto QDT2, en la que predomina el concepto MINUTENESS. En ella se asocian lexemas que, si bien comparten un sema que indica pequeñez, en el campo de la oncología, su significado básico no gira en torno de este sema. Este es el caso de los lexemas en cursiva, cuyo significado básico queda modificado por su cotexto, de forma que se encuadrarán en las áreas entre paréntesis:

Fractionation (RADIOTHERAPY)

Bronchial *compression* (SYMPTOMS)

Microscopic examination (DIAGNOSIS)

En cuanto al formato de presentación de los resultados de *Hesperus*, opinamos que su exhaustividad plantea una dificultad a la hora de plasmarlos en soporte papel. Por ello, los adaptaremos a nuevos formatos en 8.3.

Por todo esto concluimos que, si bien el programa ha sido de gran utilidad, hemos encontrado algunos puntos débiles, que intentaremos subsanar con una presentación de los resultados más informativa y visual para el lector y modificando las cadenas léxicas identificadas de acuerdo con las áreas conceptuales de la oncología y con los datos aportados por *Wordsmith Tools*.

8.3. Etapas para el análisis extensivo de la cohesión

Teniendo en cuenta las limitaciones expuestas, proponemos un análisis de la cohesión que destaque los factores que intervienen en la creación de la misma: la textualidad, las relaciones conceptuales y la repetición. A partir de estas, sugerimos cinco etapas que aplicaremos al análisis cohesivo de seis textos sobre el tratamiento del cáncer de pulmón (véase capítulo 11).

- a. Índice del texto
- b. Áreas conceptuales reconocidas a partir de listas de frecuencia lematizadas
- c. Representación modificada de los resultados de *Hesperus*
- d. Propuesta definitiva de cadenas léxicas
- e. Cuadro contrastivo de las diferencias entre la cohesión detectada por un ordenador y por un humano

Los resultados de *Hesperus* se volcarán en la tercera etapa, aunque adaptados al formato que hemos propuesto en aras de la claridad y la concisión. La cuarta etapa integra los resultados de las que la preceden.

8.3.1. Índice del texto

Todos los textos analizados comienzan con un índice con hipervínculos hacia las diferentes secciones en que se divide cada uno de ellos. Como hemos visto, existe una relación entre la cohesión, por una parte, y la estructura y segmentación de textos, por otra (Berber Sardinha 1997). Por eso, sospechamos que las secciones que aparecen en el índice de un texto apuntan hacia las grandes áreas conceptuales explotadas en el mismo y materializadas en cadenas léxicas. En definitiva, al observar el índice, nos familiarizamos con la faceta textual de la cohesión.

8.3.2. Áreas conceptuales reconocidas a partir de listas de frecuencia lematizadas

Cuando se somete una lista de frecuencia creada automáticamente a lematización se ve un claro predominio de determinados lexemas. Intuimos que estos son los que sirven de hitos en el campo conceptual de la oncología y en sus posibles marcos. En

consecuencia, para cada texto elaboraremos una lista con las áreas conceptuales bajo las que se agrupan las unidades léxicas y terminológicas con una frecuencia relativa mayor o igual al 0,1%. Dado que no incluye todas las palabras y términos del texto, esta lista será provisional. Para cada texto, la modificaremos en las secciones *Propuesta definitiva de cadenas léxicas* del capítulo 11.

Para identificar las grandes áreas conceptuales de cada texto, se exportarán los datos de las listas de frecuencia creadas con *Wordsmith Tools* a la hoja de cálculo *Microsoft Excel*, de forma que se ubiquen en la misma columna los lexemas relacionados semánticamente. Por ejemplo, en el texto QDT2, los lemas que denominan alguna parte del cuerpo tanto a nivel anatómico como histológico, aparecerían en la misma columna. Su porcentaje con respecto a las demás áreas conceptuales indica de un modo preliminar la importancia de los conceptos relacionados con alguna PARTE DEL CUERPO.

TABLA 33: Activación de términos que designan partes del cuerpo con frecuencia mayor al 0,1% en QDT2 (*Microsoft Excel*).

LEMA	%
LUNG	1,12
CELL	0,79
CHEST	0,25
BRONCHUS	0,23
BRAIN	0,17
LYMPH NODES	0,16
PLEURA	0,14
LOBE	0,11
TOTAL	2,97

La suma de todos los porcentajes de la columna nos indicaría aproximadamente la representación del área conceptual PARTS OF THE BODY en ese texto (2,97%). Un diagrama de sectores con los porcentajes atribuidos inicialmente a cada área conceptual ilustrará los marcos conceptuales preliminares del texto (Véase, por ejemplo, 11.1.2.2. *Áreas conceptuales reconocidas a partir de listas de frecuencia lematizadas*).

8.3.3. Representación modificada de los resultados de *Hesperus*

Se aprehenderán buena parte de las relaciones conceptuales que contribuyen a la cohesión con la ayuda de la *perspectiva general*, el *perfil* y la *representación de las cadenas léxicas* generadas por *Hesperus*.

En primer lugar, a partir de la *perspectiva general* y el *perfil*, describiremos cuantitativamente la *calidad cohesiva* del texto. Sabremos cuántas palabras plenas hay en el texto y cuántas categorías del tesoro *Roget*, y atribuiremos al texto un coeficiente de cohesión¹⁰⁸ dividiendo la puntuación bruta (*raw score*) entre el número total de palabras del texto. Asimismo, especificaremos las categorías del tesoro que representan más del 1% de la cohesión del texto. Muchas de estas categorías habrán ya sido identificadas en la segunda etapa, *Áreas conceptuales reconocidas a partir de listas de frecuencia lematizadas*. Estas categorías son el nexo de unión entre los distintos lexemas que constituyen las cadenas identificadas por *Hesperus*.

Seguidamente, analizaremos cada una de estas cadenas con el fin de confeccionar tablas donde se indica la frecuencia absoluta de las formas léxicas vinculadas. Las tablas tienen la siguiente presentación (véase Apéndice VII e)¹⁰⁹:

TABLA 34: Representación tabular de las cadenas reconocidas por *Hesperus* (Cadena 4 del texto QDT2)

NODES	7
NODE	4
NODULES	3
BONE	1

CHEST	1
NODULE	1
OSSEOUS	1

¹⁰⁸ Como se ha explicado con anterioridad, este valor es dependiente de la extensión del texto, por lo que para no distorsionar la información que proporciona sobre la fuerza de las cadenas y su extensión, siguiendo el consejo del profesor Ellman, lo hemos dividido entre el número total de palabras.

¹⁰⁹ La unidad léxica en negrita es la que inicia la cadena.

Asignaremos a cada cadena una etiqueta que refleja el contenido semántico que comparten los lexemas que la componen. Para ello, se examinarán las categorías conceptuales de cada cadena (por ejemplo, [respiratory-disease 651 4192 n](#)) y las listas de frecuencia lematizadas. Con estos datos y guiados por nuestra intuición, hemos etiquetado la mayoría de las cadenas a excepción de algunas que incluyen palabras relacionadas semánticamente de una forma indirecta. En estas cadenas dudosas, las relaciones de *identidad* son menos frecuentes que las de *categoría* o las de *grupo*, al contrario de lo que ocurre en las cadenas que hemos podido etiquetar, en las que predomina la relación de *identidad* o *categoría*. Además, vinculan normalmente lexemas que pertenecen al metadiscurso científico. A esas cadenas las denominaremos *cadena difusas*, las etiquetaremos con el signo '?' y para cada uno de sus elementos, explicitaremos la categoría del tesoro que justifica su inclusión en esa cadena. Intentaremos subsanar las inevitables incoherencias y las agrupaciones erróneas de lexemas derivadas de la ambigüedad y la polisemia.

Comentaremos estas cadenas, aunque nos centraremos, más que en los lexemas, en los conceptos activados y su porcentaje de activación. Para representarlos, se elaborarán tablas similares a esta:

TABLA 35: Representación de los conceptos identificados por *Hesperus* en las cadenas 4, 7 y 11 de QDT2: porcentaje de activación.

CONCEPTOS ACTIVADOS					
Cadena 4 (<i>Nodules</i>)		Cadena 7 (<i>Lymph Nodes</i>)		Cadena 11 (<i>Pleura</i>)	
SWELLING	1,56	BLOOD	0,68	LATERAL (Pleura)	0,76
HARD	0,13	SANGUINEOUS	0,22	LATERALITY	0,1
BOSOM	0,1	FLUID	0,16		
HARDNESS	0,1				
TOTAL	+1,69	TOTAL	1,06	TOTAL	+0,76
↓		↓		↓	
HARD PARTS OF THE BODY		BODY FLUIDS		PLEURA	

Como se puede apreciar, hemos agrupado estas cadenas en bloques de dos o tres en aquellas ocasiones en las que exista una relación entre ellas. En esta representación junto

al número de cadena concedido por *Hesperus* (Cadena 4), está en cursiva el lexema que inicia la cadena (*nodules*). Abajo, están las categorías del tesoro y su porcentaje de activación. En la casilla que sigue la flecha, está la etiqueta que hemos atribuido a cada cadena. Según esta tabla, habría en el texto tres cadenas en torno al concepto BODY, al que *Hesperus* atribuye un 3,51%. Este porcentaje no dista demasiado del que se infería de la lista de frecuencia lematizada (2,97%) del apartado 8.3.2.

8.3.4. Propuesta definitiva de cadenas léxicas

Con toda la información recabada en los apartados anteriores y teniendo en cuenta la forma en que se estructuran los términos en el campo de la oncología, creemos conveniente hacer una formulación definitiva de las cadenas léxicas presentes en cada texto. Para llegar a esta propuesta, se eliminarán los lexemas incluidos erróneamente, se ubicarán en cadenas nuevas, se añadirán unidades léxicas nuevas a las cadenas ya existentes o fusionaremos cadenas y crearemos nuevas cadenas y subcadenas. Asimismo, prestaremos una especial atención a las unidades lexicalizadas y fraseológicas que aparecen.

En las cadenas, no especificaremos la forma canónica de las unidades léxicas ya que nuestra aproximación a la terminología es descriptiva y nos interesa el uso de los términos en nuestro corpus y las distintas variaciones que se producen con respecto al término más frecuente. Es decir, nuestro interés está en el uso contextual de los términos y no en su uso prescriptivo y estandarizado. En consecuencia, en muchas ocasiones los verbos estarán en la forma *'-ed'*, en lugar de en infinitivo, y los sustantivos en plural. Esto nos indica que la forma canónica no tiene por qué ser la más frecuente.

Por otra parte, bajo la misma entrada, a veces introduciremos verbos, sustantivos, adjetivos o adverbios. Esto es así porque la cohesión se crea independientemente de la categoría gramatical de las unidades léxicas vinculadas y depende principalmente de la repetición conceptual y léxica. Y así, uno de los componentes de la cadena LOCATION IN THE HUMAN BODY del texto QDT2 es el lema *regional* bajo el que se agrupan las formas con las que establece cohesión mediante repetición exacta (*regional-regional*) y repetición variada sintáctica (*regional-region; regional-regionally*). Se especifican las UF que forma

en el texto, que se relacionan mediante repetición sintagmática exacta (*regional lymph node-regional lymph node*) y simple (*regional lymph node-regional lymph nodes*).

TABLA 36: Casilla perteneciente a una de las *cadena léxicas definitivas* de QDT2

REGIONAL Regional 9 ~ lymph node(s) 3 Region 1, regionally 1	11
---	----

La creación de nuevas cadenas obedece a que tanto las listas lematizadas como el perfil del documento apuntan hacia áreas conceptuales significativas en el texto y en el campo de la oncología, pero que *Hesperus* no ha reconocido. Hemos seguido una serie de criterios que justifiquen la creación, agrupación o eliminación de cadenas:

1. Se creará una cadena siempre *Hesperus* la haya creado en otro texto relacionado del corpus, con tal de que el porcentaje asignado al concepto por el perfil del documento lo respalde.
2. Se tendrá en cuenta las áreas conceptuales activadas en el evento médico oncológico (Faber 1997 y Tercedor 1999).
3. Son candidatos a formar cadena tanto los diez primeros conceptos en el perfil del documento como los diez primeros lexemas de las listas de frecuencia que no se puedan integrar dentro de otros.

8.3.5. Cuadro contrastivo (*Hesperus* – Análisis combinado)

Como última etapa, se apreciarán las diferencias entre la cohesión susceptible de ser detectada por un ordenador con tesoro y la detectada por un lector. La primera es principalmente explícita (repetición léxica) y asocia lexemas de la lengua general. En cuanto a la segunda, se basa no sólo en la repetición léxica sino también en el conocimiento especializado, que comprende conocimientos de terminología médica, del método científico y de los principales recursos textuales en ese subdominio. Se presentarán estas diferencias cuantitativamente en una tabla que compara las cadenas identificadas por *Hesperus* y nuestra propuesta. Esta tabla incluye el número de lexemas

vinculados, el número de formas distintas y la activación conceptual correspondiente a cada cadena reconocida por *Hesperus* (columna izquierda) o por nosotros (columna derecha). Esta información va a veces entre corchetes.

TABLA 37: Fragmento del cuadro contrastivo correspondiente al texto QDT2.

CADENAS LÉXICAS RECONOCIDAS POR <i>HESPERUS</i>		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 3: TREATMENT, DIAGNOSIS, MEDICINE [105-23-21,39] H 1: DRUGS [6-2-4]			0. TREATMENT [143-30 -23,6] 4. DIAGNOSIS [53-24-8,74]
Lexemas vinculados	111	196	Lexemas vinculados
Formas distintas	25	54	Formas distintas
Activación conceptual ajustada (%)	29,52	32,34	Activación conceptual (%)
ERA (0,65)			11. TIME [8-8-1,32]
Activación conceptual ajustada (%)	0,76	1,32	Activación conceptual (%)
[...]			[...]
TOTAL			
Lexemas vinculados	345	606	Suma de elementos en las cadenas
Formas distintas	80	552	Lexemas vinculados ¹¹⁰
Activación conceptual según <i>Hesperus</i>	86,01	192	Formas distintas ¹¹¹
Activación conceptual ajustada (%)	100	99,96	Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QPT3			
Nº total de palabras: 1255		Categorías conceptuales del tesoro: 83	
Palabras que crean cohesión (<i>Hesperus</i>): 562		Coeficiente de cohesión (<i>Hesperus</i>): 12,21	
		Nº de oraciones: 66	

Ha sido preciso ajustar los porcentajes conceptuales de *Hesperus* porque la suma de estos no llegaba al 100%. También se han computado los porcentajes de conceptos que no forman cadena pero que aparecían en el perfil del documento. Este es el caso del concepto ERA, que no forma cadena pero que tiene una activación conceptual (0,76%) no muy distinta a la que le hemos dado nosotros (1,32%).

¹¹⁰ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹¹¹ De estas, hay 21 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 54 lexemas adicionales a los 552 lexemas vinculados.

Los porcentajes que aportamos como resultado de nuestro análisis se calculan teniendo en cuenta el número de lexemas de cada cadena léxica (143 en la cadena *TREATMENT*) en relación con la suma de los elementos de todas las cadenas (606). Esta suma es superior al número de lexemas vinculados (552) ya que, como veremos, hay lexemas susceptibles de ser clasificados en varias cadenas. La diferencia entre las dos indica el número de lexemas que hacen de nodos de más de una cadena.

Por último, se añaden datos estadísticos obtenidos mediante *Wordsmith Tools* (número total de palabras, nº de oraciones) y *Hesperus* (palabras que crean cohesión, categorías conceptuales del tesoro *Roget* y coeficiente de cohesión). El número de palabras que crean cohesión según *Hesperus* (562) no es muy distinto al que proponemos en la casilla de la derecha, *Lexemas vinculados* (552).

9. Selección del corpus

9.1. Criterios de selección del corpus

Desde los comienzos de la lingüística de corpus se ha especulado sobre los criterios que hay que tener en cuenta a la hora de recopilar un corpus. Renouf (1984: 10) cita los aspectos considerados al elaborar el corpus en el que se basaría el diccionario COBUILD: fecha de publicación, inclusión de textos escritos de amplia difusión en diversos medios, variación dialectal, género y tema. Estos criterios se han ido refinando posteriormente para atender también a criterios como sexo, edad y nacionalidad del autor del texto, factualidad del texto, participantes en la comunicación, etc. (Atkins *et al.* 1992, Clear 1993, Sinclair 1991), aunque referidos principalmente a corpóra de la lengua general.

En relación a los corpóra de la lengua general han surgido debates en torno a si en la selección del corpus debe primar la cantidad o la calidad, entendida como *equilibrio* en la composición del corpus (Church y Mercer 1993: 17-18). Este debate se ha orientado hacia el predominio de la calidad, gracias al iluminador artículo de Biber (1993), en el que se expone que el principal criterio es la *representatividad*. Esta está relacionada con el grado en el que la muestra incluye todo el espectro de variabilidad de una población. Dicho espectro abarca los tipos textuales de una lengua y la distribución de sus recursos lingüísticos. Por ello, para Biber, el muestreo debería seguir criterios estadísticos. Sin embargo, como apunta Kennedy (1998: 74), la población suele ser tan amplia y los tipos textuales tan difícilmente delimitables, que un muestreo aleatorio es poco factible.

Asimismo, aunque se tenga en mente el criterio de la representatividad, el investigador debe tener presente que los resultados obtenidos pueden ser significativos desde dos puntos de vista: con referencia a un corpus concreto o con referencia a la lengua como conjunto (Hanks 2000: 313). Esto último es bastante inusual. Somos conscientes de que, dada la profusión de textos sobre medicina tanto en soporte papel como electrónico, difícilmente podemos recopilar un corpus representativo. Por ello, creemos que hay que añadir otros criterios que reflejen que estamos ante un corpus especializado, que va a ser analizado con herramientas informáticas y que pretendemos que sea útil desde el punto de vista de la traducción y de la terminología. Entre estos

criterios podemos citar la función que va a tener el corpus (el estudio de la cohesión en textos biomédicos), los usuarios (el investigador y el público que visite el servidor *web* del sistema bilingüe de información y recursos oncológicos *Oncoterm*), la inclusión de textos escritos, la disponibilidad de textos en soporte electrónico, la reutilizabilidad y la necesidad de colocar los textos recopilados en contexto (Hockey y Walker 1993: 239-240), en concreto, en el marco de referencia que hemos llamado *EVENTO ONCOLÓGICO*. Añadimos a estos criterios dos más propuestos por Tercedor (1999: 117): la variedad comunicativa en cuanto a intenciones comunicativas y tipo de lector, y la necesidad de incluir tanto textos originales como traducidos.

9.2. Limitaciones del corpus utilizado

La función del corpus es la de reflejar usos del lenguaje biomédico para sistematizar y computar los recursos cohesivos. Aún cuando se disponga de un corpus en soporte electrónico, este objetivo requiere trazar relaciones cohesivas no entre las unidades léxicas y terminológicas de todo el corpus, sino en cada texto como unidad. Esta premisa impone que se tome como unidad de muestreo (*sampling unit*) el texto completo, en el que se objetivarán las instancias de cohesión. Al analizar la cohesión dentro de cada texto, cualquier inferencia que se haga requiere la comparación de los datos de cada texto individual con el resto, una perspectiva nada usual dentro de la lingüística de corpus.

Por otra parte, la cohesión de un texto es extremadamente sensible a las variaciones de contenido y de audiencia, por lo que para comparar diferencias cohesivas ha sido necesario encontrar textos sobre exactamente el mismo tema. Como el factor que más influye en la cohesión es el conocimiento compartido del emisor y el receptor, ha sido preciso trabajar con parejas de textos sobre un tema muy restringido de la oncología. Por tanto, para el análisis extensivo, se ha decidido seleccionar exclusivamente textos sobre el tratamiento del cáncer de pulmón, lo cual ha hecho que se descarten muchos de los textos seleccionados inicialmente. En el momento de comenzar la investigación, muy pocos textos cumplían este requisito.

También ha supuesto una limitación la utilización de un generador de cadenas léxicas en fase experimental del que sólo se han podido obtener resultados gracias a la

buena voluntad de su creador, Jeremy Ellman. Para no abusar de esa buena voluntad, los textos que se enviaron para ser analizados han tenido una extensión razonable, en concreto, 106 922 palabras.

Por último, la disponibilidad de textos traducidos sobre el cáncer ha sido menor de lo que pensábamos en un principio. Hoy en día el mercado profesional de traducciones del inglés al español es cada vez más reducido porque se supone que los especialistas leen las publicaciones internacionales en inglés. Lo habitual es que estos contraten al traductor para que redacte en inglés el resultado de sus investigaciones o haga de revisor de la traducción. La proporción mayor de textos traducidos la constituye la versión española de sitios *web* sobre oncología de Estados Unidos.

A pesar de estas limitaciones, reconocemos la validez de aplicar al estudio de la cohesión los datos auténticos que aporta un corpus formado por textos completos, en lugar de servirnos de oraciones aisladas o párrafos, como se ha hecho en muchos estudios previos sobre cohesión.

9.3. Composición del corpus

En esta tesis se ha trabajado tanto con un corpus comparable de textos originales sobre cáncer de pulmón en inglés y en español como con un corpus paralelo sobre el mismo tema. De esta forma se pueden delimitar las diferencias con respecto a la cohesión que responden al proceso traductor en sí. En la composición del corpus, los textos destinados al análisis extensivo forman una unidad aparte.

El análisis extensivo se centra en un corpus de textos sobre cáncer de pulmón que cuentan en su forma final con un total de 106 922 palabras, aunque antes de modificarlos para el análisis contaban con más¹¹³. Han sido extraídos de *CancerNet*¹¹⁴, la página de Internet del *National Cancer Institute*, que es la principal organización norteamericana contra el cáncer. Esta página contiene información sobre el cáncer en forma de folleto explicativo en las áreas del tratamiento, prevención, detección precoz,

¹¹³ El corpus contaba con 153097 palabras inicialmente. Sin embargo, para utilizar *Hesperus* hemos tenido que suprimir las referencias bibliográficas de los textos para oncólogos porque distorsionaban los datos. También distorsionaban los índices y la información sobre el PDQ que aparecía en todos y cada uno de los textos, por lo que hemos tenido que eliminarlos de los textos.

¹¹⁴ La dirección es <http://cancernet.nci.nih.gov>

cuidado médico, datos estadísticos y oncología experimental (ensayos clínicos)(véase Apéndice II, *Textos extraídos del PDQ*).

En concreto, hemos trabajado con una sección de resúmenes llamada *Physician Data Query (PDQ)* por una serie de motivos. En primer lugar, el *National Cancer Institute* proporciona a un número muy elevado de visitantes una cantidad ingente de información, que se actualiza constantemente con rigurosidad, y cuenta con el prestigio que confiere el hecho de que es referencia obligada para muchos oncólogos. En segundo lugar, los textos de *CancerNet* tienen el interés de estar redactados con miras a dos tipos de lectores con un conocimiento distinto del tema: médicos especialistas y pacientes o familiares de pacientes. Por eso, esta página tiene dos secciones: *CancerNet Health Professionals* y *CancerNet Patients*. Teniendo en cuenta que la cohesión de un texto es extremadamente sensible a las variaciones de contenido y de audiencia, es un privilegio disponer de textos sobre exactamente el mismo tema en los que las variaciones con respecto a la cohesión se deben sobre todo a cambios en el conocimiento de los lectores de los mismos. Por otra parte, estos textos aparecen en la misma página *web* traducidos al castellano, lo cual nos permitirá comparar cambios con respecto a la cohesión dependientes de las lenguas con las que se trabaja y que son resultado de la traducción. La separación temporal entre los textos originales y traducidos oscila entre un mes y tres meses.

Se someterán 36 textos del *PDQ* al programa *Hesperus* para localizar las principales cadenas cohesivas de los mismos. De estos, 18 están destinados a especialistas (*Cancernet Health Professionals*) y otros tantos a pacientes y a familiares de pacientes (*Cancernet Patients*). Como diferencias más significativas entre unos y otros podemos destacar que los primeros presentan una mayor extensión y complejidad léxica, las cuales ilustran una mayor complejidad conceptual. De ahí que en su forma original contaban con bibliografía sobre cáncer de pulmón. Ambas secciones comparten dos de los apartados en los que se agrupa la información sobre cáncer de pulmón: uno, centrado en el tratamiento (*Cancer Treatment Information Summaries*) y otro, sobre cuidados médicos (*Supportive Care and Advocacy Issues*). Los textos sobre tratamiento están más centrados en el cáncer de pulmón, mientras que los textos sobre cuidados

médicos tratan sobre trastornos y patologías asociadas, que no se restringen exclusivamente a este tipo de tumor.

Por otra parte, los ejemplos que se presentan a lo largo de la tesis para ilustrar las funciones léxicas, para la clasificación de categorías de cohesión léxica, para el análisis intensivo y en el capítulo sobre cohesión y traducción han sido extraídos de las fuentes que se especifican en la tabla 38. En total computan 1 397 515 palabras, 1241780 correspondientes al corpus comparable y 155 735 de los textos paralelos que no están incluidos en la cifra anterior (en realidad, el corpus paralelo cuenta con 417 822). Dado que los recursos en soporte electrónico son más numerosos en lengua inglesa, la parte en inglés (893068) ocupa una proporción mayor del corpus bilingüe en comparación con el español (509 447).

En la selección de las fuentes, se ha intentado cubrir el espectro definido por Lévy-Leblond (1996) y Tercedor (1999: 172). Es decir, distinguimos entre *intercambio especializado*, donde se incluyen las publicaciones especializadas, las páginas *web* destinadas a profesionales de la salud y los manuales; *intercambio público*, que aglutina las publicaciones de divulgación semiespecializada y general y los folletos de salud pública en Internet; e *intercambio familiar*, donde se han incluido los correos electrónicos de la lista de distribución *LUNG-ONC*, en la que participan pacientes de cáncer de pulmón y sus familiares (LUNG-ONC@LISTSERV.ACOR.ORG).

TABLA 38: Composición del corpus

CORPUS COMPARABLE		
	INGLÉS (821 456 palabras)	ESPAÑOL (420 324 palabras)
Intercambio especializado	Publicaciones especializadas: <i>British Medical Journal</i> <i>Cancer</i> <i>CANCERLIT</i> <i>C-A. A Cancer Journal for Clinicians</i> <i>Lancet</i> <i>Medline</i> <i>The New England Journal of Medicine</i>	Publicaciones especializadas: <i>Archivos Bronconeumológicos</i> <i>Asociación Española de Endoscopia Respiratoria</i> <i>Medicina Clínica</i> <i>Neoplasia</i> <i>Revisiones en cáncer</i>
	Páginas web destinadas a profesionales de la salud: <i>CancerBacup</i> <i>Physician Data Query (PDQ)</i> <i>PSL</i> <i>Virtual Hospital</i>	Páginas web destinadas a profesionales de la salud: <i>Atheneum / Jano, Medicina y Humanidades.</i> <i>Diario Médico</i> <i>Sociedad Iberoamericana de Información Científica</i>
	Manuales: <i>Cancer: Principles and Practice of Oncology</i>	Manuales: <i>Medicina Interna</i> <i>Oncología Médica. Guía de Oncología Médica</i>
Intercambio público	Publicaciones de divulgación especializada: Monográficos <i>Scientific American (monográfico)</i>	
	Publicaciones de divulgación general: Monográficos <i>TIME</i> Noticias sobre oncología: <i>Daily Telegraph</i> <i>Oncolink (CNN Health and Food)</i> <i>Reuters Health</i> Publicadas en <i>Bacup</i> y la web del <i>NCI</i>	Publicaciones de divulgación general: Monográficos <i>Blanco y negro</i> <i>El Semanal</i> <i>Muy interesante</i> <i>QUO</i> Noticias sobre oncología: <i>El Mundo (CD y versión on-line)</i> <i>Ideal (Suplemento Campus)</i>
	Folletos de Salud Pública en Internet para pacientes y familiares: <i>Alcase</i> <i>American Cancer Society</i> <i>Cancer Bacup</i> <i>American Lung Association</i> <i>CancerHelp UK</i> <i>Imperial Cancer Research Fund</i> <i>PDQ</i> <i>P. Connelly Lung Cancer Page</i> <i>Oncolink</i> <i>Virtual Hospital</i>	Folletos de Salud Pública en Internet para pacientes y familiares: <i>Ciencia y Cáncer</i> <i>Instituto Madrileño de Oncología</i> <i>PDQ</i> <i>Sarenet</i>

Int. familiar	Correos electrónicos de la lista de distribución <i>LUNG-ONC</i>	
CORPUS PARALELO		
INGLÉS (71 612 palabras)		ESPAÑOL (84 123 palabras)
Intercambio especializado	Publicaciones especializadas <i>Medicina Clínica (abstract)</i> <i>Neoplasia (abstract)</i>	Publicaciones especializadas <i>Medicina Clínica (abstract)</i> <i>Neoplasia (abstract)</i>
	Páginas web destinadas a profesionales de la salud: <i>American Cancer Society</i> <i>American Lung Association</i> <i>PDQ</i>	Páginas web destinadas a profesionales de la salud: <i>American Cancer Society</i> <i>American Lung Association</i> <i>PDQ</i>
	Manuales: <i>Cancer: Principles and Practice of Oncology</i> Merck Manual	Manuales: <i>Cancer: Principios y Práctica de Oncología</i> <i>Manual Merck</i>
Intercambio público	Publicaciones de divulgación especializada: Monográficos <i>Scientific American</i>	Publicaciones de divulgación especializada: Monográficos Investigación y Ciencia
	Enciclopedias <i>Encarta</i>	Enciclopedia <i>Encarta</i>
	Folletos de Salud Pública en Internet para pacientes y familiares: <i>American Cancer Society</i> <i>American Lung Association</i> <i>PDQ</i>	Folletos de Salud Pública en Internet para pacientes y familiares: <i>American Cancer Society</i> <i>American Lung Association</i> <i>PDQ</i>

9.4. Tratamiento del corpus: eliminación de palabras forma y lematización

Para extraer las *palabras plenas (content words)* que crean cohesión, serán de utilidad las listas de frecuencia correspondientes al corpus de 36 textos sacados del *PDQ* (en adelante, *Corpus PDQ*). Para ello, es necesario eliminar de estas listas las palabras forma y aglutinar las diversas variantes morfológicas derivadas de la misma raíz (lematización). Con este fin se ha creado una *stoplist* de 256 palabras, es decir, una lista con las *palabras forma* que son muy frecuentes tanto en nuestro corpus como en córpora de la lengua general. Teniendo en cuenta que nuestros textos están escritos en inglés norteamericano, la hemos elaborado a partir de las 150 palabras más frecuentes del *Corpus Brown* (véase Apéndice VIII). De estas se han seleccionado las palabras

forma y se ha comprobado que estas coinciden con las del *BNC* excepto en la frecuencia, que les confiere un orden distinto. También se han añadido las *palabras forma* que encabezaban la lista de frecuencia del *Corpus PDQ*.

Por otra parte, como en esta tesis es fundamental el estudio de la representación conceptual, se llevará a cabo un proceso de *lematización*. Se ha utilizado como base la lista de lemas elaborada por Yasumasa Someya¹¹⁵, a la que se han añadido algunos lexemas con su flexión. De todas formas, hemos preferido la lematización manual para evitar errores ocasionados al lematizar.

¹¹⁵ La autora de esta lista, cuyo correo electrónico es ysoyeya@gol.com, la ha puesto a disposición de cualquier visitante de la página *web* de Mike Scott en <http://www.liv.ac.uk/~ms2928/index.htm>.

TERCERA PARTE:
RESULTADOS

10. Análisis intensivo de una noticia sobre oncología en español

En este capítulo se aplica el análisis propuesto para textos breves en 7. *Esquemas de representación para el análisis intensivo de la cohesión* a una noticia sobre cáncer publicada en la versión diaria *on-line* de la revista semanal de medicina general para profesionales de la salud *Jano Medicina y Humanidades*. *Jano* está editada por el grupo editorial *Doyma*, cuya página *web* incluye bajo el hipervínculo *Jano Diario*¹¹⁶, noticias sobre investigaciones publicadas en las revistas con mayor índice de impacto según el *Science Citation Index*¹¹⁷.

El texto objeto de análisis es el resumen de un artículo aparecido en 1997 en *Cancer*, una publicación internacional e interdisciplinar de la *American Cancer Society* que edita John Wiley & Sons. La revista *Cancer* ocupa el puesto número 15 por índice de impacto entre las 96 revistas de oncología del *Science Citation Index*.

El género al que podríamos asociar este texto es una versión breve y algo más sofisticada de lo que Nwogu (1991) denomina *versión periodística de trabajos de investigación médicos* (véase p. 92). A diferencia de este género, el texto que analizamos está destinado a profesionales de la medicina interesados en las últimas investigaciones médicas, y se caracteriza por la brevedad y por un formato parecido al de una *noticia*. Además, este texto tiene la particularidad de combinar dos tipos de traducción: la traducción modificada por reducción y la traducción para obtener información (Sager 1997). Por este motivo, vamos a denominar a este género tan habitual en medicina *noticia – traducción para semiespecialistas derivada de abstract y/o artículo experimental en inglés*. El destinatario de este género es un lector secundario, en concreto, un médico no especialista en oncología que quiere información sobre los últimos avances internacionales, publicados en inglés, aunque no se excluye a un lector culto interesado por la salud. El texto, modificado con marcadores de oración, es el siguiente:

(65)

¹¹⁶ http://www.doyma.net.es/cgi-bin/owa/visual.lanzar_presnot. La consulta se realizó el 20 de diciembre de 1998.

¹¹⁷ El *Science Citation Index* cubre 3 300 revistas de 100 disciplinas científicas publicadas en 60 países. Según la relevancia para la comunidad científica de los artículos publicados por una revista, computada con respecto al número de veces que se citan artículos y entradas de la misma, se otorga un índice de impacto para esa publicación. El listado de todas las publicaciones científicas, ordenadas por factor de impacto es el *Journal Citation Reports*.

1 **El cáncer de pulmón se asocia a alteraciones del cromosoma 9**

2 <Jano Diario, Barcelona> Investigadores de la Universidad de Texas (Estados Unidos) han descubierto que ciertas alteraciones en el cromosoma 9 parecen ser marcadores del riesgo de cáncer pulmonar y de la agregación familiar de este tipo de tumor, que es precisamente el que parece ser en mayor medida inducido por factores ambientales.

3 Investigaciones previas del equipo que dirige el Dr. Xifeng Wu mostraban que las anomalías del cromosoma 9 en los linfocitos periféricos constituían un predictor de riesgo significativo de carcinoma pulmonar.

4 En el nuevo estudio, publicado en el último número de "Cancer", trataron de determinar si esta asociación se debía a una inestabilidad genética heredada.

5 Identificaron a 97 pacientes con cáncer de pulmón y alteraciones en el cromosoma 9, y a continuación evaluaron a sus familiares de primer grado.

6 Los resultados indican que estos familiares presentan una probabilidad 8,5 veces superior de desarrollar carcinoma pulmonar si tienen mutaciones del cromosoma 9 respecto a los que no las tienen.

7 Aunque se trata de resultados preliminares, el Dr. Wu sugiere que el estudio de estas alteraciones cromosómicas podrían utilizarse como marcadores para identificar a poblaciones de alto riesgo.

Antes del análisis, la comparación entre las estadísticas de este texto y el original inglés del que se deriva pueden ser indicativas de diferencias cohesivas.

10.1. Datos estadísticos del texto

TABLA 39: Datos estadísticos del texto "El cáncer de pulmón se asocia a alteraciones del cromosoma 9"

N	TEXTO TRADUCIDO	TEXTO ORIGINAL
Text File	CROMOS.TXT	CHROMO.TXT
Bytes	1.298	1.919
Tokens	199	285
Types	112	122
Type/Token Ratio	56.28	42.81
Standardised Type/Token		
Ave. Word Length	5.40	5.53
Sentences	7	15
Sent.length	28.43	19.00
sd. Sent. Length	12.27	10.62

El texto cuenta con 199 palabras correspondientes a 112 formas léxicas. Lo más destacado es el hecho de que, aunque el texto español tiene una extensión de dos tercios del original, el número de formas distintas (112) es prácticamente el mismo que en el original (122). La traducción-adaptación tiene la mitad de oraciones, aunque son más extensas, 28 palabras frente a las 19 del original.

En efecto, el hecho de que el *abstract* en inglés está dirigido a oncólogos y es un texto original, frente a la traducción al español en forma de noticia adaptada a semiespecialistas o lectores cultos, lleva aparejadas diferencias en cuanto a la designación de los mismos conceptos para crear cohesión. Como se verá en el apartado siguiente, el texto en inglés tiende a la repetición léxica exacta, simple y variada sintáctica de las *unidades de significación clave* del texto (*chromosome, aberration, lung cancer*), mientras que el texto en español tiende a la variedad léxica y terminológica. Con el término *unidades de significación clave (USC)* designaremos unidades léxicas y terminológicas, tanto monoléxicas como poliléxicas que juegan un papel relevante en la activación conceptual del tema sobre el que versa un texto especializado¹¹⁸.

10.2. Listas de frecuencia

En la página siguiente presentamos listas de frecuencia lematizadas de lexemas individuales y de las combinaciones de dos, tres y cuatro palabras en las que participan. Servirán de base para las matrices de repetición y para identificar unidades fraseológicas.

Según las listas, sólo un 25,7% de todas las formas participan en algún tipo de repetición, es decir, aparecen más de una vez, frente al 34,7% del texto en inglés. Podemos justificar la mayor variedad léxica del texto español en función de dos aspectos. Por una parte, el español es una lengua más flexiva que el inglés. Es muy frecuente encontrar dos formas del mismo lema (*familiar / familiares, trata / trataron*), incluso cuando se han eliminado las palabras forma (*de/del, el/la/los/las*). En segundo lugar, la adaptación en español del *abstract* no utiliza la terminología de los especialistas.

Esta mayor variedad queda patente cuando observamos que el texto tiene 76 lemas y 199 palabras, mientras que en el original hallamos sólo 77 lemas, a pesar de contar con 285 palabras. En efecto, en el texto encontramos cuatro formas para el concepto designado por las variantes terminológicas *chromosome 9 aberrations* y *chromosomal aberration: alteraciones en el cromosoma 9, alteraciones cromosómicas, anomalías del cromosoma 9 y mutaciones del cromosoma 9* (véase p. 288 para más detalle).

¹¹⁸ Adaptación del concepto *unitat de significació especialitzada (USE)* de Estopà (1999: 25) bajo el que se engloban tanto las unidades terminológicas que forman parte de una estructura conceptual especializada y se usan en discursos temáticamente e intencionadamente especializados, como otras unidades léxicas del lenguaje natural que, sin ser términos, también son pertinentes en un determinado campo profesional.

La mayor variedad léxica también reduce las posibilidades de encontrar combinaciones de palabras con un alto grado de fijación. Entre las unidades fraseológicas más fijas, sólo aparecen *carcinoma pulmonar* (1,01%) y *cáncer de pulmón* (1,01%) (véase tablas 41, 42 y 43). En inglés, eran más habituales: *lung carcinoma* y *lung carcinoma patients* (1,75%); *family history of cancer* (1,75%); *chromosome aberrations* y *spontaneous chromosome aberrations* (1,05%).

TABLA 40: Lista de frecuencia lematizada

N	Word	Frq.	%	Lemmas
1	Cromosoma	6	3,02	cromosómicas(1)
2	Pulmonar	5	2,51	pulmón(2)
3	Alteraciones	4	2,01	
4	Cáncer	4	2,01	cancer(1)
5	Familiar	3	1,51	familiares(2)
6	Riesgo	3	1,51	
7	Asociación	2	1,01	asocia(1)
8	Carcinoma	2	1,01	
9	Dr	2	1,01	
10	Estudio	2	1,01	
11	Identificar	2	1,01	identificaron(1)
12	Investigaciones	2	1,01	investigadores(1)
13	Marcadores	2	1,01	
14	Parece	2	1,01	parecen(1)
15	Resultados	2	1,01	
16	Ser	2	1,01	
17	Tienen	2	1,01	
18	Trata	2	1,01	trataron(1)
19	Wu	2	1,01	
	Repet. exacta, simple y variada sintáctica	51	25,7	
20	Agregación	1	0,50	
21	Alto	1	0,50	
22	Ambientales	1	0,50	
23	Anomalías	1	0,50	
24	Aunque	1	0,50	
25	Barcelona	1	0,50	
26	Constituían	1	0,50	
27	Continuación	1	0,50	
28	Debía	1	0,50	
29	Desarrollar	1	0,50	
30	Descubierto	1	0,50	
31	Determinar	1	0,50	
32	Diario	1	0,50	
33	Dirige	1	0,50	
34	Equipo	1	0,50	
35	Estados	1	0,50	
36	Evaluaron	1	0,50	
37	Factores	1	0,50	
38	Genética	1	0,50	
39	Grado	1	0,50	
40	Heredada	1	0,50	
41	Indican	1	0,50	
42	Inducido	1	0,50	
43	Inestabilidad	1	0,50	
44	Jano	1	0,50	
45	Linfocitos	1	0,50	
46	Mayor	1	0,50	
47	Medida	1	0,50	
48	Mostraban	1	0,50	
49	Mutaciones	1	0,50	
50	Nuevo	1	0,50	
51	Número	1	0,50	
52	Pacientes	1	0,50	
53	Periféricos	1	0,50	
54	Poblaciones	1	0,50	
55	Podrían	1	0,50	
56	Precisamente	1	0,50	
57	Predictor	1	0,50	
58	Preliminares	1	0,50	
59	Presentan	1	0,50	
60	Previas	1	0,50	
61	Primer	1	0,50	
62	Probabilidad	1	0,50	
63	Publicado	1	0,50	
64	Respecto	1	0,50	
65	Significativo	1	0,50	
66	Sugiere	1	0,50	
67	Superior	1	0,50	
68	Texas	1	0,50	
69	Tipo	1	0,50	
70	Tumor	1	0,50	
71	Último	1	0,50	
72	Unidos	1	0,50	
73	Universidad	1	0,50	
74	Utilizarse	1	0,50	
75	Veces	1	0,50	
76	Xifeng	1	0,50	

TABLA 41: Combinaciones de dos palabras

N	Word	Freq.	%
1	Cromosoma 9	5	2,51
2	En el	4	2,01
3	Del cromosoma	3	1,51
4	Alteraciones en	2	1,01
5	Cáncer de	2	1,01

6	Carcinoma pulmonar	2	1,01
7	De la	2	1,01
8	De pulmón	2	1,01
9	El cromosoma	2	1,01
10	El Dr	2	1,01

TABLA 42: Combinaciones de tres palabras

	Word	Freq.	%
1	Del cromosoma 9	3	1,51
2	Alteraciones en el	2	1,01
3	Cáncer de pulmón	2	1,01
4	El cromosoma 9	2	1,01
5	En el cromosoma	2	1,01

TABLA 43: Combinaciones de cuatro palabras

N	Word	Freq.	%
1	Alteraciones en el cromosoma	2	1,01
2	En el cromosoma 9	2	1,01

Por último, la lista de frecuencia sin lematizar del texto en español, en comparación con el texto original también parece indicar que este último se caracteriza por un mayor grado de especialización. Presentamos los 15 lexemas más frecuentes:

TABLA 44: Formas más frecuentes del texto traducido y el texto original

N	Word	Freq.	%
1	De	17	8,54
2	El	9	4,52
3	Que	8	4,02
4	A	7	3,52
5	En	6	3,02
6	Cromosoma	5	2,51
7	Del	5	2,51
8	Alteraciones	4	2,01
9	Cáncer	3	1,51
10	Los	3	1,51
11	Pulmonar	3	1,51
12	Riesgo	3	1,51
13	Se	3	1,51
14	Y	3	1,51
15	Carcinoma	2	1,01

N	Word	Freq.	%
1	Chromosome	11	3,86
2	Of	11	3,86
3	Aberrations	10	3,51
4	The	10	3,51
5	Cancer	9	3,16
6	A	8	2,81
7	And	8	2,81
8	Family	6	2,11
9	For	6	2,11
10	History	6	2,11
11	In	6	2,11
12	Carcinoma	5	1,75
13	Lung	5	1,75
14	Was	5	1,75

15	Were	5	1,75
----	------	---	------

El texto en español tiene un porcentaje menor de términos, y de hecho, las 9 *palabras forma* de la lista se encuentran entre las 10 primeras de la parte escrita del *Corpus de Referencia del Español Actual*¹¹⁹: *de, la, que, el, y, en, a, los, se, del* (véase Apéndice IX). La proporción de *palabras forma* en inglés es menor y, curiosamente, encabeza la lista el término *chromosome*, y entre las 5 primeras formas, hay tres términos. Las 8 *palabras forma* del texto se encuentran también entre las 10 primeras del *Brown Corpus*, a excepción de *were* (véase Apéndice VIII).

10. 3. Matrices de repetición de las relaciones cohesivas primarias

En la matriz de repetición de las dos páginas siguientes, se han incluido las relaciones cohesivas más significativas, que entendemos que son las más fáciles de detectar por parte de cualquier lector, ya sea lego o experto.

Hemos identificado instancias de repetición exacta, de repetición simple y de repetición variada paradigmática, sintagmática y mixta. En estos recursos cohesivos predominan la relación de sinonimia y de identidad, marcadas por las funciones léxicas **Syn** y **Aequ** respectivamente. Estas relaciones se presentan en el texto explícitamente de manera que logran que el lector asocie oraciones que comparten uno de estos elementos:

- a) las mismas unidades léxicas (repetición exacta y repetición paradigmática simple):
cromosoma 9 - cromosoma 9; Identificaron-identificar
- b) la misma base léxica sometida a sufijación (repetición paradigmática sintáctica):
investigadores-investigaciones
- c) el mismo significado o un significado muy parecido (repetición paradigmática y mixta sinonímica): *alteraciones en el cromosoma 9 – anomalías del cromosoma 9.*
- d) términos entre los que existe una relación de hiponímica, meronímica u opositiva:
cáncer de pulmón - tumor.

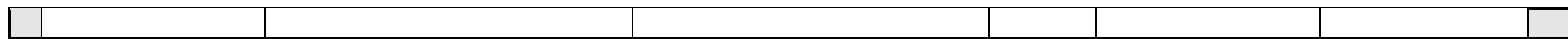
¹¹⁹ Auspiciado por la Real Academia de la Lengua Española, este corpus de lengua general constituye un banco de datos del español contemporáneo que cuenta en la actualidad con 100 millones de palabras, que se irán aumentando hasta conseguir 125 millones. Contiene textos orales y escritos producidos en todos los países de habla hispana desde 1975 hasta la actualidad. La parte escrita cuenta con aproximadamente 94 millones de palabras.

TABLA 45: Matriz de repetición de las relaciones cohesivas primarias (Tabla A)

	(1)						
2	1. Alteraciones_del_cromosoma_9 / Alteraciones_en_el_cromosoma_9 2. Cáncer_de_pulmón/ cáncer_pulmonar 3. Cáncer_de_pulmón/ Tumor	(2)					
3	4. Alteraciones_del_cr9/ Anomalías_del_cr9 5. Cáncer_de_pulmón/ Carcinoma_pulmonar	12. Alteraciones_en_el_cr9/ anomalías_del_cr9 13. Riesgo 14. Investigadores/ Investigaciones 15. Cáncer_pulmonar / carcinoma_pulmonar 16. Marcadores / predictor 17. Tumor / carcinoma_pulmonar	(3)				
4	6. Asocia / asociación	18. Inducido / debía	30. Investigaciones / estudio	(4)			
5	7. Alteraciones_del_cr9/ Alteraciones_en_el_cr9 8. Cáncer_de_pulmón	19. Alteraciones en el cr9 20. Familiar / familiares 21. Cáncer_pulmonar/ Cáncer_de_pulmón 22. Tumor /cáncer de pulmón	31. Anomalías_del_cr9 / alteraciones_en_el_cr9 32. Carcinoma_pulmonar/ cáncer_de_pulmón		(5)		
6	9. Alteraciones_del_cr9/ Mutaciones_del_cr9 10. Cáncer_de_pulmón/ Carcinoma_pulmonar	23. Alteraciones_en_el_cr9 / Mutaciones_del_cr9 24. Familiar / familiares 25. Cáncer_pulmonar/ carcinoma_pulmonar 26. Tumor /carcinoma pulmonar	33. Anomalías_del_cr9 / mutaciones_del_cr9 34. Carcinoma_pulmonar		40. Alteraciones_en_el_cr9 / mutaciones_del_cr9 41. Familiares 42. Cáncer_de_pulmón/ carcinoma_pulmonar	(6)	
7	11. Alteraciones_del_Cromosoma_9 / Alteraciones Cromosómicas	27. Alteraciones_en_el_Cromosoma_9 / Alteraciones cromosómicas 28. Marcadores 29. Riesgo	35. Anomalías_del_cromosoma9 / alteraciones cromosómicas 36. Dr. Xifeng Wu / Dr .Wu 37. Predictor / marcadores	38. Estudio 39. Trataron / trata	43. Alteraciones_en_el_Cromosoma_9/ Alteraciones cromosómicas 44. Identificaron/ identificar	45. Resultados 46. Mutaciones_del_cromosoma_9 / alteraciones cromosómicas	(7)

TABLA 46: Matriz de repetición de las relaciones cohesivas primarias (Tabla B): tipos de repetición y funciones léxicas

	(1)						
2	1. Mixta de UF (v sint) / Syn-Aequ 2. Mixta de UF (v sint) / Syn-Aequ 3. Mixta de UF y L (v léx) / Gener <i>Syn</i> \supset (cáncer de pulm)= <i>este tipo de tumor</i>	(2)					
3	4. Mixta de UF (var léx) / Syn-Aequ 5. Mixta de UF (var léx y sint) / Syn	12. Mixta de UF (var léx y sint) /Syn-Aequ 13. Exacta /Aequ 14. Parad sintáctica (var sint) / S ₂ 15. Mixta de UF (var léx y sint) /Syn-Aequ 16. Parad sinonímica / Syn \cap 17. Mixta de UF y L (var léx) / Gener <i>Syn</i> \subset (este tipo de tumor) = <i>Carcinoma pulmonar</i>	(3)				
4	6. Parad sintáctica (var sint) / S ₀	18. Paradigmática sinonímica Syn \cap ser inducido / se debía	30. Parad sinonímica Syn \cap	(4)			
5	7. Mixta de UF (var sint) / Syn-Aequ 8. Sintag exacta / Aequ	19. Sintag exacta /Aequ 20. Paradigm sintáctica / S ₀ (familiar) = familiares 21. Mixta de UF (var sint) / Syn-Aequ 22. =17. <i>Syn</i> \subset (tumor) = <i>cánc d pulmón</i>	31. Mixta de UF (var léx y sint) / Syn-Aequ 32. Mixta de UF (var léx y sint) / Syn		(5)		
6	9. Mixta de UF (var léx) / Syn 10. Mixta de UF (var léx y sint) / Syn	23. Mixta de UF (var léx y sint) /Syn-Aequ 24. Paradigm sintáctica / S ₀ (familiar) = familiares 25. Mixta de UF (var léx) / Syn-Aequ 26. Ver 17.	33. Mixta de UF (var léx) / Syn-Aequ 34. Sintag exacta /Aequ		40. Mixta de UF (var léx y sint) / Syn-Aequ 41. Exacta /Aequ 42. Mixta de UF (var léx y sint) / Syn	(6)	
7	11. Mixta de UF (var sint) / Gener <i>SynY - Aequ</i> (alterac_del_cr9) = alteraciones cromosómicas	27. Mixta de UF (var sint) / Syn-Aequ 28. Exacta /Aequ 29. Exacta /Aequ	35. Mixta de UF (var léx y sint) / Gener <i>Syn Y</i> (anomalías_del_cr9) = alterac cromosómicas 36. Paradigmática Simple / Aequ Dr Xifeng Wu – D. Wu 37. Parad sinonímica / Syn \cap	38. Exacta /Aequ 39. Parad simple /Aequ	43. Mixta de UF (var sint) / Syn Y Aequ (alterac_en_el_cr9) = alterac cromosómicas 44. Paradigm simple /Aequ	45. Exacta /Aequ 46. Mixta de UF (v léx y sint) / Gener <i>SynY Aeq</i> (mutaciones_del_cr9)= alterac cromosómicas	(7)



A continuación, comentamos los datos de esta matriz, que relaciona 7 oraciones mediante 46 conexiones, en tres apartados que se centran en la representación gráfica del número de vínculos del texto, los tipos de repetición y las funciones léxicas predominantes.

10.3.1. Cómputo de las conexiones en las que participa cada oración

En vista de las matrices de repetición, identificamos en la línea de Hoey (1991, 1995) aquellas oraciones que, por tener más conexiones, son las que contienen la información más relevante del texto. Esto se podría saber ojeando una tabla que represente las oraciones precedentes y posteriores vinculadas a cada oración.

TABLA 47: Conexiones en los que participa cada oración

Oración precedente con la que está vinculada (Nº de repeticiones)						Nº de oración	Oración posterior con la que se vincula (Nº de repeticiones)						
1	2	3	4	5	6		2	3	4	5	6	7	
						1	3	2	1	2	2	1	
3						2		6	1	4	4	3	
2	6					3			1	2	2	3	
1	1	1				4						2	
2	4	2				5					3	2	
2	4	2		3		6						2	
1	3	3	2	2	2	7							

Al transferir estos datos a una tabla del tipo de Hoey (1991), la tabla 48, se ve cómo todas las oraciones a excepción de la cuarta, quinta y sexta, están vinculadas con todas las demás mediante seis conexiones y que, por tanto, tienen igual relevancia en el texto. La que contribuye menos a la cohesión del texto y resulta menos informativa es la cuarta oración:

- (66) En el nuevo estudio, publicado en el último número de "Cancer", trataron de determinar si esta asociación se debía a una inestabilidad genética heredada.

TABLA 48: Tabla de vínculos

Oración	Cantidad de oraciones con las que se vincula (precedentes, posteriores)
1	(0, 6)
2	(1, 5)
3	(2, 4)
4	(3, 1)
5	(3, 2)
6	(4, 1)
7	(6, 0)

Este texto, a pesar de su brevedad, tiene un elevado grado de cohesión. Esto parece contradecir la opinión de Berber Sardinha (1995) de que la vinculación es más frecuente entre oraciones no contiguas. Podemos explicar esto teniendo en cuenta que el texto sintetiza la información de otro texto, lo cual hace que casi todas las oraciones contengan información muy relevante, y por tanto, estén bastante cohesionadas.

Aunque la elaboración de estas tablas puede resultar tediosa para un traductor, creemos que a medida que se perfeccionen y sean accesibles aplicaciones informáticas como las descritas en 4.8., algunos encargos de traducción se harán de una forma más efectiva. Entre estos encargos podemos mencionar los resúmenes, *gist translations* y *texts for information only*. Será posible identificar las oraciones más informativas y las palabras y términos clave sobre los que debe girar la traducción, aspectos muy relevantes en la ingeniería del lenguaje y del conocimiento y la terminología.

10.3.2. Tipos de repetición

En el texto que nos ocupa, podemos ver que el tipo de repetición más frecuente es la repetición mixta de unidades fraseológicas. Para visualizar mejor los datos de la matriz de repetición hemos elaborado la siguiente tabla en la que se especifican cuántos casos corresponden a cada tipo de repetición y un ejemplo de cada uno de ellos:

TABLA 49: Tipos de repetición

TIPO DE REPETICIÓN		Nº DE CASOS		EJEMPLOS (nº de oración)
Mixta	de Unidades Fraseológicas (sinonímica)	Variac. léxica y sintáctica	10	Cáncer de pulmón (1) / Carcinoma pulmonar (6)
		Variac. Sintáctica	7	Alteraciones del cromosoma 9 (1) / Alteraciones en el cromosoma 9 (5)
		Variac. léxica	5	Anomalías del cromosoma 9 (3) / Mutaciones del cromosoma 9 (6)
	de UF y lexema (hiponímica / sinonímica)	Variac. léxica	4	Cáncer de pulmón (1) / (este tipo de) tumor (2)
Exacta	Paradigm	6	Riesgo (2) / riesgo (7)	
	Sintagm	3	Cáncer de pulmón (2) / Cáncer de pulmón (5)	
Paradigmática sinonímica		4	Investigaciones (3) / estudio (4)	
Paradigmática sintáctica		4	Asocia (1) / asociación (4) Investigadores (2) / investigaciones (3)	
Paradigmática simple		3	Identificaron (5) / identificar (7)	
TOTAL		46		

10.3.2.1. Repetición mixta

El predominio de la repetición mixta de unidades fraseológicas, que constituye prácticamente la mitad de los casos de cohesión (22 de 46 casos) del texto, nos hace poner en entredicho los estudios mencionados en la revisión de la bibliografía. Estos estudios sólo consideraban la reiteración de palabras aisladas, sin prestar atención ni a las relaciones colocacionales de estas con otras palabras, ni a su posible pertenencia a unidades fraseológicas. De ahí que estimemos necesario complementar estos estudios con un análisis de la repetición sintagmática y mixta, si se quiere comprender mejor el comportamiento de la cohesión en textos especializados.

Dentro de la repetición mixta, la *repetición de unidades fraseológicas* relacionadas es mucho más frecuente que la de una unidad fraseológica y un lexema al que está asociado. Podemos explicar esto teniendo en cuenta dos factores.

En primer lugar, la lengua se sirve en muchas ocasiones de la repetición de dos UF que comparten algunos elementos léxicos para facilitar la comprensión de los textos. Esto es especialmente cierto cuando el lector es lego y la repetición funciona como un puente que ayuda al lector a asociar dos conceptos. Y así, en el texto que nos ocupa, un lector no experto podría inferir que las UF *cáncer pulmonar* y *carcinoma pulmonar* tienen el mismo significado si percibe la repetición del lexema *pulmonar* y el hecho de que estos términos tienen un cotexto muy parecido:

(67)

marcadores del riesgo de **cáncer pulmonar**
 predictor de riesgo significativo de **carcinoma pulmonar**.

Asimismo, cuando hay repetición de UF en lengua española, se tiende a mantener parte de la primera UF invariable al tiempo que se introduce alguna variación léxica y/o sintáctica en los elementos restantes de la segunda unidad fraseológica, algo que ocurre menos en lengua inglesa. Hallamos buena prueba de ello en el texto original inglés, en el que las unidades fraseológicas se repiten sin apenas ninguna modificación: *lung carcinoma – lung carcinoma; chromosome aberrations – chromosomal aberrations*.

En vista de que hay 22 instancias de repetición mixta de UF, que obviamente son casos de repetición variada, frente a 3 casos de repetición exacta sintagmática, se puede inferir que el español se decanta por la repetición variada. Esto queda también corroborado por el hecho de que los casos de repetición exacta solamente constituyen la quinta parte del total (9 de 46). No ocurría lo mismo en el texto original, donde de 134 instancias de cohesión, 43 son de repetición exacta paradigmática y sintagmática. Es decir, 32,1 % de todas las instancias de cohesión.

En cuanto a la *repetición mixta de UF y lexema*, al igual que ocurría con la repetición mixta de UF, su función es la de hacer referencia a una misma entidad, evitando la utilización del mismo término, debido a la ya mencionada preferencia de la lengua española por la variación léxica. En el texto que nos ocupa, hemos identificado cuatro instancias de este tipo de repetición, dos de las cuales presentamos a continuación en su cotexto. Nos referimos a la cohesión que se establece entre la primera y la segunda oración

y entre la segunda y la tercera oración, mediante las unidades léxicas que señaladas en negrita.

(68)

1 El **cáncer de pulmón** se asocia a alteraciones del cromosoma 9

2 <Jano Diario, Barcelona> Investigadores de la Universidad de Texas (Estados Unidos) han descubierto que ciertas alteraciones en el cromosoma 9 parecen ser marcadores del riesgo de **cáncer pulmonar** y de la agregación familiar de este tipo de **tumor**, que es precisamente el que parece ser en mayor medida inducido por factores ambientales.

3 Investigaciones previas del equipo que dirige el Dr. Xifeng Wu mostraban que las anomalías del cromosoma 9 en los linfocitos periféricos constituían un predictor de riesgo significativo de **carcinoma pulmonar**.

Este ejemplo tiene el valor adicional de que la relación entre la UF y el lexema viene acompañada y se basa en el vínculo intraoracional entre *cáncer pulmonar* y *tumor* en la segunda oración. En otras palabras, si *cáncer pulmonar* y *este tipo de tumor* tienen la misma referencia y *cáncer pulmonar* es sinónimo de *cáncer de pulmón* (primera oración) y de *carcinoma pulmonar* (tercera oración), entonces, es posible el vínculo entre estos dos últimos y el lexema *tumor*. En el sistema de la lengua, este vínculo es el de hiponimia. Sin embargo, la relación en el texto es de sinonimia. Haremos mención a estas funciones léxicas en el siguiente apartado.

En definitiva, cuando las distintas unidades léxicas de un texto son sinónimas o existe una relación de hiperonimia se produce el fenómeno de la variación léxica, del que podemos encontrar principalmente tres explicaciones. Una de ellas responde a la intención didáctica del autor que quiere acercar el conocimiento oncológico al médico de medicina general o incluso al lector lego. Otra se debe a la intención retórica de evitar una repetición excesiva de términos. Por último, en los textos de campos del saber consolidados y que cuentan con organismos normalizadores, como la medicina, hay pocas instancias de repetición paradigmática sinonímica o hiperonímica, que podrían causar ambigüedad, a excepción de los destinados al público lego. En el artículo experimental "Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary Chromosome 9 Aberrations", sus autores utilizan rigurosamente los descriptores o *Medical Subject Headings* (MeSH) del *Index Medicus* hasta el punto de incurrir en una

repetición léxica excesiva del término *chromosome aberrations*, mal menor cuando lo que se quiere lograr es la univocidad. Estos descriptores son un buen ejemplo del *vocabulario controlado* que imponen determinadas instituciones internacionales en aras de la claridad.

La traducción al español no utiliza de forma consistente el término normalizado *aberración cromosómica* por motivos retóricos y porque los destinatarios del texto no son oncólogos. Si vemos cómo se definen las alternativas propuestas en el tesauro de *Medline*, que contiene las categorías del *Index Medicus*, nos damos cuenta de que la traducción incurre en importantes imprecisiones fácilmente detectables por un experto. Como se ve en el árbol extraído del *MeSH*, el término *chromosome aberration* pertenece al subdominio de la genética y está relacionado conceptualmente con *chromosome deletion*, *mutation* y *gene deletion*.

CHROMOSOME ABERRATIONS:

Genetics

Cytogenetics

Aberrations

Chromosome Aberrations

DEFINITION OF TERM: Deviations from the normal number or structure of chromosomes, not necessarily associated with disease.

Chromosome Deletion

DEFINITION OF TERM: Actual loss of a portion of the chromosome.

Mutation

DEFINITION OF TERM: Any detectable and heritable change in the genetic material not caused by genetic segregation or recombination, which is transmitted to daughter cells and to succeeding generations, providing it is not a dominant lethal factor.

DNA Damage

Frameshift Mutation

Gene Deletion (+ 1)

DEFINITION OF TERM: A genetic rearrangement through loss of segments of DNA or RNA, bringing sequences which are normally separated into close proximity. This deletion may be detected using cytogenetic techniques and can also be inferred from the phenotype, indicating a deletion at one specific locus.

[...]

El término *aberración cromosómica* tiene el mismo significado en español: 'irregularidad del número o constitución de los cromosomas'¹²⁰. Sin embargo, el texto ofrece como alternativas términos con un significado distinto: *anomalías* o *alteraciones*.

Según el *MeSH*, *anomaly* hace referencia a malformaciones congénitas, que acaban manifestándose como una enfermedad y coocurre con partes del cuerpo (*coronary vessel*) o epónimos como *Ebstein*, *Pelger-Huet* o *Taussing-Bing* para designar defectos que afectan al sistema circulatorio. A continuación vemos qué defectos congénitos aparecen bajo *anomaly* y presentamos la posición de los dos primeros en el árbol conceptual.

ANOMALY:

Ebstein's Anomaly

Pelger-Huet Anomaly

Taussig-Bing Anomaly see Double Outlet Right Ventricle

Neonatal Diseases and Abnormalities (Non MeSH)

Abnormalities

DEFINITION OF TERM: Congenital malformations of organs or parts.

Chromosome Abnormalities

DEFINITION OF TERM: Defects in the structure or number of chromosomes resulting in structural aberrations or manifesting as disease.

Cardiovascular Abnormalities

Heart Defects, Congenital

Coronary Vessel Anomalies

Ebstein's Anomaly

DEFINITION OF TERM: Congenital downward displacement of the tricuspid valve with the septal and posterior leaflets being attached to the wall of the right ventricle.

Hemic and Lymphatic Diseases (Non MeSH)

Leukocyte Disorders (Non MeSH)

Pelger-Huet Anomaly

¹²⁰ Diccionario Enciclopédico Ilustrado de Medicina Dorland, 26 edic., 1985. Buenos Aires: Internamericana.

DEFINITION OF TERM: An inherited defect interfering with normal nuclear lobulation of neutrophils and eosinophils. The nuclei appear rodlike, spherical, or dumbbell-shaped and their structure is coarse and lumpy. (From Dorland, 27th ed)

Por último, comprobamos que *alterations* no es lo mismo que *aberrations* sino que se refiere normalmente a cambios metabólicos

ALTERATIONS:

Appetite Alterations see Appetite

Immunoglobulin Alterations see Immunoglobulins

Transcortin Bound Cortisol Alterations see Transcortin

Tryptophan Metabolism Alterations see Tryptophan

En conclusión, la repetición mixta está asociada a un intento estilístico de romper la monotonía del texto al tiempo que se mantiene la cohesión y a una intención didáctica del autor. La repetición mixta es más frecuente cuanto más divulgativo es el texto y está también asociada a UF cuyos elementos no están totalmente fijados. Un estudio de corpus fundamentado en listas de frecuencia y líneas de concordancia, como el del siguiente capítulo, llevaría a determinar qué UF están más fijadas, y por tanto, tienen una mayor aceptación en un subdominio determinado. También indicaría hasta qué punto la lengua española muestra una preferencia por la variación en la utilización de términos en relación con la lengua inglesa. La variación léxica funciona como una fuerza centrífuga que disipa los intentos de normalización por parte de los expertos de una comunidad de discurso.

10.3.2.2. Repetición exacta

A la repetición mixta le sigue en importancia la repetición exacta principalmente paradigmática, pero también sintagmática. En total, este tipo de repetición constituye una quinta parte de todos los ejemplos de cohesión. Si el texto estuviera destinado sólo a oncólogos, como ocurre con el texto original, esta proporción sería mayor. Esto es así porque en estos textos se da prioridad a la univocidad, aunque esta traiga consigo un estilo pobre desde el punto de vista léxico.

Los términos que participan en la repetición exacta sintagmática tienen una mayor fijación y son mucho más frecuentes que los que se han mencionado en el apartado de repetición mixta de UF. Esto ocurre, por ejemplo, con la UF *cáncer de pulmón*, que es el

término más extendido tanto en textos especializados como divulgativos sobre esta enfermedad.

10.3.2.3. Otros tipos de repetición

En este apartado se agrupan distintos tipos de repetición paradigmática. En concreto, hay cuatro instancias de repetición sinonímica, cuatro de sintáctica y dos de repetición paradigmática simple. Esto confirma que la cohesión se construye principalmente a partir de la repetición, bien de conceptos casi idénticos (sinonimia), bien de unidades léxicas casi idénticas (repetición paradigmática sintáctica y simple).

La repetición paradigmática sinonímica se establece entre palabras que en la lengua general tienen un significado casi idéntico (*investigación* y *estudio*) o entre palabras cuyo significado y cuyo cotexto es parecido. Esto ocurre en la segunda y tercera oración respectivamente, donde el cotexto hace que las palabras que hemos señalado en negrita, se conviertan en sinónimas:

(69)

Marcadores del riesgo de cáncer pulmonar

Predictor de riesgo significativo de carcinoma pulmonar

En la repetición paradigmática sintáctica y paradigmática simple también se percibe la tendencia a introducir variación, aunque, en este caso las diferencias entre las unidades de significación son simplemente meramente sintácticas o se basan en la elipsis (ejemplo 71).

(70) Asocia (*verbo*) – asociación (*sustantivo*)

(71) Dr. Xifeng Wu – Dr. Wu

Este último ejemplo no es el más representativo de la categoría repetición paradigmática simple porque contiene nombres propios pero ilustra cómo el principio de economía del lenguaje acorta los sintagmas que ya han aparecido previamente en el texto.

No obstante, según indica la tabla 51, que presenta relaciones en el texto más difusas, la activación de lexemas que se encuentran próximos en un marco conceptual también crea cohesión.

En suma, estos resultados parecen indicar el relevante papel de la fraseología en los textos científicos y cómo la cohesión depende no sólo de la repetición de lexemas aislados sino también de la repetición de UF relacionadas entre sí o relacionadas con otros lexemas. Lo interesante de la repetición de unidades fraseológicas es que en esta se ponen en funcionamiento simultáneamente tanto funciones léxicas cohesivas paradigmáticas como sintagmáticas, algo que analizamos a continuación.

10.3.3. Funciones léxicas

En vista de los datos que aportan las matrices de repetición, parece que las funciones léxicas que más contribuyen a la cohesión son paradigmáticas, en concreto, **Syn**, **Aequ** y **Gener**. Probablemente esto se deba a la necesidad de codificar en los textos redes conceptuales. No obstante, dado que las mismas UF se repiten con frecuencia en el texto, las funciones paradigmáticas suelen solaparse con funciones sintagmáticas¹²¹. Este solapamiento nos ha obligado a proponer nuevas categorías no contempladas en la *Teoría Sentido Texto* y una nueva forma de representación de las mismas que haga explícito si la función es paradigmática o sintagmática.

No basta con estudiar las funciones léxicas en el plano paradigmático. Es preciso también reconocer las funciones léxicas entre los elementos que componen la unidad fraseológica, algo que no ha recibido demasiada atención en la TST, donde las FL sintagmáticas normalmente relacionan sustantivos y verbos o adjetivos y sustantivos. En consecuencia, hemos propuesto nuevas FL que relacionan sustantivos entre sí y sustantivos y adjetivos. La flecha (→) indica que se trata de una función sintagmática.

Loc_A Adjetivo para hacer referencia a una entidad en la que se sitúa la entidad designada por la palabra clave. Su significado es similar al de un sintagma preposicional introducido por *de/en*.

Loc_A (alteración →) = *cromosómica*

Poss Entidad que posee otra entidad.

¹²¹ En las matrices de repetición no hemos podido representarlas por falta de espacio. En este apartado sí las explicitaremos.

Poss (*dieta* →) = *del paciente*

Loc_{In} Entidad dentro de la que se sitúa otra

Loc_{In} (*alteraciones* →) = *en el cromosoma 9*

Loc_{Poss} Entidad con respecto a la cual se sitúa otra, que en cierto modo la posee. El *valor* va normalmente precedido por la preposición que denota posesión (*of / de*).

Loc_{Poss} (*alteraciones* →) = *del cromosoma 9*

Nótese que la primera función (**Loc_A**) hace alusión a la categoría sintáctica del *valor*, es decir, de la palabra que acompaña a la palabra clave, porque, al fin y al cabo, no se puede desvincular la faceta semántica de los lexemas de la sintáctica

En el texto que nos ocupa, estas funciones normalmente forman UF bien por sí mismas o en combinación con la función sintagmática **Degrad**. Como en estas es importante el orden de sus elementos, se señalará si el valor precede o no a la palabra clave mediante el signo “→”, que fue el elegido para marcar las FL sintagmáticas. Estas diferencias en el orden de palabras son fundamentales cuando se traduce manejando las lenguas inglesa y española. En el ejemplo que sigue, se aprecia cómo en el paréntesis que encierra la palabra clave, esta va seguida de una flecha horizontal que, además, evita la confusión con funciones paradigmáticas en las que el valor también es un adjetivo (**Able**, **Qual**, **A₀**).

Degrad_S Loc_{Poss} (→ *cromosoma*) = *alteraciones del*

Loc_A (*alteración* →) = *cromosómica*

Loc_S (→ *aberration*) = *chromosome*

En los dos últimos ejemplos se ve cómo las diferencias entre el español y el inglés se encuentran solamente en el plano de la expresión, en el que se impone un orden y una categoría sintáctica determinadas, que hemos indicado con la flecha y los subíndices.

En la tabla que sigue, presentamos estas FL integradas dentro de las funciones paradigmáticas que en el texto vinculan UF entre sí y UF con lexemas.

TABLA 50: Funciones léxicas activadas

FUNCIONES LÉXICAS						EJEMPLOS (nº de oración)
Paradigmáticas	Freq	Sintagmáticas ¹²² - Freq				
		Degr _S	LOC Poss	LOC In	LOC A	
Aequ	12	6	1	1	1	Marcadores (2) – marcadores (7)
Syn – Aequ	11	26	10	6	6	Cáncer de pulmón (1) / Cáncer pulmonar (2)
Syn _▷ – Aequ	5	6	3	2	5	Alteraciones del cromosoma 9 (1) / alteraciones cromosómicas (7)
Syn	4	8	2	–	2	Cáncer de pulmón (1) / carcinoma pulmonar (6)
Syn _∩	4	–	–	–	–	Investigaciones (3) / estudio (4)
Syn _⊂	3	3	1	–	2	Tumor (2) / carcinoma pulmonar(3)
S ₀	3	–	–	–	–	Asocia (1) / asociación (4)
Syn _▷	1	1	1	–	–	Mutaciones del cromosoma 9 (6) / Alteraciones cromosómicas (7)
S ₂	1	–	–	–	–	Investigadores (2) / investigaciones (3)

Se ve un claro predominio de la FL **Aequ**, tanto cuando aparece sola como cuando se combina con **Syn** en la función **Syn-Aequ**, seguida de distintos tipos de sinonimia. La función **Aequ** crea cohesión sin dejar ningún resquicio a la ambigüedad, ya que consiste en la repetición de exactamente el mismo lexema. Un uso frecuente aunque no exhaustivo de esta función es recomendable para lograr que un texto científico sea unívoco. En la tabla, hemos identificado dos tipos de función **Aequ**.

Por una parte, hemos reconocido 12 casos en los que se produce la repetición exacta bien de un lexema, bien de todos y cada uno de los componentes de la UF. Estamos

¹²² Esta información no se encuentra en la matriz de repetición. Cada par relacionado paradigmáticamente puede llevar asociadas dos funciones léxicas paradigmáticas que indican la relación entre los elementos de la UF.

ante la función **Aequ**. Por ejemplo, *cáncer de pulmón*, que se repite en las oraciones primera y quinta, o *riesgo*, que aparece en la segunda y la tercera oración.

(72)

- 1 El cáncer de pulmón se asocia a alteraciones del cromosoma 9
 5 Identificaron a 97 pacientes con cáncer de pulmón y alteraciones en el cromosoma 9, y a continuación evaluaron a sus familiares de primer grado.

(73)

- 2 <Jano Diario, Barcelona> Investigadores de la Universidad de Texas (Estados Unidos) han descubierto que ciertas alteraciones en el cromosoma 9 parecen ser marcadores del riesgo de cáncer pulmonar y de la agregación familiar de este tipo de tumor, que es precisamente el que parece ser en mayor medida inducido por factores ambientales.
 3 Investigaciones previas del equipo que dirige el Dr. Xifeng Wu mostraban que las anomalías del cromosoma 9 en los linfocitos periféricos constituían un predictor de riesgo significativo de carcinoma pulmonar.

Por otra parte, cuando dos UF relacionadas mediante **Syn** comparten algunos de sus componentes, tal como *pulmonar*, que se repite en las UF *cáncer pulmonar* (2) y *carcinoma pulmonar* (3), hablaremos de la función **Syn-Aequ**. En el texto hay once casos de este tipo.

La relación de sinonimia, en sus diversas formas, es igual o más frecuente que **Aequ** y permite la continuidad temática del texto sin que se incurra en una repetición tediosa. Hemos concebido la sinonimia en un sentido amplio de forma que incluya las FL que mencionamos en 6.6. Recordamos estas añadiéndoles un breve comentario.

- | | |
|---------------|--|
| Syn | sinonimia total
Syn (<i>cáncer de pulmón</i>) = <i>carcinoma pulmonar</i> |
| Syn \supset | sinonimia parcial en la que el valor de la función es un término superordinado respecto a la palabra clave.
Syn \supset (<i>mutaciones del cromosoma9</i>) = <i>alteraciones cromosómicas</i> |
| Syn \subset | sinonimia parcial en la que el valor de la función es un término subordinado respecto a la palabra clave.
Syn \subset (<i>tumor</i>) = <i>carcinoma pulmonar</i> |
| Syn \cap | sinonimia parcial en la que tanto la palabra clave como el valor comparten el significado básico.
Syn \cap (<i>investigaciones</i>) = <i>estudios</i> |

Syn* – **Aequ** sinonimia de cualquier tipo (\emptyset / Gener / Specif / \cap) entre UF que compartan algunos de sus componentes.

Syn \emptyset - **Aequ** (*alteraciones en el cromosoma 9*) = *alteraciones del cromosoma 9*.

El cotexto hace que en el texto se establezcan estas funciones de sinonimia entre palabras que en la lengua general no las comparten. Por ejemplo, la UF *alteraciones del cromosoma 9* se presenta en el texto como sinónima de *anomalías del cromosoma 9*, cuando, en realidad, el significado de la unidad léxica *alteración* y *anomalía* no es el mismo.

El cotexto también es responsable de que se establezca una relación de sinonimia entre dos UF que contienen elementos entre los que hay una relación de hiponimia. Por ejemplo, si el término *carcinoma* es un término subordinado con respecto a *cáncer* – porque designa un tipo de cáncer de tejidos epiteliales – el hecho de que el cotexto que precede (*riesgo de*) o que sigue (*de pulmón/pulmonar*) a ambos términos sea similar en numerosas ocasiones, hace que el lector identifique *cáncer pulmonar* como sinónimo de *carcinoma pulmonar*.

La alternancia de las funciones **Syn** y **Aequ** en un texto, o su confluencia en la función **Syn-Aequ** va tejiendo redes de significado en las que los términos relacionados mediante **Aequ** hacen de nodos. Estos nodos son las palabras clave pertenecientes al lenguaje de especialidad, que son lexemas que pueden aparecer aislados (*cáncer*, *pulmón*, *alteraciones*, *cromosoma*, *riesgo*) o formando parte de una UF (*cáncer de pulmón*, *alteraciones en el cromosoma 9*, *carcinoma pulmonar*). La sección 7.2.3. *Esquemas de representación complementarios* contenía un marco que reflejaba los nodos de la red conceptual que se establece entre la primera oración y las demás del *abstract* en inglés. En 10.5., se verá el marco correspondiente al texto en español y se verá cómo se actualiza mediante *unidades de significación clave* que son más variadas en el caso del español.

Cuando en estas funciones paradigmáticas intervienen UF, sus componentes se relacionan también sintagmáticamente. Como se ve en la tabla, las FL que predominan son **Degr_S**, **Loc_{POSS}**, **Loc_A** y **Loc_{IN}**, para indicar una transformación con efectos negativos y localización respecto a la palabra clave, mediante la preposición *de*, un adjetivo o mediante la preposición *en*.

- (74) Alteraciones del cromosoma 9
 Degrad_S Loc_{POSS} (→ cromosoma 9) = alteraciones del
 Loc_{POSS} (alteraciones →) = del cromosoma 9
- (75) Alteraciones en el cromosoma 9
 Degrad_S Loc_{IN} (→ cromosoma 9) = alteraciones en
 Loc_{IN} (alteraciones →) = en el cromosoma 9
- (76) Anomalías del cromosoma 9
 Degrad_S Loc_{POSS} (→ cromosoma 9) = anomalías del
 Loc_{POSS} (anomalías →) = del cromosoma 9
- (77) Alteraciones cromosómicas
 Degrad_S (→ cromosómicas) = alteraciones
 Loc_A (alteraciones →) = cromosómicas
- (78) Mutaciones del cromosoma 9
 Degrad_S Loc_{POSS} (→ cromosoma 9) = mutaciones del
 Loc_{POSS} (mutaciones →) = del cromosoma 9
- (79) Cáncer de pulmón
 Degrad_S Loc_{POSS} (→ pulmón) = cáncer de
 Loc_{POSS} (cáncer →) = de pulmón
- (80) Cáncer pulmonar
 Degrad_S (→ pulmonar) = cáncer
 Loc_A (cáncer →) = pulmonar
- (81) Carcinoma_pulmonar ⇒ Degrad (pumonar) = carcinoma
 Degrad_S (→ pulmonar) = carcinoma
 Loc_A (carcinoma →) = pulmonar

Para las funciones **Poss**, **Loc_{POSS}** y **Loc_{IN}**, la preposición más frecuente, como confirman las listas de frecuencia correspondientes a corpóra escritos, es *de* o la contracción *del* para el español¹²³ y la preposición *of* para el inglés¹²⁴. En el texto que analizamos, *de* aparece en 17 ocasiones, *del*, en cinco y *en*, en seis ocasiones. Esto parece indicar un predominio de la función **Loc_{POSS}** sobre **Loc_{IN}**.

¹²³ Aunque algunos usos de la preposición *de*, por ejemplo, cuando introduce un complemento del verbo, quedarían excluidos del análisis, la elevada frecuencia de esta preposición y de la contracción *del*, tanto en corpóra del español (*CREA*, *Marcos Marín*) como en este texto ratifica esta opinión. En el corpus *CREA*, *de* es la forma más frecuente (6,3%) y *del* ocupa la décima posición (1,17%).

¹²⁴ Es la segunda forma léxica en el *British National Corpus* (3,24%) y en el *Brown Corpus* (3,02%).

Por otra parte, las FL paradigmáticas que siguen en importancia a los distintos tipos de **Syn** y a **Aequ** son **S₀** y **S₂**, asociadas a la repetición paradigmática sintáctica. La cohesión se crea entre un sustantivo o un verbo y un sustantivo que tiene la misma raíz. Se consigue también con esta función una mayor variedad estilística, aunque esta vez existe una relación morfológica y la variedad es meramente sintáctica. Este es el caso de *asocia* (verbo) y *asociación* (sustantivo).

Como conclusión podemos decir que de las funciones paradigmáticas, las más frecuentes son **Aequ** y **Syn**. Estas consiguen un equilibrio entre la repetición y la variedad. La función **Aequ** es imprescindible en la cohesión y la univocidad del texto, mientras que **Syn** permite un vocabulario más rico, con lo que se evita una excesiva monotonía en el texto. Asimismo, la función **Syn**, en textos divulgativos, clarifica términos de alta especificidad y establece así un puente entre el lenguaje de especialidad y la lengua común cuya finalidad es acercar el conocimiento experto al lector lego. En cuanto a las FL sintagmáticas, las más frecuentes son **Degr**, **Loc_{Poss}** y **Loc_{In}** y se reconocen al analizar los componentes de las UF.

10.4. Matriz de repetición de relaciones cohesivas secundarias

La tabla que sigue pone de manifiesto el hecho de que la cohesión también se construye sobre relaciones semánticas difusas. Estas se establecen, bien porque comparten algún sema (*cromosoma – heredada*), bien porque suelen aparecer cercanas en los textos sobre oncología o en el texto [*cáncer – poblaciones de alto riesgo*, (1,7)], bien porque la realidad extralingüística estudiada desde el prisma del método científico las vincula [*investigaciones – resultados*, (3,6)].

Seleccionemos ahora la casilla (2, 3), que vincula las oraciones 2 y 3.

(82)

2 <Jano Diario, Barcelona> Investigadores de la Universidad de Texas (Estados Unidos) han descubierto que ciertas alteraciones en el cromosoma 9 parecen ser marcadores del riesgo de cáncer pulmonar y de la agregación familiar de este tipo de tumor, que es precisamente el que parece ser en mayor medida inducido por factores ambientales.

3 Investigaciones previas del equipo que dirige el Dr. Xifeng Wu mostraban que las anomalías del cromosoma 9 en los linfocitos periféricos constituían un predictor de riesgo significativo de carcinoma pulmonar.

Además de las relaciones ya identificadas en la sección 10.3. *Matrices de repetición de las relaciones cohesivas primarias*, encontramos relaciones difusas de hiponimia (*agregación familiar – riesgo; factores ambientales – riesgo*) que se establecen a partir de los resultados de la investigación oncológica descrita en la noticia. Estas relaciones no son tan obvias pero contribuyen al entramado conceptual que recrea el lector al leer el texto. Si en la matriz anterior había 46 conexiones, aquí encontramos 74 nuevas conexiones, que juntas computan 120 instancias de cohesión.

Queda de manifiesto que un estudio de la cohesión fundamentado sólo en la repetición léxica (exacta, simple, variada sintáctica) sólo plasma una mínima parte de esa cohesión, 23 de 120 instancias (19%). Si restringimos la cohesión a las que hemos denominado relaciones cohesivas primarias, estos tres tipos de repetición contribuyen a un 50 % de la cohesión del texto. El resto de recursos cohesivos primarios identificados en el apartado anterior son menos explícitos y computan también un 19% o un 50%, según se entienda la cohesión de una forma más o menos restringida.

De nuevo, este porcentaje es bastante menor que el que proponen Benbrahim, Berber Sardinha y Ellman.

TABLA 51: [Matriz de repetición de las relaciones cohesivas secundarias](#)

10.5. Hacia una metodología empírica para la representación de marcos

A partir de las listas de frecuencia (tablas 40-43), identificamos las unidades léxicas que, por ser más frecuentes en el texto, son los nodos sobre los que se actualizan marcos conceptuales, formando redes cohesivas. El resto de unidades léxicas se vinculan con estos nodos, que podemos denominar *unidades de significación clave*, mediante funciones paradigmáticas y sintagmáticas.

Al agrupar las USC vinculadas mediante la función **Syn** o **Syn-Aequ**, cada uno de estos grupos hará mención a un concepto que se etiquetará con la USC más frecuente del grupo. Para cada concepto, especificaremos qué unidades léxicas y fraseológicas lo actualizan y cuántas veces aparecen estas en el texto.

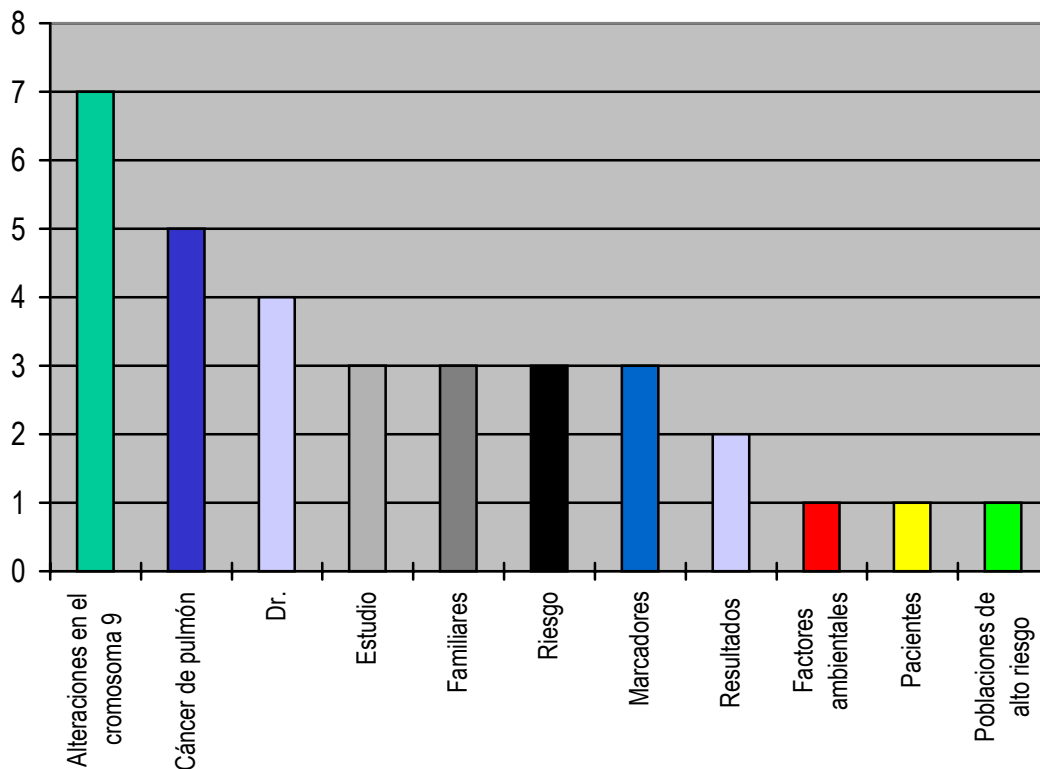
Hemos representado este tipo de análisis en una tabla en la que se destacan los conceptos más activados en el texto, los términos que los activan y la frecuencia absoluta y relativa de estos términos.

TABLA 52: Conceptos más activados en el texto

Concepto	Término	Frec. Absoluta	Frec. relativa
ALTERACIONES EN EL CROMOSOMA 9	Alteraciones en el cromosoma 9	2	1,01
	Alteraciones del cromosoma 9	1	0,5
	Alteraciones cromosómicas	1	0,5
	Anomalías del cromosoma 9	1	0,5
	Inestabilidad genética	1	0,5
	Mutaciones del cromosoma 9	1	0,5
	TOTAL	7	3,52
CÁNCER DE PULMÓN	Cáncer de pulmón	2	1,01
	Carcinoma pulmonar	2	1,01
	Cáncer pulmonar	1	0,5
	Tumor	1	0,5
	TOTAL	6	3,02
DR.	Dr. + Nombre	2	1,01
	Investigadores	1	0,5
	Equipo	1	0,5
	TOTAL	4	2,01
ESTUDIO	Estudio	2	1,01
	Investigaciones	1	0,5
	TOTAL	3	1,51
FAMILIARES	Familiares	2	1,01
	Familiar	1	0,5
	TOTAL	3	1,51
RIESGO	Riesgo	3	1,51
	TOTAL	3	1,51
MARCADORES	Marcadores	2	1,01
	Predictor	1	0,5
	TOTAL	3	1,51
RESULTADOS	Resultados	2	1,01
FACTORES AMBIENTALES	Factores ambientales	1	0,5
PACIENTES	Pacientes	1	0,5
POBLACIONES DE ALTO RIESGO	Poblaciones de alto riesgo	1	0,5

El total de alusiones a cada concepto en comparación con otros conceptos nos puede indicar su peso específico en un marco conceptual que podríamos denominar *ETIOLOGÍA DEL CÁNCER DE PULMÓN (CAUSAS GENÉTICAS)*. La representación gráfica de este peso específico aparece en el diagrama 15.

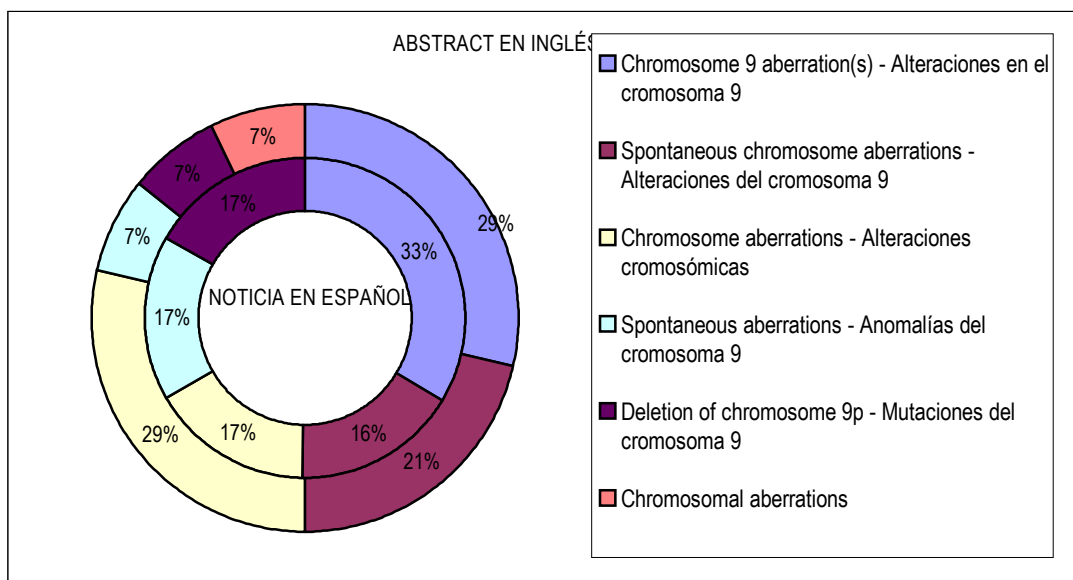
DIAGRAMA 16: Propuesta de representación del marco conceptual activado en el texto



En este diagrama de barras destacan los conceptos ALTERACIONES EN EL CROMOSOMA 9, CÁNCER DE PULMÓN y DR. como los más activados. Le siguen en importancia, por una parte, los conceptos ESTUDIO y RESULTADOS, que pertenecen a la cadena léxica que podríamos denominar *INVESTIGACIÓN*, y por la otra, tres conceptos muy frecuentes en el macromarco CÁNCER: FAMILIARES, RIESGO y MARCADORES. Por último, aparecen otros conceptos del mismo macromarco: FACTORES AMBIENTALES, PACIENTES y POBLACIONES DE ALTO RIESGO.

Para los dos conceptos más frecuentes en el marco, *alteraciones en el cromosoma 9* y *cáncer*, hemos elaborado dos diagramas de sectores en los que se desglosan y ordenan según su frecuencia los términos que hacen referencia a cada concepto en el TO y en el TT. En los anillos exteriores están los términos en inglés y en los interiores, los términos en español, que no necesariamente son su traducción. Se pone de manifiesto la ya mencionada mayor variedad terminológica del texto en español con respecto al texto en inglés.

DIAGRAMA 17: Realizaciones del concepto ALTERACIONES EN EL CROMOSOMA 9

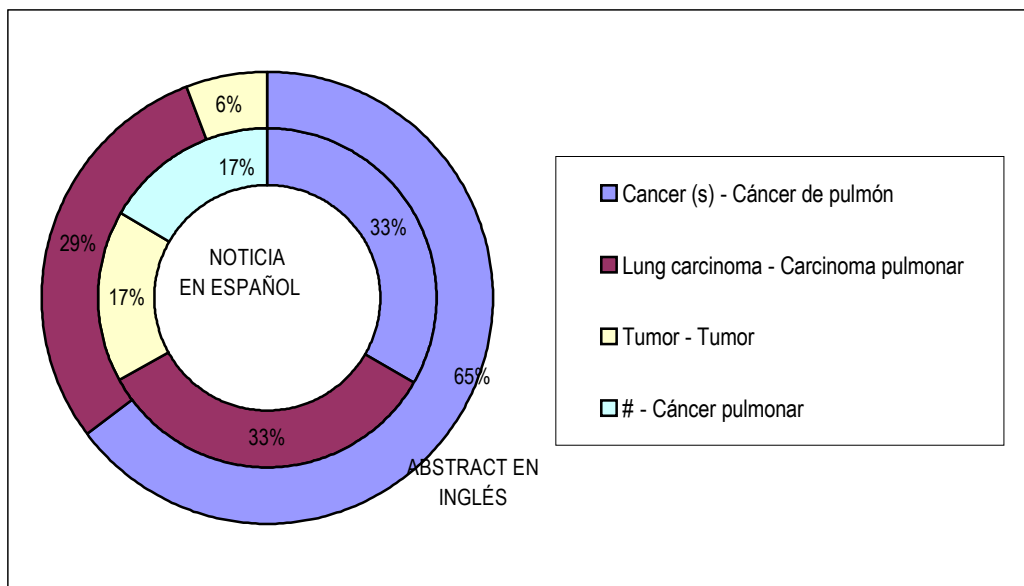


Aunque el *abstract* tenga una mayor complejidad conceptual, como cabe esperar del hecho de que está destinado a oncólogos, sólo se utiliza una forma más para designar el concepto. Además, las diferentes formas se construyen sobre la unidad terminológica *chromosome aberrations*, con la excepción de *deletion*, que, según la jerarquía del *MeSH*, es un término subordinado de *chromosome aberration*.

En cuanto a la noticia en español, encontramos no sólo variaciones sintácticas sobre la UF *alteraciones en el cromosoma 9*, sino también, otras formas léxicas de significado diferente al concepto CHROMOSOME ABERRATION.

En el diagrama 18, observamos de nuevo el mismo fenómeno. El concepto designado inglés con el término *lung cancer*, se actualiza en español con tres variantes terminológicas: *cáncer de pulmón*, *carcinoma pulmonar* y *cáncer pulmonar*.

DIAGRAMA 18: Realizaciones del concepto CÁNCER



Con estos tipos de diagrama, un simple golpe de vista puede orientar al traductor sobre qué términos de un marco conceptual se explotan más. Para ello, deberá identificar en la matriz de repetición correspondiente a cada lengua aquellas unidades léxicas relacionadas mediante la FL **Syn** o **Syn-Aequ**, que presumiblemente apuntan al mismo concepto, y por tanto, irán juntas en el mismo sector del diagrama. Después, para que el traductor sepa qué términos son los más usuales para designar un concepto en cada lengua, exportará a un diagrama de sectores los datos que sobre esos términos aportan las listas de frecuencia elaboradas con *Wordsmith Tools*.

Comparando los diagramas de sectores para la lengua inglesa y española, el traductor se cerciora de que la terminología de la traducción es la apropiada. Si además de los diagramas resultantes del análisis intensivo se elaboran diagramas a partir de las cadenas léxicas de un corpus textual extenso, la validez del análisis es aún mayor. De ahí que en el capítulo siguiente, nos adentremos en el análisis extensivo de la cohesión.

11. Análisis extensivo comparativo de textos para especialistas y textos para pacientes

En este capítulo se someterán los resúmenes del servicio *Physician Data Query* de *CancerNet* sobre cáncer de pulmón a la metodología extensiva descrita en 8.2. Se han etiquetado todos los textos destinados a oncólogos con las letras QD, a las que se añadirá una T, si los textos son sobre tratamiento, o una C si se trata de cuidados médicos. Las siglas QDT y QDC se corresponderán respectivamente con los apartados *Treatment Summaries for Health Professionals* y *Supportive Care Summaries for Health Professionals* de la página web del *National Cancer Institute*. Los textos destinados a pacientes llevan el nombre QPT y QPC y han sido extraídos de las secciones *Treatment Summaries for Patients* y *Supportive Care Summaries for Patients*.

Comentaremos *grosso modo* los resultados de todos los textos, a excepción de los dos textos sobre mesotelioma maligno: QDT1 y QPT1, cuyas cadenas léxicas analizaremos en detalle. En cuanto a los textos que versan sobre cuidados médicos, vamos a estudiar la cohesión estrictamente a partir de los datos estadísticos proporcionados por *Hesperus* y *Wordsmith Tools*.

11.1. Textos para especialistas sobre tratamiento del cáncer de pulmón

Los tres textos sobre tratamiento, que incluimos en el Apéndice II (textos 9, 10 y 11) se centran en tres tipos de cáncer que afectan al pulmón: mesotelioma (QDT1), cáncer no microcítico de pulmón (QDT2) y cáncer microcítico de pulmón (QDT3).

11.1.1. Mesothelioma (QDT1)

11.1.1.1. ÍNDICE DEL TEXTO

El índice nos indica que los conceptos que se activarán con más probabilidad son CELLULAR CLASSIFICATION, STAGE (LOCALIZED, ADVANCED AND RECURRENT MALIGNANT MESOTHELIOMA) y TREATMENT OPTIONS. Esto nos induce a pensar que algunas de las cadenas léxicas van a girar en torno a los mismos y van a incluir, por una parte, términos biomédicos (*epithelial*), y por otra, vocablos de la lengua general muy frecuentes en textos científicos (*classification*).

(83)

GENERAL INFORMATION

CELLULAR CLASSIFICATION

STAGE INFORMATION

Localized malignant mesothelioma

Advanced malignant mesothelioma

TREATMENT OPTION OVERVIEW

LOCALIZED MALIGNANT MESOTHELIOMA (STAGE I)

ADVANCED MALIGNANT MESOTHELIOMA (STAGES II, III, AND IV)

RECURRENT MALIGNANT MESOTHELIOMA

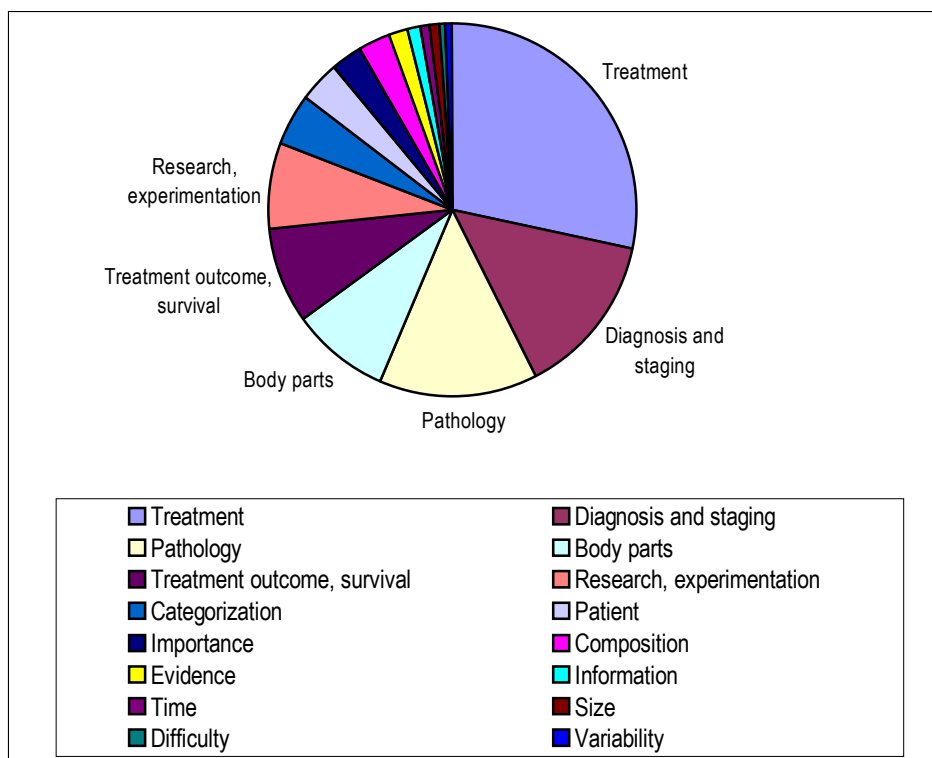
11.1.1.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTA DE FRECUENCIA LEMATIZADA

Después de observar la lista de frecuencia lematizada del texto y agrupar los lexemas con una frecuencia relativa mayor o igual al 0,1%, hemos identificado 17 áreas conceptuales, representados en una tabla y un gráfico.

TABLA 53: Identificación preliminar de áreas conceptuales activadas en QDT1 a partir de lista de frecuencia lematizada

DT1	
ÁREAS CONCEPTUALES	%
Treatment	11,5
Diagnosis and staging	5,78
Pathology	5,66
Body parts	3,54
Treatment outcome, survival	3,37
Research, experimentation	3,08
Categorization	1,78
Patient	1,48
Importance	1,14
Composition	1,12
Evidence	0,66
Information	0,47
Time	0,38
Size	0,3
Difficulty	0,28
Variability	0,19
TOTAL	40,9

DIAGRAMA 19: Representación preliminar de las áreas conceptuales activadas en QDT1



11.1.1.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

El programa *Hesperus* reconoce en el texto 15 cadenas léxicas, cuya versión sin modificar aparece en el Apéndice VIIe. Aporta también una serie de datos sobre el texto. En primer lugar, de las 1079 palabras del texto, 505 son *palabras plenas*, responsables de crear cohesión léxica mediante 141 categorías conceptuales del tesoro *Roget*. En segundo lugar, *Hesperus* atribuye al texto un valor de 10536, que, al dividirlo entre el número total de palabras¹²⁵ proporciona un valor relativo, que llamaremos *coeficiente de cohesión*, de 9,76.

Las categorías conceptuales que contribuyen a crear más del 1% de la cohesión del texto son las siguientes:

¹²⁵ Como se ha explicado con anterioridad, este valor es dependiente de la extensión del texto, por lo que para no distorsionar la información que proporciona sobre la fuerza de las cadenas y su extensión, siguiendo el consejo del profesor Ellman, lo hemos dividido entre el número total de palabras.

therapy_658_4268_n Percent: 10.05,
 change_143_892_n Percent: 6.36,
 stand_218_1417_n Percent: 6.11, (stage, staging)
 durability_113_697_n Percent: 3.67,
 disease_651_4187_n Percent: 3.23,
 medical_658_4274_a Percent: 3.03,
 laterality_239_1581_n Percent: 2.90, (pleura, pleural)
 sick_651_4205_a Percent: 2.85,
 exposure_661_4742_n Percent: 2.85,
 microscopy_196_1275_n Percent: 2.85,
 time_108_672_n Percent: 2.37,
 cancer_651_4195_n Percent: 2.18,
 sick-person_651_4202_n Percent: 2.14,
 illness_651_4186_n Percent: 2.03,
 arrangement_62_427_n Percent: 1.60,
 sullen_892_5741_a Percent: 1.56,
 study_536_3481_n Percent: 1.43,
 testing-agent_461_2962_n Percent: 1.42,
 doer_676_4377_n Percent: 1.42,
 remedial_658_4273_a Percent: 1.33,
 lateral_239_1582_a Percent: 1.21,
 place_187_1178_v Percent: 1.19,
 worry_824_5250_n Percent: 1.19,
 great_32_202_a Percent: 1.15,
 period_110_686_n Percent: 1.14,
 extrinsicity_6_38_n Percent: 1.14,
 frame_218_1425_n Percent: 1.06,
 circumstance_8_48_n Percent: 1.03,

Estas categorías son el nexo de unión entre los distintos lexemas que constituyen una cadena léxica. Veamos cómo estas categorías tejen las 15 cadenas identificadas por *Hesperus*, su porcentaje de activación en el texto y posibles asociaciones entre ellas. Comentaremos algunos fallos cometidos en la identificación de cadenas y, en la medida de lo posible, propondremos una etiqueta semántica para cada cadena.

TABLA 54: Cadenas 1, 2a y 3a (QDT1)

CONCEPTOS ACTIVADOS					
Cadena 1 (<i>Prognosis</i>)				Cadena 2a (<i>Stage</i>) y 3a (<i>Treatment</i>)	
THERAPY	10,05	DIAGNOSTIC	0,72	SHOW (<i>Staging</i> ¹²⁶)	5,69
DISEASE	3,23	PATHOLOGY	0,57	DRAMATURGY	0,53
MEDICAL	3,03	ULCER	0,57	CHANGE	6,36
SICK	2,85	SURGERY	0,34	MODIFY	0,65
CANCER	2,18	REMEDY (<i>noun</i>)	0,22		
SICK-PERSON	2,14	INFECTION	0,12		
ILLNESS	2,03	MEDICAL-ART	0 0,1		
REMEDY (<i>verb</i>)	0,98	TOTAL	29,03+	TOTAL	13,23
↓		↓		↓	
TREATMENT AND DIAGNOSIS PATHOLOGY OF MALIGNANT MESOTHELIOMA				STAGING, TREATMENT	

Comentario

A la vista de la tabla conceptual, es posible reconocer en la primera cadena un área semántica bastante amplia, que en otros textos aparece desglosada en dos: a) TREATMENT AND DIAGNOSIS y b) PATHOLOGY OF LUNG CANCER. De ahí que en nuestra propuesta definitiva de cadenas léxicas (11.1.1.4.) se contemple esta distinción. Al mismo tiempo, tendremos que integrar en esta primera cadena dos términos esenciales en el área de la oncología, en concreto, *stage* y *treatment*, que son los términos más frecuentes en las cadenas 2 y 3 respectivamente. También, se incluirán términos para designar pruebas diagnósticas, entre otros, *special stains*, que hace referencia a las sustancias colorantes que se emplean en el diagnóstico del cáncer. Dada la importancia de las cadenas 1, 2a y 3a, en nuestra propuesta definitiva de cadenas léxicas, se derivarán de ellas seis cadenas con sus respectivas subcadenas:

Chain 0: TREATMENT

Chain 2: PATHOLOGY OF MALIGNANT MESOTHELIOMA

Chain 3: DIAGNOSIS

Chain 6: STAGING

Chain 7: PATIENT

Chain 12: GENERAL MEDICINE

TABLA 55: Cadena 7 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 7 (Asbestos exposure)	
REFRIGERATION	2,85
INCOMBUSTIBILITY	0 0,1
TOTAL	2,85
↓	
RISK FACTOR	

Comentario:

La cadena 7 está muy relacionada con la primera en el sentido de que incluye el principal factor de riesgo del mesotelioma, la exposición al asbesto (*asbestos exposure*). Esta información no se deriva del tesoro, sino de nuestro conocimiento sobre oncología. En consecuencia, la etiqueta semántica que proponemos no coincide con la etiqueta REFRIGERATION. No obstante, *Hesperus* sí ha acertado al asociar los lexemas *asbestos* y *exposure*, que normalmente forman una *colocación* en textos sobre el mesotelioma maligno.

TABLA 56: Cadenas 0 y 8 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN			
Cadena 0 (Trials)		Cadena 8 (Electron microscopy)	
STUDY	1,43	MICROSCOPY	2,85
EXPERIMENT	0,72	ELEMENT	0 0,1
		MINUTENESS	0 0,1
TOTAL	2,15	TOTAL	2,85+
↓		↓	
RESEARCH, EXPERIMENTATION		LABORATORY EQUIPMENT	

¹²⁶ La importancia de este concepto se debe en cierta medida a que *stage*, que en este campo no tiene nada que ver con representaciones teatrales, aparecía en esta cadena asociado a SHOW y a DRAMATURGY (0,53 %).

Comentario:

La cadena 0 activa los conceptos STUDY y EXPERIMENT, centrales en el método científico, a los que *Hesperus* atribuye un porcentaje de 2,15. Por este motivo, hemos etiquetado esta cadena RESEARCH, EXPERIMENTATION. Los términos de esta cadena están muy relacionados con la cadena 8, que menciona algunos accesorios del instrumental de laboratorio. En esta cadena aparece *electron microscopy*, un instrumento básico en los análisis clínicos que se realizan al paciente en el diagnóstico y prognosis de su enfermedad, por lo que también se relacionará esta cadena con la ya anteriormente citada de TREATMENT AND DIAGNOSIS.

TABLA 57: Cadenas 2b, 4 y 9 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 2b (Improvement)		Cadena 4 (Survival)		Cadena 9 (Time)	
RESTORED	0,23	DURABILITY	3,7	TIME	2,37
RECUPERATION	0 0,1	PROTRACTED	0,7	PERIOD	1,14
SUPPORT	0 0,1	LIFE	0,66	ERA	0 0,1
		LASTING	0,15		
TOTAL	0,23+	TOTAL	5,21	TOTAL	3,51+
↓		↓		↓	
IMPROVEMENT		SURVIVAL		TIME	

Comentario

La cadena 2b relaciona lexemas que apuntan al concepto RESTORATION¹²⁷ con el sentido de ‘mejoría’ y ‘recuperación’ y lexemas asociados al concepto SUPPORT¹²⁸. A pesar de que en el análisis automático esta cadena no tiene demasiada representación, parece obvio que el objetivo de la medicina es *proporcionar* al paciente los medios para que *recupere* la salud.

La cadena 4 es complementaria de la 9, que gira en torno a la idea de TIEMPO. Representa el esfuerzo de la medicina por prolongar la vida en una enfermedad tan letal como el cáncer. Aunque *Hesperus* no suele identificar colocaciones, en esta cadena vincula los lexemas *long-term* y *survival*, que coocurren en el texto tres veces.

¹²⁷ Categoría 656 del Tesoro *Roget*.

(84)

1 ed patients, providing long-term survival without cure. In large r
 2 sociated with improved long-term survival include epithelial histo
 3 tients will experience long-term survival with aggressive treatmen

En cuanto a la cadena 9, en ella se perfila la lucha contra el tiempo para frenar el avance del cáncer tal y como indican las unidades léxicas hiladas en torno a los conceptos TIME, PERIOD y ERA.

TABLA 58: Cadena 3b (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 3b (Variability)	
CHANGE	6,36
MODIFY	0,65
TOTAL	7,01
↓	
CHANGE	

Comentario

La cadena 3b, que es la que se obtiene al dejar a un lado las unidades de significación clave *treatment* y *special stains*, contiene una serie de lexemas que hacen referencia al concepto CHANGE. No obstante, los conceptos CHANGE—al que *Hesperus* concede un 6,36 por ciento—y el concepto dependiente de CHANGE, MODIFY (0,65%), se materializan principalmente por los lexemas relacionados con el tratamiento de la enfermedad.

TABLA 59: Cadena 5 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 5 (Pleura)	
LATERALITY	2,9
LATERAL	1,21
TOTAL	4,11
↓	
PLEURA	

¹²⁸ Categoría 218 del Tesoro *Roget*

Comentario

En esta cadena, el término *pleura* representa la parte del cuerpo más afectada por el mesotelioma maligno. No obstante, será preciso ampliar esta cadena para incluir otras partes del cuerpo afectadas, algunas de las cuales han sido ubicadas fortuitamente como resultado del análisis automático en la cadena 11.

TABLA 60: Cadenas 12, 13 y 14 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 12 (<i>High</i>)		Cadena 13 (<i>Include</i>)		Cadena 14 (<i>System</i>)	
GREAT	1,15	JOINT	0,74	ARRANGEMENT	1,6
INCREASE	0,27	WHOLE	0,65	CLASS	0,33
CONSUMMATE	0 0,1	UNION	0,40	STRUCTURE	0,32
LARGE	0 0,1	PARTITION	0,37	DEGREE	0 0,1
GROW	0 0,1	JOINED	0,35	COMPARATIVE	0 0,1
		COMPOSING	0,11		
		ADD	0 0,1		
		CONTAIN	0 0,1		
		COMPRISE	0 0,1		
TOTAL	1,42 +	TOTAL	2,62 +	TOTAL	2,25 +
↓		↓		↓	
SIZE, INCREASE		PART-WHOLE COMPOSITION		ORDER, CATEGORIZATION	

Comentario

Los conceptos predominantes en la cadena 12, GREAT y INCREASE¹²⁹, presentan cómo en la terapia del cáncer se lucha contra el crecimiento del tumor. El concepto INCREASE está pues marcado negativamente desde el punto de vista axiológico en el subdominio de la oncología, a pesar de que normalmente se considere un valor positivo. Veamos esto en el cotexto de algunos lexemas de la cadena.

(85) Local **extension** of disease into the following: chest wall or mediastinum

The tumor **grows** through the diaphragm

Examination of the **gross** tumor at surgery

¹²⁹ El concepto INCREASE también está relacionado con la idea de cambio.

En la cadena 13, destaca la idea de TOTALIDAD / PARCIALIDAD y la idea de COMPOSICIÓN. También aparecen en la misma los lexemas *node(s)*, *combination(s)* y *combined*, cuyo significado queda restringido en textos sobre oncología. *Node(s)* designa los ganglios linfáticos y los otros dos lexemas, tratamientos de quimioterapia en los que se utilizan varios medicamentos a la vez. Incluiremos entonces estos lexemas en otras cadenas: la de PARTES DEL CUERPO y la de TRATAMIENTO.

En la cadena 14, se encadenan lexemas frecuentes en textos sobre investigación, en los que se sistematiza el conocimiento mediante una serie de categorías. Obviamente estos lexemas adquieren un sentido más restringido en textos sobre oncología, dado que coocurren con otros términos como es el caso de *cellular classification* o *staging system*, que presentamos en concordancias a continuación:

(86)

1 limited [1]. A proposed staging system based upon thoracic surger
 2 t is a modification of the older system proposed by Butchart et al
 3 utchart et al [3]. Other staging systems that have been employed,
 4 ed new international TNM staging system, are summarized by the Int

TABLA 61: Cadenas difusas: 6, 10 y 11 (QDT1)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 6 (<i>Difficult</i>)		Cadena 10 (<i>Margins</i>)		Cadena 11 (<i>Agent</i>)	
MOOT	0,65	NUMERICAL	1,51	EXAMPLE	0,81
EXAM	0,43	READING-MATTER	0,52	INTERMEDIARY	0,59
ENQUIRY	0,41	EDITION	0,15	MIDDLE	0,12
DIFFICULTY	0,10	NUMERICAL-ELEMENT	0,14	INTERJACENCY	0,1
UNCERTAINTY	0,1	TEXT-BOOK	0,1	MIDDLING	0,1
UNINTELLIGIBLE	0,1				
TOTAL	1,59 +	TOTAL	2,32 +	TOTAL	1,52 +
↓		↓		↓	
DIFFICULTY?		?		?	

Comentario

Estas tres cadenas pertenecen a una serie de cadenas que *Hesperus* ha formado con poco acierto, dado que un ordenador no es capaz de solucionar la polisemia de un lexema

mediante el contexto. No obstante, algunos de estos lexemas hacen alusión al método científico en general y, en consecuencia, pueden estar presentes en otros textos científicos cuya temática no sea la oncología. Por su frecuencia en la sección del corpus dedicada a especialistas y su pertinencia en el campo de especialidad de las ciencias empíricas y de la medicina en general, opinamos que estos lexemas de la lengua general son también unidades de significación especializada.

En la cadena 6 aparecen como lexemas de la lengua general términos del dominio de la oncología y la medicina. Este es el caso de los lexemas *acid* (*Acid-Schiff*), y *examination*, que en las cadenas léxicas definitivas quedarán asociados al concepto de TREATMENT y DIAGNOSIS. Esta cadena también contiene una asociación curiosa y disparatada entre *operative*, perteneciente a la categoría AGENCY¹³⁰, y *discussion*, perteneciente a la categoría INQUIRY¹³¹. Esta conexión se establece a través de la subdimensión de INQUIRY denominada *secret service*, donde aparecen *double agent* y *undercover agent*, que *Hesperus* asocia erróneamente a la idea de AGENCY. Se pone de relieve de nuevo el problema de la ambigüedad léxica y las dificultades de los ordenadores para detectarla. En cuanto al resto de elementos de esta cadena, constituyen una combinación bastante dispar de conceptos pertenecientes a las secciones del *Roget* INTELLECT (FORMATION AND COMMUNICATION OF IDEAS) y VOLITION (INDIVIDUAL VOLITION, *Antagonism*). A la primera sección pertenecen las categorías UNCERTAINTY (474), INQUIRY (459) y UNINTELLIGIBILITY (517), mientras que a la segunda pertenece DIFFICULTY (700), que realmente tiene importantes puntos de contacto con INTELLECT.

La cadena 10 contiene también lexemas frecuentes en las ciencias empíricas, aunque se construye sobre conexiones bastante difusas y poco acertadas. Esto queda patente en la representación que hace *Hesperus* de la cadena, en la que predominan las relaciones de *grupo* y *categoría* sobre las de *identidad*. Una de las causas de la dispersión semántica de la cadena radica en el hecho de que *margins* en textos sobre oncología significa ‘*the area or normal tissue remaining after a malignant tumor has been removed*’¹³². No tiene, por tanto, nada que ver con la edición de un libro ni con *series*, que,

¹³⁰ Categoría 173 del *Roget's thesaurus*

¹³¹ 459 del *Roget's thesaurus*

¹³² *The Cancer Dictionary*

por otra parte, tampoco está relacionada con la transmisión de ideas, sino con un grupo de pacientes que se somete a un estudio.

En vista de esta desviación de significado, interpretaremos los elementos de la cadena 10 en función de su cotexto. *Factors* hace referencia a las *circunstancias* que contribuyen a la mejoría del paciente y que tienen que ser evaluadas por el oncólogo. Por tanto, este lexema formará parte de las cadenas que denominaremos COGNITIVE PROCESSES IN DIAGNOSIS y EVALUATION. Esta interpretación parece la más correcta teniendo en cuenta que CIRCUMSTANCE tiene asociado un 1,03 %. Hay dos lexemas que también están relacionados con los pacientes que se han seleccionado como muestra: *selection* y *selected*, que añadiremos a la cadena PATIENTS. Por último, *description(s)* tiene una referencia endofórica a otros apartados del texto y constituye un elemento esencial del método científico.

La cadena 11, que se construye sobre la idea CENTRALIDAD, se convierte en difusa cuando se eliminan los lexemas que hacen referencia a partes del cuerpo con una localización central (*diaphragm, heart, wall*). El resto forma parte de la descripción de procesos desde la óptica del método científico (*median, adequate*), los casos sobre los que se fundamenta (*specimens, instances*) y sus agentes (*agents*), aunque en este último caso el sentido de *agents* se restringe al ámbito de la QUIMIOTERAPIA, a los fármacos utilizados en el tratamiento. Otro caso de restricción de significados lo encontramos en *intervention* que coocurre con *surgical*, con lo que hace mención a la CIRUGÍA. Tomaremos buena nota de esto a la hora de elaborar las cadenas léxicas definitivas.

11.1.1.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

A raíz de las tres fuentes de las que nos hemos servido para identificar la cohesión del texto, se hará una propuesta de cadenas cohesivas. Se destacarán antes de nada las cuatro cadenas identificadas en los tres apartados previos:

- TREATMENT
- STAGING
- PATHOLOGY
- CATEGORIZATION / CLASSIFICATION.

A estas se añaden las que han sido detectadas por dos o por alguno de los medios anteriormente descritos:

- DIAGNOSIS
- BODY PARTS
- RESULT OF TREATMENT/ IMPROVEMENT / SURVIVAL
- RESEARCH / EXPERIMENTATION
- PART / WHOLE
- INFORMATION
- TIME
- SIZE / SMALL
- DIFFICULTY
- VARIABILITY
- LABORATORY EQUIPMENT
- PATIENTS
- IMPORTANCE
- EVIDENCE
- RISK FACTOR

Por último, hemos creado una cadena OPTIONS que cuenta con 12 elementos y que tiene una representación conceptual de 0,16%, según el *perfil de Hesperus*.

La tabla 63 anticipa las cadenas y subcadenas que subyacen el significado del texto. Entre paréntesis indicamos la(s) cadena(s) de *Hesperus* de la(s) que se derivan, precedida(s) de la letra H y marcamos con negrita que la contribución de *Hesperus* a la cadena no es anecdótica. Cuando *Hesperus* no ha reconocido cadena, especificamos, si es el caso, el porcentaje de activación que concede al CONCEPTO relacionado con la cadena creada por nosotros. Entre corchetes se especifica el número de componentes de cada cadena.

Posteriormente, estas cadenas quedarán representadas en forma de tablas. Cada casilla contiene un lema (véase definición de lema en p. 226), que va en mayúscula, y sus correspondientes formas léxicas, ordenadas en orden de frecuencia absoluta. En estas casillas aparecerán los siguientes signos:

(i) ~

Sustituye a la forma léxica que hace de entrada léxica en las casillas. Puede aparecer como base o colocador para indicar colocaciones. En la tabla, sustituye al lexema *trials*:

TRIALS (Clinical ~)	5
AGGRESIVE ~ surgical approaches 2 ~ treatment approaches 1	3

(ii) →TREATMENT

El signo '→' acompaña al nombre de algunas subcadenas como 0.2. Properties (→ TREATMENT). Indica que los lexemas de ese apartado forman colocación con vocablos que designan el concepto TREATMENT. En dicho apartado, la presencia de este signo indica que *palliative*, *standard*, *aggressive* y *postoperative* preceden frecuentemente a TREATMENT, y por eso se mencionan en el apartado TREATMENT.

0.2. Properties (→ TREATMENT)

PALLIATIVE	5
STANDARD ~ treatment 2	4
AGGRESSIVE ~ surgical approaches 2	3

~ treatment approaches 1	
MULTIMODALITY (See11.)	2
POSTOPERATIVE (See 0.4.2.)	1
TOTAL	15

(iii) →

Indica la posición que ocupa el concepto que hila la cadena en relación al lema que sirve de entrada. En el ejemplo, el concepto sería TREATMENT. Como la flecha precede a *alone*, las referencias al concepto TREATMENT (*radiotherapy*, etc.) siempre irán delante de *alone*. En el caso de *adjuvant* ocurre lo contrario.

0.2. Properties (→ TREATMENT →)

→ ALONE Radiotherapy ~ 12 Surgery alone ~ 4 Chemotherapy ~ 2	18
ADJUVANT → ~ chemotherapy 6	12

(iv) ✖ número

Indica que el lema en cuestión proviene de la cadena de *Hesperus* indicada por el número.

COMBINATION(✖ 13) Combination 3	3
------------------------------------	---

(v) LEXEMA + LEXEMA EN VERSALES

Se utiliza cuando la entrada puede formar colocación con diversos lexemas que designan el mismo concepto (en versales). Por ejemplo, en el texto QDT1 hay cuatro UF formadas sobre el lema INTRACAVITARY: *intracavitary therapy* (2), *intracavitary chemotherapy*, *intracavitary mesothelioma*. Las representamos en una casilla de la cadena *LOCATION IN THE HUMAN BODY*, en la que se indica que el lema va seguido de unidades léxicas relacionadas con los conceptos THERAPY y DISEASE.

5.4.2.1. Chest cavity

INTRACAVITARY (+ THERAPY / + DISEASE) (See 0.2.1., 2.2.1.1, 6.1.)	4
---	---

(vi) [LEXEMA](#)

El lema [INTRACAVITARY](#) aparece como un vínculo de hipertexto porque también participa en la construcción de las cadenas *TREATMENT* y *PATHOLOGY OF MALIGNANT MESOTHELIOMA*. Como *intracavitary* designa una cavidad del cuerpo, al vincularse sintagmáticamente con los lexemas *therapy* y *chemotherapy*, que apuntan al concepto THERAPY, y con *mesothelioma* (DISEASE) restringen su significado en relación a la dimensión *LOCATION IN THE HUMAN BODY*.

Los vínculos de hipertexto presentes en algunas entradas indican que estas pueden pertenecer a más de una cadena porque su significado hace referencia al subdominio de la oncología y al mismo tiempo a un dominio más amplio. Este es el caso de [TRIAL](#), que se vincula a la cadena *TREATMENT* por ser un tratamiento experimental y también a la cadena *RESEARCH / EXPERIMENTATION*.

De estos vínculos, aquellos que contribuyen a la clasificación de un concepto desde más de una perspectiva testifican la multidimensionalidad típica del subdominio de la oncología y de la medicina en general. Entendemos la multidimensionalidad en el sentido que le dan Bowker y Meyer (1993).

... a phenomenon of classification that arises when a concept type can be subclassified in more than one way (i.e., in more than one *dimension*), depending on the conceptual characteristic that is used for the classification.

La multidimensionalidad se ve propiciada por la ingente activación de relaciones cohesivas sintagmáticas, producto de la nominalización del lenguaje de la ciencia. En efecto, cuando coocurren diversos sustantivos uno de ellos restringe el significado del otro

Como sostiene Bowker (1997: 137), las diferencias en la clasificación conceptual que crean multidimensionalidad son el resultado de diferentes exigencias sociales, diferentes niveles de percepción y cognición y, a veces, de una deficiente comprensión de las relaciones existentes entre fenómenos naturales. Luego, al estructurar el dominio de la oncología, vamos a encontrar multidimensionalidad que responde a diversos niveles de conocimiento sobre el tema (experto-lego) a diferentes exigencias sociales y, nosotros añadiríamos cognitivas (profesionales de la salud cuya actuación condiciona la curación del paciente o paciente que quiere comprender su enfermedad). Cuando los terminólogos consideran la multidimensionalidad, aseguran que las bases de datos terminológicas que elaboran pueden servir a un mayor número de usuarios. Como queremos que nuestro estudio no se restrinja al dominio de la oncología sino que sirva para la redacción de textos médicos en general, se marcarán distintos niveles que crean multidimensionalidad.

En las cadenas hemos representado la multidimensionalidad mediante vínculos de hipertexto, que reflejan áreas conceptuales solapadas. Entre estas podemos mencionar:

EVALUATION en la diagnosis y pronóstico de la enfermedad (*DIAGNOSIS*)

EVALUATION en cualquier actividad que suponga investigación, procesos cognitivos el método científico (*RESEARCH / EXPERIMENTATION*).

MODALITY relacionado con el tratamiento y con la opcionalidad (*TREATMENT y, OPTIONS*)

TRIALS , relacionado con el tratamiento y con la experimentación (*TREATMENT* y *RESEARCH*).

Nuestra propuesta es gráfica como indican las recomendaciones para tratar la dimensionalidad. Con hipervínculos y el hecho de que haya flechas que apuntan hacia los lexemas con los que forma colocación, también puede indicar multidimensionalidad. En la tabla ilustramos la multidimensionalidad del concepto THERAPY en relación a las dimensiones LOCATION IN THE HUMAN BODY y TIME, que aparecen en la primera columna con hipervínculos hacia esos conceptos.

TABLA 62: Multidimensionalidad

THERAPY	LOCATION IN THE HUMAN BODY	TIME
→ THERAPEUTIC AGENT Radiation therapy Chemotherapy	→ <u>THERAPY</u> Intracavitary therapy Intrapleural therapy Intraperitoneal therapy	→ THERAPY Postoperative radiation therapy
→ <u>LOCATION IN THE HUMAN BODY</u> Intracavitary radiation therapy Intracavitary chemotherapy Extrapleural pneumonectomy	→ STAGE OF M. MESOTHELIOMA Solitary mesothelioma Intracavitary mesothelioma	
→ <u>TIME</u> Postoperative radiation therapy	→ ANATOMIC LEVEL Pleura	
→ AIM Palliative surgery Exploratory surgery		

En nuestras cadenas léxicas, la multidimensionalidad se plasma con vínculos de hipertexto hacia las cadenas cuyo número aparece entre paréntesis.

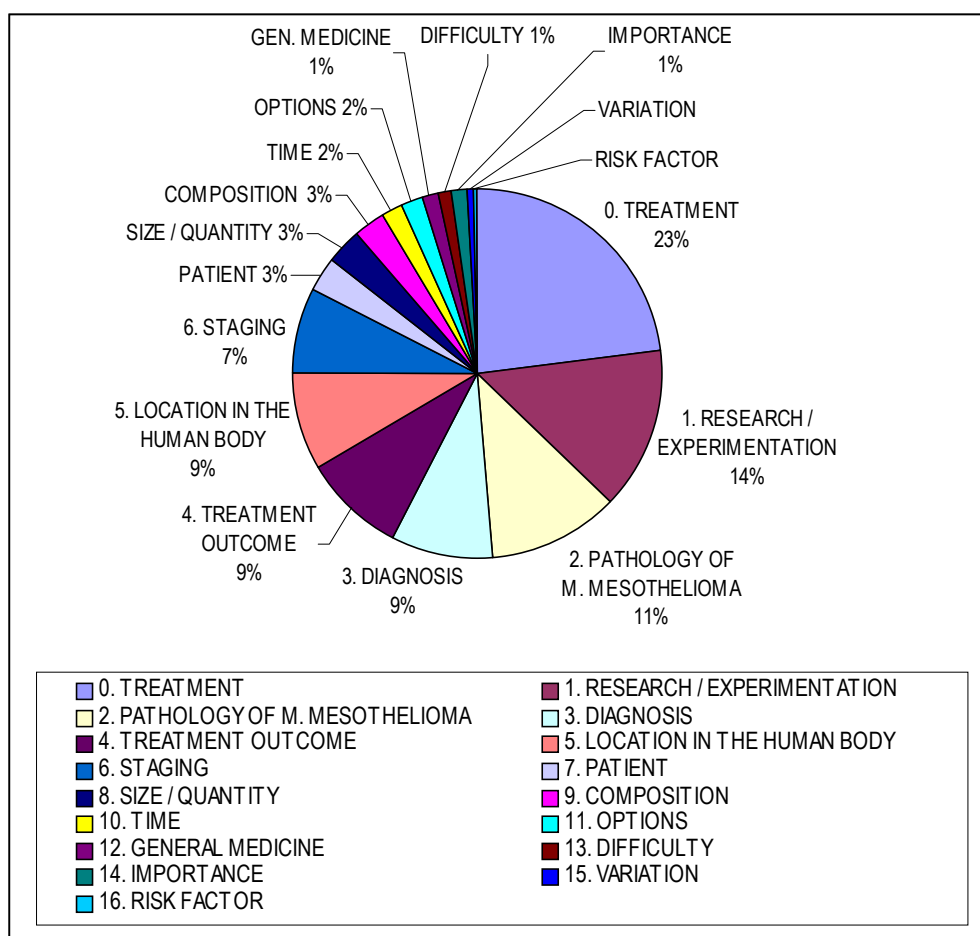
EXTRAPLEURAL (See 5.4.1.) (~ pneumonectomy)	3
--	---

TABLA 63: Cadenas léxicas en QDT1

<p>Chain 0: TREATMENT (H1, 3 a, H10, H13) [143]</p> <p>0.1. Generic treatment 0.2. Properties (→TREATMENT) 0.2.1. Location of application 0.3. Manipulation 0.4. Types 0.4.1. Chemotherapy 0.4.1.1. Chemotherapeutic agents 0.4.2. Surgery 0.4.2.1. Exploratory surgery 0.4.3. Radiotherapy 0.4.4. Other types of treatment</p> <p>Chain 1: RESEARCH / EXPERIMENTATION (H0, H10, H11, H12, H3, H6) [88]</p> <p>1.1. Study 1.2. Categorization 1.2.1. Instances 1.3. Evaluation 1.4. Evidence 1.5. Information</p> <p>Chain 2: PATHOLOGY OF MALIGNANT MESOTHELIOMA (H1) [71]</p> <p>2.1. Generic disease 2.2. Cancer 2.2.1. Types 2.2.1.1. Anatomic / Histologic classification 2.3. Metastasis</p> <p>Chain 3: DIAGNOSIS (H1) [55]</p> <p>3.1. Generic diagnosis 3.2. Cognitive processes involved 3.3. Clinical features (Symptoms) 3.4. Diagnostic tests (H1, 3, 8) 3.4.1. Special stains 3.5. Prognostic factors</p> <p>Chain 4: TREATMENT OUTCOME (H1, H2b) [55]</p> <p>4.1. Generic outcome of treatment 4.1.1. Cause</p>	<p>4.3. Existence 4.3.1. Survival 4.3.2. Death 4.4. Negative outcome</p> <p>Chain 5: LOCATION IN THE HUMAN BODY (H5, H11, H13, PLACE 1,63 %) [53]</p> <p>5.1. Generic location 5.1.1. Generic location (medicine) 5.2. Microscopic / histologic level 5.3. Anatomic level / body parts 5.4. Location in relation to specific body parts 5.4.1. Pleura 5.4.2. Thorax 5.4.2.1. Chest cavity 5.4.3. Peritoneum</p> <p>Chain 6: STAGING (H1, H2a) [46]</p> <p>6.1. Malignant mesothelioma</p> <p>Chain 7: PATIENT (H1, H6, H10) [19]</p> <p>Chain 8: SIZE / QUANTITY (H12, H14) [19]</p> <p>8.1. Generic 8.2. Minuteness 8.3. Largeness / large quantity 8.4. Increase</p> <p>Chain 9: COMPOSITION (H13, H14) [17]</p> <p>Chain 10: TIME (H9) [13]</p> <p>Chain 11: OPTIONS (MODALITY 0,16 %) [12]</p> <p>Chain 12: GENERAL MEDICINE (H1) [9]</p> <p>Chain 13: DIFFICULTY (H6) [7]</p> <p>Chain 14: IMPORTANCE (H12) [7]</p> <p>Chain 15: VARIATION (H3b) [4]</p> <p>Chain 16: RISK FACTOR (H7) [2]</p>
--	---

4.2. Positive outcome	TOTAL DE UNIDADES LÉXICAS ENCADENADAS ¹³³ : [568]
-----------------------	---

DIAGRAMA 20: Porcentaje de activación de las cadenas léxicas definitivas de QDT1



CADENA	Frec	%
0. TREATMENT	143	23,06
1. RESEARCH / EXPERIMENTATION	88	14,19
2. PATHOLOGY OF M. MESOTHELIOMA	71	11,45
3. DIAGNOSIS	55	8,87
4. TREATMENT OUTCOME	55	8,87

5. LOCATION IN THE HUMAN BODY	53	8,55
6. STAGING	46	7,42
7. PATIENT	19	3,06
8. SIZE / QUANTITY	19	3,06
9. COMPOSITION	17	2,74
10. TIME	13	2,1

¹³³ Para obtener el total no se han duplicado los lemas que pertenecen a más de una cadena.

11. OPTIONS	12	1,94
12. GENERAL MEDICINE	9	1,45
13. DIFFICULTY	7	1,13
14. IMPORTANCE	7	1,13

15. VARIATION	4	0,65
16. RISK FACTOR	2	0,32
	620	99,99

Como se observa en el diagrama, la cadena más extensa es la hilada sobre el concepto TRATAMIENTO, como cabría esperar del hecho de que el texto se haya extraído de la sección *Health Professionals Treatment Information*. Cuenta con el 23,69 % de todos los lexemas que hemos vinculado. La segunda más extensa es la compuesta principalmente por lexemas de la lengua general que hacen alusión al método científico y que hemos denominado *RESEARCH / EXPERIMENTATION*. Le siguen en importancia las cadenas en torno a los conceptos *PATHOLOGY OF MALIGNANT MESOTHELIOMA*, *DIAGNOSIS* y *TREATMENT OUTCOME*. Esta última contiene poco vocabulario específico de la oncología. Con prácticamente el mismo porcentaje que las cadenas 3 y 4, hallamos una cadena con conceptos de anatomía, *LOCATION IN THE HUMAN BODY*. La significación de esta cadena confirma la afirmación que se hace en Rector et al (1994: 226) de que la anatomía es central en la terminología médica. Con menos representación encontramos una cadena bastante relacionada con el diagnóstico de la enfermedad: *STAGING*. Por debajo del 5%, encontramos cadenas con lexemas de la lengua general que son bastante frecuentes en textos científicos y médicos. Se trata de las cadenas *PATIENT*, *SIZE*, *COMPOSITION*, *TIME*, *OPTIONS*, *DIFFICULTY*, *IMPORTANCE* y *CHANGE*. Con un 1,4% y un 0,3%, respectivamente, se encuentran lexemas que hacen alusión a la medicina general (CADENA 12) y al principal factor de riesgo para el mesotelioma maligno (CADENA 16). A continuación desglosamos las cadenas y las ordenamos según el número de ocurrencias entre las que se establece la cohesión.

11.1.1.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS

0. TREATMENT (H1, H3a, H10, H13)

0.1. Generic treatment

TREATMENT	18
-----------	----

Treatment 12, treated 5,	
--------------------------	--

treatments 1	
→ APPROACHES	5
Surgical ~ 2	
Treatment ~ 2	
Physical ~ 1	
TRIALS (See 1.1.)	5
Clinical ~ 4	
CONTROL (+ SYMPTOMS+)	4
CURE (See 4.2.)	3
Cure 2, curative 1	
MODALITIES (See 11.)	3
Modalities 2,	

combined modality 1	
OPTIONS (See 11.)	3
Treatment options 2	
Option 1	
THERAPY	3
Intracavitary ~ 2	
(See 0.4.3. RADIATION ~)	
→ PROCEDURES	1
TOTAL	45

0.2. Properties (→ TREATMENT)

PALLIATIVE	5
STANDARD	4
~ treatment 2	
AGGRESSIVE	3
~ surgical approaches 2	

combined modality 1	
MULTIMODALITY (See 11.)	2
POSTOPERATIVE (See 0.4.2.)	1
TOTAL	15

0.2.1. Location where TREATMENT is applied

(→ TREATMENT)

INTRACAVITARY (See 5.4.2.1.)	4
(+ THERAPY)	
EXTRAPLEURAL (See 5.4.1.)	3
(~ pneumonectomy)	
INTRAPERITONEAL See 5.4.3.)	1
(+ THERAPY)	

INTRAPLEURAL (See 5.4.1.)	1
(+ THERAPY)	
TOTAL	9

0.3. Manipulation

USE	4
Use 2, used 1, useful 1	
EMPLOYED	2

UTILIZES	1
TOTAL	7

0.4. Types (according to therapeutic agent)

0.4.1. Chemotherapy

CHEMOTHERAPY	7
Chemotherapy 5,	
chemotherapeutic agents 2	

AGENTS	6
Agents 3	
Chemotherapeutic ~ 2	

Agent 2	
Single-agent chemotherapy 1	
COMBINATION → (✖ 13) ~ chemotherapy 2	3

REGIMENS	2
ADMINISTRATION	1
TOTAL	19

0.4.1.1. Chemotherapeutic agents

AGENTS	6
Agents 3	
Chemotherapeutic ~ 2	
Agent 2	
Single-agent chemotherapy 1	
CISPLATIN	2
DOXORUBICIN	2
MITOMYCIN	2

CARBOPLATIN	1
CYCLOPHOSPHAMIDE	1
CYTARABINE	1
DOXORUBICIN	1
EPIRUBICIN	1
IFOSFAMIDE	1
TOTAL	12

0.4.2. Surgery

SURGICAL	11
Surgical 8 (~ resection 3), Surgery 3	
RESECTION (surgical ~ 3)	4
DECORTICATION (PLEURECTOMY +)	3
PLEURECTOMY (+ DECORTICATION)	3

PNEUMONECTOMY (Extrapleural ~)	3
MARGINS (✖ 10)	2
POSTOPERATIVE (See 0.2.)	4
TOTAL	26

0.4.2.1. Exploratory surgery

THORACOSCOPY (See 3.4.)	1
---	---

0.4.3. Radiotherapy

RADIATION THERAPY Palliative ~ 2	5
-------------------------------------	---

0.4.4. Other types of treatment

PLEURODESIS	2
-------------	---

Chest tube ~	
--------------	--

Thoracoscopic ~	
DRAINAGE OF EFFUSIONS	1
PLEURAL SCLEROSING	1

PROCEDURES	
TOTAL	4

La *cadena 0* contiene las 143 unidades léxicas que se vinculan en el texto QDT1 a través del concepto TREATMENT, lo que la convierte en la cadena más extensa del texto. De las subcadenas que hemos identificado, destacan tanto por el número total de elementos como por su variedad léxica, la que contiene unidades léxicas genéricas y las relacionadas con la quimioterapia y la cirugía.

La subcadena *CHEMOTHERAPY* es la que cuenta con un mayor número de lemas, ya que explicita los diversos fármacos utilizados. El hecho de que determinadas cadenas cuenten con una mayor variedad léxica con respecto a otras, lo cual queda reflejado visualmente con un mayor número de casillas (lemas), puede indicar una mayor complejidad en el sistema conceptual sobre el que se contruyen. En el subdominio de la quimioterapia, se aglutinan un creciente número de denominaciones de fármacos que focalizan su composición química, su nombre comercial o sus siglas. De todas formas, como vimos en la sección 10.3.2., hay otras explicaciones al fenómeno de la variación léxica en relación a un mismo concepto.

Como desde el principio nos propusimos dar relevancia a las relaciones cohesivas sintagmáticas, en la cadena hemos incluido diez lexemas que normalmente acompañan a lexemas referidos al TRATAMIENTO, tales como *palliative*, *standard* o *intracavitary*. Este último además indica la región del cuerpo donde se aplica el tratamiento, aunque con menor precisión que otros vocablos con la misma función (*extrapleural*, *intrapleural* e *intraperitoneal*). *Intracavitary* sólo especifica que el lugar donde se aplica el tratamiento es una cavidad. Es el cotexto donde se nos aclara qué cavidad: las pleuras o el peritoneo¹³⁴. Por último, hemos destacado los lexemas de la lengua general que indican MANIPULACIÓN, es decir la acción de aplicar un tratamiento. Aparte de *treated* de la

¹³⁴ Aunque las pleuras sean una doble membrana, en el texto y en sistemas de clasificación clínicos como el proyecto GALEN, se consideran una cavidad, ya que las cavidades están formadas por *hollow structures and bilayered membranes* (Rector et al. 1994: 229).

subcadena *GENERIC TREATMENT*, encontramos en el texto los lemas *use*, *employed* y *utilizes*.

1. RESEARCH, EXPERIMENTATION (H0, H10, H11, H12, H3b, H6)**1.1. Study (H0)**

REFERENCES	7
REPORTED	7
STUDIES	6
Studies 4, studied 2	
TRIALS (See 0.1.)	5
Clinical ~ 4 (See 12.)	
INTERNATIONAL	2
PHASE	2
~ I (and II) (clinical) trials 2	

PURPOSES	2
CONSISTENT	1
DATA	1
EXACT	1
MEDIAN (See 4.3.1.)	1
STATEMENT	1
TOTAL	36

1.2. Categorization (H10, H11, H14)

SYSTEM (+ PROPOSED+)	4
System 3	
~ staging system 2	
Systems 1	
BASED	3
Based 2, basis 1	
PROPOSED (+SYSTEM+)	3
DESCRIPTION	2
Description 1, descriptions 1	
CATEGORIZED	1
CLASSIFICATION	1

DEFINE	1
DESIGNATIONS	1
DETERMINATIONS	1
FORM	1
PRINCIPLES	1
REPRESENT	1
STRUCTURES	1
TOTAL	21

1.2.1. Instances (H11, EXAMPLE 0,81%)

SPECIMENS (tissue ~)	2
INSTANCES	1
TOTAL	3

1.3. Evaluation

EVALUATION (See 3.2.)	4
Clinical evaluation 2,	
Evaluated 1, evaluation 1	
DIFFERENCES	3
Differences 1, different 1,	
differentiate 1	
FACTORS (See 3.2.)	3
Factors 2, factor 1	

ASSESS (See 3.2.)	2
DISCUSSION	1
INTERPRETATION (✖ 3)	1
CONSIDERED (See 3.2.)	1
TOTAL	15

1.4. Evidence (SHOWN 0,16%)

SHOWN Shown 2, showing 1	3
APPEARANCE Appearance 1, appears 1	2
SEE	2

DEMONSTRATED	1
EVIDENCE	1
FOUND	1
TOTAL	10

1.5. Information (H6)

INFORMATION	3
-------------	---

La segunda cadena más relevante del texto abunda en lo que se denominó *vocabulario organizador del discurso* (véase p. 84). Con la excepción de términos sobre oncología como *clinical trials*, *phase I (clinical trials)* o *staging system*, los elementos de esta cadena léxica bien podrían aparecer en cualquier otro texto científico. Estos lexemas apuntan hacia aspectos fundamentales dentro de la investigación científica. Encontramos referencias a estudios, a información y a la categorización de entidades y procesos a partir de instancias concretas. También encontramos la observación empírica (*evidence*) de fenómenos y la evaluación de los mismos.

Los elementos de la subcadena *EVALUATION* pueden entenderse en un sentido más general, ya que la evaluación es parte de los procesos cognitivos inherentes al método científico, o en un sentido más específico dentro de la oncología. En los textos analizados, normalmente se evalúa el estado físico del paciente y los resultados de las pruebas diagnósticas para determinar el estadio y tratamiento de la enfermedad, que en muchos casos está en fase experimental (*under clinical evaluation*). De ahí que en los textos sobre cáncer también se mencione la evaluación de nuevos regímenes de fármacos, de variaciones en la dosis y de nuevas combinaciones de tratamientos.

(87)

1 and combination chemotherapy have been evaluated in single and combined mo
 2 ents are "standard" or "under clinical evaluation" are not to be used as
 3 tive radiation therapy. Under clinical evaluation: 1 Intracavitary chem
 4 tes for phase I and II clinical trials evaluating new biologicals, chemoth

2. PATHOLOGY OF MALIGNANT MESOTHELIOMA (H1)

2.1. Generic DISEASE

DISEASE	13
Disease 12, disease-free 1	
SYMPTOMATIC (See 3.3.)	6
Symptomatic 3, symptoms 2, symptom 1	
TOTAL	19

2.2. Cancer

MALIGNANT	7
~ mesothelioma 6	
~ local pleural disease 1	
TUMORS	7
Tumors 6, tumor 1	
CARCINOMAS	3
Carcinomas 2, carcinoma 1	
NEGATIVE (+ PART OF THE BODY)	2
POSITIVE (+ PART OF THE BODY)	1

ANAPLASTIC	1
(~ lung carcinoma)	
MASSES (Tumor ~)	1
LESIONS	1
Sessile polypoid ~	
TOTAL	23

2.2.1. Types of malignant mesothelioma

2.2.1.1. Anatomic / Histologic classification

MESOTHELIOMA	15
Malignant mesothelioma 6	
PLEURAL mesothelioma 2	
INTRACAVITARY mesothelioma 1	
EPITHELIAL MESOTHELIOMAS (2) ≈ Epithelial form (1)	3
FIBROUS OR SARCOMATOUS MESOTHELIOMA (1)	1
TOTAL	19

2.3. Metastasis

RECURRENT (See 6.1.)	4
- Recurrent 3 (+ CANCER)	
- Recurrence-free 1	
METASTATIC	2
DISTANT	1

INVASIVE	1
INVOLVEMENT	1
SITE OF ORIGIN (See 5.1.)	1
TOTAL	10

En la cadena *PATHOLOGY OF MALIGNANT MESOTHELIOMA*, incluimos tres subcadenas que hacen referencia a la enfermedad y a su evolución. En la nomenclatura de la enfermedad hay un gradiente que va desde lo menos a lo más específico. A lo largo del

texto se irán alternando términos más genéricos con los más especializados, tejiendo la cohesión del texto. En la última casilla de la siguiente tabla mostramos esta alternancia con respecto a *mesothelioma* y *disease*.

TABLA 64: Grado de especialización de los términos que designan el concepto MALIGNANT MESOTHELIOMA y su alternancia para crear cohesión.

+ Especialización ↓	Disease – symptom* Tumor* – carcinoma – anaplastic Lung carcinoma – malignant mesothelioma Pleural mesothelioma Epithelial mesothelioma – fibrous or sarcomatous mesothelioma
Posición de <i>mesothelioma</i> y <i>disease</i> (nº de palabra en el texto)	
MESOTHELIOMA: 2, 52, 191, 244, 317, 326, 416, 488, 498, 539, 661, 807, 815, 846, 887, 998, 1.002	
DISEASE: 10, 30, 114, 123, 150, 336, 356, 422, 456, 484, 520, 583, 826	

El término *malignant mesothelioma* aparece como primera mención a la enfermedad (primera y segunda palabras del texto), y en la próxima referencia al mismo concepto (décima palabra del texto), encontramos la USC genérica *disease*, que funciona de elemento catafórico para evitar la repetición del mismo término. De todas formas, hemos percibido en nuestro corpus una tendencia del lexema *disease* a coocurrir con la palabra *stage*, como veremos en las líneas de concordancia del lema *stage* (véase apartado 14.4.2.3.).

En la nomenclatura predomina la mención anatómica sobre la histológica o morfológica. No obstante, la anatómica es bastante imprecisa porque sólo menciona que el cáncer surge en una de las membranas que recubren los órganos (mesotelio). Se sobreentiende que esta membrana son las pleuras por el cotexto. El carácter metastático de

la enfermedad queda expresado a lo largo del texto mediante diez ocurrencias de los lemas *RECURRENT*, *METASTATIC*, *DISTANT*, *INVASIVE*, *INVOLVEMENT* y *SITE OF ORIGIN*, término que también pertenece a la cadena *LOCATION IN THE HUMAN BODY*, pero que hemos incluido aquí porque presupone el lugar desde donde emigran las células cancerosas.

3. DIAGNOSIS (H1)

3.1. Generic diagnosis

PROGNOSIS Prognosis 3 Better ~ 2 Prognostic 3	6
DIAGNOSIS Diagnosis 2, diagnostic 1	3

EXAMINATION	1
HISTORY OF (~ asbestos exposure)	1
TOTAL	11

3.2. Cognitive processes involved

EVALUATION (See 1.3.) Clinical evaluation 2, Evaluated 1, evaluation 1	4
FACTORS (See 1.3.) Factors 2, factor 1	3
ASSESS (See 1.3.)	2

CONSIDERED (See 1.3.)	1
TOTAL	10

3.3. Clinical features (Symptoms)

SYMPTOMATIC (See 2.1.) Symptomatic 3, symptoms 2, symptom 1	6
EFFUSIONS pleural ~ 3	5
PAIN	2

DISCOMFORT	1
TOTAL	14

3.4. Diagnostic tests (H1, H3, H8)

SPECIAL STAINS (✖ 3)	2
CYTOLOGY	1
ELECTRON MICROSCOPY (✖ 8)	1
NEEDLE BIOPSY	1

THORACOSCOPY (See 0.4.2.1.)	1
TOTAL	6

3.4.1. Special stains

CEA	1
HYALURONIC ACID	1
LEU M1	1
MUCICARMINE	1

PERIODIC ACID-SCHIFF DIASTASE	1
TOTAL	5

3.5. Prognostic factors

<u>STAGE</u> (See 6.)	42
HISTOLOGY Histology 2, histologic 1, histologically 1	4
STATUS Nodal ~ 1 Performance ~ 1	2

AGE	1
NEGATIVE LYMPH NODES	1
NEGATIVE SURGICAL MARGINS	1
TOTAL	9

La cadena *DIAGNOSIS* hila 55 unidades léxicas. Las subcadenas más relevantes son las que hemos denominado *GENERIC TERMS*, *CLINICAL FEATURES (SYMPTOMS)* y *PROGNOSTIC FACTORS*. En esta cadena aparecen lemas relacionados con procesos cognitivos que ayudan a determinar la enfermedad, con lo que se produce un solapamiento con la subcadena *I.3. EVALUATION*, perteneciente a la cadena *RESEARCH, EXPERIMENTATION*. Dentro de la subcadena *DIAGNOSTIC TESTS AND INSTRUMENTS*, surge una subdivisión que se situaría a un tercer nivel, *special stains*, donde se especifican las distintas tinciones empleadas en las pruebas citológicas. Aunque *cytology* sólo aparezca una vez, el peso conceptual de las pruebas citológicas es relevante por la cohesión que se establece entre *cytology*, *special stains* y los nombres de distintas tinciones (*CEA*, *Leu M1*, etc.). En el siguiente cuadro clasificamos el tipo de cohesión creado. El número entre paréntesis indica la posición que ocupa la unidad léxica en el texto, mientras que las líneas con trazos gruesos indican los elementos vinculados, por ejemplo, *cytology* está relacionado con todas las unidades léxicas del cuadro, y *special stains* (palabras 274-275) está relacionada con la ocurrencia *special stains* (palabras 283-4).

TABLA 65: Tipos de relaciones cohesivas a partir del término *cytology*.

<i>Cytology</i> (227)					
R. mixta asociativa (MEDIO) de lexema y unidad fraseológica				R. paradigmática hiponímica	
<i>Special stains</i> (274-5)	R. exacta	<i>Special stains</i> (283-4)	R. mixta hiponímica	<i>Hyaluronic acid</i> (295-296) <i>Leu M1</i> (300-1) <i>Periodic acid-schiff diastase</i> (292-4)	<i>CEA</i> (298) <i>Mucicarmine</i> (297)

Por esto, intuimos que, cuanto más subdivisiones se producen en una cadena, cuando hay un mayor grado de *granularidad* (Wright, 1999), es porque existe preeminencia de determinadas realidades extralingüísticas.

4. TREATMENT OUTCOME (H1, H2b, H13)

4.1. Generic outcome of treatment (H1, H2b)

RATES Rates 2, rate 1	3
RESPONSES Partial responses 2, response 1	3
RESULT	2

Result 1, results 1	
TOTAL	8

4.1.1. Cause (DOER¹³⁵ 1,42%)

CAUSED Caused 2, causes 1	3
PRODUCE Produce 1, produced 1	2

4.2. Positive outcome (H2b)

IMPROVE Improve 2, improved 2	4
CURE (See 0.1.) Cure 2, curative 1	3

BETTER	2
ENSURE	2
ALLEVIATE	1

¹³⁵ Esta es una de las denominaciones que *Hesperus* otorga al concepto AGENT.

PERFORMANCE STATUS	1
PROVIDE	1
RELIEF	1

TOTAL	15
-------	----

4.3. Existence (H4)

4.3.1. Survival

SURVIVAL	11
LONG-TERM → (~ survival)	3
LONGER →	2
OVERALL → (~ survival)	2

MEDIAN → (~ survival)	1
REMAINS	1
TOTAL	20

4.3.2. Death

MORTALITY	2
SHORT-LIVED	1
TOTAL	3

4.4. Negative outcome

FAILURE (~ rates)	1
IMPACT (✖ 13)	1
NONCONTRIBUTORY	1

TOXICITY	1
TOTAL	4

En esta cadena dominan las referencias a los resultados positivos del tratamiento. De hecho, sobre un total de 55 palabras semánticamente relacionadas con el resultado del tratamiento, sólo tres hacen referencia a la muerte. Además los lemas con mayor representación (IMPROVE, SURVIVAL) y las subcadenas con más elementos (*POSITIVE OUTCOME*, *SURVIVAL*) también indican el éxito del tratamiento.

5. LOCATION IN THE HUMAN BODY

(H5, H11, H13, PLACE 1,19%, NEAR-PLACE 0,22%, NEAR 0,22%)

5.1. Generic location

LOCAL	2
CONTIGUOUS	1
<u>SITE</u> (~ of origin) (See 2.3.)	1

TOTAL	4
-------	---

5.1.1. Generic location (medicine)

CONTRALATERAL	1
IPSILATERAL	1

5.2. Microscopic / histologic level

TISSUE	6
Tissue 2 (~ specimens)	
HISTOLOGY	3
Histology 2, histologic 1, histologically 1	
EPITHELIAL (See 2.2.1.1.)	

FIBROUS (See 2.2.1.1.)	1
CELLULAR	1
TOTAL	11

5.3. Anatomic level / body parts (H5)

PLEURAL	9
Pleural 6, pleura 3	
NODES (✖ 13)	4
Lymph node(s) 3 Nodal 1	
DIAPHRAGM (✖ 11)	3
CHEST	2
~ wall 1	
LUNG	2

PERITONEAL	2
Peritoneal 1, peritoneum 1	
HEART (✖ 11)	1
PARIETAL (Capsule of the ~ pleura)	1
PERICARDIUM	1
TOTAL	25

5.4. Location in relation to specific body parts (+ THERAPY / + BODY PART)

5.4.1. Pleura

EXTRAPLEURAL (~ pneumonectomy) (See 0.2.1.)	3
INTRAPLEURAL (+ THERAPY) (See 0.2.1.)	1

5.4.2. Thorax

EXTRATHORACIC (+ LYMPH NODE)	1
INTRATHORACIC (+ LYMPH NODE)	1

5.4.2.1. Chest cavity

<u>INTRACAVITARY</u> (+ THERAPY / + DISEASE) (See 0.2.1., 2.2.1.1., 6.1.)	4
---	---

5.4.3. Peritoneum

<u>INTRAPERITONEAL</u> (+ THERAPY) (See 0.2.1.)	1
TOTAL	11

Esta cadena contiene vocablos que designan tipos de tejidos, órganos, membranas y regiones del cuerpo y que, en su contexto, quedan vinculados también con las cadenas *PATHOLOGY*, *DIAGNOSIS* y *TREATMENT*. El solapamiento de estas cadenas es evidente, teniendo en cuenta que una forma de denominar el tipo de cáncer atiende a su localización en el cuerpo, que en el cuerpo se manifiestan síntomas que ayudan a diagnosticar la enfermedad y que el tratamiento se aplica en el cuerpo. En este último caso, los términos van precedidos por un prefijo latino que indica la posición con respecto a una región corporal que concibe como un contenedor con respecto al cual se está dentro (*intra-*) o fuera (*extra-*). Como veremos más adelante, los solapamientos entre cadenas dan prueba de la multidimensionalidad terminológica, a la que contribuye la existencia de unidades fraseológicas. Por último, la parte del cuerpo más activada, como cabría esperar, son las pleuras.

6. STAGING (H1, H2a) (+ DISEASE)

<u>STAGE</u> (✖ 2a) Stage 12, stages 3, Staging 3 ~ system 2	18
<u>STAGE I</u> + (CANCER) Stage ~	5
<u>STAGE II</u> + (CANCER)	3

<u>STAGE(S) III</u> + (CANCER)	3
<u>STAGE(S) IV</u> (CANCER)	3
<u>CONFINED</u> (✖ 1)	1
TOTAL	33

6.1. Malignant mesothelioma

LOCALIZED ~ malignant mesothelioma 2 ~ mesothelioma 1	4
SOLITARY MESOTHELIOMA	1
INTRACAVITARY MESOTHELIOMA (See 2.2.1.1.)	1
ADVANCED ~ malignant mesothelioma 2	4
RECURRENT (See 2.3.) ~ pleural effusion ~ malignant mesothelioma ~ mesothelioma	3
TOTAL	13

Existen 46 unidades léxicas relacionadas conceptualmente mediante STAGING. El lema más frecuente, STAGE, coocurre normalmente con números romanos. Las demás referencias al estadiaje de la enfermedad se realizan precediendo el sustantivo *mesothelioma* con los adjetivos *advanced*, *localized* y *recurrent*.

7. PATIENT (H1, H6, H10)

PATIENTS Selected ~ 3 ~ treated with aggressive surgical approaches 2	12
SELECTED → (See 11.) Selected patients 3, selection 1	4

SERIES (~of patients)	2
CANDIDATES (✕ 6)	1
TOTAL	19

La importancia de la cadena *TRATAMIENTO* lleva aparejada una presencia significativa de instancias del concepto PATIENT. Esto se explica por el hecho de que *patients* adquiere el rol semántico *processed experiencer* (ProcExp) en la terminología de la Gramática Funcional de Dik (1989), es decir, el paciente es una entidad que experimenta un proceso terapéutico.

Se ha incluido *selected* porque precede a *patients* en tres ocasiones, aunque en realidad pertenece a la cadena *OPTIONS*.

8. SIZE / QUANTITY (H12, H14)

8.1. Generic

SIZE (✖ 14)	1
-------------	---

8.2. Minuteness

SMALL	3
-------	---

8.3. Largeness / large quantity

HIGHER Higher 2, high 1	3
EXTENSION Extension 1, extensive 1	2
GREAT	1
GROSS	1

LARGE	1
MAJOR (See 14.)	1
WIDE	1
TOTAL	10

8.4. Increase

ADDITION (See 9.) Addition 1, additional 1	2
GROWS	1
INTEREST	1

PROGRESSION	1
TOTAL	5

El tamaño del tumor es un factor importante a la hora de determinar el tratamiento y el pronóstico de la enfermedad, de ahí que se configure una cadena en torno al concepto SIZE. En relación con el mismo encontramos el concepto QUANTITY, que se utiliza en la descripción de los tratamientos y sus resultados y que está bastante relacionado con el de IMPORTANCE.

9. COMPOSITION (H13, H14)

INCLUDE	3
ADDITION (See 8.4.) Addition 1, additional 1	2
GENERAL	2

General 1, generally 1	
INCLUDING	2
PARTIAL (~ responses)	2
COMPOSED OF	1

ELEMENTS	1
GROUP	1
MAJORITY (✖ 14)	1
MAKING	1

NATURE	1
SINGLE	1
TOTAL	17

Uno de los conceptos fundamentales en el sistema conceptual humano, la oposición entre la TOTALIDAD y la PARCIALIDAD está presente en esta cadena, que abunda en *vocabulario organizador del discurso*.

10. TIME (H9)

TIME	2
MONTHS	2
AGE	1
CURRENT	1
DURATION	1
INITIAL	1
OLDER	1

PERIOD	1
PROLONGED → ~ period of time	1
TEMPORARY	1
TRANSIENT	1
TOTAL	13

Aquí también sale a relucir uno de los conceptos fundamentales en el intelecto humano: el TIEMPO. En el texto, el TIEMPO queda asociado ante todo a la idea de supervivencia.

11. OPTIONS (MODALITY 0,16%)

SELECTED (See 7.) Selected 3, selection 1	4
MODALITIES (See 0.1) Modalities 2, Combined modality 1	3
OPTIONS (See 0.1)	3

Option 2, option 1	
MULTIMODALITY (See 0.2)	2
TOTAL	12

Esta cadena pone de manifiesto que la complejidad de la enfermedad requiere distintos tratamientos que se van acomodando a la evolución de la enfermedad. Además, para determinados cánceres y estadios, a veces existe más de un protocolo de tratamiento estándar y, si el facultativo piensa que la efectividad de esos protocolos es la misma, este solicita que el paciente elija.

12. GENERAL MEDICINE (H1)

CLINICAL (See 0.1.) ~ trials 4, ~ evaluation 2, ~ studies 2	9
---	---

Esta cadena se establece a partir de la repetición exacta del lexema *clinical*, que en la mayoría de los casos constituye la unidad fraseológica *clinical trials*. Las otras UF formadas con *clinical* también están relacionadas con el concepto CLINICAL TRIAL (ensayo clínico)¹³⁶.

13. DIFFICULTY (H6)

DIFFICULT ~ to assess 2	3
CONFUSION	1
PROBLEMS	1

UNCLEAR	1
UNKNOWN	1
TOTAL	7

Aunque poco representativa, esta cadena refleja la dificultad que encuentran los profesionales de la medicina a la hora de detectar el tipo de cáncer y de determinar su pronóstico, ya que ningún tratamiento asegura la curación en el 100% de los casos.

14. IMPORTANCE (H12)

IMPORTANT (~ prognostic factors)	2
NEW	2
MAJOR (See 8.3)	1

SIGNIFICANTLY	1
VALUABLE	1
TOTAL	7

Dado que no existe una certeza absoluta sobre el tratamiento o la curación del cáncer, como indica la gran presencia de modales, que comentamos en la sección 14.5., es conveniente hablar de los tratamientos más relevantes en cuanto a novedad o en cuanto a

¹³⁶ La definición de *clinical trial* que se propone en el *Cancer Dictionary* (1992:66) es la siguiente: "a systematic evaluation of a possible new cancer treatment conducted with cancer patients after the treatment has had some benefits in animal testing or laboratory testing".

su grado de éxito. En este sentido, el grado de éxito de un tratamiento, codificado con lexemas de la cadena *QUANTITY*, está vinculado a la cadena *IMPORTANCE*.

15. VARIATION (H3b)

VARIABILITY	2
Variability 1, various 1	
ALTERED	1
MODIFICATION	1
TOTAL	4

Los cambios biológicos que se producen durante la enfermedad y las modificaciones en el conocimiento sobre la enfermedad a raíz de nuevos descubrimientos explican la existencia de esta cadena en el texto.

16. RISK FACTOR (H7)

ASBESTOS (~ exposure)	1
EXPOSURE	1
TOTAL	2

Aunque la frecuencia de la UF *asbestos exposure* es mínima, la hemos incluido como cadena, puesto que la literatura médica lo señala como el principal factor de riesgo en el mesotelioma. En el texto no se establece la relación entre ambos de un modo explícito. Es el conocimiento del especialista el que le permite vincular la exposición al asbesto con la enfermedad, ya que, en principio, no existe ninguna relación semántica de base lingüística entre estas dos entidades. La conexión entre realidades extralingüísticas codificadas en el lenguaje hace que los expertos perciban cohesión, independientemente de lo explícitos que sean los recursos lingüísticos. En el texto no se dice "*asbestos exposure may cause cancer*", información implícita en la oración del texto:

- (88) A history of asbestos exposure is reported in about 70%-80% of all cases of mesothelioma.

11.1.1.5. CUADRO CONTRASTIVO (*HESPERUS* – ANÁLISIS COMBINADO)

Al comparar los datos que aporta el programa *Hesperus* sobre la cohesión del texto y los que obtiene el investigador al adaptar y corregir datos extraídos de forma automática (análisis combinado), afloran claras diferencias con respecto al número de palabras vinculadas en las cadenas, las variantes morfológicas de las mismas y el peso específico de cada cadena en la construcción conceptual del texto¹³⁷. En la casilla correspondiente a *Hesperus*, hemos ajustado la activación conceptual resultante de sumar dentro de cada cadena el porcentaje de sus conceptos, ya que, según el autor del programa, las cadenas sólo representan el 70% de la cohesión del texto y no siempre hemos podido encajar el 30% restante. Para el ajuste, vamos a considerar que la activación conceptual en las cadenas de *Hesperus* representa el 100% de la cohesión del texto.

TABLA 66: Descripción cuantitativa de las cadenas léxicas reconocidas.

CADENAS LÉXICAS RECONOCIDAS POR <i>HESPERUS</i>		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 1: TREATMENT AND DIAGNOSIS, PATHOLOGY OF M. MESOTHELIOMA [96-29-29] H 3 a: TREATMENT [12-1-6,4]	Lexemas vinculados 108 Formas distintas 30 Activación conceptual ajustada (%) 44,3	291 124 46,92	0. TREATMENT [143-58-23,06] 2. PATHOLOGY OF M. MESOTHELIOMA [71-27-11,45] 3. DIAGNOSIS [49-36-7,9] 7. PATIENT [19-5-3,06] 12. GENERAL MEDICINE [9-1-1,45] Lexemas vinculados Formas distintas Activación conceptual (%)
H 0: RESEARCH [5-1-2,1] H 14: ORDER, CATEGORIZATION [10-8-2,2] H 10: ? [2,3-11-8]; H 11: ? [10-7-0,8]	Lexemas vinculados 36 Formas distintas 24 Activación conceptual ajustada (%) 6,5	88 50 14,19	1. RESEARCH, EXPERIMENTATION Lexemas vinculados Formas distintas Activación conceptual (%)
H 2b: IMPROVEMENT [5-4-0,2] H 4: SURVIVAL [15-3-5,21]	Lexemas vinculados 20 Formas distintas 7 Activación conceptual ajustada (%) 6,8	55 32 8,87	4. TREATMENT OUTCOME Lexemas vinculados Formas distintas Activación conceptual (%)

¹³⁷ Esta información está en las casillas centrales y en los corchetes que llevan algunas cadenas.

H 5: PLEURA [10-3-4,1] H 11: ? [5-3-0,79]			5. LOCATION IN THE HUMAN BODY
Lexemas vinculados	15	53	Lexemas vinculados
Formas distintas	6	31	Formas distintas
Activación conceptual ajustada (%)	6,2	8,55	Activación conceptual (%)
H 2 a: STAGING [18-3-7,1]			6. STAGING
Lexemas vinculados	18	46	Lexemas vinculados
Formas distintas	3	14	Formas distintas
Activación conceptual ajustada (%)	8,9	7,42	Activación conceptual (%)
H 12: SIZE, INCREASE [18-17-1,4]			8. SIZE / QUANTITY [19-16-3,06] 14. IMPORTANCE [7-5-1,13]
Lexemas vinculados	18	26	Lexemas vinculados
Formas distintas	17	21	Formas distintas
Activación conceptual ajustada (%)	1,8	4,19	Activación conceptual (%)
H 13: PART-WHOLE, COMPOSITION [18-10-2,6]			9. COMPOSITION
Lexemas vinculados	16	17	Lexemas vinculados
Formas distintas	10	14	Formas distintas
Activación conceptual ajustada (%)	3,3	2,74	Activación conceptual (%)
H 9: TIME [5-4-3,5]			10. TIME
Lexemas vinculados	5	13	Lexemas vinculados
Formas distintas	4	11	Formas distintas
Activación conceptual ajustada (%)	4,4	2,1	Activación conceptual (%)
H 6: ? (DIFFICULTY) [20-13-1,6]			13. DIFFICULTY
Lexemas vinculados	20	7	Lexemas vinculados
Formas distintas	13	5	Formas distintas
Activación conceptual ajustada (%)	2	1,13	Activación conceptual (%)
H 3b: CHANGE [15-4-6,4]			15. VARIATION
Lexemas vinculados	15	4	Lexemas vinculados
Formas distintas	4	4	Formas distintas
Activación conceptual ajustada (%)	8	0,65	Activación conceptual (%)
H 8: LABORATORY EQUIPMENT [2-2-2,8]			3.4. DIAGNOSTIC TESTS AND INSTRUMENTS
Lexemas vinculados	2	6	Lexemas vinculados
Formas distintas	2	5	Formas distintas
Activación conceptual ajustada (%)	3,5	0,97	Activación conceptual (%)
H 7: RISK FACTOR [2-2-2,8]			16. RISK FACTOR
Lexemas vinculados	2	2	Lexemas vinculados
Formas distintas	2	2	Formas distintas
Activación conceptual ajustada (%)	3,5	0,32	Activación conceptual (%)
MODALITY ¹³⁸			11. OPTIONS
Activación conceptual (%)	0,6	12	Lexemas vinculados
Activación conceptual ajustada (%)	0,8	7	Formas distintas
		1,94	Activación conceptual (%)

¹³⁸ *Hesperus* no reconoce esta cadena, aunque el perfil del documento sí otorga un porcentaje al concepto MODALITY.

TOTAL			TOTAL
Lexemas vinculados	275	620	Suma de elementos en las cadenas
Formas distintas	122	568	Lexemas vinculados ¹³⁹
Activación conceptual según Hesperus	79,5	289	Formas distintas ¹⁴⁰
Activación conceptual ajustada (%)	99,9	99,99	Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QDT1			
Nº total de palabras: 1079		Categorías conceptuales del tesoro: 141	
Palabras que crean cohesión (<i>Hesperus</i>): 505		Coeficiente de cohesión (<i>Hesperus</i>): 9,76	
		Nº de oraciones: 63	

El hecho de que mediante el análisis combinado se reconozcan el doble de lexemas vinculados confirma que la cohesión detectada exclusivamente con medios informáticos es bastante exigua, incluso cuando estos cuentan con la base de conocimiento de un tesoro. El analista humano cuenta con un tesoro *afinado* con respecto al cambiante conocimiento sobre oncología, y con un conocimiento lingüístico y extralingüístico que le permite crear una completísima red de relaciones entre términos y la habilidad de detectar los aspectos textuales que crean cohesión. Además, el humano refleja mejor la complejidad del significado y la multidimensionalidad terminológica en forma de cadenas en las que se producen solapamientos.

No obstante, romperemos una lanza a favor del programa *Hesperus* en cuanto a la representación conceptual que el analista puede extraer a partir de los datos de *Hesperus*. A grandes rasgos, como vemos en el diagrama de sectores, las grandes áreas temáticas que hemos reconocido tras el análisis de las etapas A, B, C y D (anillo externo) no difieren demasiado con respecto a las extraídas mediante *Hesperus* (anillo interno). Las diferencias más significativas conciernen principalmente a las cadenas *RESEARCH*, *EXPERIMENTATION*, *CHANGE* y *OPTIONS*. La desviación con respecto a la cadena *RESEARCH*, *EXPERIMENTATION* puede explicarse por el hecho de que la hemos convertido en cajón de sastre del *vocabulario organizador del discurso*. Las demás diferencias reflejan las limitaciones de asociar palabras de un texto especializado con la ayuda de un tesoro de la lengua general.

¹³⁹ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

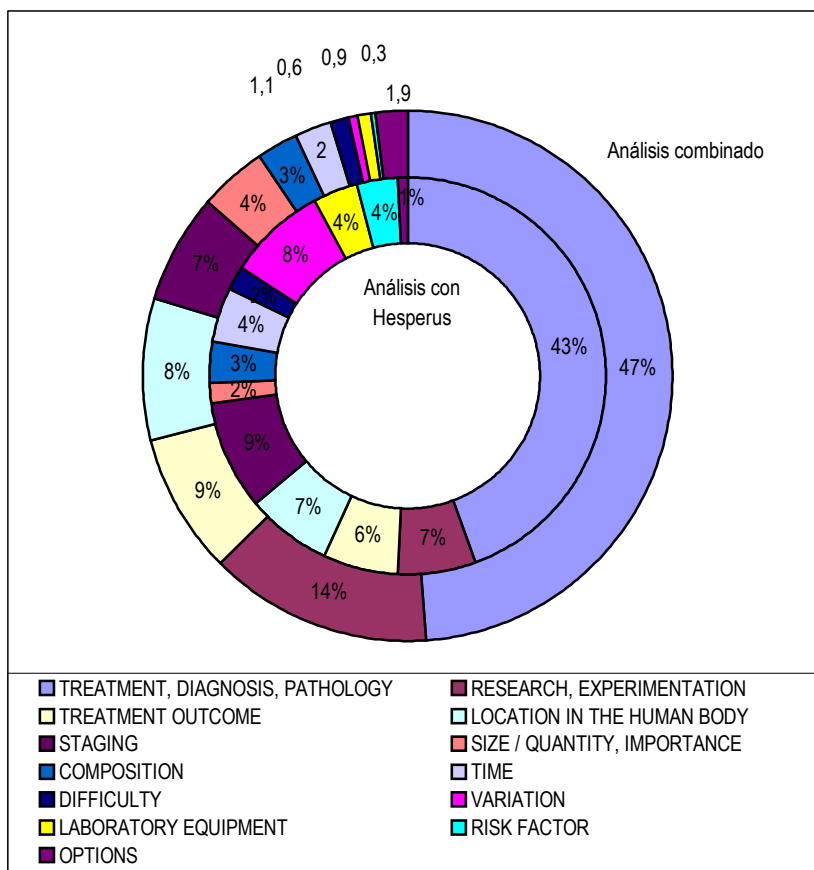


DIAGRAMA 21: Comparación de las áreas conceptuales reconocidas por *Hesperus* y las reconocidas por un humano asistido por medios informáticos (QDT1).

Por último, diremos que *Hesperus* reconocía en el texto 505 palabras plenas, susceptibles de crear cohesión léxica (46,8% del número total de ocurrencias del texto, *tokens*), mientras que nosotros hemos identificado 568 palabras vinculadas, que representan el 52,64% del número total de palabras.

Si dividimos el número de lexemas vinculados entre el número de oraciones (63), diremos que hay un promedio de 9 palabras vinculadas por oración, siendo la extensión media por oración la de 16,14 palabras. Si añadimos a esto que con 568 ocurrencias (*tokens*) de 289 formas léxicas (*word type*) y 235 lemas se consiguen activar 141 categorías

¹⁴⁰ De estas, hay 31 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 52 lexemas adicionales a los 568 lexemas vinculados.

conceptuales del *Roget's*, podemos extraer tres cifras que nos permitirán comparar este texto con el resto y sacar conclusiones con respecto a la variedad léxica. En cada 100 palabra vinculadas, encontraríamos 51 variantes léxicas de 41 lemas que activarían 25 categorías del tesoro.

En sección 11.3. se relacionarán estos datos con los de los demás textos e intentaremos interpretarlos a la luz de las características del texto.

11.1.2. Non-small cell lung cancer (QDT2)

11.1.2.1. ÍNDICE DEL TEXTO

En este apartado, aparecen en negrita los lexemas susceptibles de crear cadenas léxicas en el texto, tal y como se perfilan en el índice: *information*, *cellular classification*, *stage (TNM)*, *treatment options*, *non-small cell lung cancer*.

(88)

GENERAL INFORMATION

CELLULAR CLASSIFICATION

STAGE INFORMATION

The Revised International Staging System for Lung Cancer

TNM definitions

AJCC stage groupings

Occult carcinoma

Stage 0

Stage IA

Stage IB

Stage IIA

Stage IIB

Stage IIIA

Stage IIIB

Stage IV

TREATMENT OPTION OVERVIEW

OCCULT NON-SMALL CELL LUNG CANCER

TX, N0, M0

STAGE 0 NON-SMALL CELL LUNG CANCER

Tis, N0, M0

STAGE I NON-SMALL CELL LUNG CANCER

T1, N0, M0 or T2, N0, M0

STAGE II NON-SMALL CELL LUNG CANCER

T1, N1, M0 or T2, N1, M0 or T3, N0, M0

STAGE IIIA NON-SMALL CELL LUNG CANCER

T1, N2, M0 or T2, N2, M0 or T3, N1, M0 or T3, N2, M0

Superior sulcus tumor (T3, N0 or N1, M0)

Chest wall tumor (T3, N0 or N1, M0)

STAGE IIIB NON-SMALL CELL LUNG CANCER

T4 or N3, M0

STAGE IV NON-SMALL CELL LUNG CANCER

Any T, any N, M1

RECURRENT NON-SMALL CELL LUNG CANCER

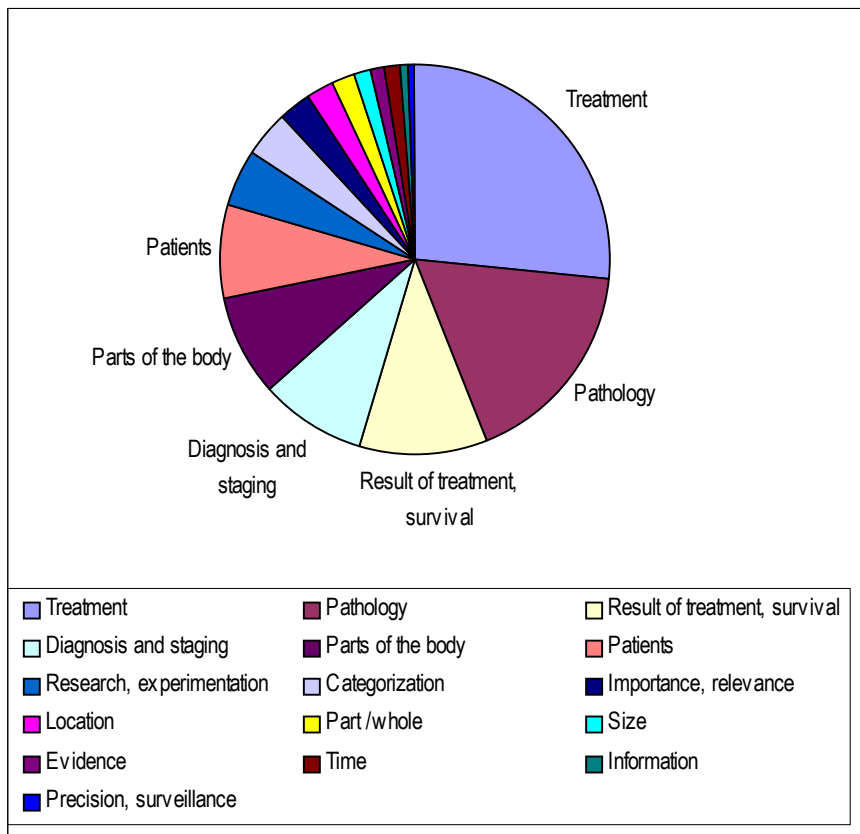
11.1.2.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTAS DE FRECUENCIA LEMATIZADAS

Al agrupar los lexemas con una frecuencia relativa mayor o igual al 0,1% según el concepto genérico al que aluden, hemos identificado 16 áreas conceptuales. Al exportar los datos a la hoja de cálculo *Microsoft Excel* obtenemos la tabla y el gráfico siguientes:

TABLA 67: Identificación preliminar de áreas conceptuales activadas en QDT2 a partir de listas de frecuencia lematizadas

DT2	
ÁREAS CONCEPTUALES	%
Treatment	9,49
Pathology	6,2
Result of treatment, survival	3,7
Diagnosis and staging	3,16
Parts of the body	2,97
Patients	2,85
Research, experimentation	1,67
Categorization	1,41
Importance	0,95
Location	0,84
Part /whole	0,68
Size	0,47
Evidence	0,42
Time	0,42
Information	0,22
Precision, surveillance	0,22
TOTAL	35,7

DIAGRAMA 22: Representación preliminar de las áreas conceptuales activadas en QDT2



11.1.2.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

Hesperus ha reconocido 14 cadenas cohesivas (véase Apéndice VIIe). De acuerdo con el perfil del texto, el número de vínculos cohesivos que se establecen es cuatro veces el del texto anterior (46217), en parte, por ser más extenso (5523 palabras) y porque las cadenas cuentan con muchos más elementos léxicos. De estos, 2470 son *palabras plenas*, que se vinculan mediante 196 categorías conceptuales del tesoro *Roget*. Le corresponde, por tanto, un coeficiente de cohesión de 8,37.

Desde el punto de vista conceptual, *Hesperus* destaca las siguientes categorías semánticas, de las cuales hemos seleccionado las que tienen un porcentaje de activación en el texto¹³⁹ mayor del 1 %.

```
therapy_658_4268_n Percent: 21.07,
sick-person_651_4202_n Percent: 7.02,
show_522_3381_v Percent: 5.69,
minuteness_196_1271_n Percent: 3.46,
surgery_658_4267_n Percent: 3.02,
medical_658_4274_a Percent: 2.72,
experiment_461_2959_n Percent: 2.64,
cancer_651_4195_n Percent: 2.56,
respiratory-disease_651_4192_n Percent: 1.90,
piece_53_370_n Percent: 1.84,
existence_1_1_n Percent: 1.73,
fundamental_156_978_a Percent: 1.67,
remedial_658_4273_a Percent: 1.66,
durability_113_697_n Percent: 1.62,
swelling_253_1658_n Percent: 1.56,
transference_272_1809_n Percent: 1.50,
quantity_26_160_n Percent: 1.23,
time_108_672_n Percent: 1.20,
small_33_224_a Percent: 1.15,
medical-art_658_4266_n Percent: 1.11,
part_53_366_n Percent: 1.09,
limit_236_1565_n Percent: 1.08,
diagnostic_658_4258_n Percent: 1.01
```

Estos conceptos son el hilo conductor de las catorce cadenas que reconoce *Hesperus*:

¹³⁹ Estos datos los hemos sacado del cálculo que hace *Hesperus* en el perfil que crea de cada documento (Ver apéndice VIIb).

TABLA 68: Cadena 0, 1 y 2a (QDT2)

CONCEPTOS ACTIVADOS					
Cadena 0 (Treatment)		Cadena 1 (Patients)		Cadena 2a (Stages)	
THERAPY	21,07	SICK-PERSON (Patients)	7,02	SHOW (Staging) ¹⁴⁰	5,69
SURGERY	3,02	CANCER	2,56	DRAMATURGY	0,53
MEDICAL	2,72	RESPIRATORY- DISEASE	1,90		
REMEDIAL	1,66	FUNDAMENTAL ¹⁴¹	1,67		
MEDICAL-ART	1,11	DISEASE	0,69		
DIAGNOSTIC	0,97	ULCER	0,51		
MEDICINE	0,94	ILLNESS	0,37		
REMEDY	0,52	PATHOLOGY	0,36		
COMPOSITION (Combination)	0,48	WOUND	0,16		
ANTIDOTE	0 0,1	CARDIOVASCULAR – DISEASE (Smoking)	0,11		
DOCTOR	0 0,1	DISEASED	0 0,1		
DRUG	0 0,1				
RELIEF	0 0,1				
TOTAL	32,49+	TOTAL	15,19+	TOTAL	6,22
↓		↓		↓	
TREATMENT AND DIAGNOSIS		PATHOLOGY OF LUNG CANCER		STAGING	

Comentario

En la cadena 0 se asocian lexemas relacionados principalmente con los conceptos THERAPY, SURGICAL, MEDICAL, REMEDIAL, MEDICAL-ART, DIAGNOSTIC. La suma de los porcentajes de activación de cada concepto es superior al 32%, lo cual nos indica la elevada aportación de esta cadena a la cohesión del texto.

A esta cadena, que cuenta con 364 lexemas, tendremos que añadir términos que, aunque están muy directamente relacionados con estos significados, por su alto grado de especialización, no son reconocidos por un tesoro de la lengua general como el *Roget*. Podemos citar, entre otros, tanto lexemas como unidades poliléxicas complejas que hacen

¹⁴⁰ La importancia de este concepto se debe en cierta medida a que *stage*, que en este campo no tiene nada que ver con representaciones teatrales, aparece en esta cadena asociado a SHOW y a DRAMATURGY (0,53 %).

¹⁴¹ Este concepto es el hiperónimo de *primary*, que aparecía originariamente en la cadena 2. El hecho de que aparezca normalmente en el radio colocacional de lexemas que designan la enfermedad nos ha llevado a incluirlo aquí.

alusión a los estadios del cáncer de pulmón (*stage, N, T, M, IV*), a fármacos utilizados en quimioterapia (*cisplatin, paclitaxel, vinorelbine*), pruebas de diagnóstico (*magnetic resonance imaging*). Algunos de los elementos de las unidades terminológicas del texto han sido identificados por *Hesperus* como unidades aisladas, tal y como *scanning* y *tomography*. Estos últimos, en textos sobre oncología, aparecen juntos formando el término *computerized tomography scanning*.

En consecuencia, esta cadena apunta principalmente hacia dos conceptos: TREATMENT, en sus diversas modalidades, y DIAGNOSIS. En nuestra propuesta definitiva mantendremos esta división.

En la cadena 1 predominan las referencias al paciente y a la patología del cáncer de pulmón. El porcentaje de activación elevado, superior al 15%, nos indica la importancia de esta cadena. Añadiremos con posterioridad términos frecuentes en este subdominio.

La cadena 2 resulta difusa puesto que agrupa lexemas en torno a conceptos bastantes variados, de ahí que la desglosemos. En la que hemos denominado 2a, *Hesperus* ha destacado la importancia de los conceptos SHOW y DRAMATURGY en la cohesión del texto y los ha asociado a los lemas SHOWED y STAGE, que, en el texto obviamente no designa un escenario sino los estadios por los que pasa la enfermedad. Dejando a un lado este error semántico, es cierto que el lexema *stage* y sus diversas variantes morfológicas computan un total de 72 ocurrencias, con lo que se convierte en el segundo lema más frecuente del texto. De hecho, según *Wordsmith Tools*, un 1,3 % de todas las variantes morfológicas del texto está representado por este lema. En cuanto al lema SHOWED, representa un 0,29%.

TABLA 69: **Cadena 3** (QDT2)

CONCEPTOS ACTIVADOS	
Cadena 3 (Small)	
MINUTENESS	3,46
QUANTITY	1,23
SMALL	1,15
LITERAL	0,47
SMALL-QUANTITY	0,21
ALMOST	00,1
INCONSIDERABLE	00,1
LITTLENESS	00,1

SLIGHTLY	00,1
TOTAL	6,52 +
↓	
MINUTENESS, QUANTITY, SIZE	

Comentario

En esta cadena es muy significativo que muchas unidades léxicas de la lengua general son parte de unidades terminológicas poliléxicas que nombran los distintos tipos de cáncer de pulmón. De hecho, *small* aparece 21 veces en *Non-small cell lung cancer* y una vez en *small cell lung cancer*. Lo mismo sucede con *cell*, que también aparece en estos dos términos y en otros cinco que designan la enfermedad (*clear cell, giant cell, large cell carcinoma, nonsquamous cell cancers y spindle cell variant*). Esta cadena también contiene dos lexemas que en el texto adquieren un sentido restringido y que vamos a incluir en la cadena *TRATAMIENTO: fraction(s), fractionated, ray* relacionadas con la *RADIOTERAPIA*; *microscopic*, que incluiremos en el *DIAGNÓSTICO* (cadena 6) de la enfermedad; *bronchial compression y shortness of breath* que son *SÍNTOMAS* de la enfermedad.

De todas formas, se aprecia una presencia importante de conceptos que indican *TAMAÑO*, en concreto, tamaño reducido. Esta categoría es la que *Roget* denomina *QUANTITY BY COMPARISON WITH A SIMILAR OBJECT*. Por este motivo, hemos incluido también el concepto *QUANTITY* en la cadena.

TABLA 70: **Cadenas 4, 7 y 11 (QDT2)**

CONCEPTOS ACTIVADOS					
Cadena 4 (Nodules)		Cadena 7 (Blood)		Cadena 11 (Pleura)	
SWELLING	1,56	BLOOD	0,68	LATERAL (Pleura)	0,76
HARD	0,13	SANGUINEOUS	0,22	LATERALITY	00,1
BOSOM	00,1	FLUID	0,16		
HARDNESS	00,1				
TOTAL	+1,69	TOTAL	1,06	TOTAL	+0,76
↓		↓		↓	
HARD BODY PARTS		BODY FLUIDS		PLEURA	

Comentario

Hemos puesto en relación la cadena 4, que hace alusión a partes del cuerpo endurecidas y/o abultadas, con las cadenas 7 y 11, que también aluden a partes del cuerpo y a fluidos corporales.

En la cadena 7 asociaremos los fluidos corporales aludidos en el texto con algunos de los lugares donde se ubican (*lymph nodes*) o generan (*bone marrow*). Hemos prescindido de seis lexemas relacionados con el concepto CLASSIFICATION, por lo que los incluiremos en la cadena CATEGORIZATION.

En la cadena 11, al igual que ocurrió en el texto anterior, la *pleura* se destaca como la parte del cuerpo más relevante en el cáncer de pulmón. No obstante, como veremos, hay otras partes del cuerpo que aparecen con una mayor frecuencia en el texto. De hecho, vemos como a la categoría a la que pertenece *pleura*, LATERAL, no le corresponde un porcentaje muy alto.

Por este motivo, añadiremos en la cadena definitiva referencias al cuerpo humano que creemos que deberían estar en la cadena BODY PARTS.

TABLA 71: Cadenas 5 y 9 (QDT2)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN			
Cadena 5 (<i>Metastasis</i>)		Cadena 9 (<i>Distant</i>)	
TRANSFERENCE	1,50	DISTANT /ALONE	0,84
		DISTANCE	0 0,1
TOTAL	1,50	TOTAL	0,84+
↓		↓	
TRANSFERENCE		DISTANCE	

Comentario

Hesperus también ha detectado la faceta del cáncer como una enfermedad que se extiende a otras partes del cuerpo, mediante el lema METASTASIS. Hay otros lexemas

relacionados con el punto de partida y el destino al que llegan las células cancerosas como resultado de esta transferencia, que aumentan la representación de este concepto.

La capacidad de las células cancerosas de producir metástasis en partes del cuerpo que pueden estar bastante alejadas del tumor primario queda también reflejada en la cadena 9, de ahí que las presentemos asociadas. Hemos eliminado de esta cadena *giant*, perteneciente al término *giant cell carcinoma*, al igual que *single* y *confused*.

TABLA 72: **Cadena 6** (QDT2)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 6 (Trials)	
EXPERIMENT-PREPARATION	2,72
LEARNING	0,11
LITERATURE	0,11
TOTAL	2,94
↓	
RESEARCH, EXPERIMENTATION	

Comentario

En un área donde se producen avances vertiginosamente y se ensayan nuevos tratamientos, el concepto de EXPERIMENTACIÓN está bastante presente, como indica esta cadena.

El término identificado automáticamente, *trial(s)*, está precedido en 27 de 34 ocasiones por *clinical*, con la que constituye la UF *clinical trial*, cuyo significado está íntimamente ligado al concepto TRATAMIENTO. Añadiremos a la cadena *RESEARCH / EXPERIMENTATION* los lexemas *testing*, *study*, *meta-analysis* y *laboratory*.

TABLA 73: Cadena 8 (QDT2)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 8 (Careful)	
SURVEILLANCE	0,65
CAREFUL	0,57
TOTAL	1,22
↓	
PRECISION, SURVEILLANCE	

Comentario

La precisión y exactitud en el diagnóstico acompañado del seguimiento y cuidado del paciente resultan esenciales en la curación del cáncer.

TABLA 74: Cadenas difusas: 10 y 12 (QDT2)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN			
Cadena 10 y 12 a (Group)		Cadena 12b (Poor)	
PIECE	1,84	UNPROSPEROUS	0,11
PART	1,09	IMPERFECT	00,1
PARTY	0,65	IMPERFECTION	00,1
UNION	0,65	MISTAKEN	00,1
PARTITION	0,22	UNEQUAL	00,1
UNIT	0,17		
TOTAL	4,62	TOTAL	+0,11
↓		↓	
PART / WHOLE COMPOSITION		IMPERFECTION	

Comentario

En la cadena 12 predominan los conceptos PART e IMPERFECT(ION). Si bien es comprensible la asociación entre TOTALIDAD y PERFECCIÓN, y entre PARCIALIDAD e IMPERFECCIÓN, creemos conveniente delimitar estos significados, de ahí que la hayamos desglosado en dos, por una parte, PART /WHOLE, COMPOSITION y por otra, IMPERFECTION.

Hemos suprimido de la cadena 12a, *involvement* que enviaremos a la cadena *METASTASIS* y *segmental* que coocurre con *resection* y, por tanto, hace alusión a un tipo de operación quirúrgica en la que sólo se extrae una pequeña parte del pulmón.

Relacionada con la cadena 12a mediante el concepto UNION, hallamos la cadena 10, que es muy difusa y contiene palabras que no tienen ninguna vinculación en el área de la oncología. Basta comentar que *Hesperus* asocia la palabra *local* con el concepto TAVERN (TABERNA), y de ahí que, en la cadena quede asociada con *joint*, perteneciente a la categoría 301 del *Roget* (FOOD) en su subdivisión *meat*. Esta especialidad culinaria de cordero o cerdo asado se puede consumir en una taberna, lo cual no está ni remotamente relacionado con el *American Joint Committee on Cancer* del texto. En vista de lo errónea de esta cadena, sólo nos queda señalar una relación difusa entre el contenido semántico UNION, presente en *joint* y *independent*, y el contenido semántico de la cadena *PART/WHOLE*.

TABLA 75: Cadena difusa: 2b (QDT2)

CONCEPTOS ACTIVADOS	
Cadena 2b	
SHOW	5,69
FUNDAMENTAL	1,67
EFFECT	0,84
EVIDENCE	0,43
PRODUCTION	0,38
EXHIBIT	0,29
CAUSE	0,23
PRODUCE	0,22
CAUSAL	0,21
MANIFESTATION	0,16
ACTING	0 0,1
CONDUCE	0 0,1
CONTAIN	0 0,1
PRODUCED	0 0,1
PRODUCT	0 0,1
REASON-WHY	0 0,1
STAGE-SHOW	0 0,1
TOTAL	9,8 +
↓	
EVIDENCE –FINDINGS, IMPORTANCE, CAUSE AND EFFECT	

Comentario

La cadena 2b consta de numerosas referencias al método científico ya que se destacan las ideas EVIDENCE, IMPORTANCE, CAUSE AND EFFECT. En el caso de EFFECT, el concepto se actualiza con alusiones a los resultados obtenidos tras el tratamiento. Por otra parte, también aparecen las formas léxicas *agent*, *combination*, y *combined* que ubicaremos en la cadena TREATMENT, *primary* y *growth*, cuyo sentido en el texto apunta a la cadena CANCER, *imaging*, que irá con DIAGNOSIS y *final*, con TIME.

TABLA 76: Cadena 13 (QDT2)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 13 (Survival)	
EXISTENCE	1,73
LONG-DURATION	0,32
BE	0,32
LASTING	0,11
TOTAL	2,48
↓	
EXISTENCE	

Comentario

Esta cadena incide en la idea de SUPERVIVENCIA. Al igual que ocurrió en el texto anterior, *Hesperus* identifica la colocación de *long-term* y *survival*.

11.1.2.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

Además de los conceptos que hemos mencionado, el *perfil* del documento parece indicar que es posible identificar más cadenas, ya que hay otros conceptos cuyo porcentaje refleja una presencia significativa en el texto. A partir de estos conceptos podremos asignar una etiqueta a los lexemas reconocidos por *Hesperus* que no encajan dentro de ninguna cadena. A continuación, incluimos los conceptos que justifican la propuesta de nuevas cadenas.

Conceptos reconocidos por Hesperus	Cadenas (% de activación conceptual)
LIMIT (1,08) → BE-NEAR (0,65) → SUPERIOR (0,49) → → TOWARDS (0,32) → SET –APART (0,32) → NEAR (0,22) → → BE-SITUATED (0,22)	LOCATION (3,3)
TIME (1,2) → DATE (0,43) → AGE (0,32)	TIME (1,95)
IMPORTANT (0,36)	IMPORTANCE (0,36)
CIRCUMSTANCE (0,26) → SITUATION (0,11)	CIRCUMSTANCE (0,37)
REQUIRE (0,26)	NEED (0,26)
CHOSEN (0,18) → MODALITY (0,11)	CHOICE (0,29)
DISSIMILAR (0,39) → DIFFERENT (0,35) → ARRANGEMENT (0,26) → MARK (0,22) → CLASSIFICATION (0,18) → COMPARISON (0,16)	CATEGORIZATION (1,56)

Con todo esto y teniendo en cuenta las áreas conceptuales normalmente activadas tanto en el *EVEN TO ONCOLÓGICO* como en los textos relacionados del corpus, propondremos estas cadenas:

TABLA 77: Cadenas léxicas en QDT2

<p>Chain 0: TREATMENT (H0, H2b, H3) [706]</p> <p>0.1. Generic treatment 0.2. Properties (→TREATMENT→) 0.2.1. Location of application 0.3. Manipulation 0.4. Process 0.5. Types (according to therapeutic agent) 0.5.1. Radiotherapy 0.5.1.1. Instruments 0.5.1.2. Substances used 0.5.1.3. Measure of the dose of radiation 0.5.2. Chemotherapy 0.5.2.1. Chemotherapeutic agents 0.5.3. Surgery 0.5.3.1. Types of surgical resection 0.5.3.1.1. Exploratory surgery 0.5.3.2. Radiosurgery 0.5.4. Laser therapy 0.5.5. Phototherapy 0.5.6. Immunotherapy 0.5.7. Chemoprevention</p> <p>Chain 1: PATHOLOGY OF NON-SMALL CELL LUNG CANCER (H1, H2b, H3) [343]</p> <p>1.1. Generic disease 1.2. Cancer 1.2.1. Properties of cancer 1.2.2. Types 1.2.2.1. Anatomic classification 1.2.2.2. Histologic classification</p> <p>Chain 2: RESEARCH, EXPERIMENTATION (H2b, H6, H8, H10, CATEGORIZATION 1,56 %) [306]</p> <p>2.1. Study 2.2. Evaluation 2.3. Evidence 2.4. Categorization 2.5. Comparison 2.6. Cause 2.7. Information</p> <p>Chain 3: OUTCOME OF TREATMENT (H2b, H10, H13) [287]</p> <p>3.1. Generic outcome of treatment 3.1.1. Cause 3.2. Positive outcome 3.3. Existence 3.3.1. Survival 3.3.2. Death 3.4. Negative result</p>	<p>Chain 4: LOCATION IN THE HUMAN BODY (H1, H4, H7, H10, H11, LOCATION 3,3%) [238]</p> <p>4.1. Generic location 4.1.1. Generic location (medicine) 4.2. Microscopic / histologic level 4.3. Anatomic level / body parts 4.3.1. Respiratory system 4.3.2. Circulatory system 4.3.3. Osseous body parts 4.3.4. Other body parts 4.4. Body fluids 4.5. Location in relation to specific body parts 4.5.1. Bronchi 4.5.2. Carina 4.5.3. Clavicle 4.5.4. Cranium</p> <p>Chain 5: PATIENT (H1) [172]</p> <p>5.1. Properties of PATIENT</p> <p>Chain 6: DIAGNOSIS (H0, H1, H8, H2b, H3) [165]</p> <p>6.1. Generic diagnosis 6.2. Sensory processes involved 6.3. Cognitive processes involved 6.4. Characteristics 6.5. Clinical features (Symptoms) 6.6. Prognostic factors 6.7. Diagnostic tests and instruments</p> <p>Chain 7: STAGING (H2a) [134]</p> <p>7.1. Stages of non-small cell lung cancer</p> <p>Chain 8: SIZE / QUANTITY (H3) [114]</p> <p>8.1. Size 8.2. Name of lung cancer 8.3. Minuteness / small quantity 8.4. Largeness / large quantity 8.5. Increase</p> <p>Chain 9: METASTASIS (H5, H9, H2b, H10) [98]</p> <p>9.1. Distance</p> <p>Chain 10: COMPOSITION (H2b, H9, H10, H12a) [53]</p> <p>Chain 11: GENERAL MEDICINE (H0) [42]</p> <p>Chain 12: TIME (H2b, TIME 1,95 %) [40]</p> <p>Chain 13: IMPORTANCE (H2b, IMPORTANCE 0,36%) [34]</p>
---	---

Chain 14: OPTIONS (CHOICE 0,29%) [34]	Chain 18. VARIATION [4]
Chain 15: PRECISION / SURVEILLANCE (H8) [34]	Chain 19. RISK FACTOR: SMOKING (H1) [3]
Chain 16: NEED (H5, NEED 0,26 %) [8]	TOTAL DE UNIDADES LÉXICAS ENCADENADAS ¹⁴² : [2574]
Chain 17: IMPERFECTION (H9, H12b) [7]	

Al igual que en el texto anterior, la cadena más importante es la de *TRATAMIENTO*, que adquiere una representación conceptual del 26,12%, seguida de *PATHOLOGY* (11,6%) y *RESEARCH / EXPERIMENTATION* (10,31%). En cuarto lugar de importancia, la cadena *TREATMENT OUTCOME*, compuesta principalmente por lexemas de la lengua general, vincula un 9,9% de todas las palabras encadenadas.

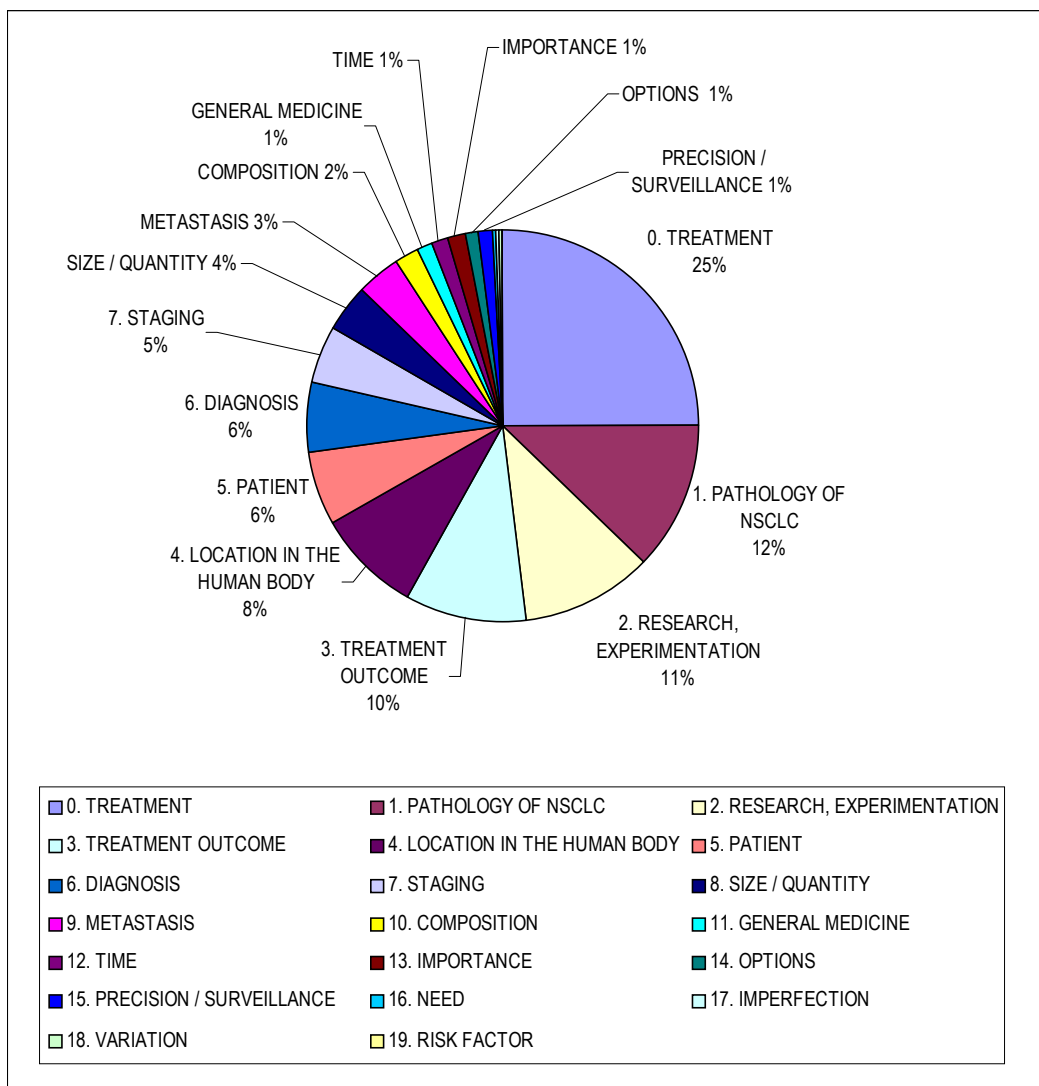
A continuación, destacan cuatro cadenas repletas de términos médicos y que en el texto sobre el mesotelioma también se encontraban entre las ocho primeras: *LOCATION IN THE HUMAN BODY*, *PATIENT*, *DIAGNOSIS* y *STAGING*. Cada una de estas cadenas tiene un porcentaje superior o próximo al 5 %. Hay once cadenas que cuentan con un porcentaje igual o menor al 4% y contribuyen al 15% de la cohesión del texto. Entre estas, destacamos las comunes a:

- los otros dos textos para especialistas: *SIZE / QUANTITY*, *COMPOSITION*, *GENERAL MEDICINE*, *TIME*, *IMPORTANCE*, *OPTIONS*.
- el texto QDT1: *RISK FACTOR*
- el texto QDT3: *NEED*, *VARIATION*.

Por último, las cadenas que sólo hemos identificado en este texto son la de *METASTASIS*, que en los demás textos hemos subsumido bajo la de *PATHOLOGY*, y las de *PRECISION / SURVEILLANCE* e *IMPERFECTION*.

¹⁴² Para obtener el total no se han duplicado los lemas que pertenecen a más de una cadena. Si sumáramos los elementos de cada cadena, obtendríamos un total de 2822.

DIAGRAMA 23: Porcentaje de activación de las cadenas léxicas definitivas de QDT2



CADENA	Frec	%
0. TREATMENT	706	25,02
1. PATHOLOGY OF NSCLC	343	12,15
2. RESEARCH / EXPERIMENTATION	306	10,84
3. TREATMENT OUTCOME	287	10,17
4. LOCATION IN THE HUMAN BODY	238	8,43
5. PATIENT	172	6,09
6. DIAGNOSIS	165	5,85
7. STAGING	134	4,75
8. SIZE / QUANTITY	114	4,04
9. METASTASIS	98	3,47
10. COMPOSITION	53	1,88

11. GENERAL MEDICINE	42	1,49
12. TIME	40	1,42
13. IMPORTANCE	34	1,2
14. OPTIONS	34	1,2
15. PRECISION / SURVEILLANCE	34	1,2
16. NEED	8	0,28
17. IMPERFECTION	7	0,25
18. CHANGE	4	0,14
19. RISK FACTOR	3	0,11
	2822	99,98

11.1.2.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS

0. TREATMENT (H0, H2b, H3)

0.1. Generic treatment

TREATMENT - Treatment 37 ~ option(s) 12 ~ planning 2 - Treated 16 (See 5.1.) - Treatments 3	56
TRIALS (See 2.1.) Trials 27, trial 7 Clinical ~ 27	34
PALLIATION - Palliation 9 ~ of symptoms 5 - Palliative 3 - Palliate 1 - Palliating 1	19
THERAPY - Therapy 9 Combined modality ~ 4 Systemic ~ 1 (See 0.5.6. ENDOSCOPIC PHOTODYNAMIC ~ ; 0.5.8. ENDOBRONCHIAL LASER ~ , LASER THERAPY) - Therapeutic 4	13
CURATIVE (See 2.1.) Curative 13 ~ intent 8 Cure 3	16
MODALITIES (See 14.) Modalities 8	14

Modality 6 Combined ~ + THERAPY 5	
COMBINED → (✖ 2b) - Combined 6 TREATMENT ~ with TREATMENT 6 ~ modality + THERAPY 5 - Combining + THERAPY + THERAPY 1	7
OPTIONS (See 14.) Options 10, option 2 Treatment ~ 11	12
CONTROL Local ~ 5 (See 4.1.) ~ of regional disease 2	8
FORM → (+ TREATMENT) Form 2 Forms 2, formal 1	5
MANAGEMENT Management 3, managed 2	5
STRATEGY (therapeutic ~)	1
TECHNIQUE (See 2.1.)	1
INTERVENTION (therapeutic ~)	1
TOTAL	192

0.2. Properties (→ TREATMENT →)

→ ALONE Radiotherapy ~ 12 Surgery alone ~ 4 Chemotherapy ~ 2	18
ADJUVANT → ~ chemotherapy 6	12
POSTOPERATIVE → (See 0.5.3., 12.) ~ (whole-brain) irradiation 4 ~ radiotherapy 3 ~ chemotherapy 1	8

COMBINED MODALITY → ~ therapy 2 ~ approaches 2 COMBINED-MODALITY (~ therapy)	5
CONVENTIONAL → + RADIOTHERAPY	5
STANDARD → ~ chemotherapy 2 ~ treatment 2	5
SYSTEMIC →	4

~ chemotherapy 3 ~ therapy 1	
ADDITIONAL →	3
CONCURRENT → ~ radiotherapy 2	3

NEOADJUVANT → ~ chemotherapy 2	3
PALLIATIVE (~ radiotherapy) →	4
TOTAL	66

0.2.1. Location of application (→TREATMENT)

ENDOBONCHIAL (See 4.5.) (~ laser therapy)	3
--	---

0.3. Manipulation

USE (+ TREATMENT) Use 7, used 6, using 3	16
GIVEN (+ TREATMENT)	6
ADMINISTRATION	2

(+ CHEMOTHERAPY)	
TOTAL	24

0.4. Process

SCHEDULE (+ TREATMENT) Schedule 2, schedules 2	4
DOSES (+ RADIOTHERAPY)	3
CYCLES (+ CHEMOTHERAPY) ~ of combination therapy	2
PROGRAMS (+ CHEMOTHERAPY)	1
TOTAL	10

0.5. Types (according to therapeutic agent)

0.5.1. Radiotherapy

RADIOTHERAPY ~ alone 14 ~ with curative intent 6 plus ~ 4 postoperative ~ 4 concurrent ~ 2 definitive ~ 2 primary ~ 2 chest ~ 1 external-beam ~ 1 Interstitial ~ 4	65
FRACTIONATION (✖ 3) - fractionation 5	9

(+ CHEMOTHERAPY)	
IRRADIATION - Chest ~ 3 - Postoperative ~ 2 - Whole-brain ~ 2	9
CONVENTIONAL → (+ RADIOTHERAPY)	5
BRACHYTHERAPY	4

DOSES	3
EXTERNAL-BEAM (~ radiotherapy)	1

INTERSTITIAL (~ radiotherapy)	1
TOTAL	88

0.5.1.1. Instruments

MEGAVOLTAGE EQUIPMENT	2
-----------------------	---

0.5.1.2. Substances used

RADIOSENSITIZERS	1
RADIOLABELED ANTIBODIES	1
TOTAL	2

0.5.1.3. Measure of the dose of radiation

cGy	6
-----	---

0.5.2. Chemotherapy

CHEMOTHERAPY	55
- CHEMOTHERAPY 54 combination ~ 7 adjuvant ~ 7 ~ regimen(s) 7 cisplatin-based ~ 6 ~ combined with THERAPY 5 ~ combined with other modalities 3 systemic ~ 3 ~ alone 2 neoadjuvant ~ 2 single-agent ~ 1	
- CHEMOTHERAPEUTIC 1	
REGIMENS	13
Regimens 8, Regimen 5 Chemotherapy ~ 7 Cisplatin-containing ~ 2	
COMBINATION → (✖ 2b)	11
Combination chemotherapy 9 Cisplatin-based combinations 2	
→ CISPLATIN-BASED (✖ 2b)	8

~ chemotherapy 6 ~ combinations 2	
AGENT (✖ 2b) Single-agent 2 Single agent 1 Agents 1	4
SYSTEMIC → (~ chemotherapy)	3
ADMINISTRATION (schedule of drug ~)	2
CYCLES ~ of () chemotherapy	2
INITIAL → (~ chemotherapy) (See 12.)	1
PROGRAMS	4
TOTAL	84

0.5.2.1. Chemotherapeutic agents

AGENT	4
Single-agent 2	
Single agent 1	
Agents 1	
DRUG	2
CISPLATIN	24
Cisplatin 13	
cisplatin-based 8	
cisplatin plus vinblastine 5	
cisplatin plus paclitaxel 3	
cisplatin plus vinorelbine 3	
cisplatin-containing 3	
PACLITAXEL (7) ≈	9
Taxol (2)	
VINORELBINE	5
CARBOPLATIN	2

CYCLOPHOSPHAMIDE	2
DOCETAXEL (1) ≈	2
Taxotere (1)	
VINDESINE	2
DOXORUBICIN	1
ETOPOSIDE	1
GEMCITABINE	1
IRINOTECAN	1
ISOTRETINOIN	1
MITOMYCIN	1
TOPOTECAN	1
VINBLASTINE	1
TOTAL	56

0.5.3. Surgery

RESECTION	56
- Resection 29	
Surgical ~ 12	
Wedge ~ 6	
Segmental ~ 3	
Complete ~ 2	
Sleeve ~ 2	
Segment ~ 1	
- Resected 11 (See 5.1.)	
- Resectable 6 (See 1.2.1.)	
- Unresectable 6 (See 1.2.1.)	
- Resections 3	
- Resect 1	
SURGERY	52
- surgery 29	
~ alone 4	
contraindication to ~ 3	
- surgical 16	
~ resection 12	
~ management 2	
- surgically 7	
(+ PATIENTS+ / + TUMOR)	
Patients treated ~ 3	
INOPERABLE	9

- <u>inoperable</u> 7	
(See 1.2.1., 5.1)	
~ patients 3,	
~ nsc1c 2,	
~ stages IIIb and IV disease	
- <u>operable</u> 2 (See 5.1.)	
~ patients 2	
<u>POSTOPERATIVE</u> (See 12.)	8
<u>LIMITED</u> → (See 8.3.)	3
~ resection(s) 2	
~ excision 1	
<u>COMPLETE</u> → (~ resection)	2
(See 10.)	
<u>EXCISION</u>	2
TOTAL	124

0.5.3.1. Types of surgical resection

LOBECTOMY	10
WEDGE RESECTION	6
SEGMENTAL RESECTION (See 10.)	3
SEGMENTECTOMY Anatomic ~ 1	2

SLEEVE RESECTION	2
PNEUMONECTOMY	2
SEGMENT RESECTION	4
TOTAL	26

0.5.3.1.1. Exploratory surgery

MEDIASTINOTOMY (See 6.7.)	1
---------------------------	---

0.5.3.2. Radiosurgery

STEREOTACTIC RADIOSURGERY	3
RADIOSURGICALLY	1
TOTAL	4

0.5.4. Laser therapy

ENDOBONCHIAL LASER THERAPY	3
LASER THERAPY	1
TOTAL	4

0.5.5. Phototherapy

ENDOSCOPIC PHOTODYNAMIC THERAPY	2
HEMATOPORPHYRIN DERIVATIVE	1
ENDOSCOPIC PHOTOTHERAPY	1
TOTAL	4

0.5.6. Immunotherapy

IMMUNOTHERAPY	3
---------------	---

0.5.7. Chemoprevention

CHEMOPREVENTION	3
-----------------	---

1. PATHOLOGY OF NON-SMALL CELL LUNG CANCER (H1, H2b, H3)

1.1. Generic disease

DISEASE STAGE + ~ 12 Disease-free survival 5 (See 3.3.1.) Advanced-stage ~ 2 Regional ~ 2	25
SYMPTOMS (See 6.5.) Symptoms 11, Symptomatic 5, asymptomatic 1	17
LESION Lesion 7, lesions 7 Primary ~ 2 New ~ 2 Obstructing ~ 2	14
PRESENT (+ DISEASE)	8

Present 6 ~ with 3 Presentation 2	
DEVELOP (See 9.) Develop 4, developing 1	5
PATHOLOGIC - Pathologic 4 ~ stage 2 - Pathologist 1	5
TOTAL	74

1.2. Cancer

TUMOR Tumor 53, tumors 17	70
< CANCER ¹⁴³ > (See 5.1.) Cancer 10, cancers 8	18
CARCINOMA	13
MALIGNANT - Malignant 4 ~ pleural effusion 2, ~ cells 1, ~ tumors 1 - malignancies 1	5
ONCOLOGY	1
TUMOR GROWTH (See 8.5.)	1
TUMOR MASS	1
TUMOR SPECIMEN	1
TOTAL	110

¹⁴³ El lema *cancer* aparece en 59 ocasiones, en 41 de las cuales constituye la UF *lung cancer*. En 1.2. sólo mencionamos las ocasiones restantes.

1.2.1. Properties of CANCER (→ CANCER) (H2b)

PRIMARY (✖ 2b)	31
SECOND	14
<u>RESECTABLE</u> (+CANCER+) (See 0.5.3.) Resectable 5, unresectable 5	10
<u>EARLY</u> (See 12.) ~ stage I tumors ~ -stage lung cancer ~ central tumors	4

~ NSCLC	
<u>INOPERABLE</u> +CANCER+ (See 0.5.3.)	3
<u>MALIGNANT</u> (~ tumors)	4
TOTAL	62

1.2.2. Types

<u>TYPES</u> (~ of CANCER) (See 2.4.)	3
---------------------------------------	---

1.2.2.1. Anatomic classification

<u>LUNG CANCER</u> Lung cancer 37 Lung cancers 5	42
<u>SUPERIOR SULCUS</u> TUMOR (S) (3) (See 4.3.1.) ≈	4

~ NSCLC	
<u>BRONCHOGENIC CARCINOMA</u>	1
TOTAL	47

1.2.2.2. Histologic classification

SMALL CELL LUNG CANCER	4
NSCLC (26) ≈ Non-small cell <u>lung cancer</u> (24) ≈ Non-small cell carcinoma (1) ≈ non-small cell lung cancers (4)	27
SQUAMOUS CELL (<u>EPIDERMOID</u>) CARCINOMA (1) ≈ <u>Epidermoid</u> or squamous carcinoma (1) (See 4.2.) ≈ Squamous cell carcinoma (1) ≈ Squamous cell <u>lung cancer</u> (1)	4
SPINDLE CELL VARIANT	1
NONSQUAMOUS CELL CANCERS	1
ADENOCARCINOMA (Adenocarcinoma 2, adenocarcinomas 1)	3
ACINAR	1
<u>BRONCHOALVEOLAR</u> CANCER (2) ≈ <u>Bronchoalveolar</u> carcinoma (2) (See 4.3.1.)	4
PAPILLARY	1
SOLID TUMOR WITH MUCIN	1
ADENOSQUAMOUS CARCINOMA	1
LARGE CELL CARCINOMA	1
CLEAR CELL	1
GIANT CELL	1
UNDIFFERENTIATED CARCINOMA	1
TOTAL	47

2. RESEARCH, EXPERIMENTATION (H2b, H6, H8, H10)

2.1. Study (H6, H10)

<u>TRIALS</u> (See 0.1.) Trials 27, trial 7 <u>Clinical</u> ~ 27 (See 11.)	34
STUDIES (✖ 10) Studies 15, study 12 <u>Clinical</u> ~ 4 (See 11.)	27
REPORTED Reported 11, report 2, Reports 2	15
RANDOMIZED ~ clinical trials 3 ~ trial(s) 3 ~ study (ies) 3	12
REFERENCES	11
DATA	5
<BASED> Based 3, basis 1	4
CONTROLLED TRIALS	4
META-ANALYSIS	4

Meta-analysis 3 (~ of patient data from NUMBER randomized clinical trials showed) Meta-analyses 1	
MEDIAN	2
TESTED Tested 1, testing 1	2
CITATIONS (✖ 2b)	1
CONTROL <u>GROUP</u> (See 10.)	1
<u>LABORATORY</u> (See 6.7.)	1
METHODS	1
PHASE (~ II studies)	1
<u>TECHNIQUE</u> (See 0.1.)	1
TOTAL	126

2.2. Evaluation (H8)

<u>EVALUATION</u> (See 6.3.) - Evaluating 8 Clinical trials ~ TREATMENT 5 - Evaluation 8 Clinical ~ Diagnostic ~ - Evaluable 1, evaluate 1, evaluated 1	19
<u>CONSIDERED</u> (✖ 8) (See 6.3., 15.)	17
<u>ASSESSED</u> (See 6.3.) Assessed 4, assessment 2	6

<u>EXAMINATION</u> (See 6.2.) Examination 2, examining 2, examined 1, examination 1	6
<u>FACTORS</u> (See 6.3.)	3
<u>JUDGE</u> (✖ 2b) (See 6.3.) Judge 1, judgement 1	2
TOTAL	53

2.3. Evidence (2b)

SHOWED Showed 8 Clinical trials ~ 3 Shown 6, show 2	16
EVIDENCE	7
DEMONSTRATE Demonstrate 3, demonstrated 2	5
OBJECTIVE	5
FOUND	4
DISCOVERED	3

Discovered 2, discovery 1	
APPARENT	1
CLEAR	1
INDICATED	1
REFLECT	1
REPRESENTS	1
TOTAL	45

2.4. Categorization (CATEGORIZATION 1,56%)

CLASSIFICATION Classification 4, classified 4	8
DEFINE Define 3, definition 3, definitions 1	7
DIVIDED	4
<u>TYPES</u> (See 1.2.2.)	3
CATEGORIES	2
DESIGNATED Designated 1, designations 1	2

NATURE Nature 1, natural 1	2
STRUCTURES (✖ 2b)	2
DESCRIBED	1
FIELD	1
QUALITY	1
SPECIMEN (✖ 2b)	1
SUBSET	1
TOTAL	35

2.5. Comparison (H2b, H10)

COMPARED (✖ 2b, 10) Compared 7, Comparison 2, Compare 1, Comparing 1	11
SIMILAR Similar 5, similarly 1	6

DIFFERENCE Difference 3, different 1	4
TOTAL	21

2.6. Cause (H2b) (See 3.1.1. Cause)

PRODUCE (See 3.1.1.) Produce 5, produced 3, produces 1	9
ENCOURAGED Encouraged 3 Encouraging 1	4
LEAD-TO (See 3.1.1.)	2
DERIVATIVE	1

DUE-TO (See 3.1.1.)	1
ESTABLISHING	1
MADE	1
RESPONSIBLE	1
TOTAL	18

2.7. Information

INFORMATION Information 4, informed 2	6
KNOWN	2
TOTAL	8

3. OUTCOME OF TREATMENT (H2b, H10, H13)

3.1. Generic outcome of treatment

RATE Rate 13, rates 9 Survival ~ 10 Response ~ 6 Mortality ~ 2	22
RESPONSE Response 15 ~ rates 4 complete ~ 4 objective ~ 3	18

Responses 3	
RESULTS Results 13, resulted 2	15
OUTCOME Outcome 2, outcomes 2	4
TOTAL	59

3.1.1. Cause (See 2.6. CAUSE)

PRODUCE (See 2.6.) Produce 5, produced 3, produces 1	9
LEAD-TO (See 2.6.)	2

DUE-TO (See 2.6.)	1
TOTAL	12

3.2. Positive outcome (H2b, H10)

IMPROVED - Improved 6 + SURVIVAL 5 - Improvement 6 + CONTROL 3 - Improve 5 + SURVIVAL 3 - Improvements 1	18
BENEFIT - Benefit 12 Survival ~ 4 Consistent ~ 3 - Benefits 4	16
<u>CURATIVE</u> (See 0.1.) Curative 13 TREATMENT + ~ intent 8 Cure 3	16
PERFORMANCE (production n) ~ status 13 Good ~ status 7	13
GOOD → - Good 8 ~ performance 7	12

- Better 2, best 2	
ACHIEVED Achieved 5, achieve 4	9
PALLIATION ~ of symptoms 5	8
APPROPRIATE Appropriate 5, appropriately 1	6
EFFECTIVE Effective 4, Effectively 1	5
EXCELLENT → ~ performance status 4	5
COMPLETE → (~ response)	4
ENHANCE ~ local control	2
OBTAIN	2
RELIEF (✖ 10)	2
TOLERATE	1
TOTAL	119

3.3. Existence (H13, EXISTENCE 2,05%)

3.3.1. Survival (H13, LONG DURATION 0,32%, LASTING 0,11 %)

SURVIVAL - Survival 45 <u>Disease-free</u> ~ 5 (See 1.1.) - Survive 1, survivors 1	47
LONG-TERM → ~ survival 6	8
REMAINS	4
MEDIAN → (~ survival)	3

ALIVE Alive 1, live 1, life 1	3
INCREASED DISEASE-FREE SURVIVAL	2
OCCUR	1
PROLONGED → (~ survival)	1
TOTAL	75

3.3.2. Death

DEATH	2
MORTALITY (postoperative ~ rate)	2
NECROTIC	1
	5

3.4. Negative outcome

RISK Risk 6, risks 2 ~ + CANCER 3 ~ OF DEATH 2	8
<u>POOR</u> (See 17.)	6
TOXIC EFFECTS	3

IMPACT (✖ 10) (~ of chemotherapy)	1
TOTAL	18

4. LOCATION IN THE HUMAN BODY (H1, H4, H7, H10, H11, LOCATION 3,3%)

4.1. Generic location

LOCAL (✖ 10) Local 15 ~ <u>control</u> 5 (See 0.1.) Localized 3, Locally 3, location 1	22
<u>DISTANT</u> (+ METASTASIS) (See 9.1.) Distant 10, distal 2	12
REGIONAL Regional 9 ~ lymph node(s) 3 Region 1, regionally 1	11
SITE	7

Site 4, sites 3	
SUPERIOR ~ sulcus tumor 3 ~ vena cava syndrome 2	7
AREA Area 1, areas 1	2
CONFINED (✖ 1)	1
SUPERFICIAL	1
TOTAL	63

4.1.1. Generic location (medicine)

IPSILATERAL + METASTASIS+ 5	6
CONTRALATERAL + IPSILATERAL +	4
PROXIMAL	3

LOCOREGIONAL (~ NSCLC)	1
TOTAL	14

4.2. Microscopic / histologic level

TISSUE Tissue 3	8
HISTOLOGIC Histologic 2, histologies 2, histologically 1	
CELL Cell 4, Cell + CANCER 30	6

Site 4, sites 3	
NODULES (✖ 4) Nodules 3, nodule 1	4
<u>EPIDERMOID</u> (See 1.2.2.2.)	2
TOTAL	20

4.3. Anatomic level / body parts (H4, H7, H11)

4.3.1. Respiratory system

LUNG - Lung cancer 42 - <Lung> 9 ----- - Pulmonary 12 ~ reserve 4 ~ function 3 impaired ~ function 2 ~ symptoms 2	21
BRONCHUS - Bronchus 6 Main ~ 4 - Bronchial 3 ~ wall 2	14
LOBE Lobe 5, lobar 1	6

PLEURA - Pleura 4 Visceral ~ 2, parietal ~ 1 - Pleural 4	8
BRONCHOALVEOLAR (See 1.2.2.2)	4
CARINA	4
HILAR	3
SUPERIOR SULCUS (~ tumors) (See 1.2.2.1.)	3
TRACHEA Trachea 1, tracheal 1	2
VOCAL CORD	1
TOTAL	66

4.3.2. Circulatory system (H4)

NODES (✖ 4) - Nodes 7, node 4 Lymph ~ 9 Regional lymph ~ 2 - Nodal 2	13
BLOOD Blood 1, bloody 1, non-bloody 1	3
VENA CAVA (✖ 1) Superior ~ syndrome (See 6.5.)	3

VESSELS Vessels 2, vascular 1	3
BONE MARROW (✖ 10)	1
HEART	1
PERICARDIUM	1
TOTAL	25

4.3.3. Osseous body parts (H4)

CHEST WALL	3
TRACHEA Trachea 1, tracheal 1	2
BONE	1
<CHEST>	1

OSSEOUS	1
VERTEBRAL (✖ 10)	1
TOTAL	7

4.3.4. Other body parts

BRAIN	14
Brain 9, whole-brain 3 ~ METASTASIS 6 Postoperative whole-brain irradiation 2	
----- CÉREBRAL 2	
MEDIASTINAL	5
Mediastinal 3, mediastinum 2	
ESOPHAGEAL	2
Esophageal 1, esophagus 1	

EYE (✖ 8)	2
ARM	1
DIAPHRAGM	1
HEPATIC	1
OVARY	1
PERITONEAL	1
SCALENE	1
TOTAL	29

4.4. Body fluids (H7)

LYMPH (~ nodes) (✖ 4)	9
BLOOD	3
Blood 1, bloody 1, non-bloody 1	
FLUID	2
BRONCHIAL WASHINGS	1

EXUDATE	1
SPUTUM	1
TOTAL	5

4.5. Location in relation to specific body parts

4.5.1. Bronchi

ENDOBONCHIAL (See 0.2.1.)	4
~ laser therapy 3 ~ lesions 1	

PERIBRONCHIAL (+ LYMPH NODES)	1
----------------------------------	---

4.5.2. Carina

SUBCARINAL (~ lymph nodes)	1
----------------------------	---

4.5.3. Clavicle

SUPRACLAVICULAR (+ LYMPH NODES)	2
------------------------------------	---

4.5.4. Cranium

EXTRACRANEAL (~ tumor)	1
------------------------	---

5. PATIENT (H1)

PATIENTS Patients 110, patient 14	124
GROUP (+ PATIENTS) (See 10.) - Group 10 ~ of patients 3 - Groups 2	12
CASES	6

CANDIDATE Candidate 4, candidates 1	5
HOST (~ factors)	1
TOTAL	148

5.1. Properties of PATIENT (→ PATIENT →)¹⁴⁴

→ WITH STAGE + NUMBER + DISEASE	10
→ TREATED (See 0.1.) (+ SURGERY / RADIOTHERAPY) ~ surgically 2	8
SELECTED → (See 14.)	7
→ WITH EXCELLENT PERFORMANCE STATUS	4
INOPERABLE + PATIENT + (See 0.5.3.)	4
RESECTED → (See 0.5.3.)	3

CANCER → (See 1.2.)	2
OPERABLE → (See 0.5.3.)	2
→ UNDERGOING (+ RESECTION)	2
→ WITH IMPAIRED PULMONARY FUNCTION	2
→ WITH NSCLC	2
APPROPRIATE →	1
TOTAL	24

6. DIAGNOSIS (H0, H1, H8, H2b, H3)

6.1. Generic diagnosis (H1)

DIAGNOSTIC Diagnostic 3, diagnosis 2, diagnosed 1	6
PROGNOSIS (+ POSITIVE / NEGATIVE +) Prognosis 3, Prognostic 2	5

HISTORY History 1, historical 1	2
TOTAL	13

¹⁴⁴ Eliminamos las formas que coocurren con *patient(s)* con una frecuencia muy elevada para no otorgar a esta cadena un peso que no se pueda respaldar textualmente.

6.2. Sensory processes involved

<u>EXAMINATION</u> (See 2.2.) Examination 2, examining 2, examined 1, examination 1	6
PRESENCE Presence 3, presentation 2	5
OBSERVATION Observation 2, observed 1	3

EXPLORATION	1
IDENTIFIED	1
SIGNS (✖ 2b)	1
TOTAL	17

6.3. Cognitive processes involved (H8)

<u>EVALUATION</u> (See 2.2.) Evaluating 8 Clinical trials ~ TREATMENT 5 Evaluation 8 Clinical ~ 2 Diagnostic ~ 2 Evaluable 1, evaluate 1, evaluated 1	19
<u>CONSIDERED</u> (✖ 8) (See 1.5., 2.2.)	17
<u>ASSESSED</u> (See 2.2.) Assessed 4, Assessment 2 (careful preoperative ~)	6
DETERMINED	6

Determined 2, determining 1, determine 1, determinations 1, Determination 1	
<u>JUDGE</u> (✖ 2b) (See 2.2.) Judge 1, judgement 1	2
ESTABLISHING (+ STAGE)	1
TOTAL	51

6.4. Characteristics

<u>SPECIFICITY</u> (✖ 0) (See 15.) Specificity 3, specific 1, specify 1	5
SENSITIVITY	2

<u>ACCURACY</u> (✖ 8) (See 15.)	1
TOTAL	8

6.5. Clinical features (Symptoms) (H3)

<u>SYMPTOMS</u> (See 1.1.) Symptoms 11, Symptomatic 5, asymptomatic 1	17
PULMONARY RESERVE	4
PULMONARY FUNCTION Impaired ~ 2 Poor ~ 1	3
SYNDROME <u>Superior vena cava</u> ~ (See 4.3.2.)	3
ATELECTASIS (1) ≈ Obstructive pneumonitis (1)	2

PAIN Pain 1, chest pain 1	2
PNEUMONITIS	2
BREATH (shortness of ~) (✖ 3)	1
BRONCHIAL COMPRESSION (✖ 3)	1
COUGH	1
HEMOPTYSIS	1
VOCAL CORD PARALYSIS	1
TOTAL	39

6.6. Prognostic factors (CIRCUMSTANCE 0,26%, SITUATION 0,11%)

STAGE ¹⁴⁵	56
FACTORS (See 2.2.)	3
host factors 1, <u>prognostic factors</u> 1	
MEDICAL CONDITION	2
NODAL STATUS	2
ErB-2 ONCOPROTEIN	1
INCREASED NUMBERS OF BLOOD VESSELS	1
MUTATION OF THE K-ras GENE	1
PULMONARY SYMPTOMS	1
TUMOR SIZE	1
VASCULAR INVASION	1
TOTAL	13

6.7. Diagnostic tests and instruments (H0, H2b)

BRONCHOSCOPY	4
Bronchoscopy 3, Bronchoscopic 1	
COMPUTE+TOMOGRAPHY+SCAN	4
Computed tomography scan 1 ≈ Computerized tomography (CT) scanning 1 ≈ CT 2	
RADIOLOGIC	3
Radiologic (~ studies) 2 Radiology 1	
BIOPSY	2
(CHEST) RADIOGRAPH	2
IMAGING (✖ 2b)	2
- Imaging 1 - Magnetic resonance imaging (MRI) 1	
POSITRON EMISSION	2

TOMOGRAPHY (PET) ≈ PET	
CHEST X-RAY (✖ 3)	1
CYTOPATHOLOGIC EXAMINATIONS	1
LABORATORY STUDIES (See 2.1.)	1
MEDIASTINOSCOPY	1
MEDIASTINOTOMY (See 0.5.3.1.1.)	1
MICROSCOPIC EXAMINATION (See 8.3.)	1
PRETHORACOTOMY EXPLORATION	1
TOTAL	24

¹⁴⁵ Las 56 ocurrencias de *stage* no se computan en este apartado para no otorgar una representación excesiva a este apartado.

7. STAGING (H2a)

STAGE (✖ 2a) Stage 56, Staging 10, Stages 2, Staged 4, Overstaging 1, understaging 1	72
STAGE + II + (LUNG CANCER)	7
STAGE + IV + (LUNG CANCER)	7
STAGE IIIA + (LUNG CANCER)	6
STAGE IIIB + (LUNG CANCER)	5
N	5
T	5

STAGE IIB + (LUNG CANCER)	3
STAGE III + (LUNG CANCER)	3
TIS	3
STAGE IA + (LUNG CANCER)	2
STAGE IB + (LUNG CANCER)	2
STAGE IIA + (LUNG CANCER)	2
M	1
TOTAL	123

7.1. Non-small cell lung cancer

ADVANCED (+ LUNG CANCER) ~ non-small cell lung cancer 3	6
OCCULT (+ LUNG CANCER) ~ carcinoma ~ lung cancer ~ non-small cell lung cancer	3

RECURRENT (~ non-small cell lung cancer) (See 9.)	2
TOTAL	11

8. SIZE / QUANTITY (H3)

8.1. Size

SIZE (+TUMOR)	8
CENTIMETERS Centimeters 4, centimeter 1	5
VOLUME (+TUMOR) Target ~ 2	4
DIMENSION	2
TOTAL	19

8.2. Name of lung cancer

NON-SMALL ~ cell lung cancer 21 ~ cell lung cancers 1 ~ cell carcinoma 1	23
GIANT (~ cell)	1
LARGE (~ cell carcinoma)	1
SMALL	1
TOTAL	26

8.3. Minuteness / small quantity

<SMALL> Small 9, smaller 2	11
LIMITED (See 0.5.3.)	6
MODEST (+ SURVIVAL) Modest 3 Modestly increased disease-free survival 2	5

REDUCED	3
SHORT Short 1, shorter 1	2
MICROSCOPIC (See 6.7.)	1
TOTAL	28

8.4. Largeness / large quantity

HIGHER Highly 5 (See 13.) Higher 5, high 4	14
GREATER Greater 4, greatest 2, great 1	7
EXTEND (See 9.) Extend 2, extends 1, extension 1, extensive 1	5

<LARGE> Large 3, largest 2	5
TOTAL	31

8.5. Increase

ADDITION (See 10.) Addition 3, additional 3	6
INCREASED	3
GROWTH (See 1.2.)	1
TOTAL	10

9. METASTASIS (H5, H9, H2b, H10)

METASTASIS Metastasis 15 Distant ~ 4 + IPSILATERAL + 5 Metastases 10 Distant ~ 6 Brain ~ 2 Metastatic 3 Metastasizing 1	29
SECOND ~ cancers 3 ~ primary cancers 3 ~ lung cancers 2 ~ primary tumor(s) 2	13

INVADES Invades 3, invasive 3, invasion 2, invade 1, noninvasive 1	10
INVOLVE (✖ 2b, 10) Involvement 6, involve 1, involves 1,	8
RECURRENCE Recurrence 5, recurrences 2, Recurrent 2, recur 1 (See 7.1.)	10
EXTEND (See 8.4.) Extend 2, extends 1, extension 1 Extensive 1	5
EXTENT	5

DEVELOP (See 1.1.) ~ regional or distant metastases 2 ~ recurrence 1	3
SOLITARY Solitary + PART OF THE BODY +	3

METASTASIS	
TOTAL	86

9.1. Distance (H9)

DISTANT → (See 4.1.) Distant (+ METASTASIS) 10 Distal 2	12
---	----

10. COMPOSITION (H2b, H9, H10, H12a)

GROUP (See 2.1., 5.) Group 10, groups 2, grouping 1	13
COMPLETE (See 0.5.3.1.) Complete 6, completely 2, completed 1	9
SINGLE (✖ 9) (~-agent 2)	7
ADDITION (See 8.5.) Addition 3, additional 3	6
INCLUDES (✖ 2b) Includes 3, include 2	5
CONSIST-OF	2
ELEMENT	2

METASTASIS	
MINORITY	2
COMMITTEE (✖ 10)	1
COMPONENT	1
INDEPENDENT (✖ 10)	1
JOINT (✖ 10)	1
MAJORITY	1
PARTIAL	1
SEGMENTAL (See 0.5.3.1.)	1
TOTAL	53

11. GENERAL MEDICINE (H0)

CLINICAL ~ trial(s) 22 (See 0.1.) ~ staging, ~ stage 3 ~ studies 3 ~ study 1 (See 2.1.) ~ evaluation 2	35
MEDICAL Medical 5 ~ contraindications 3 ~ condition 2 Medically 2	7
TOTAL	42

12. TIME (H2b, TIME, DATE, AGE: 1,95 %)

YEAR Year 6, years 5	11
-------------------------	----

METASTASIS	
POSTOPERATIVE (See 0.2.)	8

DAILY ~ doses 2 ~ fractions 2	6
<u>EARLY</u> (+ TUMOR) (See 1.2.1.) ~ stage I tumors 1 ~ -stage lung cancer 1 ~ central tumors ~ NSCLC	4
RETROSPECTIVE (✖ 2b)	4
TIME	3
CURRENT Current 2, currently 1	3

<u>INITIAL</u> (See 0.5.2.) Initial 2, initiating 1	3
DURATION Duration 1, durations 1	2
DATE	1
FINAL	1
PERIOD	1
PRIOR TO	1
TOTAL	40

13. IMPORTANCE (H2b, IMPORTANCE 0,36%)

SIGNIFICANT Significant 6, significantly 3	9
CRITICAL	8
<u>HIGHLY</u> (See 8.4.)	5
MAIN (~ bronchus)	4
NEW New 1, newer 1, newly 1	3

CENTRAL	1
IMPORTANT	1
MAJOR	1
PRIMARILY	1
PRINCIPAL	1
TOTAL	34

14. OPTIONS (CHOICE 0,29%)

<u>MODALITIES</u> (See 0.1.) Modalities 8 Modality 6 Combined ~ + THERAPY 5	14
<u>OPTIONS</u> (See 0.1.) Options 10, option 2 Treatment ~ 11	12

<u>SELECTED</u> (+ PATIENT) (See 5.1.)	7
CHOICE	1
TOTAL	34

15. PRECISION / SURVEILLANCE (H8)

<u>CONSIDERED</u> (See 2.2., 6.3.)	17
CAREFUL Careful 5 ~ preoperative treatment 2 ~ treatment planning 2 Carefully 4 ~ staged 3	9
<u>ACCURACY</u> (See 6.4.)	3

Accuracy 1, accurate 1, accurately 1	
<u>SPECIFICITY</u> (✖ 0) (See 6.4.)	3
PRECISE	2
TOTAL	34

16. NEED (H5, NEED 0,26)

NEEDED	4
REQUIRES (✖ 5) Requires 3, require 1	4
TOTAL	8

17. IMPERFECTION (H9, H12b)

POOR (See 3.4.)	2
ADVERSE (+ PROGNOSIS+) Adverse 1, adversely 1	2
CONFUSED (✖ 9)	1

ERRORS	1
NEGATIVE	1
	7

18. VARIATION

VARIATION Variation 1, variety 1, vary 1	3
CHANGE	1
TOTAL	4

19. RISK FACTOR (H1)

SMOKE Smoke 1, smokers 2	3
TOTAL	3

11.1.2.5. CUADRO CONTRASTIVO (HESPERUS – ANÁLISIS COMBINADO)

En este texto de 5523 palabras, ha sido posible vincular 2574 palabras plenas, con lo que podríamos sostener que el 46,6 % de las palabras del texto contribuyen a la cohesión léxica. Teniendo en cuenta los datos de *Hesperus* este porcentaje sería de 44,7 %.

Hay un promedio de 21,3 palabras por oración, 10 de las cuales están vinculadas. Las 2574 palabras vinculadas activan 196 categorías conceptuales del tesoro *Roget*. Estas unidades léxicas se materializan mediante 620 formas léxicas y 417 lemas. Por tanto, de cada 100 palabra vinculadas, encontraríamos 24 variantes léxicas asociadas a 16 lemas, las cuales activarían 8 categorías conceptuales del tesoro.

TABLA 78: Descripción cuantitativa de las cadenas léxicas reconocidas.

CADENAS LÉXICAS RECONOCIDAS POR HESPERUS		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 0: TREATMENT AND DIAGNOSIS [369-39-32,49] ¹⁴⁶			0. TREATMENT [706-125-25,02] 6. DIAGNOSIS [165-83 -5,85] 11. GENERAL MEDICINE [42-3 -1,49]
Lexemas vinculados	369	913	Lexemas vinculados
Formas distintas	39	211	Formas distintas
Activación conceptual ajustada (%)	34,62	32,49	Activación conceptual (%)
H 1: PATHOLOGY OF NON-SMALL CELL LUNG CANCER [144-18-8,17]			1. PATHOLOGY OF NON-SMALL CELL LUNG CANCER
Lexemas vinculados	144	343	Lexemas vinculados
Formas distintas	18	57	Formas distintas
Activación conceptual ajustada (%)	8,7	12,15	Activación conceptual (%)
H 6: RESEARCH [34-2-2,94] H 2b: ? (CAUSE AND EFFECT, EVIDENCE) [160-70-8,13] CATEGORIZATION [1,56 %]			2. RESEARCH, EXPERIMENTATION
Lexemas vinculados	194	306	Lexemas vinculados
Formas distintas	72	94	Formas distintas
Activación conceptual ajustada (%)	13,46	10,84	Activación conceptual (%)
H 13: EXISTENCE [75-11-2,48] H 2 b: ? CAUSE AND EFFECT [1,88]			3. TREATMENT OUTCOME
Lexemas vinculados	75	287	Lexemas vinculados
Formas distintas	11	57	Formas distintas
Activación conceptual ajustada (%)	4,64	10,17	Activación conceptual (%)
H 4: HARD BODY PARTS [18-7-1,69] H 7: BODY FLUIDS [26-13-1,06] H 11: PLEURA [9-3-0,76]			4. LOCATION IN THE HUMAN BODY
Lexemas vinculados	53	238	Lexemas vinculados
Formas distintas	23	86	Formas distintas
Activación conceptual ajustada (%)	3,74	8,43	Activación conceptual (%)

¹⁴⁶ La leyenda de los números entre corchetes es la siguiente: [Lexemas vinculados – formas distintas – activación conceptual].

H 1: PATHOLOGY OF NON-SMALL CELL LUNG CANCER (PATIENT) [130-3-7,02]			5. PATIENT
Lexemas vinculados	130	172	Lexemas vinculados
Formas distintas	3	20	Formas distintas
Activación conceptual ajustada (%)	7,48	6,09	Activación conceptual (%)
H 2 a: STAGING [71-4-6,22]			7. STAGING
Lexemas vinculados	71	134	Lexemas vinculados
Formas distintas	4	22	Formas distintas
Activación conceptual ajustada (%)	6,63	4,75	Activación conceptual (%)
H 3: MINUTENESS, QUANTITY, SIZE [108-24-6,52]			8. SIZE / QUANTITY
Lexemas vinculados	108	114	Lexemas vinculados
Formas distintas	24	32	Formas distintas
Activación conceptual ajustada (%)	6,94	4,04	Activación conceptual (%)
H 5 – H 9: TRANSFERENCE - DISTANCE [28-7- 2,34]			9. METASTASIS
Lexemas vinculados	28	98	Lexemas vinculados
Formas distintas	7	26	Formas distintas
Activación conceptual ajustada (%)	2,49	3,47	Activación conceptual (%)
H 10 - H12 a: PART / WHOLE, COMPOSITION [17-9-4,62]			10. COMPOSITION
Lexemas vinculados	17	53	Lexemas vinculados
Formas distintas	9	22	Formas distintas
Activación conceptual ajustada (%)	4,92	1,88	Activación conceptual (%)
TIME [1,95 %]			12. TIME [40-17-1,42]
Activación conceptual ajustada (%)	2,08	1,39	Activación conceptual (%)
H 2b: ? (IMPORTANCE) [33-3-1,67]			13. IMPORTANCE
IMPORTANCE [0,36 %]			
Lexemas vinculados	33	34	Lexemas vinculados
Formas distintas	3	13	Formas distintas
Activación conceptual ajustada (%)	2,16	1,2	Activación conceptual (%)
CHOICE [0,29%]			14. OPTIONS [34-6-1,2]
Activación conceptual ajustada (%)	0,31	1,2	Activación conceptual (%)
H 8: PRECISION AND SURVEILLANCE [19-5-1,22]			15. PRECISION / SURVEILLANCE
Lexemas vinculados	19	34	Lexemas vinculados
Formas distintas	5	8	Formas distintas
Activación conceptual ajustada (%)	1,3	1,2	Activación conceptual (%)
NEED [0,26 %]			16. NEED [8-3-0,28]
Activación conceptual ajustada (%)	0,28	0,28	Activación conceptual (%)
Cadena 12b: IMPERFECTION [4-3-0,11]			17. IMPERFECTION
Lexemas vinculados	4	8	Lexemas vinculados
Formas distintas	3	6	Formas distintas
Activación conceptual ajustada (%)	0,11	0,28	Activación conceptual (%)
			18. VARIATION [4-4 -0,14]
	0	0,14	Activación conceptual (%)

H 1: PATHOLOGY OF NSCLC (SMOKING) [2-2-0,11]	Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	2 2 0,11	3 2 0,11	19. RISK FACTOR Lexemas vinculados Formas distintas Activación conceptual (%)
TOTAL				TOTAL
Activación conceptual según Hesperus	93,85		2882	Suma de elementos en las cadenas
Lexemas vinculados	1247		2574	Lexemas vinculados ¹⁴⁷
Formas distintas	223		620	Formas distintas ¹⁴⁸
Activación conceptual ajustada (%)	99,97	99,98		Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QDT2				
Nº total de palabras: 5523		Categorías conceptuales del tesoro: 196		
Palabras que crean cohesión (<i>Hesperus</i>): 2470		Coeficiente de cohesión (<i>Hesperus</i>): 8,37		
		Nº de oraciones: 251		

Al trasladar estos resultados a un diagrama de sectores donde el anillo exterior representa el resultado de nuestro análisis y el interior, los datos aportados por *Hesperus*, encontramos bastantes coincidencias en el porcentaje otorgado a la mayoría de los macroconceptos. Lo significativo es que haya coincidencia también entre macroconceptos relacionados con la medicina y las cadenas identificadas:

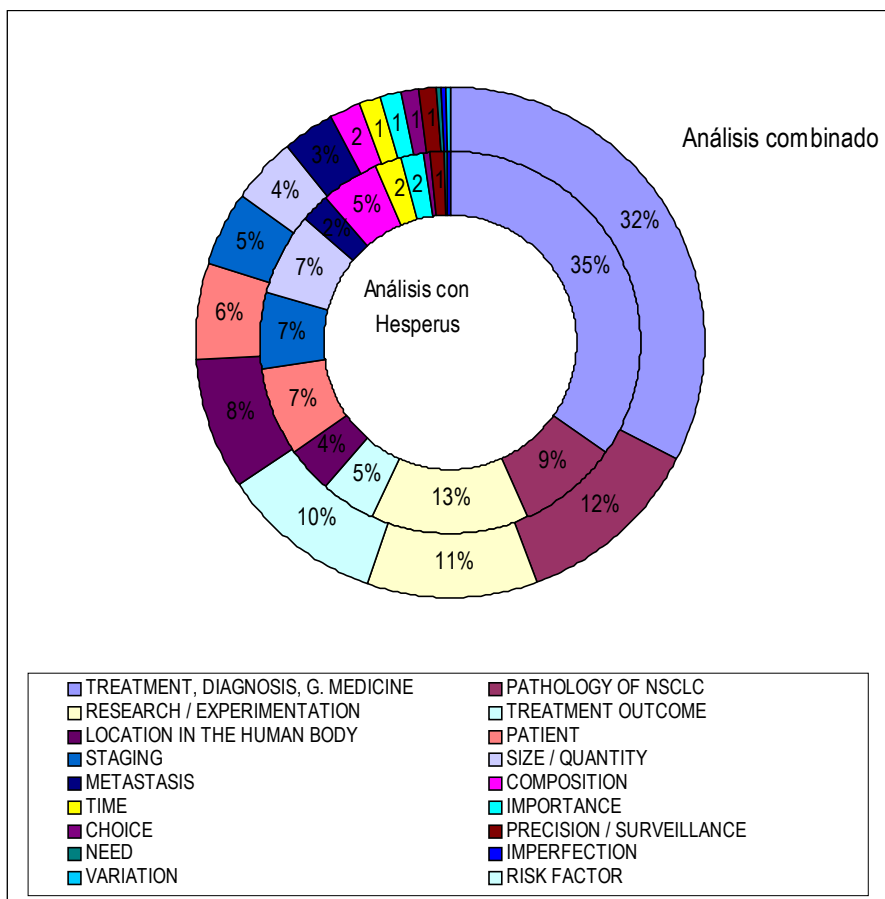
- *TREATMENT, DIAGNOSIS y GENERAL MEDICINE*
- *PATHOLOGY OF NON-SMALL CELL LUNG CANCER*
- *RESEARCH / EXPERIMENTATION*
- *PATIENT*
- *STAGING*

Estas coincidencias se deben en parte al hecho de que hemos agrupado cadenas léxicas separadas, como *TREATMENT* y *DIAGNOSIS*, para acomodarlas a las cadenas producidas por *Hesperus*.

¹⁴⁷ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹⁴⁸ De estas, hay 62 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 248 lexemas adicionales a los 2574 lexemas vinculados.

DIAGRAMA 24: Comparación de las áreas conceptuales reconocidas por *Hesperus* y las reconocidas por un humano asistido por medios informáticos (QDT2).



11.1.3. Small cell lung cancer

11.1.3.1. ÍNDICE DEL TEXTO

Al igual que en los textos anteriores, en el índice se destacan los conceptos CELLULAR CLASSIFICATION, STAGE, TREATMENT OPTIONS y LUNG CANCER. Los estadios del cáncer microcítico de pulmón adquieren la nomenclatura de *recurrent*, *limited* y *extensive*, estos dos últimos relacionados con los conceptos LOCALIZATION y INCREASE. El lexema *small* como elemento de la unidad fraseológica *small cell lung cancer* también se destaca en el índice.

(89)

GENERAL INFORMATION
CELLULAR CLASSIFICATION
 STAGE INFORMATION
 Limited stage
 Extensive stage
TREATMENT OPTION OVERVIEW
LIMITED STAGE SMALL CELL LUNG CANCER
EXTENSIVE STAGE SMALL CELL LUNG CANCER
RECURRENT SMALL CELL LUNG CANCER

11.1.3.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTA DE FRECUENCIA LEMATIZADA

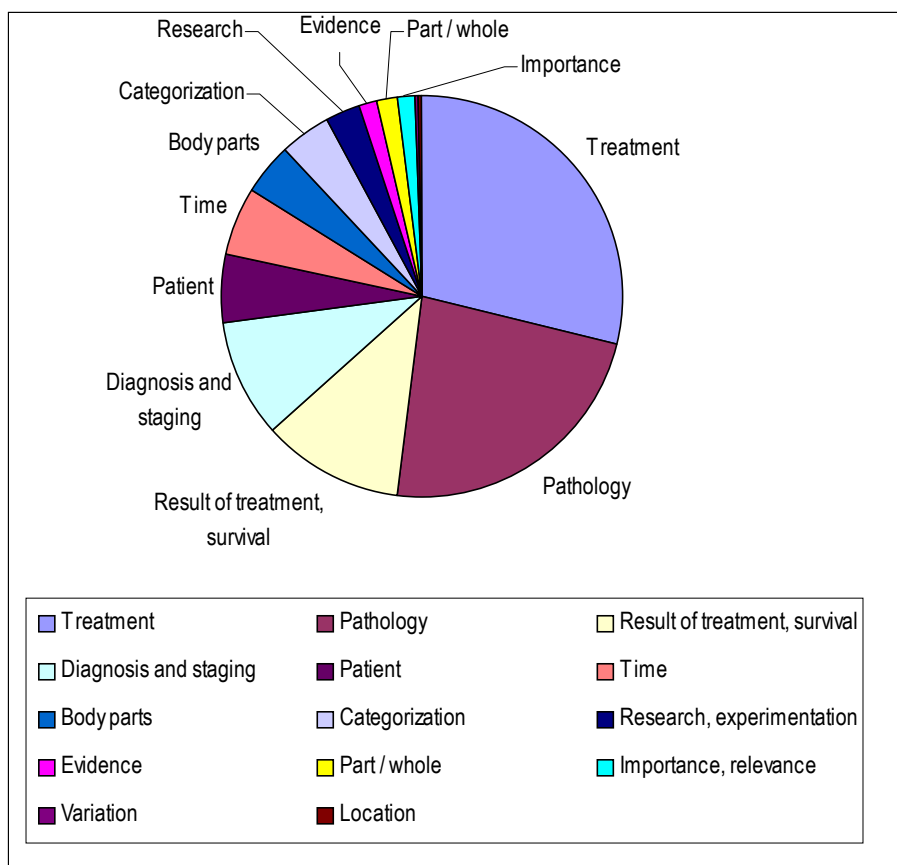
A partir de los lemas más significativos de la lista de frecuencia lematizada elaborada con *Wordsmith Tools*, se pueden anticipar las áreas conceptuales más activadas.

TABLA 79: Identificación preliminar de las áreas conceptuales activadas en QDT3 a partir de lista de frecuencia lematizada

DT3	
ÁREAS CONCEPTUALES	%
Treatment	12,1
Pathology	9,56
Result of treatment, survival	4,76
Diagnosis and staging	3,96
Patients	2,29
Time	2,28
Parts of the body	1,82
Categorization	1,66

Research, experimentation	1,18
Evidence	0,65
Part /whole	0,65
Importance	0,53
Variation	0,15
Location	0,12
TOTAL	41,71

DIAGRAMA 25: Representación preliminar de las áreas conceptuales activadas en QDT3



11.1.3.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

Hesperus extrae sólo ocho cadenas léxicas (véase Apéndice VIIe) en el texto sobre el cáncer microcítico de pulmón, con lo que se convierte en el texto para especialistas con menos cadenas. El *perfil* del documento nos indica que de 3408 palabras, 1648 son plenas y se vinculan mediante 184 categorías del tesauro. El *coeficiente de cohesión*, que lo

obtenemos de dividir el valor otorgado por *Hesperus* (30223) entre el número total de palabras es de 8,87.

Especificamos a continuación las categorías conceptuales activadas que aportan más de un 1% de la cohesión del texto.

```
therapy_658_4268_n Percent: 14.78,
small_33_224_a Percent: 8.45,
minuteness_196_1271_n Percent: 8.05,
sick-person_651_4202_n Percent: 7.29,
dramatize_594_3818_v Percent: 5.01,
life_360_2362_n Percent: 2.15,
disease_651_4187_n Percent: 2.08,
cancer_651_4195_n Percent: 2.04,
respiratory-disease_651_4192_n Percent: 2.03,
anaesthetic_375_2486_n Percent: 1.99,
drug_658_4263_n Percent: 1.88,
tobacco_388_2560_n Percent: 1.66,
insides_224_1464_n Percent: 1.65,
arrangement_62_427_n Percent: 1.52,
extensive_32_203_a Percent: 1.40,
medical_658_4274_a Percent: 1.38,
medicine_658_4257_n Percent: 1.20,
often_139_875_r Percent: 1.18,
study_536_3481_n Percent: 1.03,
important_638_4094_a Percent: 1.02,
advantage_34_233_n Percent: 1.00,
party_708_4561_n Percent: 0.99,
sequence_65_444_n Percent: 0.99,
remedy_658_4275_v Percent: 0.99.
```

La mayoría de estos conceptos subyacen las ocho cadenas reconocidas por *Hesperus*. Después de estudiar estos conceptos a la luz de las unidades léxicas que se actualizan en el texto, y, por tanto, después de aclarar el sentido de estos en el texto, hemos otorgado una etiqueta orientativa del hilo conductor de cada cadena.

TABLA 80: Cadena 0, 2, 3 a (QDT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 0 (<i>Treatment</i>)		Cadena 2 (<i>Lung cancer</i>)		Cadena 3a (<i>Stage</i>)	
THERAPY	14,78	SICK-PERSON	7,29	DRAMATIZE	5,01
DRUG	1,88	DISEASE	2,08	(<i>stage, produce</i>)	
MEDICAL	1,38	CANCER	2,04	DRAMATURGY	0,66
MEDICINE	1,20	RESPIRATORY-DISEASE	2,03		
REMEDY	0,99	PATHOLOGY	0,83		
DIAGNOSTIC	0,58	SICK	0,46		
INSTRUMENTALITY	0,33	ILL-HEALTH	0,41		
REMEDIAL	0,32	INSTRUMENTALITY	0,33		
SURGERY	0,31	ULCER	0,31		
INSTRUMENT	0,25	DISEASED	0,25		
DOCTOR	0,10	INSTRUMENT	0,25		
MEDICAL-ART	0,01	ILLNESS	0,18		
PROPHYLACTIC	0,01	INFECTION	0,01		
TOTAL	22,12+	TOTAL	16,46+	TOTAL	5,67
↓		↓		↓	
TREATMENT AND DIAGNOSIS		PATHOLOGY OF LUNG CANCER		STAGING	

Comentario

En la cadena 0 destaca la idea de TRATAMIENTO, presente tanto en lexemas genéricos (*treatment, treated, cure*) como en especializados (*radiotherapy, chemotherapy*). En el tratamiento participan personas (*pathologist, physicians*), agentes curativos (*drugs, regimens*) e instrumentos (*bronchoscope*). Por otra parte, también aparece el concepto DIAGNÓSTICO. Esta cadena es la más extensa de las reconocidas por *Hesperus* y cuenta con 225 palabras. Esto queda corroborado por el porcentaje que se le atribuye en relación a la cohesión, que es superior al 22 %.

En la cadena 2, el énfasis se pone en el paciente y la enfermedad en sí, que queda asociada a enfermedades del sistema respiratorio. A palabras genéricas como *disease, cancer, lesions* o *cases*, se añaden otras que constituyen UFs, como veremos en nuestra propuesta definitiva de cadenas léxicas. El concepto DIAGNOSIS también está presente en esta cadena, lo cual nos indica un solapamiento entre el *dominio* TRATAMIENTO y ENFERMEDAD.

Al igual que en los otros textos, hemos vinculado la *cadena 3a* al concepto STAGING y no al de DRAMATURGY, que nada tiene que ver con el cáncer de pulmón. La

importancia del estadiaje en la cura de esta enfermedad y en el texto, justifica la creación de una cadena aparte para este concepto.

TABLA 81: **Cadena 1** (QDT3)

CONCEPTOS ACTIVADOS	
Cadena 1 (Small)	
SMALL	8,45
MINUTENESS	8,05
WEAK	0,93
DEGREE	0,31
INCONSIDERABLE	0,19
MICROSCOPY	0,18
SMALL-THING	0,11
INSUFFICIENT	0 0,1
WEAKNESS	0 0,1
SLIGHTLY	0 0,1
SMALL-QUANTITY	0 0,1
TOTAL	18,22 +
↓	
MINUTENESS, SIZE	

Comentario

En esta cadena el lexema *small* es excesivamente frecuente y siempre forma parte del término que da nombre a la enfermedad, al igual que *cell*. También se agrupan una serie de palabras que adquieren un sentido en el contexto que no se acomoda demasiado al concepto SMALL. Por ejemplo, se incluye el lexema *type*, que en el texto designa la variantes patológicas del cáncer, y se asocia al verbo *type* (*teclear*), y de ahí al concepto LETTER, con lo que se consigue una dispersión de significados en la cadena. Otros lexemas que también se introducen de un modo erróneo son *median*, *impaired*, *fold*, *detail*, *approximately*, *requires*, *mixed*, y *suggestion*.

Por otra parte, los hay relacionados con la idea de pequeñez pero que, por su especificidad, es conveniente ubicarlos en otras cadenas cuya etiqueta incluimos en paréntesis:

Poor, failing (RESULT OF TREATMENT)

Limited (STAGE)

Cell (PART OF THE BODY)

Decline (DIAGNOSIS, SYMPTOMS)

Electron microscopy (DIAGNOSTIC TESTS)

Compression (SYMPTOMS)

Fractionation (RADIOTHERAPY)

(Neuroendocrine) granules (DIAGNOSIS)

Obstructing (SYMPTOMS)

Por último, en la cadena aparecen algunos lexemas que indican tamaño mediano: *median, moderate*.

TABLA 82: Cadena 3b: Difusa (QDT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 3b1 (<i>Shown</i>)		Cadena 3b2 (<i>Additional</i>)		Cadena 3b3 (<i>Include</i>)	
DRAMATIZE (<i>Stage, produce, shown</i>)	5,01	INCREASING	0,57	COMPOSITION (<i>incorporation</i>)	0,37
TRUE	0,84	INCREASE	0,37	JOIN	0,35
DRAMATURGY	0,66	GROW	0,27	PORTION	0,25
EVIDENCE	0,55	AUGMENT (<i>prolongs</i>)	0,1	CONSUMMATE (<i>total, complete</i>)	0,24
MANIFEST	0,33			CONTAIN	0,18
CLEAR	0,1			COMPOSING	0,1
TOTAL	7,39+	TOTAL	1,21+	TOTAL	1,39+
↓		↓		↓	
EVIDENCE		INCREASE (Spread of disease, improvement, survival)		PART-WHOLE	
CONCEPTOS ACTIVADOS					
Cadena 3b4 (<i>Produce</i>)		Cadena 3b5 (<i>Review</i>)		Cadena 3b6 (<i>Variation</i>)	
CAUSE	0,5	ARRANGEMENT	1,52	REVERSION	0,36
DOER	0,5	STUDY	1,03	UNSTABLE	0,17
CAUSAL	0,33	MEASUREMENT	0,54		
PRODUCER	0,2	CLASSIFICATION	0,33		
		DEGREE	0,31		
TOTAL	1,53	TOTAL	3,73	TOTAL	0,53
↓		↓		↓	
CAUSE		RESEARCH, CATEGORIZATION		CHANGE	

Comentario

Como vemos en la tabla, los términos de la cadena que hemos denominado 3b, es decir, aquellos no asociados a STAGING, quedan vinculados a través de seis conceptos que formarán cadenas separadas en nuestra propuesta posterior. Estos términos son palabras de la lengua general frecuentes en el lenguaje de la ciencia. Mediante este lenguaje se describen *cambios* en los que participan elementos que forman *parte* de un sistema que se puede *conceptualizar*, se explican las *causas* que originan estos cambios y se expresa la *evidencia* que confirma esta causalidad.

Quedan al margen de esta cadena *live* y *period*, que irán en las cadenas *SURVIVAL* y *TIME* respectivamente; y *combine* y *cycle*, que ubicaremos en la cadena *TREATMENT*.

TABLA 83: Cadena 4 (QDT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 4 (<i>Survival</i>)	
LIFE	2,15
PERIOD	0,89
AGE	0,83
EXISTENCE	0,32
LONG-DURATION	0,22
BE (<i>occur, continue</i>)	0,17
LASTING	0,15
BEGINNING	0,12
BE-LEFT / LAST (<i>remain</i>)	0 0,1
SURVIVOR	0 0,1
TOTAL	4,85 +
↓	
SURVIVAL, TIME, EXISTENCE	

Comentario

La cadena 4 pone en contacto conceptos muy relacionados en el curso de la enfermedad: la vida y la muerte, el tiempo y la lucha por continuar existiendo. Por otra parte, se han reubicado una serie de lexemas en las cadenas que indicamos entre paréntesis:

Series (PATIENTS)

Origin (METASTASIS)

Obtain (POSITIVE RESULT)

Presence (EVIDENCE, SENSORY PROCESSES INVOLVED IN DIAGNOSIS)

TABLA 84: Cadena 5b y 6 (QDT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN			
Cadena 5b (<i>Central</i>)		Cadena 6 (<i>Extensive, high</i>)	
IMPORTANT	1,02	EXTENSIVE	1,4
RELEVANT	0,12	GREAT	0,23
IMPORTANTLY	0 0,1	GREATLY	0,22
CHIEF-THING	0 0,1	EXORBITANT	0,14
		ABSOLUTE	0 0,1
		EXTREMELY	0 0,1
		GREATNESS	0 0,1
TOTAL	1,14 +	TOTAL	1,99 +
↓		↓	
IMPORTANCE		LARGENESS	

Comentario

Las cadenas *IMPORTANCE / RELEVANCE* y *LARGENESS* tienen una representación parecida en el texto y están conceptualmente relacionadas teniendo en cuenta que en muchas ocasiones, lo importante está relacionado con lo que es grande o extenso. Hemos fragmentado la cadena 5 en dos porque cuenta con lexemas que designan partes del cuerpo que ocupan una posición central, de ahí que *Hesperus* los haya incluido en la cadena *IMPORTANCE*. Entre los lexemas que no encajan, destacamos *prolonge*, que pasaremos a la cadena 4 y *rare/rarely* y *unusual*, que se asocian a la idea de GRANDEZA desde el punto de vista de que lo inusual suele ser destacado y valioso. De hecho, *Hesperus* los asocia a los conceptos GREATLY y REMARKABLE.

TABLA 85: Cadena 5a (QDT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN	
Cadena 5a (<i>Liver</i>)	
INSIDES	1,65
INSIDES (<i>chest / abdomen</i>)	0 0,1
INTERIORITY (<i>marrow</i>)	0 0,1
INTERIOR (<i>intravenous</i>)	0 0,1
TOTAL	1,65 +
↓	
BODY PARTS	

Comentario

Aunque en este texto las referencias a partes del cuerpo son bastante frecuentes, en parte porque la nomenclatura de la enfermedad incluye la palabra *lung*, será preciso completar la *cadena 5a* con el fin de especificar la ubicación de posibles metástasis. En consecuencia, esta cadena está relacionada conceptualmente con la que denominaremos *METASTASIS*.

TABLA 86: **Cadena 7** (QDT3)

CONCEPTOS ACTIVADOS	
Cadena 7 (<i>Cigarette</i>)	
TOBACCO	1,66
TOTAL	1,66
↓	
TOBACCO	

Comentario

En esta cadena queda representado el principal factor de riesgo de la enfermedad.

11.1.3.1.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

En los apartados anteriores han surgido áreas conceptuales que es preciso estructurar. A las mismas podemos añadir las que han aparecido en otros textos para especialistas y las que aparecen en el *perfil del documento*. Entre estas, hemos encontrado el concepto ESTIMATE (0,54%), que constituirá la cadena *COGNITIVE PROCESSES INVOLVED IN DIAGNOSIS*; LIMIT (0,59%) y SPACE (0,21%), que formarán la cadena *LOCATION* junto con los lexemas que designan partes del cuerpo; y MODALITY (0,66%), al que asociaremos la cadena *OPTIONS*. Los conceptos SEQUENCE (0,99 %), PERIOD (0,89%), AGE (0,83%) y BEGINNING (0,12 %) ponen en relación una serie de lexemas que formarán la cadena *TIME*. Otros conceptos que avalan la inclusión de otras cadenas son los de INSTRUMENTALITY (0,33%), INSTRUMENT (0,25%) y MEASUREMENT (0,54 %).

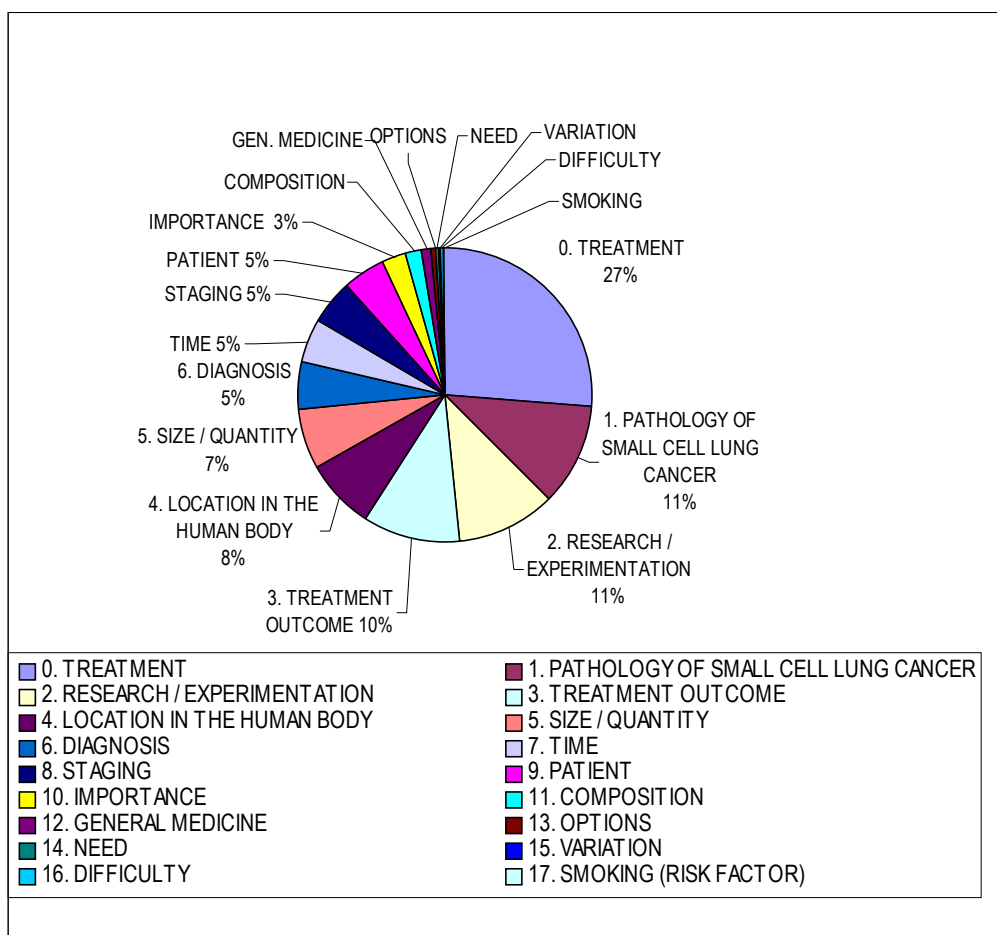
En la tabla siguiente adelantamos las cadenas que hemos propuesto y que desglosaremos posteriormente.

TABLA 87: Cadenas léxicas en QDT3

<p>Chain 0: TREATMENT (H0, H3b3) [478]</p> <p>0.1. Generic treatment 0.2. Properties (→ TREATMENT) 0.2.1. Location of application 0.3. Manipulation 0.4. Process 0.5. Types 0.5.1. Chemotherapy 0.5.1.1. Drugs 0.5.1.2. Regimens 0.5.2. Radiotherapy 0.5.3. Surgery 0.5.4. Other types of treatment</p> <p>Chain 1: PATHOLOGY OF SMALL CELL LUNG CANCER (H1, H2, H3b, H6) [200]</p> <p>1.1. Generic disease 1.2. Cancer 1.2.1. Types 1.2.1.1. Anatomic / histologic classification 1.3. Metastasis</p> <p>Chain 2: RESEARCH, EXPERIMENTATION (H3b1, H3b4, H3b5) [192]</p> <p>2.1 Study 2.2. Evidence 2.3. Categorization 2.3.1. Classification 2.3.2. Definition 2.4. Comparison 2.5. Evaluation 2.6. Cause 2.7. Measurement 2.8. Information</p> <p>Chain 3: OUTCOME OF TREATMENT (H2, H3b2, H3b4, H4) [190]</p> <p>3.1. Generic outcome 3.1.1. Cause 3.2. Positive outcome 3.3. Existence 3.3.1. Survival 3.3.2. Death 3.4. Negative outcome</p> <p>Chain 4: LOCATION IN THE HUMAN BODY (H1, H5 a, LIMIT 0,59%, SPACE 0,21%) [142]</p>	<p>4.3. Anatomic level / body parts 4.3.1. Respiratory system 4.3.2. Other body parts / systems 4.4. Location in relation to specific body parts 4.4.1. Bronchi 4.4.2. Clavicle 4.4.3. Thorax 4.4.4. Cranium 4.4.5. Veins</p> <p>Chain 5: SIZE / QUANTITY (H1, H3b2, H6)[119]</p> <p>5.1. Name of lung cancer 5.2. Minuteness / small quantity 5.3. Medium size 5.4. Largeness / large quantity 5.5. Increase</p> <p>Chain 6: DIAGNOSIS (H0, H2) [95]</p> <p>6.1. Generic diagnosis 6.2. Sensory processes involved 6.3. Cognitive processes involved 6.4. Clinical features (Symptoms) 6.5. Diagnostic tests and instruments 6.6. Prognostic factors</p> <p>Chain 7: TIME (H4, SEQUENCE 0,99 %, PERIOD 0,89%, AGE 0,83%, BEGINNING 0,12 %) [88]</p> <p>Chain 8: STAGING (H3 a, H6) [87]</p> <p>8.1. Small cell lung cancer</p> <p>Chain 9: PATIENT (H2) [85]</p> <p>Chain 10: IMPORTANCE (H5b) [51]</p> <p>10.1.Importance 10.2. Centrality 10.3. Unusualness</p> <p>Chain 11: COMPOSITION (H3b3, H6) [32]</p> <p>Chain 12: GENERAL MEDICINE (H0) [19]</p> <p>12.1. Health professionals</p> <p>Chain 13: OPTIONS (MODALITY 0,66%) [11]</p> <p>Chain 14: NEED [6]</p> <p>Chain 15: VARIATION (H3b6)[4]</p> <p>Chain 16: DIFFICULTY [2]</p> <p>Chain 17: SMOKING (RISK FACTOR) (H7) [2]</p>
---	---

4.1. Generic location 4.1.1. Generic location (medicine) 4.2. Microscopic / histologic level	UNIDADES LÉXICAS ENCADENADAS: [1733]
--	---

DIAGRAMA 26: Porcentaje de activación de las cadenas léxicas definitivas de QDT3



CADENAS LÉXICAS	Frec	%	CADENAS LÉXICAS	Frec	%
0. TREATMENT	478	26,51	10. IMPORTANCE	51	2,83
1. PATHOLOGY OF SMALL CELL LUNG CANCER	200	11,1	11. COMPOSITION	32	1,77
2. RESEARCH, EXPERIMENTATION	192	10,65	12. GENERAL MEDICINE	19	1,05
3. OUTCOME OF TREATMENT	190	10,54	13. OPTIONS	11	0,61
4. LOCATION IN THE HUMAN BODY	142	7,87	14. NEED	6	0,33
5. SIZE / QUANTITY	119	6,6	15. VARIATION	4	0,22
6. DIAGNOSIS	95	5,27	16. DIFFICULTY	2	0,11
7. TIME	88	4,88	17. SMOKING	2	0,11
8. STAGING	87	4,82	TOTAL	1803	99,98
9. PATIENT	85	4,71			

De las 18 cadenas del texto, las cuatro que vuelven a despuntar con porcentajes conceptuales parecidos a los que adquieren en los textos anteriores son las de *TREATMENT* (26,51%), *PATHOLOGY* (11,1%), *RESEARCH / EXPERIMENTATION* (10,65%) y *OUTCOME OF TREATMENT* (10,54%). Se confirma de nuevo que la temática de un texto (TRATAMIENTO DEL CÁNCER DE PULMÓN) condiciona las cadenas léxicas que se establecen en el texto. De hecho, la cadena *TREATMENT* representa una cuarta parte de la activación conceptual de este texto y de los dos precedentes.

Le siguen en importancia otras cadenas que también se hallaban entre las primeras:

- *LOCATION IN THE HUMAN BODY*, que ocupa la quinta posición tanto en este texto como en el QDT2, y una sexta posición en el QDT1.
- *SIZE / QUANTITY*, cuya importancia se debe a que la nomenclatura de la enfermedad se construye sobre el lexema *small*.
- Con un porcentaje próximo al 5% cada una, encontramos cadenas con relevancia similar en los otros dos textos: *DIAGNOSIS*, *STAGING* y *PATIENT*. A la cadena *TIME* también le corresponde un 5%, aunque esta cadena es más relevante en QDT3 que en los otros textos.

El resto de las cadenas son comunes a los textos anteriores y sólo cambia levemente el porcentaje atribuido a las mismas: *IMPORTANCE*, *COMPOSITION*, *GENERAL MEDICINE*, *OPTIONS*, *NEED*, *VARIATION*, *DIFFICULTY*, *SMOKING (RISK FACTOR)*.

11.1.3.1.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS**0. TREATMENT (H0, H3b3)**

0.1. Generic treatment

TREATMENT Treatment 35, treated 18, treatment-related 2, Treatments 1, pretreatment 1, untreated 1	56
THERAPY - Therapy 14 Combined modality ~ 3 Palliative ~ 1 (See 0.5.2. RADIATION ~ ; 0.5.4. ENDOBRONCHIAL LASER ~) - Therapeutic 1	15
PALLIATION Palliation 5 ~ of symptoms 3 Palliative 4 ~ benefit 2 Palliated 1	10
TRIALS (See 2.1.) Clinical ~ 5	9

Randomized ~ 3	
SCHEDULES (+ RADIOTHERAPY/ CHEMOTHERAPY +)	6
CONTROL + CANCER + Control 2, controlled 2	4
MODALITY (See 13.) Combined ~ therapy 3 Combined ~ treatment 1	4
OPTIONS (Treatment ~) (See 13.)	4
STRATEGY Strategy 2, strategies 1	3
CURE (See 3.2.)	1
ANESTHESIA	1
TOTAL	113

0.2. Properties (→ TREATMENT)

START ~ of (+ THERAPY)	11
STANDARD	10
→ ALONE (Chemotherapy ~)	5
ADDITIONAL (See 5.5.)	4
INITIAL (See 7.)	4
ADJUVANT	3
AGGRESSIVE ~ clinical course 1	3

Randomized ~ 3	
CONCURRENT Concurrent 1, concurrently 1	2
PALLIATIVE	2
SALVAGE	2
TOTAL	44

0.2.1. Location of application (→ TREATMENT)

ENDOBRONCHIAL (See 4.4.1.)	3
--	---

~ laser therapy 2 ~ stents 1	
<u>INTRAVENOUS</u> (~ regimens) (See 5.4.5.)	1
<u>ORAL</u> (~ regimens)(See 4.3.1.)	1

0.3. Manipulation

<u>ADMINISTRATION</u> (+ CHEMOTHERAPY / PCI) Administration 10, administered 2	12
<u>USED</u> Used 8, use 2, using 2	12
<u>GIVEN</u>	5

<u>PROVIDE</u> Provide 1, provisions 1	2
TOTAL	31

0.4. Process

<u>DOSES</u> (+ CHEMOTHERAPY / RADIOTHERAPY) Doses 11, low-dose 2	13
<u>PROGRAMS</u> (+ CHEMOTHERAPY) Programs 3, program 1	4
<u>CYCLE</u> (✖ 3)	1

TOTAL	14
-------	----

0.5. Types

0.5.1. Chemotherapy (H3b3)

<u>CHEMOTHERAPY</u> Chemotherapy 51, chemotherapeutic 1	52
<u>REGIMENS</u> Regimens 23 Chemotherapy ~ 4 Drug ~ 4 Regimen 1	24
<u>COMBINATION</u> (✖ 3) - Combination 12 ~ chemotherapy 8 - Combinations 4	16
<u>DRUG</u> Drug 11, drugs 4	15
<u>ADMINISTRATION</u> Administration 10, administered 2	12
<u>AGENTS</u> - Agents 7 New ~ 3 Chemotherapeutic ~ 2	12

<u>DOSES</u> - Doses 8 Drug ~ 3 - Low-dose regimens 2	10
<u>PROGRAMS</u> (+ CHEMOTHERAPY) Programs 3, program 1	4
<u>ALTERNATING</u> → ~ different combinations of chemotherapy 1 ~ chemotherapy regimens 1 ~ sequences 1	3
<u>CYCLE</u> (✖ 3) (~ of chemotherapy administration)	1
TOTAL	131

0.5.1.1. Drugs

<u>DRUG</u>	15
Drug 11, drugs 4	
<u>AGENTS</u>	12
Agents 7, single-agent 4, agent 1	
ETOPOSIDE	11
CISPLATIN	6
CYCLOPHOSPHAMIDE	6
VINCRIStINE	5
DOXORUBICIN	3
CARBOPLATIN	2

LOMUSTINE	2
METHOTREXATE	2
TENIPOSIDE	2
TOPOTECAN	2
<u>ANESTHESIA</u>	1
IFOSFAMIDE	1
PACLITAXEL	1
TOTAL	44

0.5.1.2. Regimens

<u>REGIMENS</u>	24
Regimens 23, regimen 1	
EC (etoposide + carboplatin)	2
CAE (cyclophosphamide + doxorubicin + etoposide)	1
CAV (cyclophosphamide + doxorubicin + vincristine)	1
CEV (Cyclophosphamide + etoposide + vincristine)	1

ECV (Etoposide + cisplatin + vincristine)	1
EP (etoposide + cisplatin)	1
ICE (ifosfamide + carboplatin + etoposide)	1
TOTAL	8

0.5.2. Radiotherapy

<u>RADIOTHERAPY</u>	21
Chest ~ 8	
Whole-brain ~ 2	
<u>IRRADIATION</u>	15
Chest ~ 6, cranial ~ 6, whole brain ~ 2	
PCI 12 ≈	14
PROPHYLACTIC CRANIAL IRRADIATION 2	
cGy	5
<u>RADIATION</u>	5
Chest ~ 1	
~ therapy 1	

thoracic ~ 1	
<u>DOSES</u> (+ cGy)	3
Administration of PCI in ~ of NUMBER cGy.	
BRACHYTHERAPY	2
EXTERNAL-BEAM RADIOTHERAPY	1
FRACTIONATION (✖ 1)	1
TOTAL	64

0.5.3. Surgery

SURGICAL	10
Surgical 8 ~ resection 5 Nonsurgical 1, surgery 1	
RESECTION (~ resection)	5
EXCISION	2

TRANSPLANT(S) (Bone marrow ~)	2
ANESTHESIA	1
TOTAL	20

0.5.4. Other types of treatment

ENDOBRONCHIAL LASER THERAPY (See 0.2.)	2
STENTS (PALLIATION)	2
TOTAL	4

1. PATHOLOGY OF SMALL CELL LUNG CANCER (H1, H2, H3b, H6)

1.1. Generic disease

DISEASE	21
Disease 18, disease-free 3	
PRESENT (+ DISEASE) (See 2.2.)	6
Present 4, presence 1, presenting 1	
DEVELOPING (~ central nervous system metastases)	4
IMPAIRED	4
Impaired 2 Impairment 2 (central nervous system ~)	
PATHOLOGIC	4

Pathologic 2, pathologically 1, Pathologist 1 (See 12.1.)	
SYMPTOMS (See 6.4.)	4
AFFECTED (See 6.4.)	2
LESIONS (See 6.4.)	2
Endobronchial (obstructing) ~	
NEUROLOGIC ABNORMALITIES (~ present)	2
TOTAL	49

1.2. Cancer

<CANCER>	5
TUMOR	18
Tumor 16, tumors 2	
CARCINOMA	16
Carcinoma 13 Small cell ~ 11 Carcinomas 3	
CARCINOID	5
Carcinoid 4, carcinoids 1	

PRIMARY → (~ tumor)	2
MALIGNANT	1
NONBENIGN	1
NEOPLASTIC	1
ONCOLOGIST (See 12.1.)	1
TOTAL	50

1.2.1. Types

TYPE (✖ 1) (+ CANCER) (See 2.3.2.) Type 2, types 2	4
--	---

1.2.1.1. Anatomic / histologic classification

<LUNG CANCER ¹⁴⁷ >	5
BRONCHOGENIC CARCINOMAS	1
NEUROENDOCRINE CARCINOMAS OF THE LUNG	1
SMALL CELL LUNG CANCER (36) ≈ Small cell cancers (1)	37
SMALL CELL CARCINOMA (11) ≈ "Oat" cell subtype (1)	12
COMBINED SMALL CELL CARCINOMA	1
MIXED SMALL CELL/LARGE CELL CARCINOMA	1
WELL-DIFFERENTIATED NEUROENDOCRINE CARCINOMA OF THE LUNG (1) ≈ Malignant carcinoid (1) ≈ Metastasizing bronchial adenoma (1) (See 4.3.1.) ≈ Pleomorphic carcinoid (1) ≈ Nonbenign carcinoid tumor (1) ≈ Atypical carcinoid (1) (See 10.3.)	6
BRONCHIAL CARCINOIDS	1
NON-SMALL CELL LUNG CANCER (1) ≈ Non-small cell cancers (1)	2
TOTAL	67

1.3. Metastasis

METASTASES Metastases 13 (See 4.3.2.) Risk of developing central nervous system ~ 4 Metastatic 4, metastasizes 1, metastasizing 1	19
DISSEMINATED Disseminated 2, dissemination 1	3
ORIGIN (Confined to the PART OF THE BODY + of origin)	3
RECURRENT (See 8.1.)	3

- Recurrent ~ small cell lung cancer 2 - recurrences 1	
DISTANT (See 4.1.) ~ metastases	2
INVOLVEMENT SPREAD (✖ 3b) (See 5.5.)	2 1
TOTAL	33

¹⁴⁷ En realidad, encontramos la UF *lung cancer* en 42 ocasiones.

2. RESEARCH, EXPERIMENTATION (H3b1, H3b4, H3b5)

2.1 Study (H3b5)

STUDIES Studies 11, study 2, studied 1	14
TRIALS (See 0.1.) Clinical ~ 5 Randomized ~ 3	9
REFERENCES	7
PROSPECTIVE → (See 7.) - ~ randomized ~ randomized trials 2 ~ randomized studies 1 - ~ studies 2	5
RANDOMIZED - ~ trials 3 - ~ study(ies) 2 - prospective ~ 3	5

RETROSPECTIVE → (See 7.) ~ studies 2	3
PHASE ~ I 1 ~ II 1	2
REPORTED	2
REVIEW (✖ 3)	2
TESTING	2
META-ANALYSES	1
MODELS (✖ 3)	1
PROTOCOL	1
STATE-OF-THE-ART	1
TOTAL	55

2.2. Evidence (H3b1)

SHOWN Shown (STUDY +) 8 Showed 2, showing 1	11
PRESENT (See 1.1.) Present 4, presence 1, presenting 1	6
APPEAR Appear 3, appears 2	5

CLEARLY (✖ 6) Clearly 4, clear 1	5
EVIDENCE	3
DEMONSTRATED	1
TOTAL	31

2.3. Categorization

CONSISTENT Consistent 2, consistently 2	4
DEVELOPED	3
SYSTEM (+ STAGE) Stage ~ 2 Staging ~ 1	3
DESIGNATED Designated 1, designations 1	2

IDENTIFICATION Identification 1, identifies 1	2
SITUATION Situation 1, situations 1	2
TOTAL	16

2.3.1. Classification

TYPE (✖ 1)(+ CANCER) (See 1.2.1.) Type 2, types 2	4
CLASSIFICATION Classification 2, classifying 1	3
SUBTYPE Subtype 2, subtypes 1	3
CASES	1
TOTAL	11

2.3.2. Definition

DEFINE Define 2, definition 2	4
MEANS	3
DETAIL (✖ 1)	1
ROLE	1
TOTAL	9

2.4. Comparison

COMPARED Compared 7, comparable 1	8
DIFFERENT Different 5, differences 1, differential 1, differentiate 1	8
SIMILAR	5
TOTAL	21

2.5. Evaluation (ESTIMATE 0,54%)

EVALUATION (See 6.3.) Evaluation 9, evaluated 1	10
CONSIDERED (See 6.3.)	4
SUGGEST (✖ 1) Suggest 2, suggestion 1	3
FACTORS (See 6.6.)	1
TOTAL	18

2.6. Cause (H3b4)

PRODUCE (✖ 3b4) (See 3.1.1.) Produce 6, produces 2	8
YIELD (✖ 3b4) (See 3.1.1.) Yield 2, yields 1	3
DERIVE (See 3.1.1.)	1

LED (See 3.1.1.)	1
MADE (See 3.1.1.)	1
TOTAL	14

2.7. Measurement (MEASUREMENT 0,54%)

RATES (See 3.1.) - Rates Response ~ 5 Survival ~ 1 - Survival rate 1	7
MEDIAN (See 3.3.1.)	6
TOTAL	13

2.8. Information

INFORMATION	2
KNOWN	2
TOTAL	4

3. OUTCOME OF TREATMENT (H2, H3b2, H3b4, H4)

3.1. Generic outcome

RESPONSE Response 7 (~ rates 5), Responses 6, respond 4, responsive 2, responded 1	20
RATES (See 2.5.) - Rates Response ~ 5 Survival ~ 1 - Survival rate 1	7

RESULTS Results 4, resulting 1	5
OUTCOMES Outcomes 3, outcome 1	4
TOTAL	36

3.1.1. Cause (H3b4)

PRODUCE (✖3b4) (See 2.6.) Produce 6, produces 2	8
YIELD (✖3b4) (See 2.6.) Yield 2, yields 1	3
DERIVE (See 2.6.)	1

LED (See 2.6.)	1
MADE (See 2.6.)	1
TOTAL	14

3.2. Positive outcome

IMPROVEMENT Improve (+ SURVIVAL+) 7 Improve 4, improvements 2, improved 1	14
EFFECTIVE Effective 3, effect 2, Effectiveness 2, effectively 1, effects 1	9
BENEFIT	6
ACHIEVED Achieved 4, achieve 1	5
BEST Best 2, better 2	4
COMPLETE RESPONSE	4
PERFORMANCE (✖3b2)	4

~ status	
APPROPRIATELY Appropriately 2 Appropriate 1	3
TOLERATE Tolerate 2, tolerable 1	3
CONTRIBUTE-TO (✖3b2)	2
OBTAIN	2
OPTIMAL	2
COMPLETE REMISSION	2
CURE (See 0.1.)	1
EXCELLENT	1
TOTAL	62

3.3. Existence (H2, H4)

3.3.1. Survival

SURVIVAL - Survival 30, survivals 1 Median ~ 6 long-term ~ 3 prolonged ~ 2 - Survivors 1 Long-term ~ 2	32
MEDIAN → (~ survival) (See 2.5.)	6
LONG-TERM → (+ SURVIVAL)	5
PROLONGED → (+ SURVIVAL) (See 5.5.)	4

~ status	
OCCUR Occur 1, occurred 1, occurs 1	3
ENTITY	1
LIVE LONGER	1
REMAIN (✖4)	1
TOTAL	53

3.3.2. Death

DIE	4
Die 3, dying 1	
→ INCREASED (+ MORTALITY)	3
MORBIDITY	2

MORTALITY	2
TOTAL	11

3.4. Negative outcome

POOR (✖ 1)	8
- Poor 6	
~ prognosis 3	
~ performance status 2	
- Poorly 2	
TOXICITY	3
Toxicity 2, toxic 1	

FAILING (✖ 1)	1
NONRESPONDING	1
UNSATISFACTORY	1
TOTAL	14

4. LOCATION IN THE HUMAN BODY (H1, H5 a, LIMIT 0,59%, SPACE 0,21%)

4.1. Generic location

CONFINED	4
LOCAL	4
Local 3, localized 1	
AREAS (✖ 6)	3
DISTANT (See 1.3.)	2
~ metastases	

SUPERIOR (~ vena cava syndrome)	2
SITES	2
TOTAL	17

4.1.1. Generic location (medicine)

CONTRALATERAL (+ LYMPH NODES)	1
IPSILATERAL (+ LYMPH NODES)	1
LOCOREGIONAL	1

(~tumor involvement)	
TOTAL	3

4.2. Microscopic / histologic level (H1)

CELL (✖ 1)	5
Cell 4	
Cell + CANCER 53	
Cellular 1	
SQUAMOUS	1
LYMPHOCYTE-LIKE	1
TOTAL	7

4.3. Anatomic level / body parts (H5a)

4.3.1. Respiratory system

CHEST (✖ 5)	17
LUNG	13
- <lung> 7	
- pulmonary 6	
THORACIC	3
Thoracic 2, thorax 1	
BRONCHIAL (See 1.2.1.1.)	2
AIRWAYS	1

(malignant ~ obstruction)	
HILAR (~ lymph nodes)	1
ORAL (~ regimens)(See 0.2.1.)	1
PLEURAL	1
TOTAL	39

4.3.2. Other body parts / systems

CENTRAL NERVOUS SYSTEM	11
~ metastases 4 (See 1.3.)	
~ impairment 2	
≈ CNS 1	
BRAIN	8
CRANIAL	8
NEUROPSYCHOLOGICAL ≈	8
NEUROPSYCHOLOGIC	
NODES	4
Lymph ~ 2	
Supraclavicular ~ 2	
BONE (~ metastases 2)	3
(See 1.3.)	
BONE MARROW (✖ 5)	3

NEUROENDOCRINE	3
EPIDURAL	2
HEMATOLOGIC	2
MEDIASTINUM	2
NEUROLOGICAL	2
SUPERIOR VENA CAVA	2
(~ syndrome)(See 6.4.)	
ABDOMEN (✖ 5)	1
GLANDULAR	1
LIVER (✖ 5)	1
ORGANS	1
TOTAL	62

4.4. Location in relation to specific body parts (+ THERAPY, + TUMOR)

4.4.1. Bronchi

ENDOBONCHIAL (See 0.2.1.)	5
~ laser therapy 2	
~ lesions 2	
~ stents 1	

4.4.2. Clavicle

SUPRACLAVICULAR	4
~ (lymph) nodes 3	

4.4.3. Thorax

HEMITHORAX (~ of origin)	2
INTRATHORACIC (~ tumor)	1

4.4.4. Cranium

INTRACRANIAL (~ metastases)	1
-----------------------------	---

4.4.5. Veins

INTRAVENOUS (~ regimens) (See 0.2.1.)	1
	14

5. SIZE / QUANTITY (H1, H3b2, H6)

5.1. Name of lung cancer

SMALL ~ cell lung cancer 36 ~ cell cancers 1 ~ cell carcinoma 1 Combined ~ cell carcinoma 1	39
NON-SMALL ~ cell lung cancer 2 ~ cell cancers 1	3

LARGE (~ cell carcinoma)	1
TOTAL	43

5.2. Minuteness / small quantity (H1)

<SMALL>	7
LOW	2
MINIMAL	2
REDUCED	2
LIGHT MICROSCOPY	1

(See 6.5.)	
MODEST	1
SIMPLE	1
TOTAL	16

5.3. Medium size

MODERATE Moderate 1, moderately 1	2
--------------------------------------	---

5.4. Largeness / large quantity (H6)

EXTENSIVE	11
HIGH	5
High 2, highly 2, higher 1	
EXCEEDS	3
Exceeds 2, exceedingly 1	
GREATER	3
Greater 2, greatly 1	
EXTREME	3
Extreme 2, extremely 1	

LARGE	2
CONSIDERABLE (See 10.1.)	1
EXTENSIVELY	1
LONG	1
MASSIVE	1
POSITIVE	1
STRONGLY	1
TOTAL	33

5.5. Increase (H3b2)

ADDITIONAL	9
Additional 6 (See 0.2.)	
Addition 2, adding 1	
INCREASED	4
Increased 3, increasing 1	
PROLONGED (+ SURVIVAL)	4
(See 3.3.1.)	
Prolonged 2, prolongation 1,	
Prolongs 1	
EXCEEDS	2
INTENSIVE	2

Intensive 1, intensity 1	
PROGRESSED	2
Progressed 1, progressive 1	
MATURE	1
~ results of prospective randomized trials	
SPREAD (✕ 3b) (See 1.3.)	1
TOTAL	25

6. DIAGNOSIS (H0, H2)

6.1. Generic diagnosis

DIAGNOSIS	10
Time of ~ 5	
PROGNOSIS	10
Prognosis 9	
Poor ~ 3	
Better ~ 2	

Intensive 1, intensity 1	
STATUS (Neurological ~)	2
TOTAL	22

6.2. Sensory processes involved

DETECTABLE	5
Detectable 2, detectably 2,	
detect 1	
IDENTIFICATION	2
Identification 1, identifies 1	

EXAMINATION	1
TOTAL	8

6.3. Cognitive processes involved (ESTIMATE 0,54%)

EVALUATION (See 2.4.)	10
Evaluation 9 Clinical ~ 6 Evaluated 1	
CONSIDERED (See 2.4.)	4
TOTAL	14

6.4. Clinical features ([Symptoms](#))

FUNCTION	9
Pulmonary ~ 3 Neuropsychologic ~ 2 Neuropsychological ~ 2	
DECLINE (✖ 1)	6
Decline 4, declining 2	
NEUROPSYCHOLOGIC	6
- Neuropsychologic ~ function 2 - Neuropsychological ~ function 2 ~ abnormalities present 2	
SYMPTOMS (See 1.1.)	4
AFFECTED (See 1.1.)	2
LESIONS (See 1.1.)	2

Endobronchial (obstructing) ~	
OBSTRUCTING	2
Obstructing 1 Malignant airways obstruction 1	
SUPERIOR VENA CAVA SYNDROME (See 4.3.2.)	2
NEUROENDOCRINE GRANULES (✖ 1)	1
PLEURAL EFFUSION	1
EXTRINSIC COMPRESSION (✖ 1)	1
TOTAL	36

6.5. Diagnostic tests and instruments (INSTRUMENTALITY 0,33%, INSTRUMENT 0,25%)

TOMOGRAPHIC	2
- Computed ~ scans 1 - Computerized ~ scans 1	
MICROSCOPY (✖ 1)	2
Electron microscopy 1 Light microscopy 1 (See 5.2.)	
NEUROPSYCHOLOGIC TESTING	2

MAGNETIC RESONANCE IMAGING SCANS	1
RADIONUCLIDE BONE SCANS	1
BRONCHOSCOPE	1
TOTAL	9

6.6. Prognostic factors

PROGNOSTIC FACTORS (See 2.4.)	1
STAGE (See 8.)	35
PERFORMANCE STATUS	4

AGE (See 7.)	1
TOTAL	6

7. TIME (H4, SEQUENCE 0,99 %, PERIOD 0,89%, AGE 0,83%, BEGINNING 0,12 %)

YEARS	14
START	13
Start 11, starting 2	
CURRENT	11
Current 9, currently 2	
MONTHS	8
TIME	8
Time 7, timing 1	
INITIAL	7
Initial 4 (See 0.2.)	
Initially 2, initiating 1	
DURATION	5
PROSPECTIVE (See 3.1.)	5
- ~ randomized	
- ~ randomized trials 2	

~ randomized studies 1	
- ~ studies 2	
DAILY	4
PREVIOUSLY	3
RETROSPECTIVE (See 3.1.)	3
~ studies 2	
BIWEEKLY	2
CONTINUE (✖ 4)	2
PERIOD (✖ 3)	1
AGE (See 6.6.)	1
SHORT-TERM	1
TOTAL	88

8. STAGING (H3 a, H6)

STAGE (✖ 3) (See 6.6.)	40
Stage 35	
Limited ~ 19	
Extensive ~ 9	
~ disease 9	
Staging 5 (~ procedures 2)	
RELAPSE	7
Relapse 5, relapsed 1,	

~ randomized studies 1	
CONFINED	4
TNM	1
TOTAL	52

8.1. Small cell lung cancer

LIMITED (✖ 1)	19
- ~ stage small cell lung cancer 7	
- ~ disease 6	
- ~ stage patients 2	
EXTENSIVE (✖ 6)	11
~ stage small cell lung cancer 6	
~ disease 2	
~ stage disease 2	
RECURRENT (See 1.3.)	3
- Recurrent	

~ randomized studies 1	
- ~ studies 2	
OCCULT METASTATIC DISEASE	2
TOTAL	35

9. PATIENT (H2)

PATIENTS	78
- Patients 75	
~ with small cell lung cancer	
7	
~ with limited stage CANCER 6	
~ treated with	
(+ RADIO / CHEMOTHERAPY) 6	
~ treated for (+ CANCER) 2	
- Patient 3	

GROUP (+ PATIENT) (See 11.)	5
Group 3, groups 2	
CANDIDATES	1
SERIES (✖ 4)	1
TOTAL	85

10. IMPORTANCE (H5b)

10.1. Importance

NEW	8
SIGNIFICANT	5
Significant 4, significantly 1	
PRIMARILY	4
Primarily 2, primary 2	
IMPORTANT	3
REPRESENTATIVE	3
Representative 2, represent 1	
CORNERSTONE	2
SUPERIOR (~ to)	2

CONSIDERABLE (See 5.4.)	1
INTRINSIC	1
MAJOR	1
SUBSTANTIAL (✖ 6)	1
UNEQUIVOCALLY (✖ 6)	1
UNIVERSALLY (✖ 6)	1
WIDESPREAD (✖ 6)	1
TOTAL	34

10.2. Centrality

CENTRAL	10
---------	----

10.3. Unusualness

RARELY	5
Rarely 3, rare 2	
ATYPICAL (~ carcinoid)	1
(See 1.2.1.)	
UNUSUAL	1
TOTAL	7

11. COMPOSITION (H3b3, H6)

INCLUDE Include 6, included 2, including 1 inclusion 1	10
COMPLETE (✖ 6)	8
GROUP Group 3, groups 2	5
MAJORITY (✖ 6)	5

COMPOSED OF Composed 1, components 1	1
INCORPORATION	1
TOTAL (✖ 6)	1
TOTAL	32

12. GENERAL MEDICINE (H0)

CLINICAL - Clinical 13 ~ evaluation 8 ~ trials 5 - Preclinical 1	14
MEDICAL	2

Medical 1, medically 1	
TOTAL	16

12.1. Health professionals

ONCOLOGIST (See 1.2.)	1
PATHOLOGIST (See 1.1.)	1
PHYSICIANS	1
TOTAL	3

13. OPTIONS (MODALITY 0,66%)

MODALITY (See 0.1.)	4
OPTIONS (See 0.1.) Options 3, option 1	4
SELECTION Selection 2, selected 1	3
TOTAL	11

14. NEED

NEEDED	4
NECESSARY	1
REQUIRES (✖ 1)	1
TOTAL	6

15. VARIATION (H3b6)

VARIATION (✖ 3) Variation 2, variety 1, varying 1	4
--	---

16. DIFFICULTY

DIFFICULT	2
-----------	---

17. SMOKING (RISK FACTOR) (H7)

CIGARETTE	1
SMOKERS	1

11.1.3.5. CUADRO CONTRASTIVO (*HESPERUS* – ANÁLISIS COMBINADO)

En el texto se vinculan 1733 palabras de las 3408 que lo componen, es decir, el 50,85% de las palabras del texto crean cohesión, un porcentaje no muy distinto al 48,35 % que se deduce de los resultados de *Hesperus*.

Por otra parte, hay un promedio de 12 palabras vinculadas por oración. Las 3408 palabras vinculadas se relacionan mediante 184 categorías conceptuales y adquieren forma mediante 478 formas distintas basadas en 337 lemas. En 100 palabras vinculadas, hay 27 formas léxicas distintas de 19 lemas y se activan 11 categorías conceptuales.

TABLA 88: Descripción cuantitativa de las cadenas léxicas reconocidas.

CADENAS LÉXICAS RECONOCIDAS POR <i>HESPERUS</i>		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H0: TREATMENT AND DIAGNOSIS [225-27-22,12] ¹⁴⁸			0. TREATMENT [478-94-26,51] 6. DIAGNOSIS [95-38-5,27] 12. GENERAL MEDICINE [19-7-1,05]
Lexemas vinculados	225	592	Lexemas vinculados
Formas distintas	27	139	Formas distintas
Activación conceptual ajustada (%)	23,57	32,83	Activación conceptual (%)
H 2: PATHOLOGY OF LUNG CANCER [201-23-9,17]			1. PATHOLOGY OF SMALL CELL LUNG CANCER
Lexemas vinculados	201	200	Lexemas vinculados
Formas distintas	23	58	Formas distintas
Activación conceptual ajustada (%)	9,77	11,1	Activación conceptual (%)
H 3b1: EVIDENCE [16-5-7,39] H 3b5: RESEARCH [5-4-3,73] H 3b4: CAUSE [8-2-1,53]			2. RESEARCH, EXPERIMENTATION
Lexemas vinculados	29	192	Lexemas vinculados
Formas distintas	11	75	Formas distintas
Activación conceptual ajustada (%)	13,48	10,65	Activación conceptual (%)
H 4: EXISTENCE, SURVIVAL [55-17-4,85] H 3b2: IMPROVEMENT, SURVIVAL [15-5-0,37]			3. OUTCOME OF TREATMENT
Lexemas vinculados	55	190	Lexemas vinculados
Formas distintas	17	70	Formas distintas
Activación conceptual ajustada (%)	5,56	10,54	Activación conceptual (%)
H 5 a: BODY PARTS [22-6-1,65] LIMIT [0,59%], SPACE [0,21%]			4. LOCATION IN THE HUMAN BODY
Lexemas vinculados	22	142	Lexemas vinculados
Formas distintas	6	49	Formas distintas
Activación conceptual ajustada (%)	2,61	7,87	Activación conceptual (%)
H 1: MINUTENESS, SIZE [192-39-18,22] H 6: LARGENESS [56-26-1,99] H 3b2: INCREASE [24-14-0,84]			5. SIZE / QUANTITY
Lexemas vinculados	272	119	Lexemas vinculados
Formas distintas	79	43	Formas distintas
Activación conceptual ajustada (%)	22,43	6,6	Activación conceptual (%)
SEQUENCE [0,99%], PERIOD [0,89%], AGE [0,83%], BEGINNING [0,12%] ¹⁴⁹ = 2,83 %			7. TIME
Activación conceptual ajustada (%)	3,02	4,88	Lexemas vinculados
			Formas distintas
			Activación conceptual (%)

¹⁴⁸ La leyenda de los números entre corchetes es la siguiente: [Lexemas vinculados – formas distintas – activación conceptual].

¹⁴⁹ *Hesperus* no reconoce esta cadena, aunque el perfil del documento sí otorga un porcentaje a conceptos relacionados con TIME.

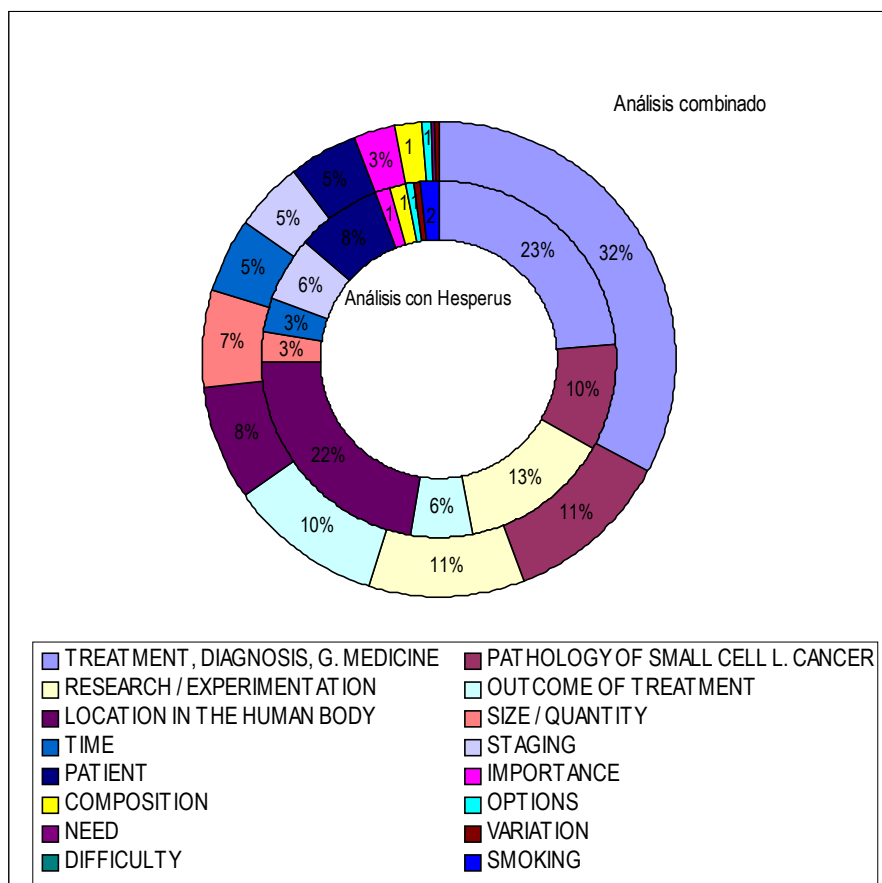
H 3 a: STAGING [41-2-5,67]			8. STAGING
Lexemas vinculados	41	87	Lexemas vinculados
Formas distintas	2	12	Formas distintas
Activación conceptual ajustada (%)	6,04	4,82	Activación conceptual (%)
H 2: PATHOLOGY OF LUNG CANCER (PATIENT) [78-2-7,29]			9. PATIENT
Lexemas vinculados	78	85	Lexemas vinculados
Formas distintas	2	6	Formas distintas
Activación conceptual ajustada (%)	7,77	4,71	Activación conceptual (%)
H 5b: IMPORTANCE [31-13-1,14]			10. IMPORTANCE
Lexemas vinculados	31	51	Lexemas vinculados
Formas distintas	13	21	Formas distintas
Activación conceptual ajustada (%)	1,21	2,83	Activación conceptual (%)
H 3b3: PART – WHOLE [22-6-1,39]			11. COMPOSITION
Lexemas vinculados	22	32	Lexemas vinculados
Formas distintas	6	13	Formas distintas
Activación conceptual ajustada (%)	1,48	1,77	Activación conceptual (%)
MODALITY [0,66%]			13. OPTIONS [11-5-0,61]
Activación conceptual ajustada (%)	0,71	0,61	Activación conceptual (%)
	0	0,33	14. NEED [6-3-0,33]
			Activación conceptual (%)
H 3b6: CHANGE [1-1-0,53]			15. VARIATION
Lexemas vinculados	1	4	Lexemas vinculados
Formas distintas	1	3	Formas distintas
Activación conceptual ajustada (%)	0,56	0,22	Activación conceptual (%)
	0	0,11	16. DIFFICULTY [2-1-0,11]
			Activación conceptual (%)
H 7: SMOKING [3-3-1,66]			17. SMOKING (RISK FACTOR)
Lexemas vinculados	3	2	Lexemas vinculados
Formas distintas	3	2	Formas distintas
Activación conceptual ajustada (%)	1,77	0,11	Activación conceptual (%)
TOTAL			TOTAL
Activación conceptual según <i>Hesperus</i>	93,83	1803	Suma de elementos en las cadenas
Lexemas vinculados	980	1733	Lexemas vinculados ¹⁵⁰
Formas distintas	190	478	Formas distintas ¹⁵¹
Activación conceptual ajustada (%)	99,98	99,98	Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QDT3	Categorías conceptuales del tesoro: 184		
Nº total de palabras: 3408	Coeficiente de cohesión (<i>Hesperus</i>): 8,87		
Palabras que crean cohesión (<i>Hesperus</i>): 1648	Nº de oraciones: 142		

¹⁵⁰ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹⁵¹ De estas, hay 43 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 70 lexemas adicionales a los 1733 lexemas vinculados.

Representamos estos datos en un diagrama de sectores doble donde el anillo exterior designa los resultados de nuestro análisis y el interior, los datos obtenidos a partir de *Hesperus*. Las diferencias entre estos dos son más significativas en QDT3 que en los otros dos textos para especialistas, sobre todo en la cadena *TREATMENT*, dado que esta contiene muchos términos sobre oncología, tal y como *lobectomy* o *pneumonectomy*, que obviamente no están contenidos en el tesauro integrado en *Hesperus*.

DIAGRAMA 27: Comparación de las áreas conceptuales reconocidas por *Hesperus* y las reconocidas por un humano asistido por medios informáticos (QDT3).



11.2. Textos para pacientes sobre tratamiento del cáncer de pulmón

11.2. Textos para pacientes sobre tratamiento del cáncer de pulmón

En este apartado se analiza la versión de divulgación de los textos precedentes, que el *National Cancer Institute* pone a disposición de los pacientes y sus familiares (Textos 12, 13 y 14 del Apéndice II). En el primer texto (QPT1), anticiparemos algunas de las diferencias entre dichas versiones, que se derivan del cambio en el parámetro *destinatario*. El resto de las diferencias se comentarán en la cuarta parte de la tesis, a partir del capítulo 12.

11.2.1. Mesothelioma (QPT1)

11.2.1.1. ÍNDICE DEL TEXTO

El índice adelanta que los conceptos marcados en negrita van a ser centrales en el texto 12.

(90)

DESCRIPTION

What is **malignant mesothelioma**?

STAGE EXPLANATION

Stages of malignant mesothelioma

Localized malignant mesothelioma

Advanced malignant mesothelioma

Recurrent malignant mesothelioma

TREATMENT OPTION OVERVIEW

How malignant mesothelioma is treated

Treatment by stage

LOCALIZED MALIGNANT MESOTHELIOMA (STAGE I)

ADVANCED MALIGNANT MESOTHELIOMA (STAGES II, III, AND IV)

RECURRENT MALIGNANT MESOTHELIOMA

TO LEARN MORE

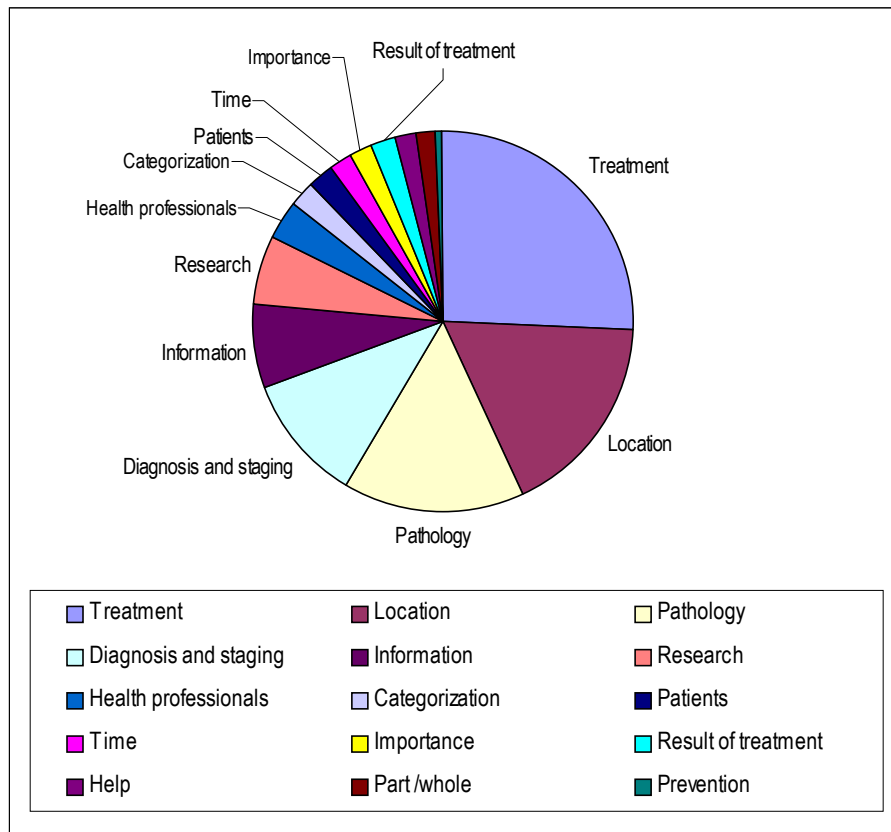
11.2.1.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTA DE FRECUENCIA LEMATIZADA

Algunos de los conceptos mencionados en 11.2.1.1. quedan representados en la tabla y el gráfico que hemos derivado de la lista de frecuencia lematizada elaborada con *Wordsmith Tools*.

TABLA 89: Identificación preliminar de áreas conceptuales activadas en QPT1 a partir de lista de frecuencia lematizada

PT1	
ÁREAS CONCEPTUALES	%
Treatment	10,42
Location	7,03
Pathology	6,16
Diagnosis and staging	4,35
Information	2,92
Research	2,35
Health professionals	1,3
Categorization	0,93
Patients	0,86
Time	0,86
Importance	0,81
Result of treatment	0,81
Help	0,68
Part /whole	0,68
Prevention	0,25
TOTAL	40,41

DIAGRAMA 28: Representación preliminar de las áreas conceptuales activadas en QPT1



11.2.1.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

Las cadenas que *Hesperus* analiza son siete (véase apéndice VIIe). Según el perfil del documento, de las 1621 palabras del texto, 704 son plenas y están vinculadas mediante 109 categorías del tesoro. El cociente de dividir el valor otorgado por *Hesperus* (17351) y el número total de palabras es de 10,7.

Las categorías conceptuales activadas que aportan más de un 0,5% de la cohesión del texto son las siguientes:

badness_645_4141_n	Percent: 15.26,
agency_173_1101_n	Percent: 8.48,
hospital_658_4269_n	Percent: 6.79,
deadly_362_2384_a	Percent: 5.49,
therapy_658_4268_n	Percent: 5.48,
box_194_1240_n	Percent: 5.08,
diagnostic_658_4258_n	Percent: 3.78,
knowledge_490_3165_n	Percent: 3.27,
divergence_294_1963_n	Percent: 3.15,
representation_551_3579_n	Percent: 2.78,
rostrum_539_3499_n	Percent: 2.22,
nonadhesive_49_343_a	Percent: 2.18,
medical_658_4274_a	Percent: 2.13,
experiment_461_2959_n	Percent: 1.82,
receptacle_194_1234_n	Percent: 1.78,
maw_194_1236_n	Percent: 1.76,
separate_46_319_v	Percent: 1.75,
scission_46_315_n	Percent: 1.73,
resident_191_1203_n	Percent: 1.36,
middle_70_473_n	Percent: 0.97,
health_650_4181_n	Percent: 0.96,
region_184_1162_n	Percent: 0.96,
distant_199_1297_a	Percent: 0.86,
truth_494_3189_n	Percent: 0.86,
know_490_3172_v	Percent: 0.77,
corporation_708_4565_n	Percent: 0.74,
respiration_352_2324_n	Percent: 0.72,
component_58_403_a	Percent: 0.69,
native_191_1204_n	Percent: 0.66,
detention_747_4854_n	Percent: 0.64,
materials_631_4046_n	Percent: 0.58,
tube_263_1726_n	Percent: 0.58,
testing-agent_461_2962_n	Percent: 0.55,
known_490_3171_a	Percent: 0.54,
aiding_703_4539_a	Percent: 0.52,
compartment_194_1237_n	Percent: 0.51,
removed_199_1298_a	Percent: 0.51,
testee_461_2963_n	Percent: 0.50,
operate_173_1103_v	Percent: 0.49,
produce_164_1056_v	Percent: 0.49.

Nos aproximamos ahora a comentar los conceptos activados en las cadenas reconocidas por *Hesperus*.

TABLA 90: **Cadena 0** (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 0 (Malignant)	
BADNESS	15,26
DEADLY	5,49
KILL	0,3
HARMFUL	0 0,1
TOTAL	+ 20,75
↓	
DISEASE, DEATH, HARM, BADNESS	

Comentario

Esta cadena es la más representativa del texto. Se destacan los aspectos letales de la enfermedad (DEADLY) y cómo el tratamiento está destinado a matar las células cancerosas.

TABLA 91: **Cadena 1** (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 1 (Screening)	
HOSPITAL	6,79
THERAPY	5,48
DIAGNOSTIC	3,78
MEDICAL	2,13
HEALTH	0,96
DOCTOR	0,2
DRUG	0,2
MEDICINE	0,18
MEDICAL-ART	0,13
REMEDY	0,13
TOTAL	19,98
↓	
TREATMENT AND DIAGNOSIS, HOSPITAL	

Comentario

En la segunda cadena más importante, se antepone a los conceptos TERAPIA y TRATAMIENTO, el concepto más genérico HOSPITAL, aunque realmente el lexema *hospital* sólo aparece tres veces.

TABLA 92: **Cadena 2** (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 2 (Sac)	
BOX	5,08
RECEPTACLE	1,78
MAW	1,76
COMPARTMENT	0,51
CELLULAR	0,29
BLADDER	0 0,1
TOTAL	+ 9,42
↓	
LOCATION, PARTS OF THE BODY	

Comentario

Con cerca de un 9,5%, la activación de la cadena *LOCALIZACIÓN* resulta clave en textos sobre oncología no especializados. Los conceptos identificados por *Hesperus* (BOX, RECEPTACLE, COMPARTMENT) inciden en la idea de los órganos como contenedor, lo cual constituye una metáfora sistemática en el campo de la medicina.

TABLA 93: **Cadena 3** (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 3 (Treatment)	
AGENCY (Treatment)	8,48
OPERATE	0,49
USE	0 0,1
TOTAL	+ 8,97
↓	
TREATMENT, AGENCY, OUTCOME OF TREATMENT	

Comentario

Esta cadena incide en la idea de ACCIÓN. Por ese motivo, asocia los lemas WORK, JOBS, EFFECTIVENESS Y TREATMENT, aunque se ubicará este último en una cadena con esa etiqueta.

TABLA 94: Cadena 4 (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 4 (Radiation)	
DIVERGENCE	3,15
REPRESENTATION?	2,78
IMAGE?	0,43
	0 0,1
TOTAL	6,36
↓	
THERAPY, DISPERSION	

Comentario

Esta cadena, en la que predomina el concepto DISPERSION, puede desglosarse en dos áreas conceptuales, por una parte la RADIOTERAPIA y, por otra, la DISPERSIÓN de las células cancerosas por el cuerpo.

TABLA 95: Cadena 5 (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 5 (Learn)	
KNOWLEDGE	3,27
KNOW	0,77
KNOWN	0,54
INSTRUCTED	0,33
TOTAL	4,91
↓	
KNOWLEDGE, INSTRUCTION	

Comentario

A juzgar por el 4,91 % que *Hesperus* otorga a la cadena 5, construida sobre KNOWLEDGE, KNOW, KNOWN y INSTRUCTED, el conocimiento es la mejor arma para afrontar la enfermedad.

TABLA 96: Cadena 6 (QPT1)

CONCEPTOS ACTIVADOS	
Cadena 6 (Institute)	
ROSTRUM (<i>stage</i>) ¹⁵⁴	2,22
CORPORATION	0,74
PRODUCE	0,49
PARTY (<i>cells</i>)	0,45
CAUSAL-MEANS	0,28
START	0 0,1
REASON-WHY	0 0,1
CLASSROOM	0 0,1
PRODUCED	0 0,1
TOTAL	+ 4,18
↓	
? (CAUSE, STAGE, BODY)	

Comentario

Como hemos señalado con el signo de interrogación, en esta cadena se concentran una disparidad de conceptos, de entre los que destaca el de CAUSALIDAD. Ubicaremos los lexemas de esta cadena difusa en las cadenas *LOCATION*, *STAGING*, *DIAGNOSIS*, *TREATMENT*, *INSTRUMENT* y *TIME*.

11.2.1.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

Además de las siete cadenas mencionadas, hemos señalado otras en torno a algunos conceptos del perfil del documento, como se ilustra a continuación.

Conceptos reconocidos por Hesperus	Cadenas (% de activación conceptual)
REGION (0,96) → DISTANT (0,86) → SPACE (0,44) → MIDDLE (0,97)	LOCATION IN THE HUMAN BODY (3,23)
EXPERIMENT (1,82) → TESTING-AGENT (0,55) → TESTEE (0,5) → QUESTION (0,66) → ENQUIRY (0,38) → ENQUIRE (0,19)	RESEARCH / EXPERIMENTATION (4,1)
SEPARATE (1,75) → COMPONENT (0,69) → WHOLE (0,29) → SEPARATION (0,25) → CONTENTS (0,22)	COMPOSITION (3,2)

En la tabla 97 adelantamos las cadenas que hemos reconocido en QPT1 y que desglosaremos posteriormente.

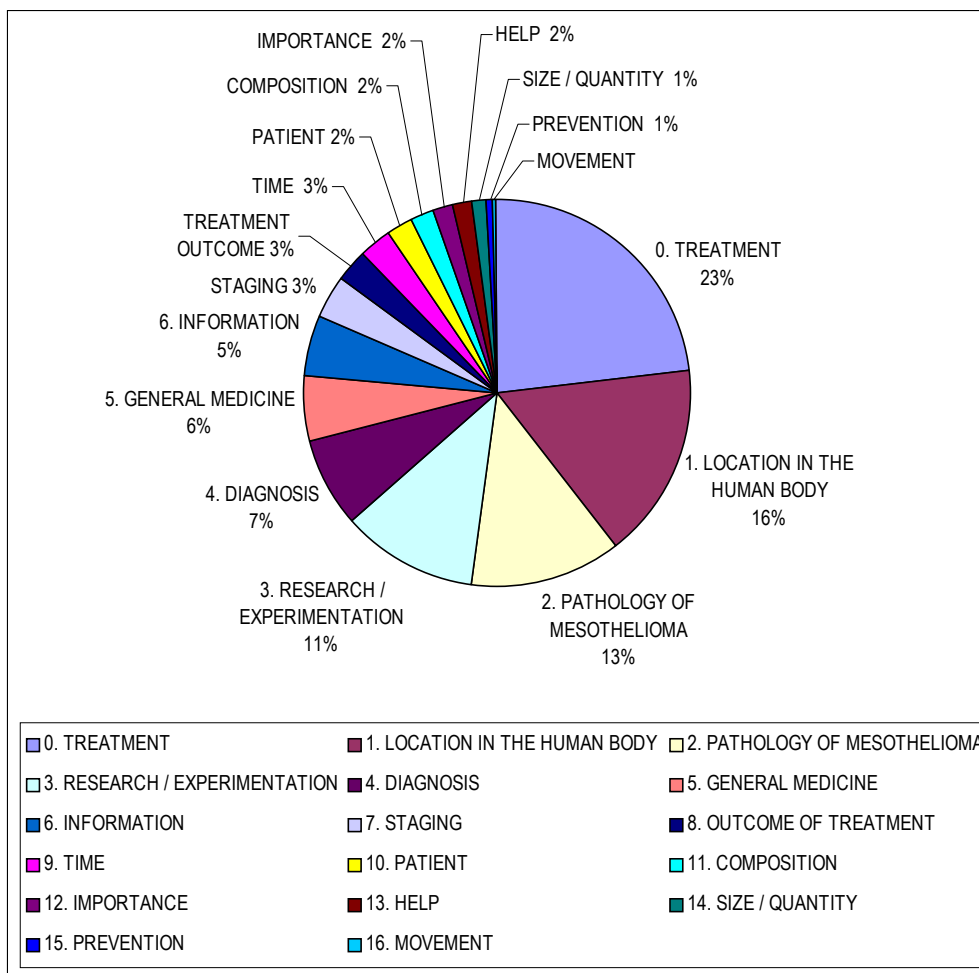
TABLA 97: Cadenas léxicas en QPT1

<p>Chain 0: TREATMENT (H1, H3, H4, H6) [182]</p> <p>0.1. Generic treatment 0.2. Properties (→ TREATMENT) 0.3. Manipulation 0.3.1. Manipulation involved in surgery 0.4. Process 0.5. Types 0.5.1. Radiation therapy 0.5.1.1. Radiation agents 0.5.2. Chemotherapy 0.5.2.1. Drugs 0.5.3. Surgery 0.5.4. Other types of treatment 0.6. Instruments</p> <p>Chain 1: LOCATION IN THE HUMAN BODY (H2, H6) [125]</p> <p>1.1. Generic location 1.2. Microscopic / histologic level 1.3. Anatomic level / body parts 1.3.1. Respiratory system 1.4. Location in relation to specific body parts (+ THERAPY) 1.5. Definition of body parts 1.6. Body fluids</p> <p>Chain 2: PATHOLOGY OF MALIGNANT MESOTHELIOMA (H0, H4) [99]</p> <p>2.1. Generic disease 2.2. Cancer 2.2.1. Types 2.3. Spread of cancer 2.4. Asbestos (risk factor)</p> <p>Chain 3: RESEARCH, EXPERIMENTATION (H6) [89]</p> <p>3.1. Study 3.2. Categorization 3.3. Communication</p>	<p>3.4. Evidence 3.5. Causality 3.6. Normality</p> <p>Chain 4: DIAGNOSIS (H1, H0, H6) [57]</p> <p>4.1. Generic diagnosis 4.2. Sensory / cognitive processes involved 4.3. Clinical features (symptoms) 4.4. Diagnostic tests and instruments</p> <p>Chain 5: GENERAL MEDICINE (H5, H6) [43]</p> <p>5.1. Health professionals</p> <p>Chain 6: INFORMATION (H5) [40]</p> <p>Chain 7: STAGING (H6) [27]</p> <p>Chain 8: OUTCOME OF TREATMENT (H0, H3) [23]</p> <p>8.1. Positive outcome 8.2. Negative outcome</p> <p>Chain 9: TIME (H5) [22]</p> <p>Chain 10: PATIENT [17]</p> <p>Chain 11: COMPOSITION [16]</p> <p>Chain 12: IMPORTANCE [13]</p> <p>Chain 13: HELP (H3) [13]</p> <p>Chain 14: SIZE [10]</p> <p>Chain 15: PREVENTION [4]</p> <p>Chain 16: MOVEMENT [2]</p> <hr/> <p>TOTAL DE UNIDADES LÉXICAS ENCADENADAS ¹⁵⁵ : [749]</p>
--	---

¹⁵⁴ Esta etiqueta es equivalente a la de DRAMATURGY, que aparece en otros textos asociada al lema STAGE.

¹⁵⁵ Para obtener el total no se han duplicado los lemas que pertenecen a más de una cadena. La suma de la frecuencia absoluta de palabras en cada una de las cadenas es de 781.

DIAGRAMA 29: Porcentaje de activación de las cadenas léxicas definitivas de QPT1



CADENAS LÉXICAS	Frec	%	CADENAS LÉXICAS	Frec	%
0. TREATMENT	182	23,33	10. PATIENT	17	2,18
1. LOCATION IN THE HUMAN BODY	125	16,03	11. COMPOSITION	16	2,05
2. PATHOLOGY OF M. MESOTHELIOMA	99	12,69	12. IMPORTANCE	13	1,66
3. RESEARCH / EXPERIMENTATION	89	11,41	13. HELP	13	1,66
4. DIAGNOSIS	57	7,31	14. SIZE	10	1,15
5. GENERAL MEDICINE	43	5,51	15. PREVENTION	4	0,51
6. INFORMATION	40	5,13	16. MOVEMENT	2	0,26
7. STAGING	27	3,46	TOTAL	782	100
8. OUTCOME OF TREATMENT	23	2,94			
9. TIME	22	2,82			

Por el momento, simplemente señalaremos que la cadena *TRATAMIENTO* es la más significativa con un 25%, seguida de la cadena que indica *LOCALIZACIÓN EN EL CUERPO HUMANO*. Esta última, junto con la de *PATOLOGÍA*, representan un 25 % de la activación conceptual.

La cadena *RESEARCH / EXPERIMENTATION*, repleta de vocabulario organizador del discurso, representa un 11%, un porcentaje parecido al que adquiere en los otros textos sobre tratamiento del *PDQ*. Este vocabulario se encuentra a caballo entre el metalenguaje científico y la lengua general, en cuanto que apunta hacia los procedimientos implicados en la construcción del conocimiento.

Por encima del 5%, encontramos cadenas con términos médicos (*DIAGNOSIS, GENERAL MEDICINE, STAGING*) y una relacionada con la cognición (*INFORMATION*). La cadena *INFORMATION* contiene numerosas palabras de la sección introductoria de los textos del *Physician Data Query* en la que se explica en qué consiste este servicio de información sobre el cáncer. Esta sección, que inicialmente estaba en todos los textos y que cuenta con 408 palabras, sólo la hemos dejado en el texto sobre mesotelioma maligno y hace referencia no sólo a la necesidad de información sino también a los profesionales que ayudan a afrontar la enfermedad, de ahí que en este texto la cadena *GENERAL MEDICINE* y su subcadena *HEALTH PROFESSIONALS* sea mucho más representativa que en los demás textos para pacientes y que *Hesperus* elaborara la cadena 1 sobre el concepto *HOSPITAL*.

11.2.1.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS

0. TREATMENT (H1, H3, H4, H6)

0.1. Generic treatment

TREATMENT (✖3) Treatment 27, treatments 7, treated 3, treat 1	38
TRIALS (See 3.1.) Trials 11, trial 7 Clinical trial(s) 16 (See 5.)	18
→ WAYS (~ to treat cancer)	2
CURED (See 8.1.)	1
FIGHT (~ cancer)	1
→ KINDS (~ of treatment)	1

(See 3.2.)	
PROGRAMS	1
<THERAPY> (See 0.5.1., 0.5.3.)	1
OPERATION (✖3)	1
→ TYPE (~ of treatment) (See 3.2.)	1
TOTAL	65

0.2. Properties (→TREATMENT)

STANDARD →	3
SYSTEMIC → (~ treatment)	1

0.2.1. Location where TREATMENT is applied

<u>INTRAPLEURAL</u> (See 1.4.) (~ chemotherapy)	1
<u>EXTERNAL</u> (~ radiation therapy) (See 1.1.)	1
<u>INTERNAL</u> (~ radiation therapy) (See 1.1.)	1

0.3. Manipulation (H6)

USED	13
GIVEN	5
TAKING	4

MADE (✖ 6)	3
<u>INJECTED</u> (+ DRUGS)	1
TOTAL	26

0.3.1. Manipulation involved in surgery

REMOVE Remove 6, removed 4	10
CUT	2

OPENING (✖ 2)	1
TOTAL	13

0.4. Process

PROGRAMS	1
----------	---

0.5. Types

0.5.1. Radiation

<u>RADIATION</u> ~ therapy 9 external beam ~ 1 external ~ 1 internal ~ 1	11
--	----

0.5.1.1. Radiation agents (H4, H6)

X-RAYS (✖ 4) X-rays 2, x-ray 1 (See 4.4.)	3
→ HIGH-ENERGY (✖ 6) (+ RAYS) (See 15.)	2
LIGHT (✖ 4) Light (v) 1, light (n) 1	2
RADIOISOTOPES ≈ Materials	2

that produce radiation	
BEAM	1
RAYS (✖ 4)	1
→ HIGH-DOSE (~ rays) (See 15.)	1
TOTAL	12

0.5.2. Chemotherapy

CHEMOTHERAPY	11
INJECTED (+ DRUGS)	1
TOTAL	12

0.5.2.1. Drugs

DRUGS Drugs 5, drug 3	8
ANESTHETIC	2
PILL	1
TOTAL	11

0.5.3. Surgery

SURGERY	13
---------	----

0.5.4. Other types of treatment (H2)

DRAIN (✖ 2)	2
INTRAOPERATIVE PHOTODYNAMIC THERAPY	1
PARACENTESIS	1

THORACENTESIS	1
SUCTION	1
TOTAL	6

0.6. Instruments (H6)

INSTRUMENT	1
NEEDLE	1
MACHINE (✖ 6)	1

TOOL (✖ 6)	1
TUBE	1
TOTAL	5

En los textos destinados a pacientes sobre tratamiento (QPT), la cadena *TRATAMIENTO* abunda en expresiones genéricas (*treatment, fight cancer, cure*) que explican el significado de términos terapéuticos desconocidos o conocidos muy vagamente por el lector. La estructura de la cadena es bastante similar a la versión correspondiente a expertos, a excepción de las subcadenas *PROCESS, MANIPULATION INVOLVED IN SURGERY* y *INSTRUMENTS*. Estas responden a la necesidad de aclarar o parafrasear términos del especialista, como se ve en estos ejemplos:

(91)

TÉRMINOS ESPECIALIZADOS

Pleurectomy

Pneumonectomy

PARÁFRASIS EN TEXTOS PARA PACIENTES

surgery to remove sections of the pleura

surgery to remove the pleura and the tissue near it

Con respecto a los textos para especialistas, se percibe un mayor uso de lexemas de la lengua general y de términos semiespecializados: *radiation therapy* en lugar de *radiotherapy*; *drain the fluid* en lugar de *drainage of effusions*; *drugs* en lugar de *agents*; *internal* en lugar de *intracavitary*.

1. LOCATION IN THE HUMAN BODY (H2, H6)

1.1. Generic location (H6)

BODY (✖ 6)	7
FAR (how ~)	3
EXTERNAL (See 0.2.)	2
LOCALIZED (See 6.1.)	2
LOCAL	2
PART(S) OF THE BODY (See 11.)	2
AREA	1
CENTER	1

DISTANT (See 2.3.)	1
INTERNAL (See 0.2.)	1
NEARBY	1
PLACE	1
SIDE	1
TOTAL	25

1.2. Microscopic / histologic level (H6)

CANCER CELLS (✖ 6)	10
TISSUE Tissue 4, tissues 1	5
TOTAL	15

1.3. Anatomic level / body parts

ABDOMEN Abdomen 14, abdominal 1	15
DIAPHRAGM	4
HEART	3
LYMPH NODES	2
MUSCLE	2

VEIN	2
ORGANS	1
PERITONEUM	1
RIBS	1
TOTAL	31

1.3.1. Respiratory system

<CHEST>	20
PLEURA	6
LUNG Lung 4, lungs 1	5

CHEST CAVITY	2
CHEST WALL	2
TOTAL	35

1.4. Location in relation to specific body parts (+ THERAPY)

INTRAPLEURAL (See 0.2.) (~ chemotherapy)	1
---	---

1.5. Definition of body parts

LINING (+ CHEST / + PLEURA)	7
SAC LINING THE CHEST	1
SAC LINING THE ABDOMEN	1
TOTAL	9

1.6. Body fluids

FLUID	8
BLOODSTREAM	1
TOTAL	9

La cadena *LOCALIZACIÓN* es la segunda más importante del texto. Los vocablos que designan tejidos, órganos, membranas y regiones del cuerpo van a menudo acompañados, por una parte, por lexemas genéricos que los definen (*part of the body, area, sac lining the chest / abdomen*), y por otra, por miembros de las cadenas *PATHOLOGY, DIAGNOSIS* y *TREATMENT*.

Comentamos ahora las alternativas que se proponen a algunos términos médicos de la sección para especialistas. Domina la sustitución de términos con prefijo latino (*intra-*, *extra-*, *ipsi-*, *contra-*) por sintagmas preposicionales introducidos por *in*, *inside* o *outside*.

(92)

TÉRMINOS ESPECIALIZADOS	DESIGNACIÓN EN TEXTOS PARA PACIENTES
Histologic	Tissue
(resection of ...) contiguous structures	(remove...) and tissue around it
Peritoneum	Abdominal lining
Mediastinum	Center of the chest
Intrathoracic lymph nodes	Lymph nodes in the chest
Diaphragm	muscle below the lungs
Intracavitary	Inside the chest cavity
Extrathoracic or contralateral lymph node involvement	Nearby lymph nodes

El último caso ilustra un fenómeno habitual en textos que acercan el conocimiento especializado al lector lego. Es la pérdida de precisión para facilitar la comprensión por parte del lector. El término *contralateral* especifica que en el estadio III, la enfermedad ha pasado a los ganglios linfáticos del lado opuesto similar a la zona de la pleura donde surgió el tumor. Este significado es mucho más específico que el que aporta el adjetivo *nearby*, que puede designar cualquier ganglio cercano.

2. PATHOLOGY OF MALIGNANT MESOTHELIOMA (H0, H4)

2.1. Generic disease

SYMPTOMS (See 4.3.)	5
DISEASE	3
TOTAL	8

2.2. Cancer

CANCER	50
Cancer 49	
~ cells 10	
~ information 5	
National ~ Institute	
Cancer-related 1	

CASES (✖ 2)	1
<MALIGNANT>	1
TUMORS	1
TOTAL	53

2.2.1. Types

MALIGNANT MESOTHELIOMA	19
Malignant mesothelioma 17, Mesothelioma 2	
→ KIND (~ of cancer) (See 3.2.)	2

Kind 1, kinds 1	
→ FORM (✖6) (~ of cancer) (See 3.2.)	1
TOTAL	22

2.3. Spread of cancer (H4)

SPREAD (✖4)	7
RECURRENT (See 6.)	4
Recurrent 3 ~ malignant mesothelioma 2	
Recurred 1	

COME BACK	3
DISTANT (+ SPREAD) (See 1.1.)	1
TOTAL	15

2.4. Asbestos (risk factor)

ASBESTOS	1
----------	---

En la cadena *PATHOLOGY OF MALIGNANT MESOTHELIOMA*, se han señalado cuatro subcadenas referentes a la enfermedad, a su evolución y a los factores de riesgo. La nomenclatura de la enfermedad sólo atiende a la localización anatómica (*malignant mesothelioma*) y no aparece clasificación según el tipo de tejido donde se asienta el cáncer.

En la evolución de la enfermedad, con la excepción del término *recur* se utilizan unidades de significación de la lengua general *spread of cancer*, *come back*. No aparecen ni una sola vez los términos agrupados bajo los lemas METASTASIS y INVOLVEMENT, que son muy frecuentes en los textos sobre cáncer para especialistas.

3. RESEARCH / EXPERIMENTATION (H6)

3.1. Study

TRIALS (See 0.1.)	18
Trials 11, trial 7 Clinical trials 16	
FOUND (See 4.2.) (✖6)	11
Found 6, find 4, finds 1	
TEST (See 4.4.)	6
Test 4, tests 1, testing 1	

STUDY	4
ANSWER	2
Answer 1, answers 1	
BASED	2
LISTED	2
QUESTIONS	2
CHECK	1

DESIGNED	1
DISCOVERED	1
FIELDS	1
LABORATORY (✖6)	1
PLAN (✖6)	1

REVIEWED	1
SCIENTIFIC	1
SPECIALIZE	1
TOTAL	56

3.2. Categorization

CALLED	12
KINDS (See 0.1., 2.2.1.)	3
FORM (See 2.2.1.)	1
MEANS (✖6)	1

DESCRIPTION	1
NAMES	1
TYPE (See 0.1.)	1
TOTAL	20

3.3. Communication

DISCUSS	1
EXPLANATION (✖6)	1

3.4. Evidence

FACTS	1
SHOW	1

3.5. Cause

DEPENDS	4
Depends 3, depending 1	
CAUSES	1
PRODUCE (✖6)	1

REASONS	1
TOTAL	7

3.6. Normality

COMMON	1
NORMAL	1

La cadena 3 cuenta con gran parte del vocabulario organizador del discurso mencionado en la cadena *RESEARCH / EXPERIMENTATION* de QDT1 para especialistas, aunque no es tan técnico.

Las subcadenas tienen etiquetas parecidas. Hemos añadido las subcadenas *CAUSE*, *COMMUNICATION* y *NORMALITY* pero no hemos encontrado lexemas para formar las

subcadenas *INSTANCES*, *EVALUATION*. Explicamos esto teniendo en cuenta que la evaluación se expresa en estos textos con verbos modales.

Por último, aunque los lexemas relacionados con el concepto *INFORMATION* bien podrían venir en una subcadena dependiente de la cadena 3, por su frecuencia los hemos incluido en una cadena aparte, la cadena 6.

4. DIAGNOSIS (H1, H0, H6)

4.1. Generic diagnosis

FACTORS	1
GENERAL HEALTH	1
PROGNOSIS	1
TOTAL	3

4.2. Sensory and cognitive processes involved (H5, H6)

FOUND (✖ 6) (See 3.2.)	6
LOOK	4
SEE	2

CONSIDERED	2
DETECTION (✖ 5)	1
TOTAL	15

4.3. Clinical features (symptoms) (H0, H2)

SYMPTOMS (See 2.1.)	5
BREATH	3
Shortness of breath 1 (See 15.), breathed 1, breathing 1	
COLLECTION (+ FLUID)	3
PAIN (✖ 0)	3

FEELING	2
ACCUMULATING (+ FLUID)	1
DISCOMFORT	1
SWELLING (✖ 2)	1
TOTAL	19

4.4. Diagnostic tests and instruments

TEST (See 3.1.)	6
Test 4, tests 1, testing 1	
BIOPSIES	2
Biopsies 1, biopsy 1	
MICROSCOPE	2
PERITONEOSCOPE	2
PERITONEOSCOPY	2
THORACOSCOPE	2
THORACOSCOPY	2

SCREENING	1
X-RAY (See 0.5.1.1.)	1
TOTAL	20

Los 57 lexemas de esta cadena se agrupan en torno a las subcadenas *GENERIC DIAGNOSIS*, *SENSORY/COGNITIVE PROCESSES INVOLVED*, *CLINICAL FEATURES (SYMPTOMS)* y *DIAGNOSTIC TESTS AND INSTRUMENTS*. Son muy relevantes—más que en el texto para especialistas—las referencias a síntomas porque, al fin y al cabo, afectan muy directamente al paciente. También los nombres de diversas pruebas diagnósticas, que aparecen definidas mediante el término superordinado *test*. Se excluyen las pruebas citológicas, que quizás puedan confundir al lector

Como en el texto se pretende asumir la perspectiva del paciente, predominan las alusiones a los procesos sensoriales implicados en el diagnóstico sobre los procesos cognitivos. Los primeros pueden ser percibidos por el paciente, los segundos, no.

Los factores pronósticos no están lexicalizados en el texto para pacientes, a excepción de la UF *general health*, de ahí que no hayamos creado una subcadena. El resto de los factores se presenta en forma de paráfrasis del término empleado en la sección para especialistas:

(93)

where the cancer is → histology

how far it has spread → stage and nodal status

5. GENERAL MEDICINE (H5, H6)

CLINICAL (~ trials)(See 0.1.)	16
HOSPITAL Hospital 2, hospitals 1	3
NATIONAL CANCER INSTITUTE (✖ 6) (2) = NCI (1)	3
HEALTH CARE	2
TOTAL	24

5.1. Health professionals

DOCTOR Doctor 10, doctors 3	13
PROFESSIONALS (✖ 5) Health care ~ 2	3

Health ~ 1	
EXPERTS (✖ 5)	1
NURSES	1
SPECIALIST	1

TOTAL	19
-------	----

La extensión de la quinta cadena nos indica que en todo el evento médico se produce la interacción DOCTOR – PACIENTE en un lugar (HOSPITAL). En este texto hay más referencias a la medicina en general que en el resto porque contiene una sección que explica qué es el servicio PDQ del *National Cancer Institute* de los Estados Unidos.

6. INFORMATION (H5)

INFORMATION (cancer ~)	14
PDQ	13
LEARN	6
KNOW	2

OVERVIEW	2
SUMMARY	2
TRAINED	1
TOTAL	40

La identificación de esta cadena se explica porque la sección sobre el arriba mencionado PDQ contiene muchos lexemas relacionados con el concepto INFORMATION, que destacan la importancia de la información para hacer frente al cáncer.

7. STAGING (H6)

STAGE (✳6) Stage 8, stages 4, staging 1	13
STAGE + II + LUNG CANCER	2
STAGE + III + LUNG CANCER	2
STAGE + IV + LUNG CANCER	2
RECURRENT (See 2.3.) ~ malignant mesothelioma 2	3

ADVANCED (~ malignant mesothelioma)	2
LOCALIZED (~ malignant mesothelioma) (See 1.1.)	2
EARLY (~ stages) (See 9.)	1
TOTAL	27

La cadena 7 trata sobre la clasificación por estadios de la enfermedad, tanto la que incluye números romanos, como la que se sirve de los adjetivos *localized*, *advanced* y *recurrent*.

8. OUTCOME OF TREATMENT (H0, H3)

8.1. Positive outcome

KILL CANCER CELLS (✖ 0)	5
RELIEVE (~ symptoms)	4
BETTER	3
CURED (See 0.1.)	1
EFFECTIVENESS (✖ 3)	1
IMPROVE	1
REDUCE (~ discomfort)	1

RECOVERY	1
RESPONDS	1
SHRINK (~ tumor) (See 14.)	1
WELL	1
WORK (✖ 3)	1
TOTAL	20

8.2. Negative outcome

RISKS	1
SIDE EFFECTS	1

En la cadena 8 se ilustra la cautela con la que los textos para pacientes se aproximan a los resultados del tratamiento. En ningún caso, se quiere dar falsas expectativas de supervivencia ni se quiere mencionar la muerte, dos aspectos que eran bastante destacados en la sección para especialistas, a juzgar por sus lemas relacionados. Por eso, sólo se alude a los efectos constatables del tratamiento: alivio de síntomas, mejoría en el estado general y disminución del tamaño del tumor.

Para suplir la ausencia del lema SURVIVAL e infundir un poco de ánimo a los pacientes, en el texto, el cáncer no es el que mata sino el tratamiento, que ataca a las células cancerosas.

9. TIME (H5)

CURRENT (✖ 5) Current 4, currently 1	5
UP-TO-DATE	3
AGE	2
PAST	2
UPDATED	2
DAYS	1
EARLY (See 7.)	1

MONTH	1
ONCE	1
PERIOD	1
→ SHORT (~ period of time)	1
START (✖ 6)	1
TIME	1
TOTAL	22

De los lemas en torno al concepto TIEMPO, destacan los que denotan actualidad (CURRENT, UP-TO-DATE, UPDATED) porque continuamente se producen avances terapéuticos y es preciso actualizar el conocimiento y los tratamientos. Por lo dicho en la cadena anterior, el tiempo no se asocia a la supervivencia, como ocurría en QDT1.

El lexema *short* para indicar un periodo corto de tiempo nos llama la atención porque muestra cómo en la cognición humana hay un transvase del dominio léxico ESPACIO al dominio léxico TIEMPO, como sostiene Sweetser (1990:21): "spatial vocabulary universally acquires temporal meaning rather than the reverse".

10. PATIENT

PATIENTS	12
Patients 8	
Patients with cancer 3	
Patient 2, patient's 2	
PEOPLE	2
FAMILIES	1

PERSON	1
→ RECEIVED (TREATMENT + PATIENT + ~)	1
TOTAL	17

El concepto PACIENTE adquiere menos relevancia en los textos para pacientes que en los destinados a médicos, en parte porque el destinatario de los textos es el mismo paciente. No aparece el lema SELECTED, que coocurría con PATIENTS, porque haría pensar en una posible discriminación. La presencia de las formas léxicas *people, families* y *person* refuerza el lado humano de la enfermedad.

11. COMPOSITION

PART (See 1.1.)	9
Part 8, parts 1	
COLLECTED (See 4.3.)	3
Collected 2, collection 1	
INCLUDING	1

PARTICIPATION	1
PIECE	1
SECTIONS	1
TOTAL	16

La oposición PARTE / TODO vuelve a conectar un 2,74 % de los lexemas del texto. Se produce un solapamiento con la cadena *LOCATION IN THE HUMAN BODY* mediante el lema PART (*parts of the body*) y con la subcadena *SYMPTOMS* al coocurrir los lemas COLLECTION con FLUID.

12. IMPORTANCE

NEW	9
-----	---

SPECIAL	4
TOTAL	13

La novedad de un tratamiento proporciona relevancia al mismo. *Special* se utiliza en las paráfrasis de términos especializados. Por ejemplo, en el texto se presenta un toracoscopio como “a special instrument called a thoracoscope”.

13. HELP (H3)

SERVICE	4
HELP	3
Help 1, helpful 1, helps 1	
PROVIDE	2

SUPPORTIVE CARE	2
JOBS	1
WORKED	1
TOTAL	13

En esta cadena hay unidades léxicas provenientes de la introducción explicativa de lo que es el servicio PDQ, que hemos eliminado de QPT2 y QPT3. Apunta hacia la necesidad que tienen los pacientes de sentirse ayudados.

14. SIZE / QUANTITY

HIGH (+ ENERGY RAYS) (See 0.5.1.1.)	3
SHRINK (See 8.1.)	1
SIZE	1
REDUCE	1

LARGER	1
SHORTNESS (See 4.3.)	1
SMALL	1
THIN	1
TOTAL	10

El tamaño del tumor y las posibilidades de reducirlo es una de los principales objetivos del tratamiento. Este concepto está muy relacionado con el de QUANTITY.

15. PREVENTION

PREVENTS	4
Prevents 2, prevention 2	

Esta cadena no aparecía en el texto para profesionales de la medicina. El sustantivo se refiere a la prevención del cáncer en general, mientras que el verbo

coocurre con lexemas que designan la acumulación de fluidos en la cavidad torácica o en el peritoneo.

16. MOVEMENT

ENTERS	1
TRAVELS	1

En la lengua general y en la medicina, el cuerpo se entiende como un contenedor en el que se introducen alimentos, medicamentos, etc. y también como un espacio físico con accidentes geográficos. Sirvan como ejemplo algunas partes de la anatomía humana: el torrente sanguíneo, el monte de Venus, el puente, en el tronco del encéfalo, el valle y el acueducto de Silvio¹⁵⁶. No es de extrañar pues encontrar verbos de movimiento. En el texto, se describe el movimiento de los agentes quimioterapéuticos cuando entran en el cuerpo y viajan por el torrente sanguíneo (*bloodstream*), que es una metáfora de la sangre.

(94) Chemotherapy is called a systemic treatment because the drug **enters** the bloodstream, **travels** through the body, and can kill cancer cells throughout the body.

11.2.1.5. CUADRO CONTRASTIVO (*HESPERUS* – ANÁLISIS COMBINADO)

Presentamos algunos datos cuantitativos del texto que han servido para comentar algunos de los rasgos cohesivos de los textos para pacientes. Con posterioridad, se relacionarán con los de otros textos con la esperanza de encontrar tendencias estadísticas que informen acerca de la cohesión de un texto.

De las 1621 palabras del texto, ha sido posible vincular 749 palabras plenas, con lo que podríamos sostener que el 46,21 % de las palabras del texto contribuyen a la cohesión léxica. Teniendo en cuenta los datos de *Hesperus* este porcentaje sería de 43,43 %.

Hay un promedio de 17 palabras por oración, 8 de las cuales están vinculadas. Con 749 ocurrencias (*tokens*) de 252 formas léxicas (*word type*) y 218 lemas se consiguen activar 109 categorías conceptuales. De ahí inferimos que en 100 palabras

¹⁵⁶ Ejemplos tomados del suplemento del número 222 de la revista *Muy Interesante* (noviembre 1999), titulado "¡Vaya cuerpo!: Las curiosidades, récords, anécdotas y cifras más insólitas y divertidas del cuerpo humano". Para una aproximación más exhaustiva a las metáforas en la medicina, véase Van Rijn-Van Tongeren (1997).

vinculadas, encontraríamos 34 variantes léxicas de 29 lemas que activarían 14 categorías del tesoro.

TABLA 98: Descripción cuantitativa de las cadenas léxicas reconocidas.

CADENAS LÉXICAS RECONOCIDAS POR <i>HESPERUS</i>		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 1: TREATMENT AND DIAGNOSIS, HOSPITAL [74-33-19,98] H 3: (TREATMENT) [38-4-8,48] H 4: THERAPY, DIVERGENCE [28-6-6,36] Lexemas vinculados 140 Formas distintas 43 Activación conceptual ajustada (%) 40,95		0. TREATMENT [182-58- 23,33] 4. DIAGNOSIS [57-29-7,31] Lexemas vinculados 239 Formas distintas 86 Activación conceptual (%) 30,64	
H 2: BODY PARTS [54-11-9,42] H 6: ? (BODY) [18-2-0,45] REGION (0,96), DISTANT (0,86), SPACE (0,44), MIDDLE (0,97)= 3,23 Lexemas vinculados 72 Formas distintas 13 Activación conceptual ajustada (%) 15,41		1. LOCATION THE HUMAN BODY Lexemas vinculados 125 Formas distintas 38 Activación conceptual (%) 16,03	
H 0: DISEASE, DEATH, HARM, BADNESS [20,75] Lexemas vinculados 81 Formas distintas 5 Activación conceptual ajustada (%) 24,41		2. PATHOLOGY OF MALIGNANT MESOTHELIOMA Lexemas vinculados 99 Formas distintas 17 Activación conceptual (%) 12,69	
H 6: ? (CAUSE) [1,51] EXPERIMENT (1,82), TESTING-AGENT (0,55), TESTEE (0,5), QUESTION (0,66), ENQUIRY (0,38), ENQUIRE (0,19)= 4,1 Lexemas vinculados 21 Formas distintas 14 Activación conceptual ajustada (%) 6,49		3. RESEARCH / EXPERIMENTATION Lexemas vinculados 89 Formas distintas 41 Activación conceptual (%) 11,41	
H 5: KNOWLEDGE, INSTRUCTION [4,91] Lexemas vinculados 32 Formas distintas 9 Activación conceptual ajustada (%) 5,77		5. GENERAL MEDICINE [43-12-5,51] 6. INFORMATION [40-7-5,13] Lexemas vinculados 83 Formas distintas 19 Activación conceptual (%) 10,64	
H 6: ? (STAGE) [2,22] Lexemas vinculados 12 Formas distintas 2 Activación conceptual ajustada (%) 3,38		7. STAGING Lexemas vinculados 27 Formas distintas 10 Activación conceptual (%) 3,46	
H 3: (AGENCY) [0,49] Lexemas vinculados 9 Formas distintas 6 Activación conceptual ajustada (%) 0,58		8. OUTCOME OF TREATMENT Lexemas vinculados 22 Formas distintas 14 Activación conceptual (%) 2,82	

SEPARATE (1,75), COMPONENT (0,69), WHOLE (0,29), SEPARATION (0,25), CONTENTS (0,22) = 3,2 Activación conceptual ajustada (%)	3,77	2,05	11. COMPOSITION [16-8-2,05] Activación conceptual (%)
		81 41 10,24	9. TIME [22-14-2,82] 10. PATIENT [17-5-2,18] 12. IMPORTANCE [13-2-1,66] 13. HELP [13-8-1,66] 14. SIZE / QUANTITY [10-8-1,15] 15. PREVENTION [4-2-0,51] 16. MOVEMENT [2-2-0,26] Lexemas vinculados Formas distintas Activación conceptual (%)
TOTAL Activación conceptual según Hesperus Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	85,01 367 92 99,99	781 749 252 99,98	TOTAL Suma de elementos en las cadenas Lexemas vinculados ¹⁵⁷ Formas distintas ¹⁵⁸ Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QPT1 Nº total de palabras: 1621 Palabras que crean cohesión (<i>Hesperus</i>): 704		Categorías conceptuales del tesoro: 109 Coeficiente de cohesión (<i>Hesperus</i>): 10,7 Nº de oraciones: 92	

¹⁵⁷ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹⁵⁸ De estas, hay 22 formas léxicas que aparecen en más de una cadena.

11.2.2. Non-small cell lung cancer (QPT2)

11.2.2.1. ÍNDICE DEL TEXTO

La repetición de ciertas USE en el índice nos anticipa los conceptos que van a adquirir un importante peso en el texto 13 del Apéndice II.

(95)

DESCRIPTION

What is **non-small cell lung cancer**?

STAGE EXPLANATION

Stages of non-small cell lung cancer

Occult stage

Stage 0

Stage I

Stage II

Stage III

Stage IV

Recurrent

TREATMENT OPTION OVERVIEW

How non-small cell lung cancer is treated

OCCULT NON-SMALL CELL LUNG CANCER

STAGE 0 NON-SMALL CELL LUNG CANCER

STAGE I NON-SMALL CELL LUNG CANCER

STAGE II NON-SMALL CELL LUNG CANCER

STAGE III NON-SMALL CELL LUNG CANCER

STAGE IV NON-SMALL CELL LUNG CANCER

RECURRENT NON-SMALL CELL LUNG CANCER

TO LEARN MORE

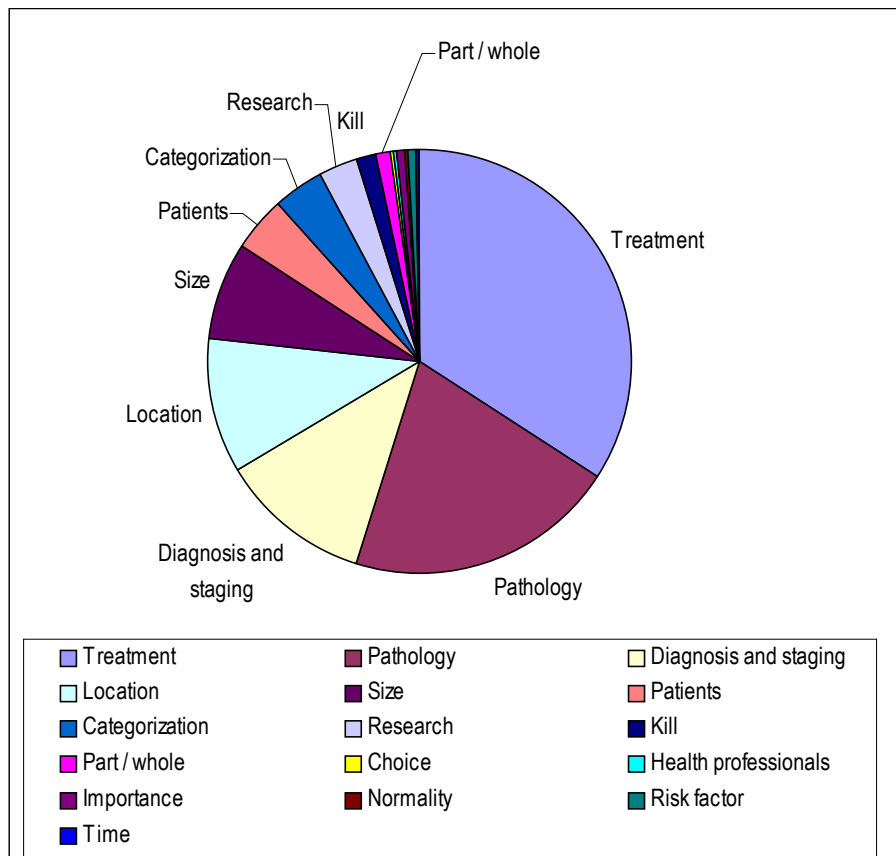
11.2.2.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTA DE FRECUENCIA LEMATIZADA

Algunos de los conceptos citados en 11.2.2.1. quedan representados en la tabla y el gráfico que hemos derivado de la lista de frecuencia lematizada elaborada con *Wordsmith Tools*.

TABLA 99: Identificación preliminar de áreas conceptuales activadas en QPT2 a partir de lista de frecuencia lematizada

PT2	
ÁREAS CONCEPTUALES	%
Treatment	13,93
Pathology	8,43
Diagnosis and staging	4,7
Location	4,16
Size	3,1
Patients	1,75
Categorization	1,52
Research	1,27
Kill	0,56
Part / whole	0,4
Choice	0,16
Health professionals	0,16
Importance	0,16
Normality	0,16
Risk factor	0,16
Time	0,16
	40,8

DIAGRAMA 30: Representación preliminar de las áreas conceptuales activadas en QPT2



11.2.2.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

Hesperus ha reconocido siete cadenas (véase apéndice VIIe). El perfil del documento indica que, de las 1258 formas léxicas del texto, 553 son plenas. La vinculación entre estas se establece mediante 80 categorías del tesoro. El cociente de dividir el valor de cohesión otorgado por *Hesperus* (17424) y el número total de palabras es de 13,85.

Las categorías conceptuales activadas que contribuyen a más de un 0,5% de la cohesión del texto son:

therapy_658_4268_n	Percent: 32.72,
minuteness_196_1271_n	Percent: 7.72,
respiratory-disease_651_4192_n	Percent: 6.07,
degree_27_165_n	Percent: 5.45,
cancer_651_4195_n	Percent: 5.18,
disease_651_4187_n	Percent: 4.48,
surgery_658_4267_n	Percent: 3.91,
tobacco_388_2560_n	Percent: 2.87,
photography_551_3582_n	Percent: 2.54,
literal_558_3619_a	Percent: 2.14,
small_33_224_a	Percent: 2.08,
medical_658_4274_a	Percent: 1.99,
arrange_62_430_v	Percent: 1.81,
refrigerate_382_2543_v	Percent: 1.72,
star_321_2161_n	Percent: 1.72,
space_183_1154_n	Percent: 1.46,
grow_36_253_v	Percent: 1.46,
sick-person_651_4202_n	Percent: 1.23,
doctor_658_4276_v	Percent: 1.08,
indication_547_3541_n	Percent: 1.06,
pathology_651_4203_n	Percent: 0.91,
prohibit_757_4900_v	Percent: 0.83,
call_547_3548_n	Percent: 0.80,
opportunity_137_857_n	Percent: 0.72,
produce_164_1056_v	Percent: 0.72,
be-distant_199_1299_v	Percent: 0.70,
experiment_461_2959_n	Percent: 0.56,
initiate_68_463_v	Percent: 0.52,

Estas etiquetas semánticas están presentes en las 7 cadenas reconocidas por *Hesperus*, que comentamos brevemente:

TABLA 100: **Cadena 0** (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 0 (<i>Learn</i>)	
LEARN	0 0,1
TOTAL	
↓	
LEARN	

Comentario

El concepto LEARN está implícito en una página de información para pacientes. Antes de modificar el texto para su análisis¹⁵⁹, los conceptos LEARNING e INFORMATION eran más representativos.

TABLA 101: **Cadena 1** (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 1 (<i>Treated</i>)	
THERAPY	32,72
SURGERY	3,91
MEDICAL	1,99
DOCTOR	1,42
DIAGNOSIS	0,34
TOTAL	40,38
↓	
TREATMENT AND DIAGNOSIS	

Comentario

En esta cadena, que es la más importante del texto, a juzgar por su elevado porcentaje, *Hesperus* pone en contacto diversos aspectos de la práctica de la medicina. Destaca el concepto THERAPY (*therapy, treatment*) y, de entre las distintas formas de terapia, la quimioterapia y la cirugía, a la que *Hesperus* atribuye un porcentaje relativamente alto (3,91). También hay una leve representación del diagnóstico, que se actualiza mediante las formas léxicas *tests* y *symptoms*.

¹⁵⁹ Nos referimos a los párrafos iniciales que describían la función del *Physician Data Query*. Sólo hemos dejado este párrafo introductorio en el texto QPT1.

TABLA 102: Cadena 2 (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 2 (<i>Lung cancer</i>)	
RESPIRATORY-DISEASE	6,07
CANCER	5,18
DISEASE	4,48
SICK-PERSON	1,23
PATHOLOGY	0,91
ILLNESS	0,43
	00,1
TOTAL	+ 18,3
↓	
DISEASE, PATHOLOGY	

Comentario

Con más de un 18%, esta cadena se perfila como la segunda más importante. Aglutina conceptos relacionados con la patología del cáncer. Incluye las unidades léxicas *diagnosis* y *prognosis*, que ubicaremos en la cadena *DIAGNOSIS*.

TABLA 103: Cadena 3 (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 3 (<i>Small</i>)	
MINUTENESS	7,72
DEGREE	5,45
LITERAL	2,14
SMALL	2,08
GROW	1,46
TOTAL	18,85
↓	
SIZE, MINUTENESS	

Comentario

Lo elevado del porcentaje que atribuye *Hesperus* a esta cadena se debe a que *small* está presente en todas las referencias al cáncer no microcítico de pulmón. Dado que en la enfermedad es importante el tamaño y la extensión del tumor, tienen también un porcentaje significativo los conceptos *DEGREE* y *GROW*. Por último, como indica el Apéndice VIIe, *Hesperus* incluye en la tercera cadena la unidad léxica *neck* que, aunque

nombre a una entidad de reducida anchura, debería aparecer en una cadena que nombre las partes del cuerpo donde se puede localizar la enfermedad.

TABLA 104: **Cadena 4** (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 4 (Places)	
SPACE	1,46
BE-DISTANT	0,7
TOTAL	2,16
↓	
LOCATION	

Comentario

Muy difusamente se perfila en esta cadena la idea de localización y distancia, factores muy decisivos en el pronóstico del cáncer de pulmón.

TABLA 105: **Cadena 5** (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 5 (Sputum)	
BLOOD	0,26
FLUID	0,15
TOTAL	0,41
↓	
FLUID	

Comentario

En esta cadena se agrupan lexemas que nombran distintos fluidos corporales.

TABLA 106: **Cadena 6** (QPT2)

CONCEPTOS ACTIVADOS	
Cadena 6 (Smoke)	
TOBACCO	2,87
TOTAL	2,87
↓	
TOBACCO	

Comentario

En esta cadena, se atribuye un 2,87% al principal factor de riesgo del cáncer de pulmón.

11.2.2.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

El perfil del documento nos indica que existen conceptos que, aunque no han sido identificados en las cadenas, tienen relevancia conceptual. Entre estos, destacamos:

- ARRANGE (1,81%), EXPERIMENT (0,56%), BASIS (0,17%), que bien podrían formar parte de una cadena *RESEARCH / EXPERIMENTATION*
- BOSOM (0,3%), que hace referencia al cuerpo humano
- PRODUCE (0,72%), KILL (0,48%) y PRODUCED (0,1%), que son los conceptos subyacentes a palabras que en QPT1 hemos incluido bajo la cadena *OUTCOME OF TREATMENT*.
- INITIATE (0,52%) y EARLY (0,26%), relacionados con el concepto TIME

TABLA 107: Cadenas léxicas en QPT2

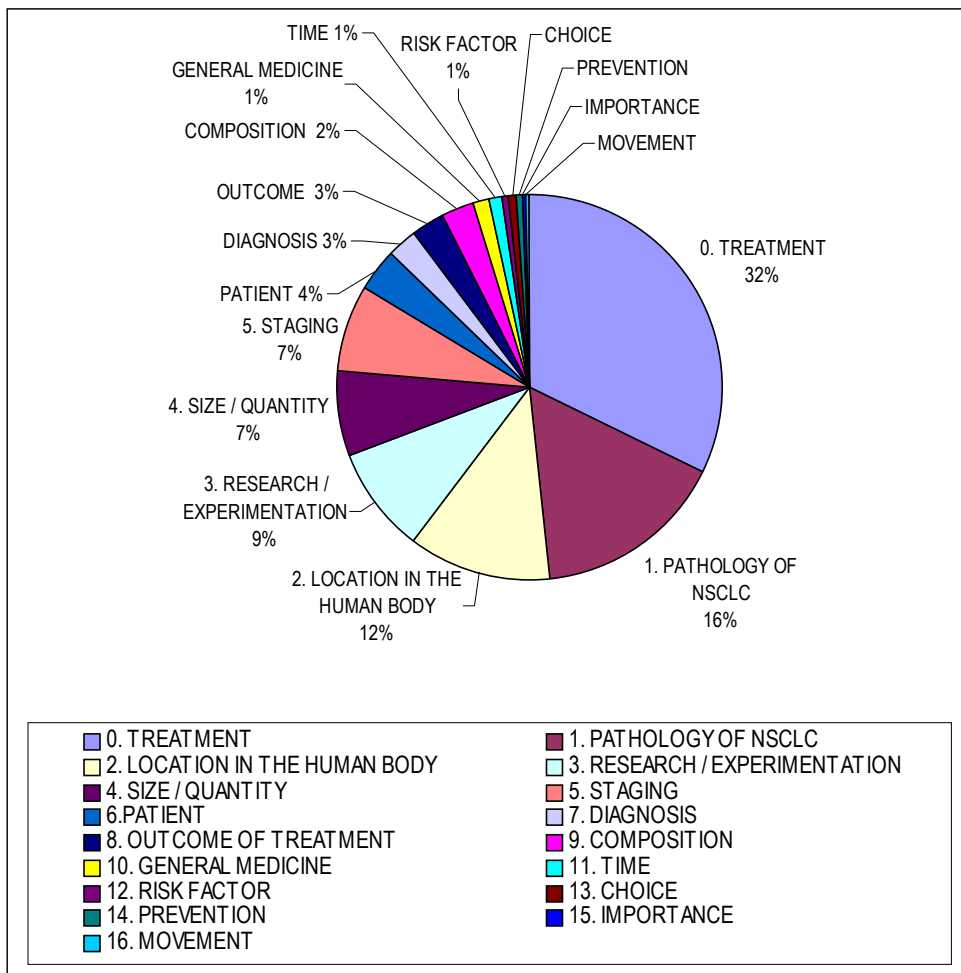
<p>Chain 0: TREATMENT (H1) [200]</p> <p>0.1. Generic treatment</p> <p>0.2. Properties (↔ TREATMENT)</p> <p> 0.2.1. Location where TREATMENT is applied</p> <p>0.3. Manipulation</p> <p>0.4. Process</p> <p>0.5. Types</p> <p> 0.5.1. Radiation therapy</p> <p> 0.5.1.1. Radiation agents</p> <p> 0.5.2. Surgery</p> <p> 0.5.2.1. New techniques</p> <p> 0.5.3. Chemotherapy</p> <p> 0.5.3.1. Drugs</p> <p> 0.5.4. Other types of treatment</p> <p>0.6. Instruments</p> <p>Chain 1: PATHOLOGY OF NON-SMALL LUNG CANCER (H2) [100]</p> <p>1.1. Generic disease</p> <p>1.2. Cancer</p> <p> 1.2.1. Types</p> <p> 1.2.1.1. Anatomic classification</p> <p> 1.2.1.2. Histologic classification</p> <p>1.3. Spread of cancer</p> <p>Chain 2: LOCATION IN THE HUMAN BODY (H4, H5) [73]</p> <p>2.1. Generic location</p> <p>2.2. Microscopic / histologic level</p> <p>2.3. Anatomic level / body parts</p> <p> 2.3.1. Respiratory system</p> <p>2.4. Body fluids</p>	<p>Chain 3: RESEARCH / EXPERIMENTATION (H0) [55]</p> <p>3.1. Study</p> <p>3.2. Categorization</p> <p>3.3. Cognition</p> <p>3.4. Cause</p> <p>3.5. Normality</p> <p>3.6. Communication</p> <p>Chain 4: SIZE / QUANTITY (H3) [46]</p> <p>4.1. Size</p> <p>4.2. Name of lung cancer</p> <p>4.3. Increase</p> <p>4.4. Large quantity</p> <p>4.5. Minuteness</p> <p>Chain 5: STAGING [44]</p> <p>Chain 6: PATIENT (H2) [22]</p> <p>Chain 7: DIAGNOSIS (H2, H1) [17]</p> <p>Chain 8: OUTCOME OF TREATMENT [17]</p> <p>Chain 9: COMPOSITION [16]</p> <p>Chain 10: GENERAL MEDICINE (H1) [9]</p> <p>10.1. Health professionals</p> <p>Chain 11: TIME [7]</p> <p>Chain 12: RISK FACTOR (H6) [4]</p> <p>Chain 13: CHOICE [3]</p> <p>Chain 14: PREVENTION [3]</p> <p>Chain 15: IMPORTANCE [2]</p> <p>Chain 16: MOVEMENT [2]</p> <p>TOTAL DE UNIDADES LÉXICAS ENCADENADAS ¹⁶⁰ : [562]</p>
---	---

¹⁶⁰ Para obtener el total no se han duplicado los lemas que pertenecen a más de una cadena. La suma de la frecuencia absoluta de palabras en cada una de las cadenas es de 620.

A la vista del gráfico queda patente de nuevo el énfasis conceptual en el TRATAMIENTO, que en este texto adquiere un porcentaje de 32,57 %, muy superior con respecto a los demás textos. Esto determinará que al contrastarlas con otros textos, la representación conceptual de otras cadenas (*LOCALIZATION*, *RESEARCH/EXPERIMENTATION*) sea menor. La cadena *PATHOLOGY* también adquiere relevancia por la abundancia de unidades de significación especializada que designan la enfermedad y la capacidad del cáncer para reaparecer.

El concepto *LOCALIZATION IN THE HUMAN BODY* hila la tercera cadena y no la segunda, como ocurre en los otros dos textos. La cuarta cadena en importancia vuelve a ser la de *RESEARCH / EXPERIMENTATION*. Por último, con un porcentaje del 7% cada una, encontramos las cadenas *SIZE / QUANTITY* y *STAGING*.

DIAGRAMA 31: Porcentaje de activación de las cadenas léxicas definitivas de QPT2



CADENAS LÉXICAS	Frec	%	CADENAS LÉXICAS	Frec	%
0. TREATMENT	200	32,26	9. COMPOSITION	16	2,58
1. PATHOLOGY OF NSCLC	100	16,13	10. GENERAL MEDICINE	9	1,45
2. LOCATION IN THE HUMAN BODY	73	11,77	11. TIME	7	1,13
3. RESEARCH / EXPERIMENTATION	55	8,87	12. RISK FACTOR	4	0,65
4. SIZE / QUANTITY	46	7,42	13. CHOICE	3	0,48
5. STAGING	44	7,1	14. PREVENTION	3	0,48
6. PATIENT	22	3,35	15. IMPORTANCE	2	0,32
7. DIAGNOSIS	17	2,74	16. MOVEMENT	2	0,32
8. OUTCOME OF TREATMENT	17	2,74	TOTAL	620	99,99

11.2.2.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS

0. TREATMENT (H1)

0.1. Generic treatment

TREATMENT	27
Treatment 17 Cancer ~ 5 Treated 5, treat 3, treatments 2	
CLINICAL TRIALS	6
<THERAPY> (See 0.5.1., 0.5.4.)	5
CURE (See 8.) Cure 2, cured 1	3
→ CHOICE (+ TREATMENT+)	2

(See 13.) Choice 1, choices 1	
CONTROL (+ SYMPTOMS)	1
→ OPTION (Treatment ~) (See 13.)	1
TOTAL	45

0.2. Properties (→ TREATMENT)

→ ALONE (THERAPY ~)	6
SYSTEMIC → (~ treatment)	1
TOTAL	7

0.2.1. Location where TREATMENT is applied

INTERNAL → (~ radiation therapy) (See 2.1.)	4
EXTERNAL → (~ radiation therapy) (See 2.1.)	2

0.3. Manipulation (H1)

USED	17
Used 11, uses 5, using 1	
TAKEN OUT	7
Taken out 4, take out 1, takes out 1, taking out 1	
OPERATED	5
Operated 5, operation 3	
DONE	3

PUT (✖ 1)	2
Put 1, putting 1	
FOCUSED (✖ 1) (~ radiation)	1
REMOVE	1
TAKEN	1
TOTAL	37

0.4. Process

<u>PLAN</u> (+ TREATMENT) Plan 1, planned 1	2
<u>DOSE</u>	1

0.5. Types

0.5.1. Radiation therapy

RADIATION THERAPY Internal ~ 4 External ~ 2	25
<RADIATION>	5

RADIOTHERAPY	1
TOTAL	31

0.5.1.1. Radiation agents (H3)

LIGHT (✖ 3)	2
RADIOISOTOPES ≈ Materials that produce radiation	2
X-RAYS	2
BEAM (~ of light)	1
→ <u>HIGH-DOSE</u> (~ x-rays)	1

(See 4.4.)	
→ <u>HIGH-ENERGY</u> (~ x-rays) (See 4.4.)	1
TOTAL	9

0.5.2. Surgery

SURGERY	16
LOBECTOMY	1
PNEUMONECTOMY	1

WEDGE RESECTION	1
TOTAL	19

0.5.2.1. New techniques

RADIOSURGERY	4
CRYOSURGERY	2

0.5.3. Chemotherapy

CHEMOTHERAPY	18
--------------	----

0.5.3.1. Drugs

DRUGS	3
Drugs 2, drug 1	
PILL	1
TOTAL	4

0.5.4. Other types of treatment

LASER THERAPY	4
PHOTODYNAMIC THERAPY	4
CHEMOPREVENTION (See 14.)	2

CRYOTHERAPY	1
TOTAL	11

0.6. Instruments

TUBES	2
NEEDLE	1
MACHINE	1

TOTAL	4
-------	---

1. PATHOLOGY OF NON-SMALL CELL LUNG CANCER (H2)

1.1. Generic disease

DISEASE	2
SYMPTOMS (See 7.)	2
GET (+ SECOND CANCER)	1

OCCURRING (+ SECOND CANCER)	1
TOTAL	6

1.2. Cancer¹⁶¹

<CANCER>	12
Cancer 9, cancers 3	
CANCER CELLS	9
TUMOR	10

→ GROW (See 4.3.)	2
TOTAL	33

¹⁶¹ El lema CANCER (*cancer, cancers*) aparece en 58 ocasiones, aunque en la mayoría de los casos constituye UF. La UF más frecuente es *lung cancer*, que se repite en 33 ocasiones y que constituye las 21 ocurrencias de la UF *non-small cell lung cancer*.

1.2.1. Types

→ TYPE (~ of lung cancer) (See 3.2.)	2
→ KINDS (~ of non-small cell lung cancer) (See 3.2.)	1

1.2.1.1. Anatomic classification

<LUNG CANCER> Lung cancer 9, lung cancers 1	10
--	----

1.2.1.2. Histologic classification

NON-SMALL CELL LUNG CANCER	21
EPIDERMOID CARCINOMA (1) = Squamous cell carcinoma (1)	2
ADENOCARCINOMA	1
ADENOSQUAMOUS CARCINOMA	1
LARGE CELL CARCINOMA	1
UNDIFFERENTIATED CARCINOMA	1
SMALL CELL LUNG CANCER	2
TOTAL	29

1.3. Spread of cancer

SPREAD (See 4.3.)	10
RECURRENT (See 5.) Recurrent 2, recurred 1	3
SECOND	3

INVOLVE	2
COME BACK	1
TOTAL	19

2. LOCATION IN THE HUMAN BODY

2.1. Generic location (**H4**)

BODY	6
INTERNAL (See 0.2.1.) Internal 4, internally 2	6
AREA	4
NEARBY	3
PARTS OF THE BODY (See 9.)	3
LOCAL	2

EXTERNAL (See 0.2.1.)	1
PLACES	1
SIDE (✖ 3)	1
TOP (✖ 0)	1
TOTAL	30

2.2. Microscopic / histologic level

CELL	12
Cell 25	
SIZE + ~ + LUNG CANCER	
Cells 12	
Cancer ~ 9	

TISSUE	3
LAYERS (~ of cells)	1
TOTAL	16

2.3. Anatomic level / body parts

LYMPH NODES	5
BRAIN	2
DIAPHRAGM	1
MEDIASTINUM	1
MUSCLE	1

NECK (✖ 3)	1
VEIN	1
TOTAL	12

2.3.1. Respiratory system

<LUNG ¹⁶² >	8
Lung 6, lungs 2	
LUNG CANCER 34	
<CHEST>	2
CHEST WALL	1
LINING (~ of the lung)	1
LOBE	1
TOTAL	13

2.4. Body fluids (H5)

BLOODSTREAM	1
SPUTUM	1

3. RESEARCH / EXPERIMENTATION

3.1. Study

FOUND (See 7.)	10
Found 8, find 1, find out 1	
CLINICAL TRIALS (See 0.1., 10.)	6

TESTS (See 7.)	2
OVERVIEW	1
TOTAL	19

¹⁶² Quedan excluidas las ocurrencias de *lung* dentro de UF. El lema LUNG aparece en 43 ocasiones.

3.2. Categorization

CALLED	7
TYPE (See 1.2.1.) Type 6, types 1	7
DIVIDED	3
DIFFERENT Different 1, differently 1	2

DESCRIPTION	1
KINDS (See 1.2.1.)	1
TERM	1
TOTAL	22

3.3. Cognition (H0)

INFORMATION	1
KNOW	1
LEARN	1
THINK	1

SUMMARY	1
TOTAL	5

3.4. Cause

DEPEND Depend 1, depends 1	2
ASSOCIATED	1
PRODUCE	1
TOTAL	4

3.5. Normality

NORMAL (~ tissue)	2
COMMON	1
TOTAL	3

3.6. Communication

DISCUSSED (✖0)	1
EXPLANATION	1

4. SIZE / QUANTITY (H3)

4.1. Size

SIZE	1
------	---

4.2. Name of lung cancer

NON-SMALL (~ cell lung cancer)	21
SMALL (~ cell lung cancer)	2
LARGE (~ cell carcinoma)	1
TOTAL	24

4.3. Increase

SPREAD (See 1.3.)	10
GROW (See 1.2.)	2
TOTAL	12

4.4. Large quantity

AMOUNT	1
HIGH-DOSE (~ x-rays) (See 0.5.1.1.)	1
HIGH-ENERGY (~ x-rays) (See 0.5.1.1.)	1

4.5. Minuteness

<SMALL>	2
LITTLE	1
NARROW	1
SHRINK (✖ 1) (See 8.)	1

THIN	1
TOTAL	6

5. STAGING

STAGE	26
Stage 21, stages 4, staging 1	
STAGE + II + (LUNG CANCER)	3
STAGE + III (+ LUNG CANCER)	3
STAGE + IIIA (+ LUNG CANCER)	2
STAGE + IIIB (+ LUNG CANCER)	2
STAGE + IV (+ LUNG CANCER)	2
EARLY (+ CANCER / STAGE)	2

RECURRENT (+ CANCER) (See 1.3.)	2
OCCULT (~ non-small cell lung cancer)	1
CARCINOMA IN SITU	1
TOTAL	44

6. PATIENT (H2)

PATIENTS (✖ 2)	16
Patients 13, patient 3	
GROUP (+ PATIENT) (See 9.)	6
TOTAL	22

7. DIAGNOSIS (H2, H1)

FIND (See 3.1.)	9
Found 8, find 1	
FIND OUT	1
TESTS (See 3.1.)	2
DIAGNOSIS (✖ 2)	1
PAIN (✖ 2)	1

PROGNOSIS (✖ 2)	1
GENERAL HEALTH (✖ 1)	1
MEDICAL PROBLEMS (✖ 1)	1
TOTAL	17

8. OUTCOME OF TREATMENT

KILL	7
Kill cancer cells 6	
Kill (+ TUMOR) 1	
CURE (See 0.1.)	3
Cure 2, cured 1	
FREEZES (~ the tumor)	1
IMPROVE	1
PREVENT (~ cancer)	1

RECOVERY	1
RELIEVE (~ pain)	1
SHRINK (~ the cancer)(✖ 1) (See 4.5.)	1
WORKS	1
TOTAL	17

9. COMPOSITION

GROUP (See 6.)	6
PARTS (See 2.1.)	4
Parts of the body 3, part 1	
WHOLE	2
SEPARATE	2

INCLUDES	1
SECTION	1
TOTAL	16

10. GENERAL MEDICINE

CLINICAL (~ trials) (See 0.1., 3.1.)	6
MEDICAL	1
TOTAL	7

10.1. Health professionals (H1)

DOCTOR (✖1)	2
-------------	---

11. TIME

EARLY	2
FOLLOWING	2
ONCE	1
PREVIOUS	1

PRIOR	1
TOTAL	7

12. RISK FACTOR (H6)

SMOKING	2
→ PASSIVE (~ smoking)	1
RADON EXPOSURE	1
TOTAL	4

13. CHOICE

CHOICE (See 0.1.) Choice 1, choices 1	2
OPTION (See 0.1.)	1

14. PREVENTION

CHEMOPREVENTION (See 0.5.4.)	2
PREVENT	1

15. IMPORTANCE

MAIN	2
NEW	1

16. MOVEMENT

ENTERS	1
--------	---

TRAVELS	1
---------	---

11.2.2.5. CUADRO CONTRASTIVO (*HESPERUS* – ANÁLISIS COMBINADO)

Presentamos ahora algunos datos estadísticos que ilustran las diferencias entre la cohesión percibida por métodos informáticos y la que detecta el investigador con la ayuda de estos.

En este texto, que tiene una extensión de 1258 palabras, ha sido posible vincular 562 instancias de 207 variantes morfológicas que crean cohesión. Estas variantes están asociadas a 173 lemas y activan 80 categorías conceptuales. Por tanto, el 44,67 % de las ocurrencias léxicas del texto crean cohesión o, ciñéndonos a los datos de *Hesperus*, el 43,95%. Esta proporción se mantiene dentro de la oración, donde un promedio de 6 palabras contribuyen a la cohesión léxica, siendo la media de palabras por oración la de 13,2 palabras.

Asimismo, de cada 100 palabras vinculadas, encontramos 37 variantes correspondientes a 31 lemas, que activarían 14 categorías conceptuales. Por último, hemos señalado con hipervínculos 58 palabras que pertenecen a más de una cadena.

TABLA 108: Descripción cuantitativa de las cadenas léxicas reconocidas en QPT2

CADENAS LÉXICAS RECONOCIDAS POR <i>HESPERUS</i>		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 1: TREATMENT AND DIAGNOSIS [40,38]			0. TREATMENT [200-60-32,26] 7. DIAGNOSIS [17-9-2,74] 10. GENERAL MEDICINE [9-3-1,45]
Lexemas vinculados	130	226	Lexemas vinculados
Formas distintas	27	72	Formas distintas
Activación conceptual ajustada (%)	45,87	36,45	Activación conceptual (%)
H 2: DISEASE, PATHOLOGY			1. PATHOLOGY OF NSCLC
Lexemas vinculados	85	100	Lexemas vinculados
Formas distintas	10	26	Formas distintas
Activación conceptual ajustada (%)	20,48	16,13	Activación conceptual (%)
H 4: LOCATION [5-2-2,16]			2. LOCATION IN THE HUMAN BODY
H 5: FLUID [6-2-0,41]			
BOSOM [0,3]			
Lexemas vinculados	11	73	Lexemas vinculados
Formas distintas	4	30	Formas distintas
Activación conceptual ajustada (%)	3,26	11,77	Activación conceptual (%)

H 0: LEARN [0,1] ARRANGE [1,81], EXPERIMENT [0,56], BASIS [0,17] Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	1 1 3	55 26 8,87	3. RESEARCH / EXPERIMENTATION Lexemas vinculados Formas distintas Activación conceptual (%)
H 3: SIZE, MINUTENESS [18,85] Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	70 11 21,41	46 14 7,42	4. SIZE / QUANTITY Lexemas vinculados Formas distintas Activación conceptual (%)
PRODUCE [0,72], KILL [0,48], PRODUCED [0,41] Activación conceptual ajustada (%)	1,83	2,77	8. OUTCOME OF TREATMENT [17-10- 2,74] Activación conceptual (%)
INITIATE [0,52], EARLY [0,26] Activación conceptual ajustada (%)	0,88	1,13	11. TIME [7-5- 1,13] Activación conceptual (%)
H 6: SMOKING [2,87] Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	2 1 3,26	4 3 0,65	12. RISK FACTOR Lexemas vinculados Formas distintas Activación conceptual (%)
		92 31 14,83	5. STAGING [44-12-7,1] 6. PATIENT [22-3 -3,55] 9. COMPOSITION [16-7-2,58] 13. CHOICE [3-3-0,48] 14. PREVENTION [3- 2 -0,48] 15. IMPORTANCE [2-2-0,32] 16. MOVEMENT [2-2 -0,32] Lexemas vinculados Formas distintas Activación conceptual (%)
TOTAL Activación conceptual según Hesperus Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	88,03 299 54 99,99	620 562 207 99,99	TOTAL Suma de elementos en las cadenas Lexemas vinculados ¹⁶³ Formas distintas ¹⁶⁴ Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QPT2 Nº total de palabras: 1258 Palabras que crean cohesión (<i>Hesperus</i>): 553		Categorías conceptuales del tesauro: 80 Coeficiente de cohesión (<i>Hesperus</i>): 13,85 Nº de oraciones: 85	

¹⁶³ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹⁶⁴ De estas, hay 10 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 58 lexemas adicionales a los 562 lexemas vinculados.

11.2.3. Small cell lung cancer (QPT3)

11.2.3.1. ÍNDICE DEL TEXTO

El índice adelanta que los conceptos marcados en negrita van a ser centrales en el texto.

(96)
 DESCRIPTION
 What is **small cell lung cancer**?
 STAGE EXPLANATION
 Stages of small cell lung cancer
Limited stage
Extensive stage
Recurrent stage
 TREATMENT OPTION OVERVIEW
 How small cell lung cancer is treated
 Treatment by stage
 LIMITED STAGE SMALL CELL LUNG CANCER
 EXTENSIVE STAGE SMALL CELL LUNG CANCER
 RECURRENT SMALL CELL LUNG CANCER
 TO LEARN MORE

11.2.3.2. ÁREAS CONCEPTUALES RECONOCIDAS A PARTIR DE LISTA DE FRECUENCIA LEMATIZADA

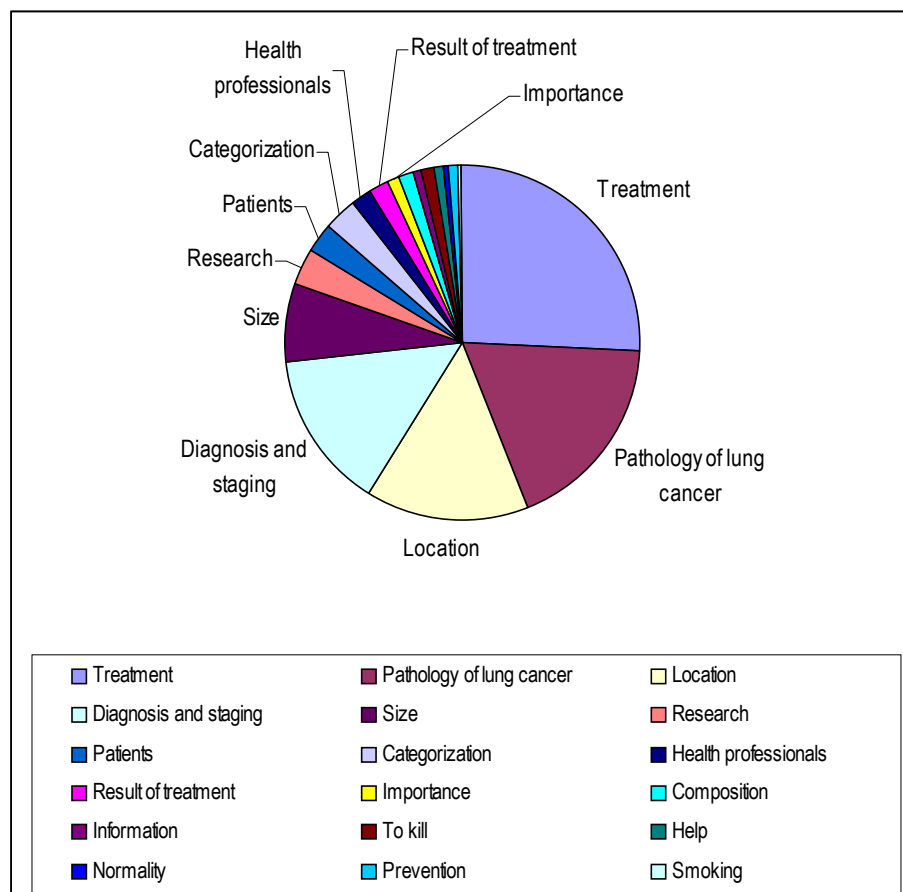
Algunos de los conceptos a los que apunta el índice quedan representados en la tabla y el gráfico que hemos derivado de la lista de frecuencia lematizada elaborada con *Wordsmith Tools*.

TABLA 109: Identificación preliminar de las áreas conceptuales activadas en QPT3 a partir de lista de frecuencia lematizada

PT3	
Treatment	10,06
Pathology of lung cancer	6,94
Location	5,76
Diagnosis and staging	5,6
Size	2,73
Research	1,36
Patients	1,04
Categorization	1,2
Health professionals	0,72
Result of treatment	0,64
Importance	0,48

Composition	0,48
Information	0,4
To kill	0,4
Help	0,32
Normality	0,24
Prevention	0,24
Smoking	0,16
TOTAL	38,77

DIAGRAMA 32: Representación preliminar de las áreas conceptuales activadas en QPT3



11.2.3.3. REPRESENTACIÓN MODIFICADA DE LOS RESULTADOS DE *HESPERUS*

En los resultados de *Hesperus* es posible reconocer diez cadenas (véase apéndice VIIe). Asimismo, el perfil del documento nos indica que, de las 1255 palabras del texto, 562 son plenas. La vinculación se establece mediante 83 categorías del tesoro. El cociente

de dividir el valor de cohesión otorgado por *Hesperus* (15327) y el número total de palabras es de 12,21.

Las categorías conceptuales activadas que aportan más de un 0,5% de la cohesión del texto son las siguientes:

therapy_658_4268_n	Percent: 14.51,
minuteness_196_1271_n	Percent: 10.33,
small_33_224_a	Percent: 7.35,
disease_651_4187_n	Percent: 6.61,
respiratory-disease_651_4192_n	Percent: 5.19,
serial-place_73_488_n	Percent: 4.84,
tobacco_388_2560_n	Percent: 4.05,
poison_659_4279_n	Percent: 4.00,
insides_224_1464_n	Percent: 2.57,
diagnostic_658_4258_n	Percent: 2.49,
sick-person_651_4202_n	Percent: 2.11,
joint_45_300_n	Percent: 2.02,
fluid_335_2238_n	Percent: 1.98,
frequent_139_873_a	Percent: 1.96,
small-coin_33_222_n	Percent: 1.96,
space_183_1154_n	Percent: 1.71,
hardness_326_2202_n	Percent: 1.63,
divergence_294_1963_n	Percent: 1.59,
topmost_213_1387_a	Percent: 1.43,
hospital_658_4269_n	Percent: 1.40,
chief-part_52_358_n	Percent: 1.36,
region_184_1162_n	Percent: 1.19,
prophylactic_658_4259_n	Percent: 1.15,
focus_76_508_n	Percent: 1.11,
medical_658_4274_a	Percent: 0.99,
cardiovascular-disease_651_4193_n	Percent: 0.98,
preserve_666_4322_v	Percent: 0.97,
intellect_447_2864_n	Percent: 0.89,
respiration_352_2324_n	Percent: 0.84,
opportunity_137_857_n	Percent: 0.68,
pungent_388_2561_a	Percent: 0.67,
stream_350_2304_n	Percent: 0.67,
whole_52_356_n	Percent: 0.66,
medicine_658_4257_n	Percent: 0.66,
era_110_687_n	Percent: 0.65,
auspicate_68_464_v	Percent: 0.61,
loudness_400_2607_n	Percent: 0.57,
microscopy_196_1275_n	Percent: 0.56,

Estas etiquetas semánticas son el fundamento de las 10 cadenas presentadas en el Apéndice VIIe, a las que añadiremos un breve comentario.

TABLA 110: Cadena 0 (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 0 (Learn)	
INTELECT	0,89
LEARN	0 0,1
TOTAL	+ 0,89
↓	
LEARN	

Comentario

Al igual que en los otros textos destinados a pacientes, con esta cadena se destaca la importancia de la información a la hora de afrontar la enfermedad.

TABLA 111: Cadenas 1, 3 y 6 (QPT3)

CONCEPTOS ACTIVADOS / PORCENTAJE DE ACTIVACIÓN					
Cadena 1 (Drugs)		Cadena 3 (Bronchoscope)		Cadena 6 (Doctor)	
POISON	4	THERAPY	14,51	PRESERVE (Doctor)	0,97
		DIAGNOSTIC	2,49		
		HOSPITAL	1,4		
		PROPHYLACTIC	1,15		
		MEDICAL	0,99		
		MEDICINE	0,66		
REMEDY	0,19				
TOTAL	4	TOTAL	21,39	TOTAL	0,97
↓		↓		↓	
DRUGS		TREATMENT, DIAGNOSIS, MEDICINE		HEALTH CARE	

Comentario

La relación semántica que vincula los medicamentos (cadena 1) con el tratamiento y el diagnóstico de la enfermedad (cadena 3) y con el agente que efectúa el tratamiento y diagnóstico, nos ha llevado a asociar estas tres cadenas. Además de estos conceptos, aparecen otros conceptos médicos genéricos: HOSPITAL, PROPHYLACTIC, MEDICAL y MEDICINE.

TABLA 112: **Cadena 2** (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 2 (Testing)	
ENQUIRING	0 0,1
ATTEMPTING	0 0,1
TOTAL	+ 0,1
↓	
EXPERIMENT	

Comentario

Esta cadena pasa casi desapercibida y hace mención a la constante experimentación con nuevos fármacos que se produce en el campo de la oncología.

TABLA 113: **Cadena 4** (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 4 (Small)	
MINUTESS	10,33
SMALL	7,35
SMALL-COIN	1,96
PUNGENT	0,67
MICROSCOPY	0,56
COMPRESSION	0 0,1
CONTRACTION	0 0,1
SMALL-QUANTITY	0 0,1
TOTAL	+ 20,87
↓	
MINUTENESS, SIZE	

Comentario

La frecuencia del término *small cell lung cancer* puede explicar el elevado porcentaje atribuido a esta cadena. La idea SMALL también está presente en el lexema *microscope*, que realmente estaría más relacionado con la cadena *DIAGNOSTIC TESTS*.

TABLA 114: Cadena 5 (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 5 (Lung cancer)	
DISEASE	6,61
RESPIRATORY-DISEASE	5,19
SICK-PERSON	2,11
CARDIOVASCULAR-DISEASE	0,98
PATHOLOGY	0,39
ILLNESS	0,25
TOTAL	15,53
↓	
PATHOLOGY	

Comentario

La patología del cáncer microcítico de pulmón representa un 15,53% de los conceptos activados en el texto. Al igual que en todos los textos que hemos visto, pone en relación los conceptos DISEASE, SICK-PERSON, ILLNESS y PATHOLOGY y afina para incluir dos subdimensiones dentro del dominio DISEASE: RESPIRATORY-DISEASE y CARDIOVASCULAR-DISEASE.

TABLA 115: Cadena 7 (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 7 (Stage)	
SERIAL-PLACE? (stage,remove)	4,84
SPACE	1,71
REGION	1,19
TOTAL	7,79
↓	
STAGING, LOCATION	

Comentario

En esta cadena predomina la idea de LOCALIZACIÓN, aunque la etiqueta semántica con mayor porcentaje (SERIAL-PLACE) se refiere en realidad al estadiaje del cáncer. El porcentaje que *Hesperus* le atribuye deja patente el papel que juega un estadiaje apropiado en la cura del cáncer. La localización del tumor está bastante relacionada con el estadiaje.

TABLA 116: **Cadena 8** (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 8 (Lungs)	
INSIDES	2,57
RESPIRATION	0,84
LOUDNESS	0,57
INTERIORITY	0,26
TOTAL	4,24
↓	
RESPIRATORY TRACT	

Comentario

El aparato respiratorio se convierte en el concepto central de esta cadena que incluye anatomía (*lung, chest*), fisiología (*breathing*) y patología (*coughing*). Dejaremos al margen de la misma el lexema *called* que en textos no especializados introduce la nomenclatura de un concepto nuevo para el lector lego que ha sido explicado con anterioridad.

TABLA 117: **Cadena 9** (QPT3)

CONCEPTOS ACTIVADOS	
Cadena 9 (Smoke)	
TOBACCO	4,05
TOTAL	4,05
↓	
TOBACCO	

Comentario

De nuevo, *Hesperus* otorga una porcentaje significativo (4,05%) al concepto TOBACCO, por lo que deducimos que en el texto se explicita claramente la relación entre el tabaco y el cáncer, en concreto, en la primera línea de uno de los párrafos iniciales:

(97) Small cell lung cancer is usually found in people who smoke or who used to smoke cigarettes. (QPT3, tercer párrafo)

Curiosamente, los lexemas *smoke* y *smoker* aparecen en la categoría RESPIRATORY–DISEASE del Tesouro *Roget's*.

11.2.3.4. PROPUESTA DEFINITIVA DE CADENAS LÉXICAS

Además de las cadenas mencionadas, crearemos otras en función de los conceptos del perfil del documento que no formaban parte de las cadenas reconocidas por *Hesperus*.

Conceptos reconocidos por <i>Hesperus</i>	Cadenas (% de activación conceptual)
CHIEF-PART (1,36) → WHOLE (0,66) → WHOLE (0,44)	COMPOSITION (2,36)
SPACE (1,71)	LOCATION IN THE HUMAN BODY (1,71)
INTELLECT (0,81) → FORM (0,27)	RESEARCH / EXPERIMENTATION (1,08)
ERA (0,65)	TIME (0,65)
CAUSAL-MEANS (0,2) → CAUSE (0,18)	OUTCOME OF TREATMENT (0,38)

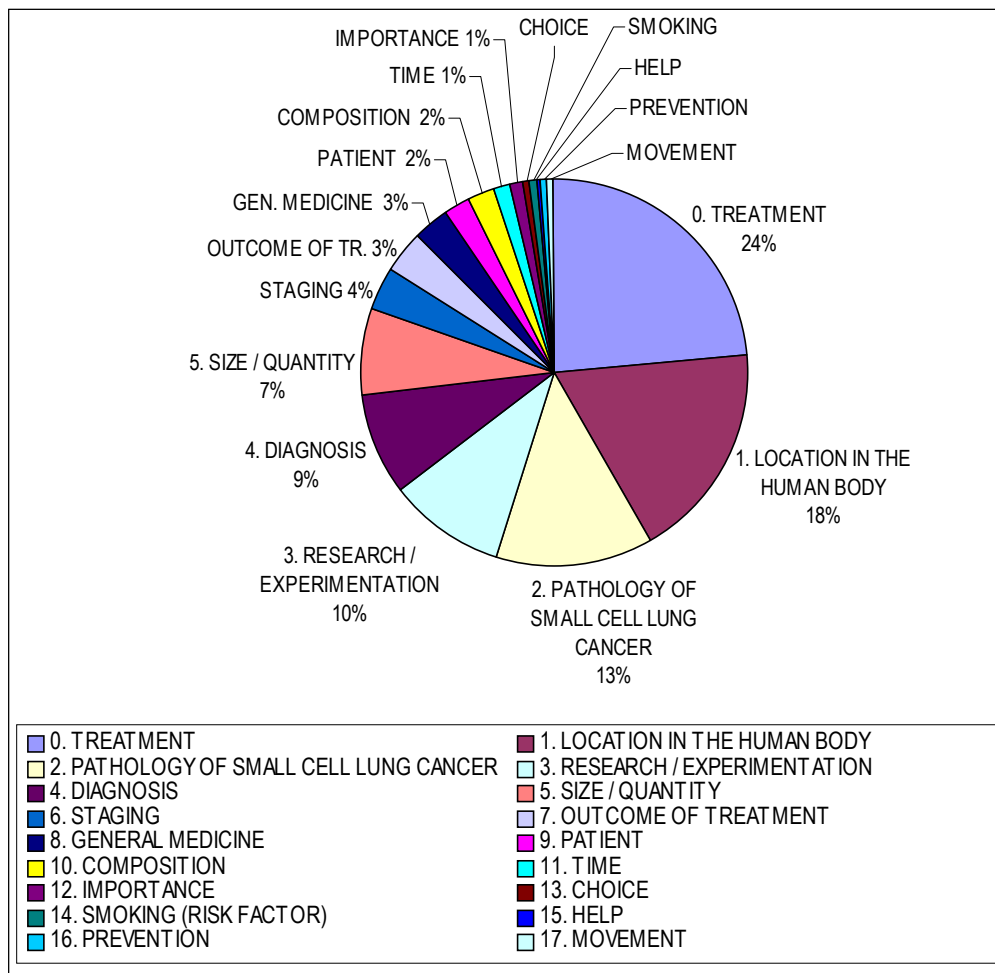
La tabla 118 adelanta las cadenas que hemos propuesto y que desglosaremos posteriormente.

TABLA 118: Cadenas léxicas en QPT3

<p>Chain 0: TREATMENT (H1, H3) [143]</p> <p>0.1. Generic treatment 0.2. Properties (→ TREATMENT) 0.3. Manipulation 0.3.1. Manipulation involved in surgery 0.4. Types 0.4.1. Radiation therapy 0.4.1.1. Radiation agents 0.4.2. Chemotherapy 0.4.2.1. Drugs 0.4.3. Surgery 0.4.4. Laser therapy 0.5. Instruments</p> <p>Chain 1: LOCATION IN THE HUMAN BODY (H7, H8) [110]</p> <p>1.1. Generic location 1.2. Microscopic / histologic level 1.3. Anatomic level / body parts 1.3.1. Respiratory system 1.3.1.1. Body processes 1.3.1.2. Substances 1.3.2. Circulatory system 1.4. Definition of body parts</p> <p>Chain 2: PATHOLOGY OF SMALL LUNG CANCER (H5) [79]</p> <p>2.1. Generic disease 2.2. Cancer 2.2.1. Types 2.2.1.1. Anatomic classification 2.2.1.2. Histologic classification 2.3. Spread of cancer</p> <p>Chain 3: RESEARCH / EXPERIMENTATION (H0, H2) [58]</p> <p>3.1. Study (H2) 3.2. Cognition (H0) 3.2.1. Information 3.3. Categorization 3.4. Causality 3.5. Normality</p>	<p>Chain 4: DIAGNOSIS (H3, H4, H5, H6) [53]</p> <p>4.1. Sensory processes involved in diagnosis 4.2. Clinical features (symptoms) 4.3. Diagnostic tests and instruments</p> <p>Chain 5: SIZE / QUANTITY (H4, H6) [44]</p> <p>5.1. Minuteness 5.2. Name of lung cancer 5.3. Increase 5.4. Large quantity</p> <p>Chain 6: STAGING (H7) [22]</p> <p>Chain 7: OUTCOME OF TREATMENT [21]</p> <p>7.1. Positive outcome 7.2. Negative outcome</p> <p>Chain 8: GENERAL MEDICINE (H3, H6)[18]</p> <p>8.1. Health professionals</p> <p>Chain 9: PATIENT (H5)[14]</p> <p>Chain 10: COMPOSITION [14]</p> <p>Chain 11: TIME [8]</p> <p>Chain 12: IMPORTANCE (H6) [7]</p> <p>Chain 13: CHOICE [3]</p> <p>Chain 14: SMOKING (RISK FACTOR) (H9)[3]</p> <p>Chain 15: HELP [3]</p> <p>Chain 16: PREVENTION [3]</p> <p>Chain 17: MOVEMENT [3]</p>
<p>TOTAL DE UNIDADES LÉXICAS ENCADENADAS ¹⁶⁵ : [552]</p>	

¹⁶⁵ Para obtener el total no se han duplicado los lemas que pertenecen a más de una cadena. La suma de las frecuencias absolutas correspondientes a cada cadena es 606.

DIAGRAMA 33: Porcentaje de activación de las cadenas léxicas definitivas de QPT3



CADENAS LÉXICAS	Frec	%	CADENAS LÉXICAS	Frec	%
0. TREATMENT	143	23,6	10. COMPOSITION	14	2,31
1. LOCATION IN THE HUMAN BODY	110	18,15	11. TIME	8	1,32
2. PATHOLOGY OF SMALL CELL LUNG CANCER	79	13,04	12. IMPORTANCE	7	1,15
3. RESEARCH / EXPERIMENTATION	58	9,57	13. CHOICE	3	0,49
4. DIAGNOSIS	53	8,74	14. SMOKING (RISK FACTOR)	3	0,49
5. SIZE / QUANTITY	44	7,26	15. HELP	3	0,49
6. STAGING	22	3,63	16. PREVENTION	3	0,49
7. OUTCOME OF TREATMENT	21	3,46	17. MOVEMENT	3	0,49
8. GENERAL MEDICINE	18	2,97	TOTAL	606	99,96
9. PATIENT	14	2,31			

En este texto, aflora la misma estructuración conceptual que en QPT2. Las únicas diferencias son que la localización y el diagnóstico ocupan el segundo y el cuarto lugar, respectivamente, como en el texto QPT1, y que aparecen dos cadenas que no estaban en QPT2: *PREVENTION* y *HELP*. Dada la alta tasa de morbilidad de este tipo de cáncer, parece necesario prevenir antes que curar.

11.2.3.4.1. DESARROLLO DE LAS CADENAS LÉXICAS DEFINITIVAS

0. TREATMENT (H1, H3)

0.1. Generic treatment

TREATMENT Treatment 13, treatments 4, treated 2, treat 1	13
CLINICAL TRIALS (See 3.1., 8.) Clinical trials 6, clinical trial 2	8
WAYS	3
→ KINDS (~ of TREATMENT) (See 3.3.)	2

<THERAPY > Therapy 1, therapies 1 (See 0.4.1., 0.4.4.)	2
→ CHOICE (~ of treatment) (See 13.)	1
CURED (See 7.)	1
→ OPTION (Treatment ~) (See 13.)	1
PLAN (~ treatment)	1
TOTAL	32

0.2. Properties (→ TREATMENT)

STANDARD → (+ TREATMENT)	3
→ ALONE (THERAPY + ~)	2
SYSTEMIC → (~ treatment)	1
TOTAL	6

0.3. Manipulation

USED Used 9, using 2, use 1, uses 1	13
GIVEN Given 2, giving 2	4
TAKE	2

Take 1, taken 1	
PUT	2
TOTAL	21

0.3.1. Manipulation involved in surgery (H4, H7)

REMOVES (✖ 7)	4
Removes 3, remove 1	
TAKE OUT	4

CUT (✖ 4)	2
OPERATION	1
TOTAL	11

0.4. Types

0.4.1. Radiation therapy

RADIATION THERAPY	13
→ EXTERNAL BEAM (~ radiation therapy)	1
PROPHYLACTIC CRANIAL	7

IRRADIATION (5) = PCI (2)	
RADIATION	2
TOTAL	23

0.4.1.1. Radiation agents

X-RAYS	2
RAYS	2
BEAM	1
→ HIGH-ENERGY (~ rays)	2

IRRADIATION (5) = PCI (2)	
→ HIGH-DOSE (~ x-rays) (See 5.4.)	1
TOTAL	8

0.4.2. Chemotherapy

CHEMOTHERAPY	10
--------------	----

0.4.2.1. Drugs (H1)

DRUGS	10
Drugs 8, drug 2	
ANESTHETIC	1

PILL	1
	12

0.4.3. Surgery

SURGERY	8
LOBECTOMY	1
PNEUMONECTOMY	1

SURGICAL IMPLANTATION	1
WEDGE RESECTION	1
TOTAL	12

0.4.4. Laser therapy

LASER THERAPY	1
---------------	---

0.5. Instruments

NEEDLE	4
DEVICES	1
INSTRUMENT	1

MACHINE	1
TOTAL	7

1. LOCATION IN THE HUMAN BODY (H7, H8)

1.2. Generic location

BODY Parts of the ~ 3 (See 10.)	11
PLACES (✖ 7) Place 2, places 1	3
LOCAL	2

NEARBY	2
BACK (~ of the throat)	1
ROOM	1
TOTAL	20

1.2. Microscopic / histologic level

CELLS Cells 14 Cancer ~ 9 SIZE + Cell 2 SIZE + Cell + LUNG CANCER 19	16
TISSUE	5
TOTAL	21

1.3. Anatomic level / body parts

BRAIN	6
CRANIAL	5
LYMPH NODES	4
BONE	1
FACE	1
MUSCLE	1

NECK (✖ 4)	1
ORGANS	1
RIBS	1
SKIN (✖ 8)	1
SPINE	1
TOTAL	23

1.3.1. Respiratory system (H8)

<LUNGS> Lungs 8, lung 4	12
BRONCHI	4
CHEST	4
THROAT	2

TUBES (+ BRONCHI)	2
AIRWAYS	1
LOBE	1
WALLS (~ of the bronchi tubes)	1
TOTAL	27

1.3.1.1. Body processes (H4)

BREATH (✖ 4) (See 4.2.) Breath 1, breathing 1	2
COUGH (See 4.2.) Cough 1, coughing 1	2

ASPIRATION	1
CALL	1
SOUND	1
WHEEZING	1

TOTAL	8
-------	---

1.3.1.2. Substances

CARBON DIOXIDE	1
OXIGEN	1
WASTE PRODUCT	1
TOTAL	3

1.3.2. Circulatory system

VEIN (✖4)	1
BLOOD	1
BLOODSTREAM	1
TOTAL	3

1.4. Definition of body parts

BEAN-SHAPED (✖4)	1
CONE-SHAPED	1
STRUCTURES	1

STORE (✖6)	1
TOTAL	4

2. PATHOLOGY OF SMALL CELL LUNG CANCER (H5)

2.1. Generic DISEASE (H3, H5)

DISEASE	4
SYMPTOMS (✖3, 5) (See 4.2.)	3
INFECTION	1
TOTAL	11

2.2. Cancer

<CANCER> ¹⁶⁶	11
CANCER CELLS	9
→ GROWING (See 5.3.)	1
MALIGNANT	1
TUMORS	1

¹⁶⁶ El lema CANCER aparece en 48 ocasiones, aunque en la mayoría de los casos constituye UF.

TOTAL	23
-------	----

2.2.1. Types

TYPE (~ of lung cancer) (See 3.3.)	2
KINDS (~ of lung cancer) (See 3.3.)	1

2.2.1.1. Anatomic classification

<LUNG CANCER>	8
---------------	---

2.2.1.2. Histologic classification

SMALL CELL LUNG CANCER	17
NON-SMALL CELL LUNG CANCER	2
TOTAL	19

2.3. Spread of cancer

SPREAD (See 5.3.)	8
RECURRENT (See 6.) Recurrent 3, recurred 1	4

COME BACK	2
TOTAL	14

3. RESEARCH, EXPERIMENTATION (H0, H2)

3.1. Study (H2)

FOUND (See 4.1.) Found 7, find 2	9
CLINICAL TRIALS (See 0.1., 8.) Clinical trials 6, clinical trial 2	8
TESTING (See 4.3.) Testing 3, test 1	4

DESIGNED	1
EXPLANATION	1
STUDIES	1
TOTAL	24

3.2. Cognition (H0)

LEARN	2
-------	---

KNOW	1
DETERMINE	1
TOTAL	4

3.2.1. Information

INFORMATION	3
PDQ	1
OVERVIEW	1

SUMMARY	1
TOTAL	6

3.3. Categorization

CALLED	7
KINDS (See 2.2.1.)	3
TYPE (See 2.2.1.)	2
DESCRIPTION	1

MEANS	1
TOTAL	14

3.4. Cause

DEPEND	2
Depend 1, depends 1	
PRODUCE	2
Produce 1, product 1	

CAUSES	1
REASONS	1
TOTAL	6

3.5. Normality

STANDARD	3
COMMON	1
TOTAL	4

4. DIAGNOSIS (H3, H4, H5, H6)

PROGNOSIS (✖5)	1
GENERAL STATE OF HEALTH	1

4.1. Sensory processes involved

FOUND (See 3.1.)	9
Found 7, find 2	
SEE	5
See 4, seen 1	

LOOK	4
CONSIDERED	2
TOTAL	20

4.2. Clinical features (symptoms) (H5)

SYMPTOMS (✖3, 5) (See 2.1.)	3
FEELING Feeling 2, felt 1	3
DISCOMFORT	3
PAIN (✖5)	3
COUGH (✖5) (See 1.3.1.1.) Cough 1	2

Coughing up blood 1	
CONDITION (✖5)	1
HOARSENESS	1
SHORTNESS OF BREATH (See 5.1., 1.3.1.1.)	1
SWELLING (✖5)	1
TOTAL	19

4.3. Diagnostic tests and instruments (H3, H4, H6)

TEST (✖6)(See 3.1.) Testing 3, test 1	4
MICROSCOPE (✖4)	3
BIOPSY	2

BRONCHOSCOPE	2
BRONCHOSCOPY (✖3)	1
TOTAL	12

5. SIZE / QUANTITY (H4)

5.1. Minuteness

<SMALL>	5
LIMITED (See 6.1.)	2
SHORT Short 1, shortness 1 (See 4.2.)	2

PIECES (See 10.)	1
SHRINK (See 7.1.)	1
TOTAL	11

5.2. Name of lung cancer

SMALL (~ cell lung cancer)	17
NON-SMALL (~ cell lung cancer)	2
TOTAL	21

5.3. Increase

SPREAD (See 2.3.)	8
GROWING (See 2.2.)	1

TOTAL	9
-------	---

5.4. Large quantity (H6)

HIGH-ENERGY (~ rays) (See 0.4.1.1.)	2
HIGH-DOSE (~ x-rays) (See 0.4.1.1.)	1

TOTAL	3
-------	---

6. STAGING (H7)

STAGE	15
Stage 10, stages 4, staging 1	
RECURRENT (See 2.3.)	3
~ stage	
~ disease	
~ small cell lung cancer	
EXTENSIVE STAGE (~ small	2

cell lung cancer)	
LIMITED STAGE (~ small cell lung cancer) (See 5.1.)	2
TOTAL	22

7. OUTCOME OF TREATMENT

7.1. Positive outcome (H 5)

KILL CANCER CELLS (✖5)	5
EFFECT	2
Effect 1, effects 1	
PREVENT	2
(~ spread of the cancer)	
REDUCE (~ discomfort)	2
RELIEVE	2
~ symptoms	

cell lung cancer)	
BETTER	1
CURED (See 0.1.)	1
EFFECTIVENESS	1
RECOVERY	1
SHRINK (~ tumors) (See 5.1.)	1
TOTAL	17

7.2. Negative outcome

→ AFFECT (~ brain function)	1
PRESSURE (✖4)	1
SIDE EFFECTS	1

8. GENERAL MEDICINE (H3, H6)

CLINICAL (See 0.1., 3.1.)	8
HOSPITAL (✖3)	1
TOTAL	9

8.1. Health professional

DOCTOR (✖6)	9
-------------	---

9. PATIENT (H5)

PATIENT Patient 9, patients 4	13
PEOPLE	1
TOTAL	14

10. COMPOSITION

PART Part 3, parts of the body 3 (See 1.2.)	6
ENTIRE	2
CONTAIN	1
INCLUDING	1
MAKE UP	1
PARTITION	1

PIECES (✖ 4) (See 5.1.)	1
SECTION	1
TOTAL	14

11. TIME

AGE	1
BEGAN	1
ONCE	1
PAST	1
PERIOD	1

→ SHORT (~ period of time)	1
TIME	1
UP-TO-DATE	1
TOTAL	8

12. IMPORTANCE

NEW	6
SPECIAL	1
TOTAL	7

13. CHOICE

DECIDE	1
CHOICE (See 0.1.)	1
OPTION (See 0.1.)	1
TOTAL	3

14. SMOKING (RISK FACTOR) (H9)

SMOKE (✖9)	2
CIGARETTES (✖9)	1
TOTAL	3

15. HELP

HELP	2
SERVICE (✖6)	1
TOTAL	3

16. PREVENTION

PREVENT (~ spread of the cancer 2)	3
---------------------------------------	---

17. MOVEMENT

ENTERS	1
SLIDES	1
TRAVELS	1
TOTAL	3

11.2.3.5. CUADRO CONTRASTIVO (*HESPERUS* – ANÁLISIS COMBINADO)

En la tabla 118 quedan resumidos los datos cuantitativos acerca de cómo la cohesión léxica contribuye a la activación conceptual dentro del texto QPT3.

El texto cuenta con 1255 ocurrencias léxicas, entre las que hemos identificado 552 palabras plenas vinculadas, que representan un 43,98 % del total. Según *Hesperus*, el porcentaje de palabras vinculadas es de 44,78 %. El texto cuenta con 66 oraciones que tienen una extensión media de 17,2 palabras. Hay un promedio de 6 palabras por oración que establecen relaciones cohesivas.

Teniendo en cuenta que las 552 palabras vinculadas se corresponden con 192 formas léxicas y con 190 lemas, estimamos que en 100 palabras vinculadas podemos encontrar 35 variantes morfológicas asociadas a 34 lemas, las cuales activarían 15 categorías conceptuales.

TABLA 119: Descripción cuantitativa de las cadenas léxicas reconocidas en QPT3

CADENAS LÉXICAS RECONOCIDAS POR HESPERUS		NUESTRA PROPUESTA DE CADENAS LÉXICAS	
H 3: TREATMENT, DIAGNOSIS, MEDICINE [105-23-21,39] H 1: DRUGS [6-2-4]			0. TREATMENT [143-30 -23,6] 4. DIAGNOSIS [53-24-8,74]
Lexemas vinculados	111	196	Lexemas vinculados
Formas distintas	25	54	Formas distintas
Activación conceptual ajustada (%)	29,52	32,34	Activación conceptual (%)
H 8: RESPIRATORY TRACT [26-10-4,24] H 7: LOCATION [7-4-2,95] SPACE [1,71]			1. LOCATION IN THE HUMAN BODY
Lexemas vinculados	33	110	Lexemas vinculados
Formas distintas	14	48	Formas distintas
Activación conceptual ajustada (%)	10,35	18,15	Activación conceptual (%)
H 5: PATHOLOGY [13,42]			2. PATHOLOGY OF SCLC
Lexemas vinculados	78	79	Lexemas vinculados
Formas distintas	13	16	Formas distintas
Activación conceptual ajustada (%)	15,6	13,04	Activación conceptual (%)
H 0: LEARN [2-2- 0,89] H 2: EXPERIMENT [15-4-0,1] INTELLECT (0,81), FORM (0,27) = 1,08			3. RESEARCH / EXPERIMENTATION
Lexemas vinculados	17	58	Lexemas vinculados
Formas distintas	6	27	Formas distintas
Activación conceptual ajustada (%)	2,41	9,57	Activación conceptual (%)
H 4: MINUTENESS / SIZE [20,87]			5. SIZE / QUANTITY
Lexemas vinculados	73	44	Lexemas vinculados
Formas distintas	13	12	Formas distintas
Activación conceptual ajustada (%)	24,26	7,26	Activación conceptual (%)
H 7: STAGING [4,84]			6. STAGING
Lexemas vinculados	14	22	Lexemas vinculados
Formas distintas	2	6	Formas distintas
Activación conceptual ajustada (%)	5,63	3,63	Activación conceptual (%)
CAUSAL-MEANS (0,2), CAUSE (0,18) = 0,38			7. OUTCOME OF TREATMENT [21-14- 3,46]
Activación conceptual ajustada (%)	0,44	3,46	Activación conceptual (%)
H 6: HEALTH CARE [0,97]			8. GENERAL MEDICINE
Lexemas vinculados	17	18	Lexemas vinculados
Formas distintas	6	3	Formas distintas
Activación conceptual ajustada (%)	1,13	2,97	Activación conceptual (%)
H 5: SICK-PERSON [2,11]			9. PATIENT [14-3 -2,31]
Activación conceptual ajustada (%)	2,45	2,31	Activación conceptual (%)

CHIEF-PART (1,36), WHOLE (0,66), WHOLE (0,44) = 2,36 Activación conceptual ajustada (%)	2,74	2,31	10. COMPOSITION [14-9-2,31] Activación conceptual (%)
ERA (0,65) Activación conceptual ajustada (%)	0,76	1,32	11. TIME [8-8-1,32] Activación conceptual (%)
H 9: TOBACCO [4,05] Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	2 1 4,71	3 2 0,48	14. SMOKING (RISK FACTOR) Lexemas vinculados Formas distintas Activación conceptual (%)
		19 11 3,11	12. IMPORTANCE [7-2-1,15] 13. CHOICE [3-3-0,49] 15. HELP [3-2-0,49] 16. PREVENTION [3-1-0,49] 17. MOVEMENT [3-3-0,49] Lexemas vinculados Formas distintas Activación conceptual (%)
TOTAL Activación conceptual según Hesperus Lexemas vinculados Formas distintas Activación conceptual ajustada (%)	86,01 345 80 100	606 552 192 99,96	TOTAL Suma de elementos en las cadenas Lexemas vinculados ¹⁶⁷ Formas distintas ¹⁶⁸ Activación conceptual (%)
DATOS ADICIONALES SOBRE EL TEXTO QPT3 Nº total de palabras: 1255 Palabras que crean cohesión (<i>Hesperus</i>): 562			Categorías conceptuales del tesoro: 83 Coeficiente de cohesión (<i>Hesperus</i>): 12,21 Nº de oraciones: 66

¹⁶⁷ Sólo se ha computado una vez los lexemas que aparecen en más de una cadena.

¹⁶⁸ De estas, hay 21 formas léxicas que pertenecen a más de una cadena y que llevan asociadas 54 lexemas adicionales a los 552 lexemas vinculados.

11.3. Datos estadísticos relacionados con el estudio de la cohesión y sus limitaciones

En esta sección se comentan las estadísticas correspondientes a los textos para profesionales de la salud y para pacientes, comentadas en 11.1. y 11.2., para ver si la cohesión se construye de forma distinta en distintos tipos textuales definidos con respecto al parámetro "grado de conocimiento del lector". También se establece la relación entre el formato con el que hemos presentado las cadenas léxicas y los tipos de repetición propuestos en el capítulo 6.

11.3.1. Representación de las cadenas léxicas y tipos de repetición

El formato de las tablas incluidas en las secciones tituladas *Desarrollo de las cadenas léxicas definitivas* de este capítulo son indicativas del tipo de relación cohesiva establecida, de acuerdo con una serie de consideraciones. Las aplicamos a la subcadena *GENERIC TREATMENT* de QDT1 (tabla 120):

- Para detectar las instancias de *repetición exacta* en una cadena/subcadena, se le resta al TOTAL el número de formas que tienen una frecuencia absoluta de 1. En la subcadena *GENERIC TREATMENT* encontraríamos 40 instancias de repetición exacta, resultantes de restar a 45, las 5 variantes morfológicas que sólo aparecen una vez (*treatments, curative, modality, option, procedures*). Para todo el texto, a partir de una lista de frecuencia en la que se han suprimido las palabras forma, sólo basta con restar al número de ocurrencias (*tokens*), el número de formas que tienen una frecuencia igual a 1.
- El encontrar formas distintas (*types*) en una casilla indica que entre ellas existe *repetición simple* y *variada sintáctica*. Las formas *treatment* y *treatments* están relacionadas mediante repetición simple y estas lo están a su vez con *treated* mediante repetición variada sintáctica. Para computar los casos de repetición simple y variada sintáctica en una subcadena, se resta al número de formas distintas (14) el número de lemas (9). Por tanto, en esta cadena habría 5 casos entre repetición simple (*treatment – treatments; modalities – modality; options – option*) y variada sintáctica (*treatment – treated; cure – curative*).
- El número de lemas de una cadena/subcadena queda reflejado en el número de casillas. El número de lemas indica el mínimo de instancias de *repetición*

paradigmática semántica, dado que todos esos lemas están relacionados mediante el concepto que etiqueta la cadena. En la subcadena de la tabla 120 hay nueve lemas y relaciones paradigmáticas sinonímicas (TREATMENT – THERAPY, TRIALS – TREATMENT), hiponímicas (APPROACH – THERAPY, PROCEDURES – THERAPY) y asociativas circunstanciales basadas en la relación causa-efecto (TREATMENT– CURE, TREATMENT– CONTROL).

- Cada vez que aparece una unidad poliléxica o los signos ~ o →, seguidos de un número igual o mayor que 2 nos encontramos ante casos de *repetición sintagmática* o *repetición mixta*. En la tabla 120, podemos señalar *surgical approaches, treatment approaches, clinical trials, treatment options, intracavitary treatment*.

TABLA 120: Subcadena *TREATMENT* del texto QDT1.

0. TREATMENT (H1, H3a, H10, H13)

0.1. Generic treatment

TREATMENT Treatment 12, treated 5, treatments 1	18
→ APPROACHES Surgical ~ 2 Treatment ~ 2 Physical ~ 1	5
TRIALS (See 1.1.) Clinical ~ 4	5
CONTROL (+ SYMPTOMS+)	4
CURE (See 4.2.) Cure 2, curative 1	3
MODALITIES (See 11.) Modalities 2, combined modality 1	3
OPTIONS (See 11.) Treatment options 2 Option 1	3
THERAPY Intracavitary ~ 2 (See 0.4.3. RADIATION ~)	3
→ PROCEDURES	1
TOTAL	45

11.3.2. Estadísticas de textos sobre tratamiento del cáncer de pulmón

Si en los textos sobre tratamiento para profesionales, QDT1, QDT2 y QDT3, el porcentaje medio de palabras que contribuyen a la cohesión del texto (52,64 - 46,6 - 50,85) es de 50,03%, en los textos para pacientes, QPT1 (46,21%), QPT2 (44,67%) y QPT3 (43,98%), este es ligeramente menor, en concreto, 44,95%. Esto se debe principalmente a dos factores.

Por una parte, en inglés, la nominalización típica de textos científicos para especialistas toma cuerpo en *sintagmas nominales compuestos* (Bhatia 1993: 149). La estructura de estos sintagmas nominales es la de *modificador(es) + núcleo*. La posición que precede al sustantivo nuclear queda ocupada por una sucesión lineal de sustantivos, entre los que se puede intercalar algún adjetivo. Veamos algunos ejemplos sacados de los textos QDT2 y QDT3:

- (98) Non-small cell lung cancer (QDT2)
- (99) Squamous cell lung cancer (QDT2)
- (100) Ipsilateral hilar lymph nodes (QDT2)
- (101) Endobronchial laser therapy (QDT3)
- (102) Magnetic resonance imaging scans (QDT3)

Esta estructura es muy frecuente en las unidades fraseológicas de los textos para profesionales de la salud y hace que haya una mayor profusión de palabras plenas, y por tanto, un porcentaje mayor de palabras que contribuyen a la cohesión del texto, ya que los sintagmas nominales se construyen sin la necesidad de preposiciones o pronombres relativos.

Por otra parte, en los textos para especialistas, nuestros cálculos indican que la repetición léxica exacta y simple es más habitual que en los textos para pacientes. Como estos recursos son totalmente explícitos, se han computado todos los casos, con lo que se obtiene un porcentaje elevado de palabras cohesionadas, algo que no podemos afirmar en el caso de la repetición variada léxica, más frecuente en textos para pacientes pero más difícil de detectar.

En los textos para especialistas se observa una mayor repetición exacta, destinada a evitar la ambigüedad, mientras que existe una mayor variedad léxica en los textos para pacientes, de acuerdo con los cálculos que hicimos en torno al número de variantes morfológicas y de lemas cada 100 palabras. Se dejará a un lado el texto QDT1, en el que se produce una variedad léxica inusual debido a su estilo esquemático, sobre todo en el apartado *Treatment options*.

En QDT2 y QDT3, por cada 100 palabras, hay respectivamente 24 y 27 formas léxicas distintas correspondientes a 16 y 19 lemas distintos. La variedad léxica es mayor en la versión para pacientes, en concreto, por cada 100 palabras hay una media de 35,3 formas léxicas distintas asociadas a 31,3 lemas. Esto se explica por el fenómeno que cita Meyer (1991: 5) por el que en textos de contenido científico destinados a un lector lego, en este caso a pacientes, la cohesión se logra mediante sinónimos y términos subordinados y superordinados, y no tanto mediante la repetición de la misma forma léxica. Profundizaremos más sobre esta idea en el capítulo 13.

Otra explicación a la variedad léxica y al elevado porcentaje de unidades léxicas que contribuyen a la cohesión la hallamos en la utilización de paréntesis en la versión para pacientes, un recurso muy sintético con el que se activan un gran número de conceptos con el mínimo de palabras, como se verá también en el capítulo 13.1.

(103) The prognosis (chance of recovery) [QPT2]

(104) The area that separates the two lungs (mediastinum) [QPT2]

Con el paréntesis se consigue un aumento de la *type-token ratio*, y por eso, este índice es anormalmente elevado en los textos para pacientes.

A partir del cómputo que hace *Hesperus* no sólo de las categorías del *Roget* activadas sino también del número de palabras que crean cohesión en los textos analizados, se puede hallar un promedio de las categorías conceptuales relacionadas con cada 100 palabras vinculadas. Este promedio no difiere apenas del calculado en 11.1. y 11.2. a partir de nuestro análisis:

Categ. conceptuales por cada 100 palabras activadas

QDT1 - QDT2 - QDT3:	QPT1 - QPT2 - QPT3:
25 – 8 – 11 (11.1. y 11.2.)	14 – 14 – 15 (11.1. y 11.2.)
28 – 8 – 11 (<i>Hesperus</i>)	15 – 14 – 15 (<i>Hesperus</i>)

La obtención de este cálculo mediante métodos informáticos permite el estudio de la densidad conceptual correspondiente a un mismo número de palabras plenas. Además, permite su estudio no sólo en seis textos sobre el mismo tema que se diferencian respecto al parámetro *destinatario* sino en un corpus de 36 textos. Como se verá en 11.3.3., en los textos destinados a un lector lego, se necesitan menos palabras plenas para activar las principales áreas conceptuales que aparecen en los textos para especialistas. Estas palabras plenas son en su mayor parte unidades léxicas de la lengua general que "arropan" aquellos términos que pueden crear dificultad al lector.

11.3.3. Estadísticas generales en el resto de los textos

Si se extienden al resto del corpus del *Physician Data Query* algunos de los cálculos aplicados a los textos sobre tratamiento, es posible identificar tendencias estadísticas que puedan confirmar diferencias cohesivas entre textos para especialistas y textos para lectores legos. Claro está, aparecen las limitaciones que impone la composición del corpus analizado. La tabla de datos estadísticos incluye los siguientes parámetros:

- a) *type / token ratio (x100)*: indica el número de formas distintas por cada 100 palabras de texto, es decir, el porcentaje entre variantes morfológicas y total de palabras. Su valor depende en gran medida de la extensión del texto, de forma que, cuanto mayor sea su extensión, su *type / token ratio* es menor. Para que no quede distorsionado, en la casilla correspondiente habrá dos cálculos. El primero se basará en una lista de frecuencia sin modificar y el segundo, en una lista en la que se han eliminado las palabras forma con la ayuda de una *stoplist* que contiene 256 palabras. Como resultado de esto, en el corpus para especialistas se han eliminado 130 palabras forma (1,88% del total), mientras que en el destinado a pacientes se ha prescindido de 208 palabras (6,14% del total). Asimismo, para todo el corpus se ha calculado la *standardised*

- type/token ratio*¹⁶⁹ por cada 300 palabras de texto, puesto que el corpus cuenta con algunos textos que superan escasamente las 600 palabras.
- b) *Palabras vinculadas por cada 100 palabras de texto*: para este cálculo, nos serviremos del primer valor que otorga *Hesperus* en la perspectiva general (p. 707).
 - c) *Coefficiente de cohesión*: es el cociente que se obtiene al dividir entre el total de palabras, la cifra que *Hesperus* otorga a cada texto en función del número de conceptos del *Roget* activados (segunda cifra entre paréntesis del Apéndice VIIa). Como esta cifra depende de la extensión del texto, estimamos que la fiabilidad de este cálculo es escasa.
 - d) *Palabras vinculadas en cada oración*: dividiremos el número de palabras vinculadas entre el número de oraciones para ver cuántas palabras hay vinculadas por oración. Se relacionarán estos cálculos con la extensión de las oraciones, que es mayor en textos para especialistas (22,15 palabras) que en textos para pacientes (17,71).
 - e) *Categorías conceptuales del tesoro Roget por cada 100 palabras vinculadas*: con este cálculo se pretende detectar qué textos son más informativos en relación con el número de conceptos que activan y las palabras plenas que los activan.
 - f) *Número de cadenas*: no se prestará atención a este cálculo porque existe una gran disparidad entre los textos. La mayor extensión de los textos QD hace que *Hesperus* reconozca en ellos más cadenas, una media de 26,5 cadenas, en comparación con el promedio de 7,83 cadenas que reconoce en QP.

El coeficiente de cohesión y el número de categorías conceptuales por cada 100 palabras vinculadas es sólo orientativo. Somos conscientes de las limitaciones de estos cálculos ya que el tesoro es de la lengua general y no permite la identificación de unidades léxicas exclusivas de la medicina, y por tanto, excluye relaciones conceptuales que se dan en el texto.

¹⁶⁹ Este cálculo es la media de las *type/token ratios* que se obtienen a partir de n palabras consecutivas de texto. Si se fija n con el valor de 300, se estima la *type/token ratio* de las primeras 300 palabras, después de las 300 siguientes, y así sucesivamente, y luego se hace una media que es a lo que se le denomina la *standardised type/token ratio*.

TABLA 121: Estadísticas correspondientes al corpus analizado en el capítulo 11.

	PROFESIONALES DE LA SALUD (QD)					PACIENTES (QP)				
	Type/token ratio(x100) (con / sin pal. forma)	Palabras vinculadas (x100) / n. palabras	Coefficiente de cohesion	Palabras vinculadas/ n. de oraciones	Cat. conceptuales por 100 pal. vinculadas	Cat. conceptuales por 100 pal. vinculadas	Palabras vinculadas/ n. oraciones	Coefficiente de cohesion	Palabras vinculadas (x100) / n. palabras	Type/token ratio(x100) (con/ sin pal. forma)
OVERALL	Normal 9,68 9,5	46,55	9,73	10,71	13,41	17,26	8,69	11,93	45,86	Normal 9,50 8,92
	Estánd 56,04									Estánd 50,22
Malignant mesothelioma (T1)	38,92 <u>32,44</u>	505/ 1079= 46,8	9,76	505/ 63= 8,01	141/ 505= <u>27,92</u>	109/ 704= 15,48	704/ 92= 7,65	10,7	704 / 1621= 43,42	24,80 19,31
Non-small cell lung cancer (T2)	18,23 <u>16,15</u>	2470/ 5523= 44,72	8,37	2470/ 251= 9,84	196/ 2470= <u>7,93</u>	80/ 553= 14,46	553/ 85= 6,50	13,83	553 / 1258= 43,95	23,37 <u>16,93</u>
Small cell lung cancer (T3)	21,27 18,37	1648/ 3408= 48,35	8,87	1648/ 142= 11,6	184/ 1648= 10,93	83/ 562= 14,77	562/ 66= 8,51	12,21	562 / 1255= 44,78	26,29 19,6
Anxiety	29,38 26,23	1489/ 3267= 45,58	10,84	1489/ 166= 8,97	210/ 1489= 14,1	147/ 931= 15,79	931 / 104= 8,95	12,42	931/ 1954= 47,64	28,61 23,59
Constipation	26,59 23,93	2159/ 4709= 45,85	12,39	2159/ 244= 8,85	272/ 2159= 12,6	139/ 679= 20,47	679 / 77= 8,82	12,62	679/ 1457= 46,6	33,97 27,18
Delirium	35,58 30,17	759/ 1793= 42,33	7,69	759/ 76= 9,98	158/ 759= 20,82	98/ 339= 28,91	339 / 40= 8,47	12,67	339 / 749= 45,26	38,99 28,84
Fatigue	27,92 25,19	1987/ 4005= 49,61	11,11	1987/ 207= 9,6	252/ 1987= 12,68	155/ 1198= 12,94	1198 / 128= 9,36	11,03	1198/ 2390= 50,12	26,19 21,92
Fever	35,79 31,04	838/ 1917= 43,71	9,73	838/ 92= 9,11	167/ 8380= 19,93	116/ 512= 22,65	512 / 65= 7,87	14,07	512/ 1148= 44,59	35,37 27,79
Hypercalcemia	21,82 20,01	3025/ 7093= 42,65	8,94	3025/ 287= 10,54	293/ 3025= 9,68	159/ 990= 16,06	990 / 127= 7,79	9,28	990/ 2179= 45,43	28,09 23,36
Loss, grief, and bereavement	23,69 21,1	2871/ 6413= 44,77	9,83	2871/ 257= 11,17	295/ 2871= 10,27	199/ 2039= <u>9,76</u>	2039 / 225= 9,06	9,92	2039/ 4683= 43,54	<u>21,67</u> 18,24
Lymphedema	32,71 29,4	1717/ 3571= 48,08	9,46	1717/ 154= 11,15	249/ 1717= 14,5	172/ 1075= 16	1075 / 121= 8,88	10,23	1075/ 2352= 45,7	29,08 24,02
Nausea and vomiting	22,97 21,08	2333 / 5541= 42,19	6,95	2333/ 198= 11,78	228/ 2333= 9,77	128/ 707= 18,23	707 / 76= 9,30	12,31	707/ 1452= 48,69	30,99 25,14

		<u>42,1</u>		11,78	9,77	18,10	9,3		48,69	
	Type/token ratio(x100) (con / sin pal. forma)	Palabras vinculadas (x100) / n. palabras	Coefficiente de cohesion	Palabras vinculadas/n. de oraciones	Cat. conceptuales por 100 pal. vinculadas	Cat. conceptuales por 100 pal. vinculadas	Palabras vinculadas/n. oraciones	Coefficiente de cohesion	Palabras vinculadas (x100) / n. palabras	Type/token ratio(x100) (con/ sin pal. forma)
Nutrition	27,53 25,15	2272/ 4609= 49,29	12	2272/ 185= 12,28	285/ 2272= 12,54	181/ 1406= 12,87	1406 / 159= 8,84	14,04	1406/ 2973= 47,29	26,20 22,23
Oral complications of cancer	28,73 26,42	2180/ 4651= 46,87	8,74	2180/ 181= 12,04	292/ 2180= 13,39	191/ 1415= 13,5	1415 / 164= 8,63	9,74	1415 / 3032= 46,66	24,14 20,09
Post-traumatic stress disorder	24,37 21,42	1824 / 3992= 45,69	5,96	1824/ 114= 16	210/ 1824= 11,51	149/ 911= 16,35	911 / 95= 9,58	9,44	911/ 1884= 48,35	28,77 23,3
Pruritus	30,82 28,16	1758/ 3761= 46,74	11,24	1758/ 193= 9,11	260/ 1758= 14,79	125/ 476= 26,26	476 / 72= 6,61	12,75	476/ 1030= 46,21	39,90 32,72
Sleep disorders	30,68 28,12	1554/ 3119= 49,82	12,43	1554/ 122= 12,74	194/ 1554= 12,48	98/ 323 = <u>30,34</u>	323 / 23= 14,04	17,01	323 / 579= <u>55,78</u>	<u>45,08</u> <u>35,58</u>
Transitional care planning	29,15 25,47	1412/ 2792= 50,57	10,99	1412 / 141= 10,01	207/ 1412= 14,66	176/ 1530 = 11,5	1530 / 200= 7,65	10,43	1530/ 3683= 41,54	<u>21,04</u> 17,43
Desviación típica	5,51 4,55	2,63	1,83	1,88	4,77	5,89	1,6	2,03	3,15	6,66 5,13
Desviación típica ajustada	4,24 3,82	2,49			3,18	4,93			2,18	5,2 4,17

Al calcular la desviación típica de estos datos, se percibe que hay textos en los que los cálculos *Palabras vinculadas (x100)/ N° total de palabras* o la *type/token ratio* difieren bastante con respecto al resto. Se han indicado estas anomalías mediante subrayado y se han eliminado en la media aritmética para no distorsionar los datos.

Entre las inferencias sobre nuestro corpus extraídas de las estadísticas del texto podemos citar que el número de palabras vinculadas por cada 100 palabras es ligeramente superior en textos para especialistas. En concreto, 46,55 frente al 45,86 correspondiente a textos para pacientes. Este porcentaje es alto por el carácter sintético de los resúmenes objeto de estudio (*Summaries for health professionals and patients*). La *type/token ratio* es elevada porque el corpus analizado no es muy extenso. Aún más elevada es la *standardised type/token ratio* por cada 300 palabras de texto.

En cada oración de QD, que tiene una extensión media de 22,15 palabras, hay 10,71 palabras vinculadas mientras que en QP, de las 17,71 palabras por oración, 8,69 están vinculadas. Aunque la proporción *palabras por oración/palabras vinculadas* es muy similar, si consideramos la cohesión como fenómeno interoracional, los textos para especialistas son más cohesivos porque a mayor número de palabras cohesionadas por oración, hay un mayor número de conexiones entre oraciones.

El número de categorías conceptuales activadas por cada 100 palabras vinculadas es menor en textos para especialistas, en parte porque las USE no son reconocidas por un tesoro de la lengua general. En los textos no especializados, con menos palabras se actualizan las mismas áreas conceptuales, aunque obviamente con menor profundidad. Por eso también, el coeficiente de cohesión es mayor.

11.3.4. Recapitulación

En definitiva, a partir de cálculos estadísticos básicos, se pueden extraer algunas inferencias sobre el tipo de relación cohesiva que se establece en diferentes tipos textuales:

- a) Los datos aportados por el programa *Hesperus* a grandes rasgos no distan demasiado de los obtenidos por el analista, por lo que pueden aplicarse a un corpus más extenso.
- b) La mitad de las unidades léxicas de los resúmenes analizados contribuyen a la cohesión del texto, aunque en los textos para pacientes son suficientes menos palabras cohesionadas para activar más áreas conceptuales.
- c) La mayor extensión de las oraciones en textos para profesionales de la salud hace que haya más palabras vinculadas en cada oración con respecto a los textos para pacientes.
- d) En definitiva, la estructuración conceptual que emana de un determinado campo del saber está relacionada con la cohesión de las cadenas léxicas generadas a partir de la activación textual de los conceptos. Por eso, las cadenas léxicas y la cohesión en textos sobre el mismo tema tiene que ser igual o muy parecida.

12. Cadenas léxicas y activación conceptual

De la estructuración en cadenas léxicas de los textos analizados en el capítulo anterior se extraerán algunas conclusiones de índole léxica y conceptual.

En primer lugar, ha quedado constancia de que textos que comparten el mismo tema tienen unas cadenas léxicas parecidas. Los seis textos sobre el tratamiento del cáncer de pulmón tienen prácticamente las mismas cadenas y lo que cambia es el peso específico de las mismas. Así encontramos que tanto en QDT1 como en QPT1 es posible encontrar una cadena *LOCATION IN THE HUMAN BODY*, aunque en el primer texto tiene una representación conceptual del 8,55%, y en el segundo, del 16,03%.

En segundo lugar, existe un paralelismo entre la distribución de las cadenas léxicas y las áreas conceptuales que reciben un mayor énfasis en el texto. La preeminencia de dichas áreas conceptuales queda también reflejada en el índice que precede cada texto (véase secciones 11.n.n.1. del capítulo anterior). Este es un elemento de la superestructura del texto e incluye unidades de significación clave que sirven de hilo conductor de las principales cadenas léxicas.

El porcentaje de representación conceptual y el orden de importancia de las cadenas en textos sobre un tema parecido, en concreto, el tratamiento del cáncer, suele ser similar si estos están dirigidos a una misma audiencia. Como ilustración, basta con buscar en la tabla 122 de la página 488 las cuatro cadenas más representativas de los textos para pacientes. *TREATMENT* siempre es la más importante. *LOCATION* ocupa el segundo lugar en dos de los textos y en uno, el tercer lugar. En dos de los tres textos, *PATHOLOGY* aparece en tercer lugar, y en uno, en el segundo. *RESEARCH* siempre es la cuarta cadena más relevante.

No descartamos que el elevado grado de similitud entre cadenas y de representación conceptual se deba en parte a que los textos analizados contienen un vocabulario controlado, el que se utiliza en la sección *Physician Data Query* de la página web del *National Cancer Institute*. En esta sección se reduce la diversidad terminológica en torno al mismo concepto y se sigue una superestructura parecida para conferir consistencia. A esta consistencia también contribuye el elevado nivel de estandarización del lenguaje biomédico.

Tabla 122: [Cadenas léxicas correspondientes a los textos para especialistas \(ODT\)](#)

Tabla 123: [Cadenas léxicas correspondientes a los textos para pacientes \(OPT\)](#)

En las tablas 122 y 123, las cadenas correspondientes a cada texto se ordenan según el porcentaje de palabras plenas que contienen. Se han establecido agrupaciones entre cadenas con un porcentaje parecido, porque este nos indica la activación conceptual de las cadenas. De estas tablas se pueden sacar las siguientes conclusiones:

- 1) La cadena con más elementos en los seis textos analizados es la de *TREATMENT*. Su porcentaje de activación ronda el 25%. Esto coincide con la temática de los textos, que está presente en el título de las secciones del servicio *PDQ* de donde provienen: [Treatment summaries for health professionals](#) y [Treatment summaries for patients](#).
- 2) Hay siete cadenas que representan aproximadamente el 85% de la activación conceptual en todos los textos. Estas cadenas son: *TREATMENT*, *RESEARCH*, *PATHOLOGY*, *DIAGNOSIS*, *OUTCOME OF TREATMENT*, *LOCATION IN THE HUMAN BODY* y *STAGING*. Todas ellas tienen una activación superior al 4,75% en textos para especialistas y un 3,5% en textos para pacientes.
- 3) El 15 % restante lo constituyen cadenas cuya activación conceptual se encuentra en el intervalo 0,11 – 4,75% en textos para especialistas y 0,26 – 3,5% en textos para pacientes. Estas aglutinan lexemas que no son exclusivos de la oncología.
- 4) Se han identificado 25 cadenas léxicas distintas
- 5) Existen coincidencias entre las cadenas léxicas de textos sobre el mismo tema, que se incrementan cuando los textos están dirigidos a un mismo lector.

12.1. Diferencias conceptuales y terminológicas entre textos para especialistas y textos para pacientes

Se ha constatado que las cadenas son prácticamente las mismas en textos dirigidos a una misma audiencia. Creemos ahora necesario señalar diferencias entre el peso conceptual de cadenas léxicas asociadas a las mismas áreas conceptuales pero en textos destinados a una audiencia distinta. A estas habría que añadir diferencias terminológicas.

Para detectar diferencias conceptuales entre los textos para profesionales de la salud con los destinados a pacientes, se han elaborado dos diagramas de sectores en los que las cadenas léxicas con igual etiqueta se suman entre sí para obtener macrocadenas léxicas. A partir de estas, proponemos una representación conceptual correspondiente a textos para especialistas, por una parte, y textos para no-especialistas, por otra.

DIAGRAMA 34 : Macrocadenas léxicas en los textos para especialistas (QDT)

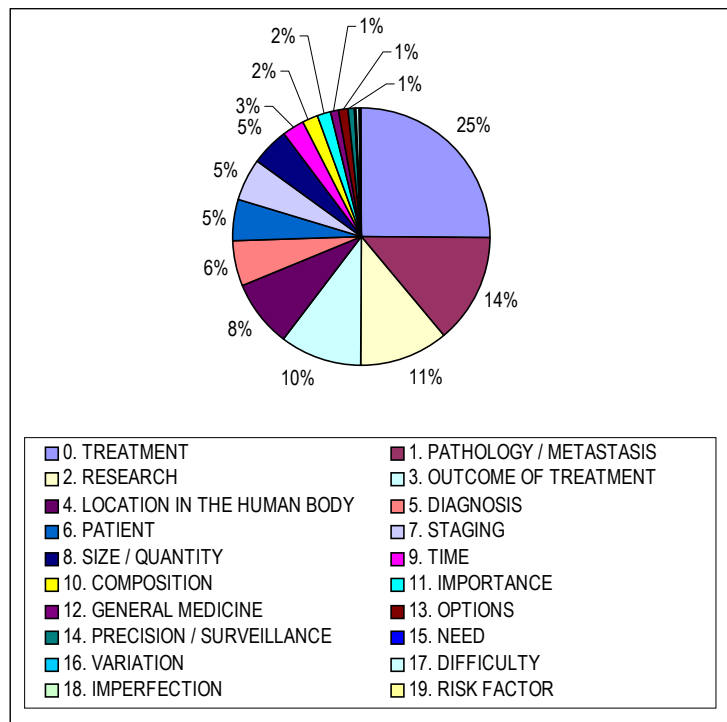
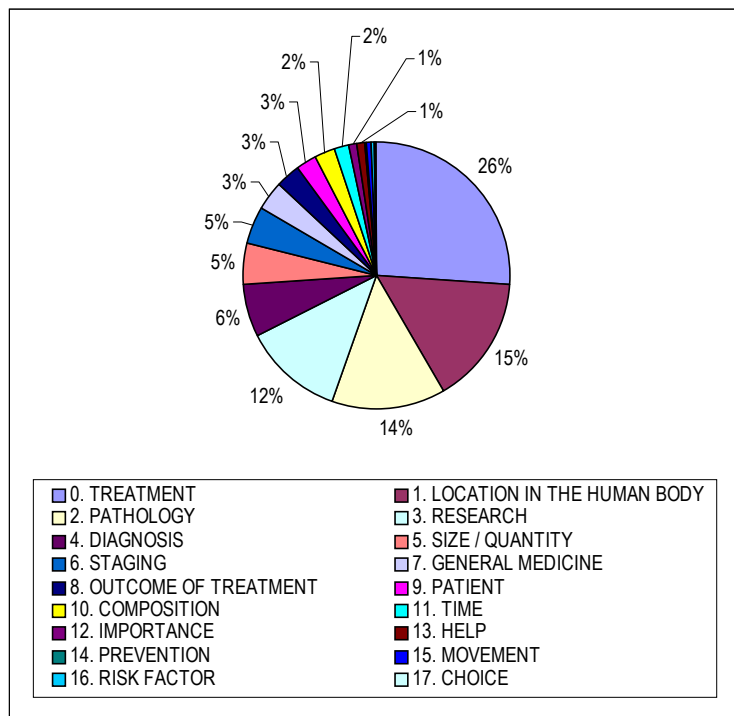


DIAGRAMA 35: Macrocadenas léxicas en los textos para pacientes (QPT)



En los diagramas, se percibe cómo en los textos para especialistas se hace un mayor hincapié en determinadas áreas conceptuales. Por ejemplo, la macrocadena *OUTCOME OF TREATMENT*, que ocupa el cuarto lugar en textos para especialistas, y el noveno en la versión para pacientes, nos indica que los primeros tienen un mayor interés por este concepto, y esto queda de manifiesto en la articulación cohesiva del texto.

TABLA 124: Lexemas encadenados en las macrocadenas de los textos para especialistas sobre TRATAMIENTO

MACROCADENAS	Lexemas encadenados (%)				
0. TREATMENT	1327	25,32	11. IMPORTANCE	87	1,66
1. PATHOLOGY / METASTASIS	712	13,58	12. GENERAL MEDICINE	70	1,33
2. RESEARCH	586	11,18	13. OPTIONS	57	1,1
3. OUTCOME OF TREATMENT	532	10,15	14. PRECISION / SURVEILLANCE	34	0,65
4. LOCATION IN THE HUMAN BODY	433	8,26	15. NEED	14	0,25
5. DIAGNOSIS	315	6,01	16. VARIATION	12	0,22
6. PATIENT	276	5,27	17. DIFFICULTY	9	0,17
7. STAGING	267	5,09	18. IMPERFECTION	8	0,15
8. SIZE / QUANTITY	252	4,8	19. RISK FACTOR	7	0,13
9. TIME	141	2,7		5241	99,96
10. COMPOSITION	102	1,94			

TABLA 125: Lexemas encadenados en las macrocadenas de los textos para pacientes sobre TRATAMIENTO

MACROCADENAS	Lex. encadenados(%)				
0. TREATMENT	525	26,18	8. OUTCOME OF TREATMENT	60	2,99
1. LOCATION IN THE HUMAN BODY	308	15,36	9. PATIENT	53	2,64
2. PATHOLOGY	278	13,86	10. COMPOSITION	46	2,25
3. RESEARCH	242	12,06	11. TIME	37	1,8
4. DIAGNOSIS	127	6,34	12. IMPORTANCE	22	1,1
5. SIZE / QUANTITY	100	4,99	13. HELP	16	0,8
6. STAGING	93	4,64	14. PREVENTION	8	0,4
7. GENERAL MEDICINE	70	3,5	15. MOVEMENT	7	0,35
			16. RISK FACTOR	7	0,35
			17. CHOICE	6	0,29
				2005	99,9

A continuación comentamos estas diferencias conceptuales en las primeras 9 macrocadenas¹⁷⁰, cuyos porcentajes oscilan entre el 26% y el 3% y toman como referencia el número de lexemas que constituyen cada macrocadena.

¹⁷⁰ Los títulos de los apartados que siguen incluyen unas coordenadas con las que se indica el lugar de las macrocadenas en los textos para especialistas y para pacientes respectivamente. La macrocadena a la que

12.1.1. *TREATMENT* (0,0)

Por su extensión, ha sido posible desglosar esta cadena en subcadenas que llegan a un nivel 4 de dependencia, es decir, que cuentan con cuatro dígitos, por ejemplo, *0.4.1.1. CHEMOTHERAPEUTIC AGENTS* del texto QDT1. En los textos para especialistas, las unidades léxicas vinculadas mediante el concepto *TREATMENT* constituyen un 25,3% del total, que se convierte en un 26,2% en la parte para pacientes. Destacamos las cinco subcadenas más importantes y los dos lemas más frecuentes en cada una:

- En textos para especialistas:

CHEMOTHERAPY (27%): *CHEMOTHERAPY, REGIMENS, CISPLATIN*

GENERIC TREATMENT (26%): *TREATMENT, CLINICAL TRIALS*

SURGERY (15%): *RESECTION, SURGERY*

RADIOTHERAPY (12%): *RADIOTHERAPY, IRRADIATION*

PROPERTIES OF TREATMENT (11%): *ALONE, STANDARD*

- En textos para pacientes

GENERIC TREATMENT (27%): *TREATMENT, CLINICAL TRIALS*

MANIPULATION (20%): *USED, GIVEN*

RADIOTHERAPY (18%): *RADIATION THERAPY, <RADIATION>¹⁷¹, PROPHYLACTIC
CRANIAL IRRADIATION*

CHEMOTHERAPY (15%): *CHEMOTHERAPY, DRUGS*

SURGERY (9%): *SURGERY, RADIOSURGERY*

Como cabría esperar, en los textos para pacientes, hay más USE genéricas, muchas de las cuales pertenecen a la lengua general, relacionadas con el TRATAMIENTO, una parcela con la que estos no están muy familiarizados. Por eso, los porcentajes de activación de las subcadenas genéricas (*GENERIC TREATMENT, MANIPULATION*) son los más elevados en estos textos. De hecho, la subcadena *GENERIC TREATMENT* tiene el mismo porcentaje que la más significativa de los textos para especialistas, *CHEMOTHERAPY*. En la subcadena *MANIPULATION*, se incluyen verbos esenciales en las explicaciones de los distintos tratamientos: *USED* y *GIVEN*. Este último lema está referido principalmente a la quimioterapia, como se ve en (105).

se asignan las coordenadas (0, 0) es la más significativa tanto en QDT (primera coordenada) como en QPT (segunda coordenada).

¹⁷¹ Se están considerando los usos en los que no forma colocación.

(105)

1 tain cancer. Radiation therapy uses x-rays or other high-energ
 2 d. Radiation therapy may also be used to prevent the cancer from
 3 ossible. Radiosurgery is usually used as treatment of tumors tha
 4 ncer is treated Chemotherapy uses drugs to kill cancer cells
 5 surgery followed by chemotherapy given inside the chest. 5 A cli

En vista de que en los textos para especialistas, a la subcadena *PROPERTIES OF TREATMENT* le corresponde un porcentaje significativo (11%), inferimos que en estos abundan las UF construidas sobre un sustantivo relacionado con la terapia del cáncer. Estas UF contribuyen a la multidimensionalidad terminológica a la que aludimos en las páginas 306-307. Por ejemplo, al construir el sistema conceptual *RADIOTHERAPY*, encontramos el término *chest radiotherapy* correspondiente a la dimensión *LOCATION OF APPLICATION*, mientras que si el énfasis está en el *OBJETIVO (AIM)*, estaremos ante el término *palliative radiotherapy*.

Presentamos más instancias de multidimensionalidad y de diferencias terminológicas en la tabla 126, que agrupa las *USE* de la subcadena *RADIOTHERAPY* en todos los textos. En la primera columna, la jerga de los especialistas se distribuye en las subcadenas: *RADIOTHERAPY*, *SUBSTANCES USED*, *INSTRUMENTS*, *DOSE OF RADIATION* y *HEALTH PROFESSIONAL*. Las subcadenas correspondientes a los textos para pacientes son sólo *RADIATION THERAPY*, *RADIATION AGENTS*, que es el homólogo genérico para *SUBSTANCES USED*, y por último, *DOSE OF RADIATION*. Es decir, aparecen menos subcadenas y estas son más genéricas porque, para no abrumar al lector, el autor del texto quiere activar las mínimas parcelas del sistema conceptual *RADIOTHERAPY*.

El término *radiotherapy* es el más utilizado por los especialistas, en oposición al *radiation therapy* de los textos para pacientes. *Radiation therapy* es más pedagógico para el lector lego en el sentido de que explica el agente terapéutico (*radiation*), algo que no queda tan claro en el vocablo *radiotherapy*. Algún lector podría interpretar este último como *terapia con la ayuda de la radio*. Por eso sólo aparece una vez.

La mayoría de *USE* vinculadas en esta subcadena se construyen sobre la raíz latina *radio* (radiación), aunque se imponen determinadas construcciones en cada tipo de texto. Los especialistas recurren más a las unidades poliléxicas que, en muchas ocasiones tienen como base un núcleo grecolatino. En QDT, *radiotherapy* está

premodificado en 43 ocasiones sobre 86, y *radiation therapy* en 3 de 6. Este 50% aumenta aún más con los lexemas *irradiation* y *radiation*. En los textos para pacientes, esta preferencia por la premodificación sólo se percibe en las formas léxicas *irradiation* y *radiation*. *Radiation therapy* y *radiotherapy* suelen ir solos.

TABLA 126: Unidades de significación especializada dentro de la subcadena *RADIOTHERAPY*

0.4.1. RADIOTHERAPY (Doctor)	0.4.1. RADIATION THERAPY (Patient)
<p><i>Generic</i> <RADIOTHERAPY> 43 / 86 <RADIATION THERAPY> 3 / 6 <IRRADIATION> 4 / 24 <RADIATION> 1 / 11 <u>FRACTIONATION</u></p>	<p><i>Generic</i> <RADIATION THERAPY> 34 / 47 <RADIATION> 7 / 54 <RADIOTHERAPY> 1 / 1 <u>IRRADIATION</u></p>
<p><i>Source of radiation</i> Internal radiotherapy Brachytherapy 4 Interstitial radiotherapy 1 External-beam radiotherapy 2</p>	<p><i>Source of radiation</i> Internal radiation therapy 5 External radiation therapy 2 External beam radiation therapy 2</p>
<p><i>Interaction with other therapies</i> Radiotherapy alone 14 Postoperative radiotherapy 4 Concurrent radiotherapy 2 Postoperative irradiation 2 Postoperative whole-brain irradiation 2</p>	<p><i>Interaction with other therapies</i> Radiation therapy alone 4</p>
<p><i>Location of application</i> PCI 12 Prophylactic cranial irradiation 2 Chest radiotherapy 9 Chest irradiation 9 <Cranial irradiation> 4 Whole-brain irradiation 2 Whole-brain radiotherapy 2 Chest radiation 1 Chest radiation schedules 1 Thoracic radiation 1 Thoracic irradiation 1</p>	<p><i>Location of application</i> Prophylactic cranial <u>irradiation</u> 5 PCI 2</p>
<p><i>Dose of radiation</i> Conventional <u>fractionation</u> 2 Standard <u>fractionation</u> 2 <u>Fractionation</u> schemes 2 Hypofractionated schemes 1</p>	<p><i>Dose of radiation</i> (High-dose x-rays 3) (High-energy x-rays 2)</p>
<p><i>Aim</i> Radiotherapy with curative intent 6 Palliative radiotherapy 2 Palliative radiation therapy 2</p>	
<p><i>Effectiveness</i> Primary ~ 2 Conventional fractionation 2 Standard fractionation 2 Conventional whole-brain radiotherapy 1 Conventional radiotherapy 1</p>	

0.4.1.1. SUBSTANCES USED	0.4.1.1. RADIATION AGENTS
	Light 4 (beam of ~ 1) Rays 3 High-dose x-rays 3 High-energy x-rays 2 X-ray(s) 2
Radiolabeled antibodies 2	Radioisotopes 2 ≈
Radiosensitizers 2	Materials that produce radiation 2 Drug that makes cancer cells more sensitive to light 1
0.4.1.2. INSTRUMENTS	0.4.1.2. INSTRUMENTS
Megavoltage equipment 2	
0.4.1.3. DOSE OF RADIATION	0.4.1.3. DOSE OF RADIATION
cGy 6 <DOSES> 2 Daily doses 2 Radiotherapy doses 1 Daily fractions 3	(High-dose x-rays 3) (High-energy x-rays 2)
0.4.1.4. HEALTH PROFESSIONAL	0.4.1.4. HEALTH PROFESSIONAL
Radiation oncologist 1	

Los elementos del sistema conceptual RADIOTHERAPY presentes en los textos para especialistas tienen un alto grado de especificidad. Por ejemplo, encontramos dos tipos de radioterapia interna, *brachytherapy* e *interstitial radiotherapy*, pero no su término superordinado (*internal radiotherapy*), que sí aparece modificado en la versión para pacientes (*internal radiation therapy*). En definitiva, si se creara una ontología sobre terapias oncológicas a partir de las USE del texto, las representaciones de los conceptos tendrían un alto grado de granularidad y pertenecerían a diversas dimensiones, es decir, apuntarían hacia las diversas perspectivas desde las que se puede considerar el concepto RADIOTHERAPY. No obstante dicha granularidad, en los textos especializados también existe una presencia elevada de unidades léxicas de la lengua general, como indica el hecho de que la subcadena *GENERIC TREATMENT* sea la segunda más representativa en QDT. Esta subcadena contiene vocablos como *treatment*, *trials*, *curative*, *use*, *given*, etc. Algunos de ellos restringen su significado en este subdominio dado que el *cotexto*, es decir, el contexto lingüístico inmediatamente anterior y posterior

a una determinada unidad léxica y terminológica¹⁷² funciona como restrictor semántico, como se verá en 14.4.

Por el contrario, en los textos para pacientes, la ontología contaría con bastantes conceptos genéricos pero quedaría vacía en los niveles inferiores y en muchas dimensiones. Esto es lo que ocurre cuando se prescinde de los tipos de radioterapia según el lugar donde se aplica (*thoracic radiation, chest radiation, etc.*) o de los tipos de radioterapia interna (*brachytherapy e interstitial radiotherapy*).

Hemos percibido también una tendencia a la variación terminográfica para referirse a un mismo concepto. Por ejemplo, el concepto RADIOTHERAPY se expresa con tres formas distintas: *radiation therapy, radiation y radiotherapy*. Este tipo de variación no debe confundirse con la de los textos para especialistas, que obedece a una mayor complejidad conceptual.

12.1.2. PATHOLOGY (1, 2)

La nomenclatura de la enfermedad recibe mayor atención en los textos para profesionales de la salud. Al nombrar el cáncer, se hace hincapié en tres aspectos: el hecho de ser una enfermedad (*disease*), su localización anatómica (*lung cancer*) y/o su localización histológica (*non-small cell lung cancer, large cell carcinoma*).

Aunque la clasificación de la enfermedad según el tipo de tejido donde se asienta el tumor no es usual en textos de divulgación general (Tercedor 1999: 262), esta sí está presente en los textos para pacientes. Esto se debe a que los pacientes han oído el nombre de su enfermedad y, aun sin saber las diferencias entre los tipos de cáncer de pulmón, son conscientes de que llevan consigo diferencias en cuanto a complicaciones clínicas, pronóstico y tratamiento.

También se hace alusión a la evolución de la enfermedad, aunque con términos diferentes: *spread of cancer* es la explicación que se da al paciente del concepto METASTASIS. Como METASTASIS era una subcadena de PATHOLOGY en todos los textos, a excepción del QDT2, hemos incluido en PATHOLOGY la cadena METASTASIS de QDT2.

¹⁷² No utilizamos *contexto* en el sentido que se le da en terminología, es decir, la oración que contiene el término en cuestión. Creemos que la restricción semántica se produce gracias a elementos léxicos que están más allá de los límites de la oración.

La cadena *PATHOLOGY* cuenta con un 13,6% de las USC vinculadas en los textos para especialistas y con un 13,8% en los textos para pacientes. Veamos las tres subcadenas más importantes, su porcentaje con respecto al total de elementos de la cadena y los lemas más frecuentes en cada una:

- En textos para especialistas:

CANCER (61%): TUMOR, LUNG CANCER, SMALL CELL LUNG CANCER, NSCLC, MALIGNANT MESOTHELIOMA

GENERIC DISEASE (20%): DISEASE, SYMPTOMS

METASTASIS (19%): METASTASIS, RECURRENT

- En textos para pacientes

CANCER (73%): CANCER, NON-SMALL CELL LUNG CANCER, MALIGNANT MESOTHELIOMA, SMALL CELL LUNG CANCER

SPREAD OF CANCER (17%): SPREAD, RECURRENT

GENERIC DISEASE (9%): SYMPTOMS, DISEASE

En *PATHOLOGY*, las USE genéricas se subordinan a las más específicas puesto que hay un interés por incluir la nomenclatura de los distintos tipos de cáncer. Los síntomas tienen más relevancia en los textos para pacientes, que son los que los sufren.

Por último, es posible perfilar claras diferencias terminológicas entre las designaciones más usuales entre los oncólogos y sus equivalentes, en cuanto que apuntan al mismo concepto, en textos para pacientes.

(106)

TÉRMINOS ESPECIALIZADOS

Metastasis
Tumor
NSCLC

DESIGNACIÓN EN TEXTOS PARA PACIENTES

spread of cancer
cancer
Non-small cell lung cancer

12.1.3. RESEARCH (2, 3)

Esta cadena incluye gran parte del vocabulario semiespecializado común a la mayoría de los textos académicos (*vocabulario organizador del discurso*). Aunque ocupe el tercer lugar en QDT y el cuarto en QPT, el porcentaje es mayor en estos últimos (12,1% frente al 11,1% en textos para profesionales de la salud).

En los textos aparecen las siguientes subcadenas, que están presentes normalmente en la mayoría de los textos científicos: *STUDY*, *INFORMATION*, *COGNITION*, *COMPARISON*, *CAUSE*, *MEASURE*, *NORMALITY*, *CATEGORIZATION*, *EVIDENCE*, *EVALUATION*, *COMMUNICATION*. Las tres subcadenas con mayor número de elementos en orden de importancia son las siguientes:

- En textos para especialistas
STUDY (37%): (CLINICAL) TRIALS, STUDIES, REPORTED
CATEGORIZATION (19%): CLASSIFICATION, DEFINE
EVIDENCE (15%): SHOWED, EVIDENCE
- En textos para pacientes
STUDY (41%): CLINICAL TRIALS, FOUND
CATEGORIZATION (23,1%): CALLED, TYPE
INFORMATION / COGNITION (22,7%): INFORMATION, LEARN

Hay coincidencia en las dos subcadenas más importantes y en el lema más frecuente de la subcadena *STUDY*: CLINICAL TRIALS. Aunque CLINICAL TRIALS pertenece también a la cadena *TREATMENT*, el carácter experimental de estos nuevos tratamientos nos ha llevado a incluirlo aquí. El resto de los lemas muestran claramente el carácter académico de los textos destinados a especialistas.

También se detecta cómo la EVIDENCIA es un valor sobresaliente entre los especialistas, y la INFORMACIÓN, entre los pacientes. En efecto, la evidencia que se extrae de las pruebas diagnósticas indica el tipo de cáncer y su estadio, y el tratamiento se fija considerando la evidencia que aportan diferentes estudios sobre la efectividad de las distintas terapias.

12.1.4. OUTCOME OF TREATMENT (3, 8)

El RESULTADO DEL TRATAMIENTO está más activado en los textos para especialistas porque son estos los que tienen que determinar el tratamiento de forma que sea lo más beneficioso posible para sus pacientes. El principal factor que indica al oncólogo el resultado exitoso del tratamiento es la supervivencia. También es fundamental el

tratamiento de los síntomas. Los 532 lexemas vinculados en esta cadena se agrupan en cuatro subcadenas, de las que indicamos su porcentaje y los lexemas más frecuentes:

POSITIVE OUTCOME (37%): IMPROVE (+ SURVIVAL), CURATIVE

EXISTENCE (31%):

SURVIVAL: (LONG-TERM) SURVIVAL

DEATH: DIE

GENERIC OUTCOME (25%): RESPONSE, RATES

NEGATIVE OUTCOME (7%): POOR, RISK, TOXICITY

En los textos para pacientes, el único resultado positivo del tratamiento que se destaca es su poder paliativo. La supervivencia y la muerte no se mencionan para no desmoralizar al paciente. Por eso, sólo se considera una perspectiva a corto plazo: el ataque a las células cancerosas, el alivio de síntomas, y la posibilidad de efectos secundarios. Las dos subcadenas que hemos identificado son:

POSITIVE OUTCOME (90%): KILL CANCER CELLS, RELIEVE (+ SYMPTOMS / PAIN)

NEGATIVE OUTCOME (10%): SIDE EFFECTS

Para suplir la ausencia de las formas léxicas del lema *SURVIVAL* e infundir ánimo a los pacientes, en estos textos, el cáncer no es el que mata sino el tratamiento (*kill cancer cells*). Esto es parte del lenguaje combativo típico de textos de divulgación sobre oncología en los que el tratamiento se concibe como una lucha, una guerra o una batalla contra el cáncer. En esta línea también está la expresión *fight cancer* que aparece en la cadena *TRATAMIENTO* de los textos para pacientes. Para ilustrar este lenguaje combativo, empezaremos citando unas palabras del periodista José Ayala:

GUERRA. Cada vez que se habla con un investigador sobre el cáncer, o con un paciente, siempre sale esta misma palabra. Quizá con ninguna otra enfermedad tenemos esta misma sensación de que luchamos contra alguien ajeno a nosotros, contra un odioso alien infiltrado. Sin embargo, el enemigo de esta contienda implacable lo producimos nosotros mismos.

(Revista *El Semanal* 25/10/1998)

Asimismo, incluimos algunas líneas de concordancia en las que subyace esta metáfora de la guerra. Se han extraído de la sección del corpus (véase p.256-257) que contiene artículos de divulgación sobre cáncer publicados en las revistas *Blanco y Negro*, *El Semanal*, *Investigación y Ciencia*, *Muy Interesante* y *Quo*, y en los periódicos *El Mundo e Ideal (Suplemento Campus)* y que cuenta con 60542 palabras. Las líneas girarán en torno a los lemas GUERRA, LUCHA, BATALLA, GANAR.

(107)

GUERRA

1 comprometió a los EE.UU. en una "guerra" contra el cáncer. Dura
2 lentador saber que, aunque en la guerra contra tan terrible morb
3 nas noticias, va a ser una larga guerra. Guerra Cada vez que
4 s, va a ser una larga guerra. Guerra Cada vez que se habla co
5 se llama metástasis. Y es una guerra masiva: cada año se diag
6 siendo nuestra mejor arma en la guerra contra el cáncer, y pasa
7 Es verdad que no hemos ganado la guerra. Pero muchas batallas va
8 velocidad que muchos querían, la guerra contra el cáncer progres
9 tigadores en el mundo entero, la guerra está aún lejos de ser ga
10 n casi todos los frentes de esta guerra a muerte que la Humanida

LUCHA

1 plo de los nuevos enfoques en la lucha contra el cáncer. Los méd
2 emos esta misma sensación de que luchamos contra alguien ajeno a
3 . Así describía los métodos para luchar contra los tumores Aulo
4 valenciana que lleva nueve años luchando valientemente contra s
5 experta Thais Monros, nueve años luchando contra la enfermedad,
6 l "equipo médico habitual" en la lucha contra los tumores. La co
7 ente en fase experimental. La lucha contra el cáncer no está
8 , que podría ser utilizada en la lucha contra el cáncer.
9 e podría resultar muy útil en la lucha contra el cáncer y empezó
10 miento. Nuevos frentes. En la lucha contra el cáncer, existen
11 que pueden ser utilizados en la lucha contra el cáncer. Desde m

BATALLA

1 os conseguidos en este frente de batalla. Dejando a un lado el c
2 de las mayores esperanzas en la batalla contra el cáncer sean l
3 nuel Nieves> CÁNCER. HACIA LA BATALLA FINAL. Aunque no a l
4 sin que nadie la pueda parar: la batalla final contra el cáncer
5 idos desde entonces se ha estado batallando contra esta enfermed
6 os ganado la guerra. Pero muchas batallas van siendo nuestras".

GANAR

1 a la vida, amándome a mí misma, ganaré a estas malditas células
2 de piel. Es verdad que no hemos ganado la guerra. Pero muchas b
3 la guerra está aún lejos de ser ganada. Las tentaciones de fals

MISCELÁNEA

1 ado nunca una auténtica ofensiva contra el cáncer: "El total de
 2 o de cáncer. **NUEVAS ARMAS CONTRA EL CÁNCER** Si el día
 3 nto al bisturí las armas básicas contra el cáncer. Las dosis tie
 4 imero, los elementos que atentan contra los genes que controlan
 5 en los diversos frentes abiertos contra esta enfermedad. Durante
 6 la defensa natural del organismo contra el cáncer son obtenidos

12.1.5. LOCATION IN THE HUMAN BODY (4, 1)

Como ya se dijo, los conceptos anatómicos están muy presentes en los textos sobre medicina. Lo curioso es que su presencia es mayor en textos para pacientes, donde la cadena *LOCATION IN THE HUMAN BODY* es la segunda más extensa (15,4%), en contraposición al quinto lugar que ocupa en textos para especialistas (8,3%). Pensamos que esto es así porque el paciente quiere conocer dónde está su enfermedad y dónde le van a aplicar el tratamiento.

En los textos para pacientes, el nivel celular e histológico recibe menor atención que el anatómico. Las únicas referencias a las células y los tejidos tienen que ver con el nombre de la enfermedad y con la UF *cancer cells*.

En el cuadro contrastivo de abajo, se presentan las diferencias terminológicas y de activación conceptual entre QDT y QPT. Aparecen alineados los términos que designan el mismo concepto en los dos tipos textuales, precedidos por un número que indica qué términos son más frecuentes. En algunos casos, no existirá ese concepto en el texto divulgativo, que tiende a la simplificación, como se aprecia en estos ejemplos, donde se codifica la misma información para un lector distinto:

(108) ----- Limited stage

Limited stage small cell lung cancer means tumor confined to the hemithorax of origin, the mediastinum, and the supraclavicular nodes [...]. [QDT3]

(109) ----- Limited stage

Cancer is found only in one lung and in nearby lymph nodes. [QPT3]

TABLA 127: Cuadro contrastivo con la activación conceptual y terminológica de la macrocadena *LOCATION IN THE HUMAN BODY*

	Activación conceptual (%) y terminológica (Doctor)		Activación conceptual (%) y terminológica (Patient)	
ANATOMIC LEVEL / BODY PARTS	58	1 lung / pulmonary 2 brain / cerebral 3 (lymph) nodes 4 chest / thorax / thoracic # peritoneum	50	1 lung 5 brain 4 lymph nodes 2 chest 3 abdominal
GENERIC LOCATION	24	1 local 2 distant 3 regional 4 ipsilateral 5 contralateral	24	3 places # far / distant 1 body 2 internal # on the other side of + BODY PART
MICROSCOPIC/ HISTOLOGIC LEVEL	9	1 histology /tissue 2 cell	17	2 tissue 1 cancer cells
LOCATION IN RELATION TO SPECIFIC BODY PARTS	8	1 endobronchial 2 supraclavicular	1	
DEFINITION OF BODY PARTS	-		4	1 lining 2 shaped
BODY FLUIDS	1	1 lymph 2 blood	4	1 fluid 2 bloodstream

En la subcadena *ANATOMIC LEVEL /BODY PARTS* de textos para especialistas, se percibe una variedad léxica mayor que en otras subcadenas. Se debe a la utilización de tecnicismos grecolatinos junto con la forma comúnmente empleada en inglés. Los tecnicismos están a veces prefijados (*intrathoracic, endobronchial*) y preceden a sustantivos relacionados con alguna dolencia o con el tratamiento, como vemos en el ejemplo (110). La premodificación nominal permite la clasificación de los tratamientos según el criterio de localización (*endobronchial laser therapy*) y contribuye a la multidimensionalidad de conceptos típica del área de la oncología (Tercedor 1999: 263).

(110) Some patients with intrinsic **endobronchial obstructing lesions** or extrinsic compression due to tumor have achieved successful palliation with **endobronchial laser therapy** (for endobronchial lesions only).

En los textos para pacientes, no aparecen los vocablos grecolatinos—a excepción de *abdominal*—y las unidades léxicas referentes al cuerpo humano suelen formar parte de sintagmas preposicionales (*outside the body, inside the lungs*). Los

sintagmas preposicionales con *inside* y *outside* son la alternativa a los términos formales *intrinsic*, *extrinsic*, *internal*, *external*. Parece pues que algunas palabras plenas de las cadenas léxicas para especialistas tienen su contrapartida en sintagmas preposicionales dentro de los textos para pacientes.

12.1.6. DIAGNOSIS (5, 4)

El diagnóstico tiene una activación un poco mayor en los textos para pacientes. Estos últimos precisan una explicación detallada de los tests diagnósticos a los que van a someterse, algunos de los cuales pueden causar incomodidad.

Las subcadenas que presentamos a continuación destacan dos perspectivas. Desde la perspectiva del profesional, el diagnóstico es el resultado de la interpretación y evaluación de una serie de pruebas. Los procesos cognitivos en la mente del oncólogo son necesarios para llegar a un diagnóstico y a un pronóstico de la enfermedad acertados. Desde la perspectiva del paciente, lo que se pone en juego en el diagnóstico son una serie de pruebas (*tests*, *biopsy*) y procesos sensoriales (*found*, *look*) que afectan a partes de su cuerpo. La perspectiva de los afectados por el cáncer también queda clara en el hecho de que el síntoma más frecuente en los textos para pacientes sea el dolor (*pain*).

Encontramos evidencia de esto en la disposición de las subcadenas y los lemas más activados:

- En textos para especialistas

CLINICAL FEATURES (28%): SYMPTOMS, PULMONARY FUNCTION
COGNITIVE PROCESSES INVOLVED (24%): EVALUATION, CONSIDERED
GENERIC DIAGNOSIS (15%): DIAGNOSIS, PROGNOSIS
DIAGNOSTIC TESTS (14%): TOMOGRAPHY, BRONCHOSCOPY

- En textos para pacientes

SENSORY PROCESSES (34%): FOUND, LOOK
CLINICAL FEATURES (31%): SYMPTOMS, PAIN
DIAGNOSTIC TESTS (27%): TESTS, MICROSCOPE, BIOPSY

12.1.7. PATIENT (6, 9)

Las alusiones al paciente son numerosas en los textos para especialistas, de forma que representan el 5,26% de las palabras plenas de los textos QDT1, QDT2 y QDT3. El porcentaje es inferior (2,64%) en su versión para pacientes. Los dos lemas que predominan en ambos tipos de texto son PATIENTS y GROUP.

El lexema *patients* coocurre en numerosas ocasiones con el participio pasado *selected*, que indica que sólo los pacientes que cumplan una serie de requisitos clínicos pueden participar en determinados tratamientos. En su cotexto también aparece la palabra *survival*. Sin embargo, el lexema *selected* no se incluye en textos para pacientes. Como dijo al hilo del texto QPT1, la idea de SELECCIÓN puede resultar al paciente contraria a la deontología médica. Además, la idea de supervivencia no aparece en los textos para pacientes.

12.1.8. STAGING (7, 6)

La clasificación por estadios del cáncer recibe una mayor atención en la literatura médica: 5,1% frente al 4,6% de la sección para pacientes. En lo que sí se encuentran coincidencias es en los lemas más frecuentes: STAGE, LIMITED, EXTENSIVE, RECURRENT.

Los términos *limited* y *extensive* son muy frecuentes y son propios del cáncer microcítico de pulmón. Podemos interpretar la elevada frecuencia de los mismos teniendo en cuenta que este tipo de cáncer es el que tiene un curso más agresivo, y por tanto, es imprescindible determinar el estadio lo más acertadamente posible.

12.1.9. SIZE / QUANTITY (8, 5)

La activación de esta macrocadena no varía apenas de un tipo de texto a otro, aunque en los textos para pacientes analizados, su rango es superior: es la sexta cadena más importante, mientras que es la novena en los textos para especialistas. Las subcadenas más relevantes son las que se indican a continuación:

- En textos para especialistas

LARGENESS, LARGE QUANTITY (29%): HIGH, GREATER

NAME OF LUNG CANCER (27%): SMALL, NON-SMALL

MINUTENESS (17%): SMALL, REDUCED

INCREASE (16%): ADDITION, INCREASED

- En textos para pacientes:

NAME OF LUNG CANCER (45%): NON-SMALL, SMALL

MINUTENESS (27%): <SMALL>, SHRINK

INCREASE (21%): SPREAD, GROW

La idea *SIZE / QUANTITY* se actualiza principalmente en la subcadena *LARGENESS* de los textos para especialistas, en la que se evalúa principalmente la efectividad de los tratamientos y en *NAME OF LUNG CANCER* correspondiente a la nomenclatura de la enfermedad, sobre todo en los textos para pacientes. En los textos para especialistas, *small* predomina sobre *non-small* porque, como ya dijimos, el término más frecuente para designar la enfermedad es el acrónimo *NSCLC*. En la versión para pacientes ocurre lo contrario.

También se percibe una preocupación por reducir las complicaciones de la enfermedad, en los textos *QDT*, o el tamaño del tumor, en los textos *QPT*. Esto justifica que los lemas más usuales sean *REDUCED* en los primeros y *SHRINK* en los segundos:

(111) The risk of developing central nervous system metastases can be **reduced** by more than 50% by the administration of PCI in doses of 2400 cGy [22]. [QDT3]

(112) Radiation therapy may be used to **shrink** the cancer and to relieve pain. [QPT2]

Asimismo, se produce una activación de los conceptos *SIZE / QUANTITY* mediante la subcadena *INCREASE*, a la que corresponde un 17% y un 21% de toda la macrocadena. Para los especialistas, el concepto *INCREASE* está en el mismo cotexto que *survival* y *mortality* principalmente, en secciones que evalúan los resultados de un determinado tratamiento, como indica el ejemplo:

(113) [...] adjuvant combination chemotherapy with cisplatin, doxorubicin, and cyclophosphamide produced modestly **increased disease-free survival**. [QDT2]

En cuanto a los textos para pacientes la subcadena *INCREASE* alude a la naturaleza expansiva de la enfermedad, y por eso los lemas más frecuentes asociados son *SPREAD* y *GROW*.

12.1.10. RESTO DE LAS MACROCADENAS

De las demás macrocadenas, que juntas representan un 15% de la activación conceptual, no añadiremos apenas nada a lo dicho en la explicación de las cadenas de las secciones 11.1. y 11.2. Solamente diremos que estas contienen o bien unidades léxicas de la lengua general o bien términos médicos bastante genéricos (*clinical, prevention*).

En la tabla se indica el lugar que ocupan en relación con otras, ya sea en textos para especialistas (DT) o en textos para pacientes (PT), algo que antes se especificaba con un par de coordenadas. Presentamos en primer lugar aquellas que contienen lexemas de la lengua general, es decir, desde *TIME* hasta *MOVEMENT* y después las relacionadas con la medicina general. La activación conceptual de estas macrocadenas oscila entre el 2,7% y el 0,1%.

TABLA 128: Macrocadenas léxicas que activan un 15% de los conceptos.

MACROCADENAS LÉXICAS	RANGO		MACROCADENAS LÉXICAS	RANGO	
	DT	PT		DT	PT
<i>TIME</i>	9	12	<i>IMPERFECTION</i>	18	-
<i>COMPOSITION</i>	10	10	<i>INFORMATION</i>	-	11
<i>IMPORTANCE</i>	11	13	<i>HELP</i>	-	14
<i>OPTIONS/CHOICE</i>	13	18	<i>MOVEMENT</i>	-	16
<i>PRECISION/SURVEILLANCE</i>	14	-			
<i>NEED</i>	15	-	<i>GENERAL MEDICINE</i>	12	7
<i>VARIATION</i>	16	-	<i>RISK FACTOR</i>	19	17
<i>DIFFICULTY</i>	17	-	<i>PREVENTION</i>	-	15

12.2. Recapitulación

El desarrollo de las cadenas, cuyos lemas hemos etiquetado con un sustantivo en la mayoría de los casos, deja patente la supremacía de los sustantivos sobre los verbos. Esta puede explicar en parte la tendencia de los diccionarios especializados a contener principalmente sustantivos y sintagmas nominales.

Parece confirmado que existe un paralelismo entre cadenas léxicas y áreas conceptuales. Intuimos que también existe un paralelismo entre las cadenas léxicas que elabora el investigador a partir de las USC del texto y la activación conceptual que se produce en la lectura. El lector recrea un texto coherente a partir de la cohesión de los lexemas del texto.

Por otra parte, las similitudes y diferencias conceptuales obedecen obviamente a la idiosincrasia del tema tratado y también al factor *destinatario del texto*. Cuando el redactor piensa en su lector, focaliza los lexemas del texto hacia determinados aspectos del subdominio. Esto tiene consecuencias terminológicas relacionadas con la cohesión y la variedad léxica. La variedad léxica en los textos para especialistas obedece a la complejidad conceptual designada y a la multidimensionalidad, mientras que en los textos divulgativos, responde al intento de definir el concepto con hipónimos, hiperónimos y comparaciones para facilitar la comprensión.

Por último, para explicar la interacción entre cadenas léxicas y activación conceptual, se nos ocurre una metáfora orográfica. El texto es una cadena montañosa que se alza sobre distintos estratos que no se ven pero que sustentan la montaña. Estos estratos serían la estructura conceptual de un determinado campo del saber. Las palabras plenas del texto son los picos montañosos que destacan y que ayudan al observador—en nuestro caso, lector—a percibir un perfil de la cordillera o del texto, donde realmente sólo hay una concentración de materia o de palabras. El perfil que construiría el lector serían las cadenas léxicas del texto.

TABLA 122 : Cadenas léxicas correspondientes a los textos para especialistas (QDT)

	MESOTHELIOMA	%	NON-SMALL CELL LUNG CANCER	%	SMALL CELL LUNG CANCER	%
Activación conceptual superior al 4,75 %	0. TREATMENT	23,06	0. TREATMENT	25,02	0. TREATMENT	26,51
	1. RESEARCH / EXPERIMENTATION	14,19	1. PATHOLOGY OF NSCLC	12,15	1. PATHOLOGY OF SMALL CELL LUNG CANCER	11,1
	2. PATHOLOGY OF MALIGNANT MESOTHELIOMA	11,45	2. RESEARCH / EXPERIMENTATION	10,84	2. RESEARCH, EXPERIMENTATION	10,65
	3. DIAGNOSIS	8,87	3. OUTCOME OF TREATMENT	10,17	3. OUTCOME OF TREATMENT	10,54
	4. OUTCOME OF TREATMENT	8,87	4. LOCATION IN THE HUMAN BODY	8,43	4. LOCATION IN THE HUMAN BODY	7,87
	5. LOCATION IN THE HUMAN BODY	8,55	5. PATIENT	6,09	5. SIZE / QUANTITY	6,6
	6. STAGING	7,42	6. DIAGNOSIS	5,85	6. DIAGNOSIS	5,27
		82,41	7. STAGING	4,75	7. TIME	4,88
			83,3	8. STAGING	4,82	
					88,24	
Activación conceptual superior al 1 %	7. PATIENT	3,06	8. SIZE / QUANTITY	4,04	9. PATIENT	4,71
	8. SIZE / QUANTITY	3,06	9. METASTASIS	3,47	10. IMPORTANCE	2,83
	9. COMPOSITION	2,74	10. COMPOSITION	1,88	11. COMPOSITION	1,77
	10. TIME	2,1	11. GENERAL MEDICINE	1,49	12. GENERAL MEDICINE	1,05
	11. OPTIONS	1,94	12. TIME	1,42		
	12. GENERAL MEDICINE	1,45	13. IMPORTANCE	1,2		
	13. DIFFICULTY	1,13	14. OPTIONS	1,2		
	14. IMPORTANCE	1,13	15. PRECISION / SURVEILLANCE	1,2		
Activ. concept. inferior al 1 %	15. VARIATION	0,65	16. NEED	0,28	13. OPTIONS	0,61
	16. RISK FACTOR	0,32	17. IMPERFECTION	0,25	14. NEED	0,33
			18. VARIATION	0,14	15. VARIATION	0,22
			19. RISK FACTOR	0,11	16. DIFFICULTY	0,11
					17. SMOKING	0,11
	99,99		99,98		99,98	

TABLA 123: Cadenas léxicas correspondientes a los textos para pacientes (QPT)

	MESOTHELIOMA	%	NON-SMALL CELL LUNG CANCER	%	SMALL CELL LUNG CANCER	%
Activación conceptual próxima al 3,5 %	0. TREATMENT	23,33	0. TREATMENT	32,26	0. TREATMENT	23,6
	1. LOCATION IN THE HUMAN BODY	16,03	1. PATHOLOGY OF NSCLC	16,13	1. LOCATION IN THE HUMAN BODY	18,15
	2. PATHOLOGY OF MESOTHELIOMA	12,69	2. LOCATION IN THE HUMAN BODY	11,77	2. PATHOLOGY OF SMALL CELL LUNG CANCER	13,04
	3. RESEARCH / EXPERIMENTATION	11,41	3. RESEARCH / EXPERIMENTATION	8,87	3. RESEARCH / EXPERIMENTATION	9,57
	4. DIAGNOSIS	7,31	4. SIZE / QUANTITY	7,42	4. DIAGNOSIS	8,74
	5. GENERAL MEDICINE	5,51	5. STAGING	7,1	5. SIZE / QUANTITY	7,26
	6. INFORMATION	5,13		83,55	6. STAGING	3,63
	7. STAGING	3,46			83,99	
		84,87				
Activación conceptual superior al 0,65 %	8. OUTCOME OF TREATMENT	2,82	6. PATIENT	3,55	7. OUTCOME OF TREATMENT	3,46
	9. TIME	2,82	7. DIAGNOSIS	2,74	8. GENERAL MEDICINE	2,97
	10. PATIENT	2,18	8. OUTCOME OF TREATMENT	2,74	9. PATIENT	2,31
	11. COMPOSITION	2,05	9. COMPOSITION	2,58	10. COMPOSITION	2,31
	12. IMPORTANCE	1,66	10. GENERAL MEDICINE	1,45	11. TIME	1,32
	13. HELP	1,66	11. TIME	1,13	12. IMPORTANCE	1,15
	14. SIZE / QUANTITY	1,15	12. RISK FACTOR	0,65		13,52
		14,34		14,84		
Activ. concept. inferior a 0,65	15. PREVENTION	0,51	13. CHOICE	0,48	13. CHOICE	0,49
	16. MOVEMENT	0,26	14. PREVENTION	0,48	14. SMOKING (RISK FACTOR)	0,49
			15. IMPORTANCE	0,32	15. HELP	0,49
			16. MOVEMENT	0,32	16. PREVENTION	0,49
				17. MOVEMENT	0,49	
		99,98		99,99	99,96	

13. El papel de la cohesión en la construcción del conocimiento: recursos cohesivos en textos de divulgación

Como se ha visto en el capítulo 12, los textos para pacientes prácticamente se refieren a los mismos conceptos que los textos para especialistas. No obstante, lo hacen mediante recursos lingüísticos que hacen asequible el vocabulario de los especialistas al lector lego.

Estos recursos son propios de los textos de divulgación científica y apoyan nuestra afirmación de que los textos analizados ilustran perfectamente los rasgos de la divulgación, en la que se intenta reducir el vacío cognitivo que separa al público lego de la ciencia.

El vocabulario de los especialistas también está presente en los textos para pacientes, arropado por las explicaciones pertinentes. La inclusión del vocabulario para expertos no es incompatible con la preferencia por términos semi-especializados. En efecto, vimos cómo en los textos para pacientes se explota mucho más el término semi-especializado *radiation therapy* que el especializado *radiotherapy*.

Aunque algunos de estos recursos cohesivos también aparezcan en QDT, se ha constatado que su cotexto no es el mismo. Básicamente, en los textos para profesionales no se produce la duplicidad *término especializado – lexema de la lengua general (término superordinado)*. Cuando hay que vincular cohesivamente dos conceptos se suele utilizar la misma palabra y no va seguida de ninguna definición, como indica el ejemplo:

- (114) **Stage** has a critical role in the selection of therapy. The **stage** of disease is based on a combination of clinical (physical examination, radiologic, and laboratory studies) and pathologic (biopsy of lymph nodes, bronchoscopy, mediastinoscopy, or anterior mediastinotomy [1]). The distinction between clinical **stage** and pathologic **stage** should be considered when evaluating reports of survival outcome. [QDT2]

Si en el texto anterior, la cohesión se crea mediante la repetición exacta del lexema *stage*, en los textos para pacientes, el término *stage* va acompañado de una explicación con unidades léxicas de la lengua general. La cohesión se crea mediante

repetición exacta (*stage-stage*), repetición variada sintáctica (*stage-staging*) y mediante paráfrasis (*stage-whether it is just in the lung or has spread to other places*).

- (115) The chance of recovery (prognosis) and choice of treatment depend on the **stage** of the cancer (whether it is just in the lung or has spread to other places), and the patient's gender and general state of health.

STAGE EXPLANATION

----- **Stages** of small cell lung cancer

Once small cell lung cancer has been found, more tests will be done to find out if cancer cells have spread from one or both lungs to other parts of the body (staging). [QPT3]

Antes del listado de recursos, presentamos un fragmento de la sección *What is small cell lung cancer* del texto QPT3. En este se ilustra la progresión del conocimiento sobre el diagnóstico del cáncer a lo largo del párrafo de forma que se intuye cómo el lector lego ancla la información médica nueva para él con información conocida. Hemos añadido números a cada oración.

- (116) **1** If there are symptoms, a doctor may want to look into the bronchi through a special instrument, called a **bronchoscope**, that slides down the throat and into the bronchi. **2** This test, called **bronchoscopy**, is usually done in the hospital. **3** Before the test, the patient will be given a **local anesthetic** (a drug that causes a loss of feeling for a short period of time) in the back of the throat. **4** Some pressure may be felt, usually with no pain. **5** The doctor can take cells from the walls of the bronchi tubes or cut small pieces of tissue to look at under the microscope to see if there are any cancer cells. **6** This is called a biopsy.

En la primera oración, para hacer asequible el término desconocido por el paciente (*bronchoscope*), este está precedido por un sustantivo de la lengua general (*instrument*) del que es hipónimo. Le sigue una explicación de las partes del cuerpo donde se aplica el broncoscopio. Una vez que el lector sabe lo que es un broncoscopio, se aprovecha la ocasión para que el lector también aprenda el nombre de la prueba diagnóstica (*bronchoscopy*). Para evitar confusiones, un término superordinado (*test*) precede de nuevo al término desconocido (*broncoscopy*).

En la tercera oración, se repite otra vez la palabra *test*, que tiene aquí un valor catafórico y además refuerza la idea de que el párrafo trata sobre pruebas a las que el

paciente va a ser sometido. De ahí el interés por la localización en el cuerpo humano y por procesos sensoriales (*feeling, felt, pain, look, see*), algo que ya se mencionó en 12.1.6. en relación con la macrocadena *DIAGNOSIS*. También en la tercera oración se incluye el término *local anesthetic* con su definición entre paréntesis.

En la quinta oración, se detallan los pasos de una prueba diagnóstica, cuyo nombre especializado (*biopsy*) aparece en la sexta oración.

A continuación se enumeran los recursos que, en los textos para pacientes, acercan el lenguaje del especialista al lector lego. Al igual que hasta ahora, se presentarán los términos que pueden crear potenciales problemas de comprensión en negrita, y con el subrayado, identificamos los lexemas de la lengua general y los hiperónimos que van a facilitar la comprensión. Los ejemplos tienen un doble formato: líneas de concordancia y, cuando es necesario incluir más contexto, oraciones y párrafos completos.

13.1. Expresiones *punte* entre el lenguaje del especialista y el del lector lego

13.1.1. Explicación + (término especializado)

En este apartado se incluyen ejemplos en los que la explicación de un término mediante palabras conocidas por el lector lego va seguida de ese término entre paréntesis. En la explicación puede aparecer un término superordinado del concepto definido (*treatment, section*).

- (117) PDQ also provides information about research on new treatments (**clinical trials**). [QPT1]
- (118) Malignant mesothelioma, a rare form of cancer, is a disease in which cancer (**malignant**) cells are found in the sac lining the chest (the **pleura**) or abdomen (the **peritoneum**). Most people with malignant mesothelioma have worked on jobs where they breathed asbestos. [QPT1]
- (119) When a whole section (**lobe**) of the lung is taken out, the operation is called a lobectomy. When one whole lung is taken out, it is called a pneumonectomy. [QPT2]
- (120) The chance of recovery (**prognosis**) [QPT3]

Es posible reseñar una variación de este formato, en la que participan dos USE. La primera es un término conocido por el lector, y la segunda una unidad fraseológica que lo contiene como base. Así, en el ejemplo 121, el término *chemotherapy* queda repetido como parte de la UF *intrapleural chemotherapy*.

- (121) In mesothelioma, chemotherapy may be put directly into the chest (**intrapleural chemotherapy**). [QPT1]
- (122) Radiation may come from a machine outside the body (**external radiation therapy**) [QPT2]

13.1.2. Término especializado + (explicación / definición)

Este recurso es muy parecido al del apartado anterior, aunque en este caso el término especializado precede a la explicación o definición del mismo, las cuales van entre paréntesis.

- (123) **Surgery** (taking out the cancer) [QPT1]
- (124) **Radiation therapy** (using high-dose x-rays or other high-energy rays to kill cancer cells) [QPT1]
- (125) **Chemotherapy** (using drugs to fight the cancer). [QPT1]
- (126) [...] a local **anesthetic** (a drug that causes a loss of feeling for a short period of time). [QPT1]
- (127) [...] **prognosis** (chance of recovery) [QPT2]
- (128) Tests are done to find the main **tumor** (cancer). [QPT2]
- (129) **Chemotherapy** (using drugs to kill cancer cells). [QPT3]
- (130) [...] the **stage** of the cancer (whether it is just in the lung or has spread to other places) [...]. [QPT3]

13.1.3. Explicación / lexema de la lengua general (término superordinado) + *called* + término especializado

En este recurso cohesivo, el lexema *called* hace de preludeo del término especializado. La explicación del mismo contiene unidades léxicas de la lengua general que, por ser

términos superordinados, facilitan la comprensión de la nomenclatura científica que salpica los textos para pacientes.

(131)

1 e product of the body's cells. Tubes called bronchi make up the inside o
 2 ncer: squamous cell carcinoma (also called epidermoid carcinoma), adeno
 3 oat and into the bronchi. This test, called bronchoscopy, is usually don
 4 e chest between two ribs. This test, called thoracoscopy, is usually don
 5 also may be removed in an operation called a pneumonectomy. Sometimes p
 6 the vein or muscle. Chemotherapy is called a systemic treatment because
 7 the vein or muscle. Chemotherapy is called a systemic treatment because
 8 One new type of radiation therapy is called radiosurgery. In radiosurger
 9 out only a small part of the lung is called a wedge resection. When a wh
 10 the vein or muscle. Chemotherapy is called a systemic treatment because
 11 . Like surgery, radiation therapy is called local treatment because it w
 12 lung is taken out, the operation is called a lobectomy. When one whole
 13 n one whole lung is taken out, it is called a pneumonectomy. Radiation

Cuando el término especializado designa un proceso complejo que implica una serie de etapas, hay una descripción del mismo, a la que sigue la fórmula: *This is called (a) + término especializado.*

- (132) If tissue that is not normal is found, the doctor will need to cut out a small piece and have it looked at under a microscope to see if there are any cancer cells. This is called a **biopsy**. Biopsies are usually done during the thoracoscopy or peritoneoscopy. [QPT1]
- (133) Once malignant mesothelioma is found, more tests will be done to find out if cancer cells have spread to other parts of the body. This is called **staging**. A doctor needs to know the stage of the cancer to plan treatment. [QPT1]
- (134) If fluid has collected in the chest or abdomen, the doctor may drain the fluid out of the body by putting a needle into the chest or abdomen and using gentle suction to remove the fluid. If fluid is removed from the chest, this is called **thoracentesis**. If fluid is removed from the abdomen, **this is called paracentesis**. [QPT1]
- (135) The doctor can take cells from the walls of the bronchi tubes or cut small pieces of tissue to look at under the microscope to see if there are any cancer cells. This is called a **biopsy**. The doctor may also use a needle to remove tissue from a place in the lung that may be hard to reach with the brochoscope. A cut will be made in the skin and the needle will be put in between the ribs. This is called a **needle aspiration biopsy**. [QPT3]
- (136) Radiation therapy may also be used to prevent the cancer from growing in the brain. This is called **prophylactic cranial irradiation (PCI)**. [QPT3]

El término genérico que precede a la expresión *called* puede estar premodificado por el lexema *special*.

(137)

- 1 to look into the bronchi through a special instrument, called a bronchoscope
- 2 look inside the chest cavity with a special instrument called a thoracoscope
- 3 the abdomen (peritoneoscopy) with a special tool called a peritoneoscope.

Además de los ejemplos del apartado anterior, se han identificado casos en los que, para no "apabullar" al lector con tantos vocablos médicos, se omite el término especializado, y en su lugar aparece la construcción "a (*special*) + término superordinado de la lengua general".

(138)

- 3 **Intraoperative photodynamic therapy** is a new type of treatment that uses special drugs and light to kill cancer cells during surgery. A drug that makes cancer cells more sensitive to light is injected into a vein several days before surgery.
- 4 as much of the cancer as possible, a special light is used to shine on t
- 6 y uses a certain type of light and a special chemical to kill cancer cel

En la quinta línea de concordancia, si quisiéramos saber a qué concepto se refiere la paráfrasis "a drug that makes cancer cells more sensitive to light", habría que ir a la versión para especialistas, en la que, obviamente, el término designado por dicha paráfrasis, *hematoporphyrin derivative*, no va acompañado de definición.

- (139) **Endoscopic phototherapy** with a **hematoporphyrin derivative** has been described as an alternative to surgical resection in carefully selected patients [1-3]. This investigational treatment seems to be most effective for very early central tumors that extend less than 1 centimeter within the bronchus [...] **Endoscopic photodynamic therapy** [...] [QDT1].

Al contrastar la versión para especialistas con la versión para pacientes, afloran de nuevo dobles terminológicos:

(140)

TÉRMINOS ESPECIALIZADOS

Endoscopic phototherapy
 Endoscopic photodynamic therapy
 Hematoporphyrin derivative

DESIGNACIÓN EN TEXTOS PARA PACIENTES

Intraoperative photodynamic therapy
 A drug that makes cancer cells more sensitive to light

13.1.4. Oración de relativo con *that / which*

Como cabría esperar, un recurso muy productivo para definir el término especializado es la oración de relativo. El lexema de la lengua general que sirve de *definiens* es generalmente el antecedente del pronombre relativo (en cursiva). Hemos modificado las líneas de concordancia para que incluyan tanto el término especializado como el lexema de la lengua común.

(141)

1 **Small cell lung cancer** is a disease in *which* cancer (malignant) cells ar
 2 y and take out **carbon dioxide**, *which* is a waste product of the body' cel
 3 **Malignant mesothelioma**, a rare form of cancer, is a disease in *which* can
 4 **lungs** are a pair of cone-shaped organs *that* take up much of the room ins
 5 therapy) or from putting materials *that* produce radiation
 (**radioisotopes**)

Este recurso aparece en algunas ocasiones en combinación con otros, como la aposición de sintagmas nominales, el paréntesis, la utilización de lexemas como *type* o *called* o de verbos copulativos que relacionan el término especializado con uno común o semiespecializado.

(142)

1 **diaphragm**, the muscle below the lungs *that* helps with breathing, is also
 2 be given a local **anesthetic** (a drug *that* causes a loss of feeling for a
 3 **Intraoperative photodynamic therapy** is a new type of treatment *that* uses
 4 special drugs and light to kill cancer cells.
 5 The operation *that* takes out only a small part of the lung is called a
 6 **wedge resection**.
 7 **Lymph nodes** are small, bean-shaped structures *that* are found throughout
 8 to the lymph nodes in the area *that* separates the two lungs
 (**mediastinum**)
 7 ial instrument, called a **bronchoscope**, *that* slides down the throat and i
 8 A **clinical trial** is a research study *that* attempts to improve current tr

13.1.5. Término especializado + *Is a* + definición

El verbo copulativo *to be* es el más usual a la hora de enlazar el término especializado que lo precede con su definición o con un término superordinado, introducidos por el determinante *a*. En la mayoría de las instancias de este recurso, el *definiens* va seguido por una oración de relativo, como vemos en las líneas de concordancia 1– 6.

(143)

1 g cancer? **Small cell lung cancer** is a disease in which cancer (malignant) and take out **carbon dioxide**, which is a waste product of the body's cells.
 2
 3 **Intraoperative photodynamic therapy** is a new type of treatment that uses
 4 F PDQ ----- What is PDQ? **PDQ** is a computer system that gives up-to-date information on cancer and its prevention, detection, treatment, and supportive care. It is a service of the National Cancer Institute (NCI)
 5 about these trials. A **clinical trial** is a research study that attempts to
 6 **mesothelioma**, a rare form of cancer, is a disease in which cancer (malignant) grows to fight the cancer). **Surgery** is a common treatment of malignant
 7
 8 carcinoma. **Non-small cell lung cancer** is a common disease. It is usually

13.1.6. *Type / kind / form*

Si bien los lexemas *type*, *kind* y *form* también aparecen en la versión para especialistas, en estos últimos su función principal es la de evitar la repetición o entrar en detalles, como se aprecia en estos dos fragmentos destinados al personal sanitario:

(144)

1 However, even these patients are at risk from dying of lung cancer (both small and non-small cell types) [2].
 2 Nonsquamous cell cancers may be more likely to recur after surgical resection of early stage I tumors than other types of non-small cell lung cancers.

En la primera línea de concordancia se utiliza *types* para no repetir de nuevo la UF *lung cancer*. En la segunda, para no citar todas las variantes de cáncer pulmonar no microcítico, se emplea el sintagma nominal *other types of non-small cell lung cancer*.

Además de este uso, en los textos para pacientes la inclusión de *type / kind / form* contribuye sobre todo a ampliar el conocimiento. Esto se consigue al especificar las categorías que se derivan de un determinado concepto médico, como ilustran las líneas de concordancia 1– 4 del ejemplo (145).

Asimismo, estos lexemas ponen en relación el término especializado con su término superordinado. Por ejemplo, la radioterapia craneal profiláctica (*PCI*) es un tipo de radioterapia (línea 5) y el mesotelioma maligno (*malignant mesothelioma*) es un tipo de cáncer (línea 11).

(145)

1 Lung cancers can be divided into two types: **small cell lung cancer** and n
 2 There are two kinds of lung cancer based on how the cells look under a
 microscope: **small cell and non-small cell**.
 3 The main kinds of non-small cell lung cancer are named for the type of
 cells found in the cancer: **squamous cell carcinoma** (also called
 epidermoid carcinoma), **adenocarcinoma**, **large cell carcinoma**,
adenosquamous carcinoma, and **undifferentiated carcinoma**.
 4 s with small cell lung cancer. Three kinds of treatment are used:
surgery
 5 Because **PCI** may affect brain function, the doctor will help the patient
 decide whether to have this kind of radiation therapy.
 6 **Intraoperative photodynamic therapy** is a new type of treatment that uses
 ion therapy). One new type of radiation therapy is called
radiosurgery
 8 Photodynamic therapy uses a certain type of light and a special chemical
 9 about current treatment of different kinds of cancer. You may find it he
 10 s with malignant mesothelioma. Three kinds of treatment are used:
surgery
 11 **Malignant mesothelioma**, a rare form of cancer, is a disease in which ca

13.1.7. *This / these*

En los textos no especializados, a la función endofórica de los determinantes *this / these* se le suma la de preceder al hiperónimo de un término especializado mencionado con anterioridad. En (147), el merónimo (*early stage*) del término especializado *occult non-small cell lung cancer* está introducido por *this*, lo cual nos indica que el carcinoma oculto es uno de los primeros estadios por los que pasa la enfermedad.

(146) *** **OCCULT NON-SMALL CELL LUNG CANCER**

(147) Tests are done to find the main tumor (cancer). Lung cancer that is found at this early stage can be cured by surgery. [QPT2]

(148) **Surgery, radiation therapy, and chemotherapy** are used to treat non-small cell lung cancer. However, these treatments often do not cure the disease. [QPT2]

(149) ***** **STAGE 0 NON-SMALL CELL LUNG CANCER**

(150) Treatment may be one of the following: 1 Surgery to cure these very early cancers. [QPT2]

- (151) If there are symptoms, a doctor may want to look into the bronchi through a special instrument, called a bronchoscope, that slides down the throat and into the bronchi. This test, called **bronchoscopy**, is usually done in the hospital. [QPT3]
- (152) Radiation therapy may also be used to prevent the cancer from growing in the brain. This is called **prophylactic cranial irradiation (PCI)**. Because PCI may affect brain function, the doctor will help the patient decide whether to have this kind of radiation therapy. [QPT3]

13.1.8. Otros recursos

Hemos identificado otros recursos que no son tan habituales en los textos estudiados, aunque sí lo son en textos divulgativos, cuyo objetivo es acercar el conocimiento al lector que no es experto en el mismo. Entre estos recursos, podemos citar:

- **Aposición:** la inclusión entre comas (,) de la explicación del término especializado (ejemplos 153–154)
- **Término especializado + *Or other* + término genérico:** con este recurso se evita la mención de nuevos términos técnicos (ejemplos 155 – 156)
- **Término especializado + *means that* + explicación** (ejemplo 157)

En relación a la aposición, cuya función es la de aportar información adicional, Meyer (*apud* Kennedy 1998: 163) sugiere que la aposición es muy utilizada en géneros en los que existe un menor conocimiento compartido entre el escritor y el lector. En nuestro corpus, el número de aposiciones queda reducido por la excesiva utilización del paréntesis.

- (153) Sometimes part of the **diaphragm**, the muscle below the lungs that helps with breathing, is also removed [QPT1]
- (154) **Malignant mesothelioma**, a rare form of cancer, is a disease in which cancer (malignant) cells are found in the sac lining the chest (the pleura) or abdomen (the peritoneum). [QPT1]
- (155) Radiation therapy (using high-dose **x-rays** or other high-energy rays [...]) [QPT1]
- (156) Radiation therapy uses **x-rays** or other high-energy rays to kill cancer cells. [QPT3]

- (157) **Recurrent disease** means that the cancer has come back (recurred) after it has been treated .[QPT1, QPT3]

13.2. Recapitulación

El objetivo de este capítulo ha sido el de revisar algunos de los recursos generadores de cohesión en los textos para pacientes analizados en 11.2. Estos recursos hacen explícita la relación cohesiva entre USE. Al mismo tiempo, estas expresiones hacen de puente entre el conocimiento del lector y el conocimiento especializado, en un movimiento que siempre va de la información que sabe el lector a la información especializada. Esto nos induce a pensar que la cohesión juega un papel esencial en la progresión *tema / rema*.

Han aflorado una serie de patrones sintácticos que pueden resultar útiles al traductor y al terminógrafo en el análisis conceptual que llevan a cabo como parte de su tarea de documentación. Estos patrones son denominados por Meyer y Mackintosh (1996: 21) *conceptually rich words and phrases*, y *knowledge probes* (sondas de conocimiento) por Ahmad y Fulford (1992). Según Meyer y Mackintosh, estos recursos presentan un enorme potencial en la recuperación de la información dentro del área de la ingeniería del conocimiento. En efecto, las fórmulas identificadas apuntan hacia la *relación de abstracción*, que es la que se da entre un concepto superordinado y su correspondiente concepto subordinado¹⁷³.

Cualquier investigación en este sentido facilitará la labor terminológica de búsqueda de sinónimos en còrpora en soporte electrónico.

¹⁷³ Definición de Arntz y Picht (1995: 108).

14. Frecuencia léxica y cohesión

El estudio de listas de frecuencia léxica dentro de un corpus de la lengua general tiene limitaciones, como afirma Geens (*apud* Johansson 1985: 124): "Frequency lists have to be handled with care [...]; no absolute pronouncements about a language can safely be made when based on frequency only".

No obstante, es innegable su papel en lexicografía a la hora de seleccionar los vocablos que se deben incluir en un diccionario, el orden de presentación de sus significados y de determinar su uso.

En relación al papel que la frecuencia puede jugar en la terminología, Meyer y Mackintosh (1996: 8) y Cabré *et al.* (1996) apuntan que ayuda a determinar qué unidades polilexemáticas¹⁷⁴ han de considerarse términos o no. La limitación que plantea es que los terminógrafos se interesan principalmente por los nuevos términos introducidos en un campo del saber, y estos apenas aparecen en los textos. Podría pensarse pues que la frecuencia no es de gran utilidad en terminografía.

Sin embargo, el papel que juega la frecuencia en la identificación de cadenas léxicas y la gran accesibilidad a textos en soporte electrónico que pueden ser analizados con programas informáticos adecuados nos hacen pensar que la frecuencia léxica es un parámetro útil en terminología. Además, nos aproxima al vocabulario que debiera introducir cualquier definición de los términos (14.1.), a la organización textual (14.1.1.) y contribuye a identificar patrones sintácticos y semánticos dominantes en el texto (14.4.).

Veamos qué aporta la frecuencia léxica al estudio de las diferencias entre textos para especialistas y textos para pacientes.

14.1. Textos para especialistas

Como resultado del análisis, hemos encontrado una correlación entre las USE más frecuentes y los conceptos que hilan las cadenas léxicas, algo que ya se constató al estudiar los lemas con una frecuencia relativa mayor al 1% en cada texto sobre tratamiento.

¹⁷⁴ Corbin (*apud* Cabré 1996).

TABLA 129: Lemas con frecuencia relativa superior al 1% en los textos sobre tratamiento para especialistas

MALIGNANT MESOTHELIOMA (QDT1)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	STAGE	18	1,67	Staging (3), stages (3)	6. STAGING
2	TREATMENT	18	1,67	Treated (5), treatments (1)	0. TREATMENT
3	MESOTHELIOMA	17	1,58	Mesotheliomas (3)	2. PATHOLOGY OF M. MESOTH.
4	DISEASE	13	1,20	Disease-free (1)	2. PATHOLOGY OF M. MESOTH.
5	PATIENTS	12	1,11		7. PATIENT
6	SURGERY	11	1,02	Surgical (8)	0. TREATMENT (0.4.2.Surgery)
7	SURVIVAL	11	1,02		4. OUTCOME OF TREATMENT (4.3.1. Survival)
NON-SMALL CELL LUNG CANCER (QDT2)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	PATIENTS	124	2.25	Patient(9), patient's(5)	5. PATIENT
2	STAGE	72	1.30	Staging(10),stages(2), staged(4)	7. STAGING
3	TUMOR	70	1.27	Tumors(17)	1. PATHOLOGY OF NSCLC
4	RADIOTHERAPY	65	1.18		0. TREATMENT (0.5.1. Radiotherapy)
5	LUNG	62	1.12	Pulmonary(12)	4. LOCATION IN THE HUMAN BODY
6	CANCER	59	1.07	Cancers(13)	1. PATHOLOGY OF NSCLC
7	TREATMENT	56	1.01	Treated(16),treatments(3)	0. TREATMENT
8	CHEMOTHERAPY	55	1.00	Chemotherapeutic(1)	0. TREATMENT (0.5.2. Chemotherapy)
SMALL CELL LUNG CANCER (QDT3)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	PATIENTS	78	2.29	Patient(3)	9. PATIENT
2	CELL	58	1.70	Cellular(1)	1. PATHOLOGY OF SCLC
3	SMALL	56	1.64		1. PATHOLOGY OF SCLC 4. SIZE / QUANTITY
4	LUNG	55	1.61	Pulmonary(6)	1. PATHOLOGY OF SCLC 5. LOCATION IN THE HUMAN BODY
5	TREATMENT	55	1.61	Treated(18), treatment-related(2), treatments(1),untreated(1)	0. TREATMENT
6	CHEMOTHERAPY	52	1.53	Chemotherapeutic(1)	0. TREATMENT (0.5.1. Chemotherapy)
7	CANCER	49	1.44	Cancers(1)	1. PATHOLOGY OF SCLC
8	STAGE	40	1.17	Staging(5)	8. STAGING
9	SURVIVAL	37	1.09	Survivors(3),survivals(1)	3. OUTCOME OF TREATMENT (3.3.1. Survival)

Según la tabla 129, las USE más repetidas pertenecen a las cadenas léxicas más extensas, en concreto, a las ocho primeras, con la excepción de las cadenas *RESEARCH /*

EXPERIMENTATION y *DIAGNOSIS*¹⁷⁵ y nos han servido para etiquetar cadenas y subcadenas. En otras palabras, los lemas más repetidos representan los marcos conceptuales más relevantes en el texto y los focos de interés del profesional. A partir de cinco lemas que encabezan la lista de frecuencia (*PATIENTS*, *STAGE*, *TREATMENT*, *CELL* y *SMALL*) es posible deducir: a) que los pacientes han de ser el principal objetivo de la medicina; b) que es fundamental determinar con exactitud el estadio de la enfermedad; c) que el tema del texto es el tratamiento del cáncer, cuya nomenclatura se construye a partir de los lexemas *cell* y *small*; d) que la identificación del tipo de células cancerosas es también vital.

No obstante somos conscientes de que estos lemas sólo representan instancias de repetición léxica exacta, simple y variada sintáctica. Por tanto, no dan cuenta de todas las variadas relaciones cohesivas sobre las que se contruyen las cadenas léxicas.

Al igual que con las palabras más frecuentes hemos etiquetado cadenas léxicas, pensamos que las palabras más frecuentes de un texto o un corpus sirven como *definiens* de los términos que hay en los mismos. El *definiens* de USE como *radiotherapy* o *endoscopic photodynamic therapy* sería el lexema *treatment*, que es la designación del concepto TRATAMIENTO más frecuente. En consecuencia, en un glosario terminológico, estas palabras constituirían el vocabulario básico definicional sobre el que se construyen todas las definiciones.

Además, se puede aducir que la repetición de términos muy frecuentes juega un papel primordial en la cohesión del texto, y por tanto, en la consecución de una activación conceptual coherente. Al detectar los lemas más frecuentes—bajo los que se agrupan variantes morfológicas que comparten la misma raíz y que se vinculan mediante repetición exacta, simple y variada sintáctica—hemos percibido las áreas conceptuales más activadas en los textos y hemos identificado posibles cadenas y subcadenas.

14.1.1. Correlación entre frecuencia léxica y organización textual

Para confirmar que estos lemas tan frecuentes crean cohesión a lo largo de los textos sobre tratamiento para especialistas, utilizamos la aplicación *dispersion plot* de *Wordsmith Tools*. Esta aplicación muestra con rayas verticales en qué sitio de los textos aparece el lema en

¹⁷⁵ Esto se debe a la gran variedad léxica con la que contaban estas cadenas y a la escasa repetición léxica exacta y simple en las mismas.

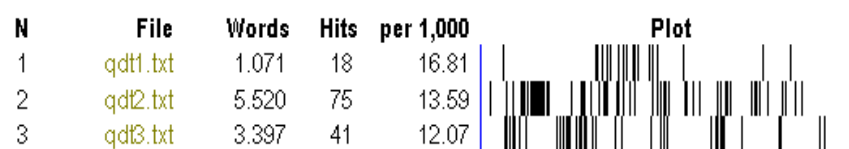
cuestión, considerando como el principio y el final del texto las líneas verticales de los extremos. En los gráficos que incluimos, la columna *Words* indica el número total de palabras del texto; *Hits*, el número de ocurrencias de la palabra estudiada; y *Hits per 1,000*, cuántas ocurrencias hay por cada 1000 palabras. En los apartados que siguen se especifican las cadenas léxicas que engarzan estos lemas.

En los gráficos quedan patentes dos aspectos. En primer lugar, las unidades léxicas elegidas se suceden a lo largo de los textos, y crean así cohesión. En segundo lugar, se detectan áreas en las que hay una mayor concentración de líneas y sospechamos que esto se puede explicar desde el punto de vista textual.

a) Distribución de los lemas más frecuentes en la cadena *STAGING*

Las mayores concentraciones de líneas, que reflejan la conjunción de unidades léxicas vinculadas en la cadena *STAGING*, se encuentran en el primer tercio, al igual que la sección *Stage Information* de los textos QDT (véase índice de los textos en 11.1.1.1., 11.1.2.1. y 11.1.3.1.). Las concentraciones más modestas coinciden con secciones que incluyen *stage* en el título, por ejemplo, *Limited stage small cell lung cancer*.

DIAGRAMA 36: Líneas de dispersión del lema *STAGE*



b) Distribución de los lemas más frecuentes en la cadena *TREATMENT*

La disposición del título y del apartado *Treatment Option Overview* en la segunda parte de QDT1 y QDT3 y en el segundo tercio de QDT2 es exactamente la misma que la de los lugares del diagrama 37 donde las líneas están más juntas.

DIAGRAMA 37: Líneas de dispersión de los lemas TREATMENT, RADIOTHERAPY, CHEMOTHERAPY y SURGERY

*Treat**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	18	16.81	
2	qdt2.txt	5.520	56	10.14	
3	qdt3.txt	3.397	54	15.90	

Radiotherapy

N	File	Words	Hits	per 1,000	Plot
1	qdt2.txt	5.520	65	11.78	
2	qdt3.txt	3.397	21	6.18	

*Chemotherap**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	7	6.54	
2	qdt2.txt	5.520	55	9.96	
3	qdt3.txt	3.397	52	15.31	

*Surg**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	11	10.27	
2	qdt2.txt	5.520	53	9.60	
3	qdt3.txt	3.397	9	2.65	

c) Distribución de los lemas más frecuentes en la cadena *PATHOLOGY*

El solapamiento de líneas se da en la primera mitad, que es donde está el apartado *Cellular classification*. Vemos cómo para un mismo texto, por ejemplo, el QDT2, los lemas TUMOR, DISEASE y CANCER se van alternando hasta tejer el entramado del texto.

DIAGRAMA 38: Líneas de dispersión de los lemas TUMOR, DISEASE, CANCER, SMALL, CELL, LUNG y MESOTHELIOMA.

*Tumor**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	7	6.54	
2	qdt2.txt	5.520	70	12.68	
3	qdt3.txt	3.397	20	5.89	

*Disease**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	13	12.14	
2	qdt2.txt	5.520	25	4.53	
3	qdt3.txt	3.397	21	6.18	

*Cancer**

N	File	Words	Hits	per 1,000	Plot
1	qdt2.txt	5.520	61	11.05	
2	qdt3.txt	3.397	49	14.42	

*Small**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	3	2.80	
2	qdt2.txt	5.520	35	6.34	
3	qdt3.txt	3.397	59	17.37	

*Cell**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	1	0.93	
2	qdt2.txt	5.520	36	6.52	
3	qdt3.txt	3.397	58	17.07	

*Lung**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	2	1.87	
2	qdt2.txt	5.520	50	9.06	
3	qdt3.txt	3.397	49	14.42	

*Mesothelioma**


N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	17	15.87	

d) Distribución de los lemas más frecuentes en las cadenas *PATIENT* y *SURVIVAL*


De acuerdo con las líneas, las secciones *General information* y *Treatment Option overview* son las que prestan mayor atención a dos conceptos muy relacionados: PACIENTE y SUPERVIVENCIA. En la primera sección, se habla del pronóstico de la enfermedad y posibilidades de supervivencia, y en la segunda, se aportan porcentajes de supervivencia de pacientes sometidos a determinados tratamientos.

DIAGRAMA 39: Líneas de dispersión de los lemas *PATIENT* y *SURVIVAL*

*Patient**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	12	11.20	
2	qdt2.txt	5.520	124	22.46	
3	qdt3.txt	3.397	78	22.96	

*Surviv**

N	File	Words	Hits	per 1,000	Plot
1	qdt1.txt	1.071	11	10.27	
2	qdt2.txt	5.520	47	8.51	
3	qdt3.txt	3.397	37	10.89	

En definitiva, parece existir una cierta coincidencia entre la disposición de los títulos en los textos sobre tratamiento y las áreas de máxima confluencia de los lemas más frecuentes. Nuestras observaciones están en la línea de los resultados de la tesis de Berber Sardinha (1997a), que afirmaba que en los límites de las secciones de un texto, la cohesión léxica es mayor (véase 4.8.1.).

14.2. Textos para pacientes. Diferencias con respecto a los textos para especialistas

Los lemas más frecuentes en los textos para pacientes pertenecen a las siete cadenas principales, con la excepción de la cadena *OUTCOME OF TREATMENT*. Los cinco lemas más frecuentes son *CANCER*, *LUNG*, *TREATMENT*, *CELL* y *THERAPY*. En la tabla 129 correspondiente a textos para especialistas, *CANCER* estaba en sexto o séptimo lugar o

simplemente no aparecía; LUNG estaba en cuarto o quinto lugar o no estaba; TREATMENT en segundo, quinto y séptimo lugar; CELL sólo se incluía en la tabla correspondiente a QDT3 y en segundo lugar; y THERAPY no contaba entre los lemas con un porcentaje superior al 1%.

Parece que el orden de los lemas más frecuentes en distintos tipos textuales aporta información sobre esos textos¹⁷⁶. Echando un vistazo al resto de lemas más frecuentes, nos aventuramos a decir que:

- a) Los lemas más frecuentes están relacionados en orden de importancia con el nombre de la enfermedad, el tratamiento ya sea convencional o experimental (CLINICAL TRIALS), el estadiaje de la enfermedad y el paciente. Estos dos últimos aspectos eran la principal prioridad en textos para especialistas.
- b) En general se prefieren lexemas menos especializados. Se prefiere CANCER a TUMOR; RADIATION THERAPY a RADIOTHERAPY. Incluso se prefieren lemas con un significado muy vago (USED).
- c) El porcentaje de términos anatómicos es mayor en textos para pacientes. Valga como ejemplo los lemas CHEST y LUNG que ocupan el tercer y el segundo lugar respectivamente, en oposición al septuagésimo octavo lugar y cuarto / quinto lugar en textos para especialistas.

¹⁷⁶ Johansson (1985: 124) afirma al respecto: "Word frequencies vary greatly depending upon the type of text, even frequencies of function words".

TABLA 130: Lemas con frecuencia relativa superior al 1% en los textos sobre tratamiento para pacientes.

MALIGNANT MESOTHELIOMA (QPT1)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	CANCER	50	3,08	cancer-related(1)	2. PATHOLOGY OF M. MESOTH.
2	TREATMENT	39	2,41	Treatments(7),treated(3), treat(2),treating(1)	0. TREATMENT
3	CHEST	24	1,48		1. LOCATION IN THE HUMAN BODY
4	MESOTHELIOMA	19	1,17		2. PATHOLOGY OF M. MESOTH.
5	MALIGNANT	18	1,11		2. PATHOLOGY OF M. MESOTH.
6	TRIALS	18	1,11	trial(7)	0. TREATMENT 1. EXPERIMENTATION
7	CLINICAL	16	0,99		5. GENERAL MEDICINE
8	INFORMATION	16	0,99		6. INFORMATION
NON-SMALL CELL LUNG CANCER (QPT2)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	CANCER	58	4,61	Cancers(3)	1. PATHOLOGY OF NSCLC
2	LUNG	43	3,42	Lungs(2)	1. PATHOLOGY OF NSCLC 2. LOCATION IN THE HUMAN BODY
3	CELL	37	2,94	Cells(12)	1. PATHOLOGY OF NSCLC
4	THERAPY	34	2,70		0. TREATMENT
5	RADIATION	30	2,38		0. TREATMENT (0.5.1. Radiation therapy)
6	TREATMENT	27	2,15	treated(5),treat(3), treatments(2)	0. TREATMENT
7	STAGE	26	2,07	stages(4), staging(1)	5. STAGING
8	NON-SMALL	21	1,67		1. PATHOLOGY OF NSCLC 4. SIZE / QUANTITY
9	CHEMOTHERAPY	18	1,43		0. TREATMENT (0.5.3.Chemotherapy)
10	SURGERY	17	1,35		0. TREATMENT (0.5.2. Surgery)
11	USED	17	1,35	Uses (5), using(1)	0. TREATMENT (0.3. Manipulation)
12	PATIENTS	16	1,27	patient's(1),patient(2)	6. PATIENT
SMALL CELL LUNG CANCER (QPT3)					
N	Word	Freq.	%	Lemmas	Lexical chains
1	CANCER	47	3,75		2. PATHOLOGY OF SCLC
2	LUNG	39	3,11	lungs(8)	1. LOCATION IN THE H. BODY 2. PATHOLOGY OF SCLC
3	CELL	35	2,79	cells(14)	2. PATHOLOGY OF SCLC
4	SMALL	21	1,67		2. PATHOLOGY OF SCLC 5. SIZE / QUANTITY
5	TREATMENT	20	1,59	Treatments(4),treated(2), treat(1)	0. TREATMENT
6	THERAPY	16	1,27	Therapies(1)	0. TREATMENT
7	STAGE	15	1,20	staging(1),stages(4)	6. STAGING
8	PATIENT	13	1,04	patient's(2),patients(4)	9. PATIENT
9	RADIATION	13	1,04		0. TREATMENT (0.4.1. Radiation therapy)
10	USED	13	1,04	using(2),uses(1),use(1)	0. TREATMENT (0.3. Manipulation)

14.3. Frecuencia léxica y grado de especialización

Desde el punto de vista terminológico y traductológico, no es suficiente con detectar las unidades léxicas más frecuentes del texto, que normalmente no plantean dificultades por su reducido grado de especialización, aunque es preciso conocer el significado y los patrones sintácticos y colocacionales que adquieren en el campo de la oncología. Habría también que discriminar en los tres textos para especialistas (*Treatment Summaries for Health Professionals*), las USE que son específicas a cada variante de cáncer, que normalmente tienen una frecuencia más destacada en los textos en cuestión que en el resto del corpus. La operación se completaría identificando el vocabulario específico del tratamiento, que no resalta tanto en los textos sobre cuidados paliativos (*Supportive Care Summaries for Health Professionals*).

Para este fin, nos hemos servido de la aplicación *Keywords* de *Wordsmith Tools*. Esta identifica las *palabras clave* de un texto, es decir, las palabras cuya frecuencia está muy por encima de la norma. Lo consigue comparando las listas de frecuencia del texto o textos en los que se quiere encontrar las palabras clave, con las listas de frecuencia de un corpus más amplio que sirve de corpus de referencia.

El resultado es una doble lista de frecuencia como la de la tabla 131 en la que aparecen en color ocre las formas que destacan (*positive keywords*) en el texto seleccionado, QDT1, con respecto al corpus sobre tratamiento para especialistas (véase sexta columna, *D123.LST%*). En rojo van aquellas formas que son menos frecuentes de lo que cabría esperar (*negative keywords*). *Wordsmith Tools* ha reconocido 26 palabras clave.

Según se aprecia en la tabla con las palabras clave del texto QDT1, las USE que marcan la diferencia con respecto a los otros textos sobre tratamiento tienen un alto grado de especificidad y pertenecen a las siguientes áreas conceptuales:

(158)

NAME OF DISEASE: *malignant mesothelioma, mesotheliomas, epithelial (mesothelioma), disease*

SYMPTOMS: *effusions*

LOCATION IN THE HUMAN BODY: *intracavitary, pleural, extrapleural, diaphragm*

TREATMENT: *approaches, pleurectomy, decortication, radiation, palliative, procedures, pneumonectomy*

RESEARCH: *reported, proposed, references*

STAGING: *advanced, localized, II, stages*

TABLA 131: Palabras clave en QDT1 (*Treatment for malignant mesothelioma*)¹⁷⁷

N	WORD	FREQ.	QDT1.TXT %	FREQ.	D123.LST %	KEYNESS	P
1	MESOTHELIOMA	14	1.30	14	0.14	29.4	0.000000
2	MALIGNANT	7	0.65	13	0.13	9.4	0.002159
3	EFFUSIONS	5	0.46	6	0.06	9.4	0.002187
4	EPITHELIAL	4	0.37	4	0.04	8.4	0.003797
5	INTRACAVITARY	4	0.37	4	0.04	8.4	0.003797
6	PLEURAL	6	0.56	11	0.11	8.2	0.004292
7	APPROACHES	5	0.46	8	0.08	7.6	0.005746
8	PLEURECTOMY	3	0.28	3	0.03	6.3	0.012197
9	DECORTICATION	3	0.28	3	0.03	6.3	0.012197
10	EXTRAPLEURAL	3	0.28	3	0.03	6.3	0.012197
11	MESOTHELIOMAS	3	0.28	3	0.03	6.3	0.012197
12	REPORTED	7	0.65	20	0.20	5.8	0.015754
13	DISEASE	12	1.11	48	0.48	5.7	0.016616
14	RADIATION	5	0.46	11	0.11	5.7	0.017074
15	PROPOSED	3	0.28	4	0.04	5.2	0.022032
16	DIAPHRAGM	3	0.28	4	0.04	5.2	0.022032
17	PALLIATIVE	5	0.46	12	0.12	5.2	0.022974
18	ADVANCED	4	0.37	8	0.08	5.0	0.025213
19	LOCALIZED	4	0.37	8	0.08	5.0	0.025213
20	II	5	0.46	13	0.13	4.7	0.030139
21	STAGES	3	0.28	5	0.05	4.4	0.035452
22	PROCEDURES	3	0.28	5	0.05	4.4	0.035452
23	PNEUMONECTOMY	3	0.28	5	0.05	4.4	0.035452
24	REFERENCES	7	0.65	25	0.25	4.1	0.042072
25	PATIENTS	12	1.11	197	1.97	4.4	0.034936
26	CHEMOTHERAPY	5	0.46	110	1.10	4.7	0.029715

Estas pertenecen a seis de las siete cadenas léxicas más importantes del texto. Esto mismo ocurre con las palabras clave de los *Treatment summaries for health professionals* con respecto a los *Supportive care summaries for health professionals*, que también pertenecen a las principales cadenas léxicas, como se aprecia en las 22 primeras palabras clave de la tabla 132 (véase Apéndice X para más detalles).

La forma léxica más significativa es *stage*, que es el principal factor que determina el tratamiento a seguir. El segundo lugar lo ocupa *lung*, que forma parte de la nomenclatura de la enfermedad. Entre las 22 primeras palabras clave, hay bastantes lexemas relacionados con el tratamiento: *radiotherapy*, *chemotherapy*, *resection*, *clinical trials*, *regimens*, *treated*, *surgical*. También encontramos otros elementos léxicos que designan la enfermedad: *cell*, *small*, *tumor*, *non-small*, *NSCLC*, *carcinoma*, *metastases*, *mesothelioma*. La supervivencia

¹⁷⁷ Para obtener palabras clave que representen la terminología correspondiente al tratamiento del mesotelioma maligno, hemos fijado un valor de p elevado. Con valores menores, obteníamos muy pocos términos.

es un elemento clave para el profesional que lee textos sobre tratamiento, como indica el cuarto lugar de *survival*.

TABLA 132: Palabras clave en los textos sobre tratamiento para especialistas (QDT)

N	WORD	FREQ. DTREAT.LST %	FREQ. PCARE.LST %	KEYNESS	P		
1	STAGE	103	1.03	17	0.03	312.4	0.000000
2	LUNG	101	1.01	17	0.03	305.1	0.000000
3	RADIOTHERAPY	86	0.86	5		301.0	0.000000
4	SURVIVAL	89	0.89	12	0.02	280.0	0.000000
5	CELL	91	0.91	26	0.04	241.7	0.000000
6	CHEMOTHERAPY	110	1.10	91	0.15	183.2	0.000000
7	SMALL	69	0.69	31	0.05	156.7	0.000000
8	TUMOR	77	0.77	50	0.08	147.5	0.000000
9	PATIENTS	197	1.97	398	0.65	139.7	0.000000
10	RESECTION	38	0.38	5		119.9	0.000000
11	TRIALS	41	0.41	12	0.02	108.0	0.000000
12	NON-SMALL	26	0.26	0		102.1	0.000000
13	NSCLC	26	0.26	0		102.1	0.000000
14	CLINICAL	57	0.57	45	0.07	97.6	0.000000
15	CHEST	30	0.30	7	0.01	84.0	0.000000
16	REGIMENS	33	0.33	14	0.02	76.6	0.000000
17	TREATED	39	0.39	26	0.04	73.6	0.000000
18	CARCINOMA	27	0.27	8	0.01	70.8	0.000000
19	SURGICAL	32	0.32	18	0.03	65.8	0.000000
20	STAGING	18	0.18	1		63.1	0.000000
21	METASTASES	23	0.23	7	0.01	59.8	0.000000
22	MESOTHELIOMA	14	0.14	0		55.0	0.000000

Se pone de manifiesto de nuevo que las cadenas léxicas, que se fundamentan en la cohesión y en la repetición léxica, son un índice que apunta hacia las áreas conceptuales de donde emana la terminología del texto. Las cadenas léxicas identificadas por el traductor, el terminólogo o el documentalista con la ayuda de un programas de análisis léxico, se convierten en un instrumento útil en la tarea indagadora de estos en el sentido de que son una puerta de entrada hacia la terminología que utilizan los especialistas.

Por otra parte, la frecuencia léxica es importante porque, según el modelo cognitivo de Bell (1991) sobre la traducción, que expusimos en 2.1., a la hora de traducir, se ponen en juego dos elementos del sistema de la memoria: una base de datos con las palabras más frecuentes (*frequent lexis store*) y un inventario con las estructuras sintácticas más frecuentes (*frequent structures store*).

14.4. El cotexto de los lemas más frecuentes como restrictor semántico

El cotexto de los lemas más frecuentes, ya sean estas unidades terminológicas o unidades léxicas de la lengua general, es una fuente de información sobre el significado de los mismos, sobre las unidades fraseológicas y colocaciones que forman y sobre los patrones sintácticos en los que encajan.

De estos aspectos, vamos a estudiar el modo en el que el radio colocacional restringe el significado de lexemas de la lengua general, con lo que estos hacen referencia a eventos, entidades y propiedades dentro del ámbito de la oncología. Es decir, estudiaremos cómo el cotexto contribuye al proceso de *terminologización* de palabras de la lengua general. Dubuc y Lauriston (1997: 81) destacan el papel del *contexto*, nombre que recibe el *cotexto* en terminología, a la hora de asociar contenido semántico al término en cuestión y determinar la relación entre este y un campo del saber.

De estos aspectos, nos centraremos en el modo en el que la reiteración en nuestro corpus de cotextos en torno a una palabra de la lengua general restringe su significado de forma que contribuye a la consolidación de unidades fraseológicas y puede llevar consigo un "movimiento de unidades de la lengua general a las especialidades" que Cabré (1999: 115) llama *terminologización*. Buena prueba de este proceso la encontramos en el lexema *administration*, que, en textos sobre oncología, suele aparecer con términos relacionados con la quimioterapia principalmente y con la radioterapia, con lo cual, su significado es más restringido que el que se refleja en un diccionario y que el étimo original latino.

administration n.

- 1 management of a business.
- 2 the management of public affairs; government.
- 3 the government in power; the ministry.
- 4 US a President's period of office.
- 5 Law the management of another person's estate.
- 6 (foll. by of) a the administering of justice, an oath, etc. b application of remedies.

Etymology ME f. OF administration or L administratio (as administrate)

Concise Oxford Dictionary. 1993. Oxford University Press.

Asimismo, el estudio del radio colocacional de los lemas más significativos refleja un fenómeno colocacional que Sinclair (1991: 112) y Louw (1993: 157) denominan *prosodia semántica*¹⁷⁸, que consiste en que ciertas palabras llevan asociadas un componente axiológico positivo o negativo que determina las unidades léxicas que pueden aparecer en su cotexto. Por ejemplo, Sinclair (1991: 112) sostiene que el lema HAPPEN está asociado a fenómenos desagradables como accidentes. Stubbs (1996: 173-4) también señala que el lema CAUSE tiene una prosodia semántica negativa y suele coocurrir con formas tales como *accident, concern, damage, death* o *trouble*.

En los apartados anteriores, la frecuencia léxica estudiada correspondía al corpus de la sección *Treatment Summaries* del *National Cancer Institute*. Como este corpus no sirve para detectar tendencias sintácticas y semánticas en el cotexto de los lemas más frecuentes, se ampliará el corpus con la sección *Supportive Care Summaries*, que incluye bastantes lexemas relacionados con el tratamiento. Se considerarán sólo los textos para profesionales de la salud.

Para mostrar una vez más las aplicaciones de la lingüística de corpus a la terminología, se pretende analizar el cotexto de algunos de los lemas más frecuentes, que se sitúan en el siguiente continuo que va desde lo más a lo menos especializado:

- términos propios la oncología (*cancer, chemotherapy, metastasis*)
- términos médicos genéricos (*patients, treatment, symptoms*), muchos de los cuales pertenecen a la lengua general o han sufrido un proceso de banalización¹⁷⁹ como consecuencia de fenómeno de divulgación de temas médicos
- términos propios del metadiscurso científico (*dose, factors, include*)
- lexemas de la lengua general que restringen su significado en función de su radio colocacional (*use, effects, stage, administration, disease*)
- lexemas de la lengua general (*include, care, increase*)

¹⁷⁸ Adoptamos el término, muy extendido en lingüística de corpus, aunque llamando la atención de que en este caso *prosodia* no tiene nada que ver con el nivel suprasegmental de la lengua.

¹⁷⁹ Para Cabré (1999: 115) este fenómeno sería el inverso al de terminologización, es decir, la utilización de términos como unidades de la lengua general.

TABLA 133: Lemas más frecuentes en textos sobre tratamiento y cuidados paliativos para profesionales de la salud

N	Word	Freq.	%	Lemmas
1	PATIENTS	856	1.20	Patient(212),patient's(49)
2	CANCER	469	0.66	Cancers(29)
3	TREATMENT	461	0.65	Treated(65),treatments(25),treating(19), treats(1),treat(12), Treatable(1),treatment-(1)
4	USE	301	0.42	Use(150),using(21),used(128)
5	SYMPTOMS	281	0.39	Symptom(40),symptomatic(24)
6	INCLUDE	253	0.36	Including(70),includes(27),included(14)
7	EFFECTS	244	0.34	Effective(69),effect(63),effectiveness(9),effectively(4)
8	THERAPY	219	0.31	Therapies(14),therapeutic(16),therapist(3),therapeutically(1)
9	CHEMOTHERAPY	217	0.30	Chemotherapeutic(12),chemoprevention(3),chemotherapy-(1)
10	CARE	203	0,28	Caregivers(9),caregiver(8),carefully(6),cared(2),caring(2), caregiver's(1)
11	DOSE	202	0,28	Doses (76), dosages(23), dosage(10), dosing (5)
12	DISEASE	182	0.26	Disease-free(12), diseases(10)
13	CAUSE	179	0.25	Cause(82), causes (45), caused(37), causing(13), causative(2)
14	INCREASED	171	0.24	Increased(91), increase(45), increases (20), increasing(14), increasingly(1)
17	DAY	165	0.23	Days(40),daily(57)
18	SURVIVAL	164	0.23	Survivors(48), surviving(9), survive(2), survivorship(2), survived(1), survivals(1)
19	ORAL	162	0.23	Orally(17)
20	STAGE	162	0.23	Staging(19),stages(18), staged(5),
21	FATIGUE	161	0.23	Fatigued(3)
22	CELL	154	0.22	Cells(31),cellular(6)
23	FACTORS	147	0.21	Factor(19)
24	TIME	144	0.20	Times(45),timing(5),time-limited(5),timely(4),time-related(1)
25	DRUGS	142	0.20	Drug(69)
26	SLEEP	142	0.20	Sleeping(4),sleep-wake(9),sleep-related(1)
27	LUNG	140	0.20	Pulmonary(22)
28	DIAGNOSIS	138	0.19	Diagnostic(22), diagnosed(18),diagnoses(2),diagnose(1),
29	HYPERCALCEMIA	135	0.19	Hypercalcemic(2)
30	DEVELOPMENT	131	0.18	Develop(34), developing(28), developed(18), develops (8), developmental(6), developments(1)
31	ASSOCIATED	130	0.18	Tumor-associated (7)
32	GRIEF	130	0.18	Grieving(25),grieve(2),grief-related(1),grievors(1)
33	PAIN	130	0.18	Painless(1),painful(13)
34	REFERENCES	130	0.18	Referring(1),referred(9),refers(4),refer(9),referral(4), reference(1),referrals(1)
35	SURGERY	128	0.18	Surgeries(1),surgical(50),surgically(7),surgery-induced(1)
36	VOMITING	128	0.18	Vomitus(1)
37	DEATH	127	0.18	Dying(22),deaths(4),died(5),dead(2),dies(2)
38	EXPERIENCE	126	0.18	Experiencing(11),experienced(25),experiences(17)
39	ANXIETY	125	0.18	Anxieties(2),anxious(3)
40	FOLLOWS	124	0.17	Following(100),followed(13),follow(3)
41	RESPONSE	124	0.17	Responses(30),respond(10),responded(3),responds(6), responder(1),responders(1)
42	CLINICAL	118	0.17	Clinically(14),clinics(2)

43	LOSS	118	0.17	Losses(11),losing(3),lost(7),lose(3),loses(2)
44	CHILDREN	114	0.16	Child(28),child's(10),children's(3),childhood(10)
45	OCCUR	114	0.16	Occurs(26), occurrence(9), occurred(7), occurring(6),
46	NAUSEA	113	0.16	
47	COMBINATION	112	0.16	Combined(27), combinations(16), combine(3), combining(2), combines(2)
48	ASSESSMENT	110	0.15	Assessed(14),assess(9),assessing(2), assessments(4)
49	FAMILY	106	0.15	Families(20),family's(6),familiar(4)
50	HIGH	105	0.15	Highest(1),higher(37)
51	SMALL	102	0.14	Smaller(2)
52	E.G.	101	0,14	
53	PERSON	100	0.14	Persons(32),person's(12)
54	WELL	100	0.14	Best(21),good(16),better(12)
55	ADMINISTRATION	99	0.14	Administered(29),administer(4), administering(2)
56	LIFE	99	0.14	Life-threatening(19),living(13),live(7),lifestyle(4),life- (1), lifelong(1),lifesaving(1)
57	NUTRITIONAL	98	0.14	Nutrition(30),nutrients(4),nutrient(2),nutritionally(2)
58	PRODUCE	96	0.13	Produced(12),products(12), produces(11), producing(8),producers(4),product(1)

De los diez lemas más frecuentes de la lista de frecuencia lematizada, el lema PATIENTS es el que ocupa el primer lugar, lo que indica la preeminencia del ser humano con *cáncer*, segunda palabra más frecuente. El enfermo adquiere el rol semántico de *sujeto paciente* que recibe un tratamiento y una serie de cuidados. A la idea de TRATAMIENTO se le asigna un porcentaje muy elevado, como indica la presencia de los lemas TREATMENT, USE, THERAPY y CHEMOTHERAPY, que ocupan respectivamente, el tercer, octavo y noveno puesto. El lema CARE ocupa el décimo lugar. También en relación con el paciente, y en quinto lugar, encontramos el lema SYMPTOMS. Por otra parte, destacan dos lemas de la lengua general bastante habituales en textos científicos, INCLUDE y EFFECTS.

Se analizarán todos estos lexemas con más detalle, centrando nuestro estudio, primero, en las unidades de significación clave del subdominio de la biomedicina y la oncología, y después, en las unidades léxicas de la lengua general, cuyo significado queda limitado en el subdominio de la oncología.

14.4.1. Elaboración de una ficha con información cotextual

El procedimiento que se seguirá se basa en una ficha que aúna los datos obtenidos con diversas aplicaciones y ventanas de *Wordsmith Tools*:

- *Wordlist*, que proporcionará, por una parte, la frecuencia absoluta (*Frec*) y relativa (%) del lema en cuestión (la forma léxica más frecuentes del lema va en mayúsculas y negrita) y, por otra, la frecuencia absoluta de cada variante morfológica (en minúscula).
- *Concord*, con el que se extraerán diez concordancias que ilustran el cotexto de la unidad de significación especializada¹⁸⁰.
- *Collocates*, que informa de los lexemas que hay en un radio colocacional de cinco palabras a la izquierda y a la derecha de la USE. Estos se presentan en la casilla superior del recuadro *Lexemas en su radio colocacional / unidades fraseológicas*, seguidos de la frecuencia con la que coocurren con la palabra clave.
- *Clusters*, que identifica colocaciones y UF cuya *base* es la palabra clave. Las UF se incluyen en la casilla inferior del recuadro *Lexemas en su radio colocacional/unidades fraseológicas*, y se desarrollan en algunos casos.

En el recuadro *Lexemas en su radio colocacional / unidades fraseológicas*, quedan reflejados aquellos que aparecen un 5% de la frecuencia absoluta de la *base* de la colocación¹⁸¹.

A continuación, presentamos la ficha que elaboraríamos para el lema ADMINISTRATION, que representa un 0,14% de las 71 243 palabras del corpus para profesionales de la salud. Las distintas formas de este lema están en la columna de la izquierda.

En la tabla se aprecia que la mayoría de los lexemas de su radio colocacional están relacionados con la quimioterapia [*intravenous(ly)*, *drug*, *dose(s)*, etc.], lo cual nos indica que en el subdominio de la medicina, *administration* hace referencia a la aplicación de un fármaco. De todas formas, también hemos encontrado una mención a la radioterapia. Algunos de estos lexemas preceden o siguen al lema y constituyen colocaciones (*administered orally*) y UF (*drug administration*).

¹⁸⁰ Cuando la USE consta de más de una palabra, como ocurre con *lung cancer*, ha sido preciso ajustar Wordsmith Tools para que en el radio colocacional se cuenten seis palabras a la derecha del primer elemento (*lung*), y por tanto, cinco palabras a la derecha de la USE (*lung cancer*).

¹⁸¹ Excepto para los lemas PATIENTS y CANCER.

TABLA 134: Ficha con el radio colocacional de *administ**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional				Frec	%
	Colocaciones / unidades fraseológicas					
ADMINISTRATION Administration 64, Administered 29, Administer 4, Administering 2	Intravenous (10), hours (9), drug (9), intravenously (8), doses (8), dose (7), intramuscularly (7), orally (6), rapid (6), route (6), schedule (6), chemotherapy (6), oral (5), nausea (5)				99	0,14
	Drug administration	8	Hours after administration	3		
	Intravenous administration	6	Route of administration	3		
	Administered orally	5	Administration schedules	3		
	Chemotherapy administration	5	Rapid administration	3		
<p>1 of modalities and schedule of drug administration remains to be determ 2 hematologic toxic effects; proper administration requires close colla 3 survival when the duration of drug administration exceeds 6 months [4, 4 st or second cycle of chemotherapy administration. The relative ef 5 more effective than the consistent administration of a single regimen 6 urvival after the duration of drug administration exceeds 3-6 months [7 y. Intrapleural or intraperitoneal administration of chemotherapeutic 8 cisplatin plus chest radiotherapy administered during the first or se 9 een provided. Some physicians have administered two of these or other 10 be reduced by more than 50% by the administration of PCI in doses of 2</p>						

14.4.2. Unidades de significación más frecuentes del subdominio de la biomedicina y la oncología

Los términos médicos más frecuentes aluden a los participantes en el evento médico y a los procesos biológicos y psicológicos que afectan al paciente. Estos términos se utilizan también en la lengua general con prácticamente el mismo significado. Es decir, han experimentado un proceso de banalización, en parte por el interés del público general por la salud y el auge de la divulgación médica. En el radio colocacional de estas USC predominan una serie de lexemas, con los que pueden formar UF, como se verá.

14.4.2.1. PATIENTS

En el radio colocacional del lema PATIENTS, el más frecuente del corpus, destacan formas léxicas que designan el cáncer y sus estadios, su tratamiento y la familia del paciente. La incidencia del tratamiento en el paciente queda en los textos prácticamente igualada con la del apoyo familiar. Las UF que hemos destacado giran en torno a los tres aspectos antes señalados.

Por una parte, tenemos las UF *cancer patients* y *patients with cancer*, que juntas representan 11'6% de todas las ocurrencias de PATIENTS. En el corpus aparecen

variaciones terminológicas del mismo concepto, *people with cancer* y *individuals with cancer*, que se mencionará bajo el lema CANCER.

El paciente como *sujeto paciente* del tratamiento está presente en la colocación *patients treated* y *patients receiving*. En estas expresiones PATIENTS tiene el rol semántico de *processed experiencer (ProcExp)*¹⁸². El lema RECEIVE es bastante representativo en el corpus (0,11%), lo cual hace que la función semántica *ProcExp*, asociada a su sujeto gramatical, esté bastante activada en los textos. Los tratamientos más usuales son la quimioterapia, la radioterapia y la cirugía. Los dos primeros coocurren tanto con *treated* como con *receiving*, mientras que las formas *surgically*, *surgical* y *surgery* aparecen sólo con *treated*, como se aprecia en las líneas de concordancia.

La incidencia del tratamiento en el paciente queda en los textos prácticamente igualada con la del apoyo familiar. La vinculación entre el paciente y sus seres queridos queda de manifiesto en la colocación *patient and family*. Otras colocaciones habituales son *selected patients*, *patients with good performance status* y *patients with small cell lung cancer*.

TABLA 135: Radio colocacional de *patient**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional	Frec	%
		Colocaciones / unidades fraseológicas		
PATIENTS Patients 595 Patient 212 Patient's 49		Cancer (148), stage (44), treatment (43), disease (35), family (33) <u>1</u> <u>2</u> <u>3</u> <u>4</u> Selected patients 11 Patients with good performance status 7 Patients with small cell lung cancer 7	856	1,20
1	CANCER PATIENTS (66) Cancer patient (11)	Fever (12), sleep (6), experience (5), survivors (5)	77	0,11
1 e this has not yet been studied in cancer patients or survivors, the 2 rotein and calories needed by each cancer patient will vary dependin 3 .) Adequate sleep may increase the cancer patient's pain tolerance. 4 ***** SLEEP DISTURBANCE IN CANCER PATIENTS Cancer patien 5 nded for other non-small cell lung cancer patients with similar stag 6 c drugs can also cause insomnia in cancer patients. Sustained use of 7 not been widely used or studied in cancer patients. Benzodiazepi 8 gement of sleep disturbances among cancer patients. The anticholiner 9 anagement of sleep disturbances in cancer patients. Barbiturates hav 10 of melatonin to treat insomnia in cancer patients is under evaluati				

¹⁸² La definición de esta función semántica introducida en Dik (1989: 102) es la siguiente: "the entity that experiences a Process".

2	PATIENTS WITH CANCER	PTSD (3), anxiety (3), symptoms (3)	22	0,03
1	patient and home care treatment of patients with cancer are a coordina			
2	traumatic stressor for individual patients with cancer. Within the mu			
3	as natural disasters or rape. For patients with cancer, the stressor			
4	traumatic stress disorder (PTSD) to patients with cancer began with the			
5	been applied to PTSD development in patients with cancer [3-5]. PTSD sym			
6	veloped specifically for use with patients with cancer and survivors			
7	[2]. Altered protein metabolism in patients with cancer may stem from			
8	py. A major cause of anorexia in patients with cancer is food odor.			
9	h reduced odors are preferable for patients with cancer. This explains			
10	taste that sometimes distresses patients with cancer, but which may			
3	PATIENTS TREATED	Radiotherapy (5), chemotherapy (4), surgically (3), survival (3), surgical (3)	22	0,03
1	consistent survival advantage for patients treated with higher doses			
2	significant survival advantage for patients treated with twice daily c			
3	respective studies have shown that patients treated with PCI do not ha			
4	of their cranial irradiation [26]. Patients treated for small cell lun			
5	e status, and histology [1,2]. For patients treated with aggressive su			
6	alleviate pain in the majority of patients treated. However, the dura			
7	onal chemotherapy. The majority of patients treated with radiotherapy			
8	e status, and histology [1,2]. For patients treated with aggressive su			
9	of pathologic stage I (T1, N0, M0) patients treated surgically, eligib			
10	reduction in local recurrence for patients treated with lobectomy com			
4	PATIENTS RECEIVING	Therapy (6), chemotherapy (6), placebo (5), radiation (5), emetogenic (5)	19	0,03
		Patients receiving placebo 4		
		Patients receiving radiation 4		
1	eous reactions including pruritus. Patients receiving antineoplastic d			
2	t has so far been most reported in patients receiving interferons [17-			
3	ancies and in approximately 40% of patients receiving chemotherapy for			
4	olon cancer, the survival time for patients receiving the hydrazine wa			
5	razine was decreased compared with patients receiving placebo [13]. Com			
6	ceiving placebo [13]. Compared with patients receiving placebo, patient			
7	d with patients receiving placebo, patients receiving hydrazine sulfat			
8	tment [1]. Approximately 10%-44% of patients receiving chemotherapy exp			
9	--- Radiation therapy: Although patients receiving radiation therap			
10	py can experience ANV, in general, patients receiving radiation to the			
4	PATIENT AND FAMILY	Perception (4), health (4), care (4)	19	0,03
1	es increased responsibility on the patient and family [2 and requires			
2	oing support and education for the patient and family. (Throughout thi			
3	Thus care needs are unique to each patient and family. Unfortunately,			
4	e status, treatment possibilities, patient and family goals and their			
5	[3]. Routine re-evaluations of the patient and family biopsychosocial			
6	biopsychosocial assessment of the patient and family [1]. This assessm			
7	h care team and community. Patient and family perception of il			
8	ily perception of illness. Patient and family perception of qu			
9	eption of quality of life. Patient and family fears and concer			
10	ut in ways that are unique to each patient and family. These feelings			

14.4.2.2. CANCER

En el 20% de las apariciones del lema CANCER, le precede el lexema *lung*. La UF resultante, *lung cancer*, también coocurre con una cierta frecuencia con los lemas SMALL y CELL, con los que forma dos UF que designan los dos principales tipos de cáncer de pulmón: *small cell lung cancer* y *non-small cell lung cancer*. En su cotexto, aparecen principalmente lexemas relacionados con los conceptos STAGE, PATIENTS y THERAPY.

PATIENTS aparece en 129 ocasiones en el radio colocacional de CANCER y forma las UF *cancer patients* y *patients with cancer*. Relacionadas conceptualmente encontramos las expresiones *people with cancer*, que en la mayoría de los casos está precedida por un porcentaje (*45% of people with cancer*), *survivors of cancer* y *individuals with cancer*.

Por último, destacamos dos colocaciones que asocian el lema CANCER al concepto THERAPY: *cancer treatment* y *cancer therapy*. Estas representan más del 6% de todas las instancias de dicho lema.

TABLA 136: Radio colocacional de *cancer**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional	Frec	%
		Colocaciones / unidades fraseológicas		
CANCER Cancer 440, cancers 29		Patients (129), lung (91), cell (71), treatment (49, small (48) <u>1</u> <u>1a</u> <u>1b</u> <u>2</u> <u>3</u> <u>4</u> Cancer treatment 20 Breast cancer 19 Survivors of cancer 11 Cancer experience 9 Cancer therapy 9 Cancer diagnosis 8 Individuals with cancer 6	469	0,66
1	LUNG CANCER Lung cancers 5	Cell (71), Small (47), stage (29), non-small (27), patients (21), Nsclc (9), limited (8), advanced (7), m(6), recurrent (5), carcinoma (5), chemotherapy (5), system (5) <u>Small cell lung cancer</u> (39) <u>Non-small cell lung cancer</u> (22)	92	0,13
1	<p>0 patients diagnosed with advanced lung cancer and advanced colon cancer. In the two studies of advanced lung cancer patients, the addition of Lung Persons with lung cancer make up 25%-35% of reproduction [1]. Hypercalcemia in lung cancer is usually not an early sign of advanced disease and/or lung cancer experienced more distress. People with lung cancer report higher levels of central nervous system metastases, lung cancer, and treatment with corticosteroids in patients with primary lung cancers. Determining whether treatment with corticosteroids is beneficial in patients with lung cancer compared with other cell types of lung cancer, small cell carcinoma patients are at risk of dying from lung cancer (both small and non-small).</p>			
1a	SMALL CELL LUNG CANCER	Stage small cell lung cancer 11 Limited stage small cell lung cancer 7 Patients with small cell lung cancer 7	39	0,05

1	thods under clinical evaluation in small cell lung cancer include addi			
2	urrent prognosis for patients with small cell lung cancer is unsatisfa			
3	inical evaluation in limited stage small cell lung cancer include new			
4	rent classification of subtypes of small cell lung cancer are:[1] s			
5	shown that long-term survivors of small cell lung cancer (>2 years fr			
6	motherapy alone in extensive stage small cell lung cancer. However, ra			
7	umor, and atypical carcinoid. Like small cell lung cancer, it occurs p			
8	erential pathologic diagnosis from small cell lung cancer may be diffi			
9	ave a major impact on treatment of small cell lung cancer since patien			
10	tone of treatment of limited stage small cell lung cancer. Combination			
1b	NON-SMALL CELL LUNG CANCER	Stage (12), patients (9) Non-small cell lung cancer (nscLC) 9	22	0,03
1	d in detail in the PDQ summary for non-small cell lung cancer, may als			
2	f choice for patients with stage I non-small cell lung cancer (NSCLC).			
3	***** GENERAL INFORMATION Non-small cell lung cancer (NSCLC)			
4	series of patients with resectable non-small cell lung cancer include			
5	. Histologic classification of non-small cell lung cancer: squa			
6	* TREATMENT OPTION OVERVIEW In non-small cell lung cancer (NSCLC),			
7	ferences> ***** OCCULT NON-SMALL CELL LUNG CANCER ----			
8	ical to that recommended for other non-small cell lung cancer patients			
9	----- STAGE 0 NON-SMALL CELL LUNG CANCER --			
10	idered for patients with stage III non-small cell lung cancer (NSCLC)			
2	CANCER PATIENTS	(See under <i>patients</i>)	77	0,11
3	PEOPLE WITH CANCER	Fatigue (7), helping (3), elderly (3), include (3), factors (3)	26	0,04
1	nt [2].It is estimated that 45% of people with cancer have sleep distu			
2	nt has not been studied in elderly people with cancer as a treatment f			
3	eat, is the most common symptom in people with cancer that may occur e			
4	een estimated that one half of all people with cancer experience cache			
5	s been estimated that up to 20% of people with cancer may die of the e			
6	quate nutritional status can help people with cancer feel and look be			
7	odor. Suggestions for helping people with cancer manage anorexia			
8	za). Suggestions for helping people with cancer manage taste cha			
9	ncer. Fatigue exists in 72%-95% of people with cancer, particularly in			
10	y decreasing muscle contractility. People with cancer who experience d			
4	PATIENTS WITH CANCER	(See under <i>patients</i>)	22	0,03

14.4.2.3. TREATMENT

TREATMENT es el tercer lema más frecuente del corpus. Lo hemos desglosado en dos para ver sus 363 realizaciones como sustantivo (*treatment**) y sus 97 apariciones como verbo (*treated, treating, treat y treats*). En su radio colocacional destacan los lexemas *patients, cancer, chemotherapy, disease, diagnosis y radiotherapy*, elementos que interactúan en todo el proceso que implica el cáncer.

Las UF que sobresalen son *cancer treatment, treatment options, standard treatment y patients treated*. Su radio colocacional pone de manifiesto una serie de síntomas asociados a la enfermedad (*hypercalcemia*), al tratamiento (*nausea*) y a ambos (*fatigue, anxiety*). El

cotexto también apunta hacia las principales terapias: *chemotherapy, radiotherapy y surgery.*

TABLA 137: Radio colocacional de *treat**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional	Frec	%
		Colocaciones / unidades fraseológicas		
TREATMENT Treatment (338), treated (65), treatments (25), treating (19), treat (12), treatable (1), treats (1)		Patients (69), cancer (63), chemotherapy (31), disease (23), diagnosis (20), radiotherapy (20)	461	0,65
TREATMENT(S)		<u>1</u> <u>2</u> Standard treatment 6	363	0,51
1	CANCER TREATMENT	Fatigue (6), common (5), factors (5), anxiety (4), undergoing (3), factors (3)	20	0,03
1	1 factors related to cancer and/or cancer treatment. Anxiety and depression to the diagnosis of cancer, cancer treatment, and hospitalizations problems are common during and after cancer treatment. Some of the reports in patients who are undergoing cancer treatment as well as in those threatening disease, of undergoing cancer treatment, and of living with undergoing any degree of autonomy (as cancer treatment often dictates), were time from the discontinuation of cancer treatment (a common research difficulty in adjusting to cancer and cancer treatment, their PTSD symptoms with the oral complications of cancer treatment should examine the and not adversely affected by the cancer treatment, enteral support i			
2	TREATMENT OPTIONS	Radiotherapy (6), chemotherapy (6), surgery (4), Standard (3), NSCLC (3), surgical (3)	20	0,03
1	line in intellectual function. Treatment options: Standard: 2 or low-dose regimens [22-24]. Treatment options: Standard: 3 second-line chemotherapy [13]. Treatment options: 1 Palliative 4 ly NSCLC remains to be proven. Treatment options: 1 Surgical re 5 resected for stage I NSCLC [13]. Treatment options: 1 Lobectomy o 6 resection should be encouraged. Treatment options: 1 Lobectomy, 7 far in the treatment of NSCLC. Treatment options: 1 Surgery alone 8 and to resect necrotic tissue. Treatment options: 1 Radiotherapy 9 tumor is completely resected. Treatment options: 1 Surgery [21 10 with or without chemotherapy. Treatment options: 1 Radiotherapy			
TREATED Treated, treating, treat , treats		Radiotherapy (11), cancer (11), chemotherapy (11), hypercalcemia (7), nausea (6) Patients treated 22 (see under <i>patients</i>)	97	0,15
1	nt survival advantage for patients treated with higher doses of chemotherapy survival advantage for patients treated with twice daily chest radiation superior vena cava syndrome are treated with combination chemotherapy patients with small cell lung cancer treated with chemotherapy with or without the studies have shown that patients treated with PCI do not have detectable logical function than patients not treated [22]. In addition, the major cranial irradiation [26]. Patients treated for small cell lung cancer Brain metastases are appropriately treated with whole-brain radiotherapy to identify and develop ways to treat patients of other cultures who strategies are perhaps more useful in treating anticipatory nausea and vomiting			

En el recuadro correspondiente al verbo *to treat*, quedan patentes las construcciones en las que suele aparecer:

- ProcExp (PATIENTS) + *treated with* + Force¹⁸³ (TREATMENT)
- Proc (DISEASE) + *treated with* + Force (TREATMENT)
- *Treat* + ProcExp (PATIENTS)
- *Treat* + Goal (DISEASE / SYMPTOMS)

14.4.2.4. SYMPTOMS

Al estudiar el contexto del lema SYMPTOMS es posible identificar las unidades de significación que designan los síntomas más habituales que padece el enfermo de cáncer: *PTSD (post traumatic stress disorder)*, *PTSD symptoms*, *anxiety*, *anxiety symptoms*, *hypercalcemia* y *gastrointestinal symptoms*. Encontramos en su horizonte textual diversas menciones a los intentos por contrarrestar estos síntomas. Por ejemplo, *treatment*, causa y solución de algunos de estos síntomas, *control*, *palliation of symptoms* y *symptomatic relief*.

TABLA 138: Radio colocacional de *symptom**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones / unidades fraseológicas			
SYMPTOMS Symptoms (212), symptom (40), Symptomatic (24), symptomatology (4), Symptomatically (1)	Patients (29), cancer (24), treatment (22), PTSD (21), anxiety (14), associated (15), control (13), hypercalcemia (13)		281	0,39
SYMPTOM (S)	Palliation of symptoms 9 Gastrointestinal symptoms 6	Anxiety symptoms 6 PTSD symptoms 6	252	0,35
<p>1 ly important role in palliation of symptoms of the primary tumor and o 2 ronchoscope, resulting in improved symptoms and pulmonary function in 3 ant tumor responses, palliation of symptoms, and short-term local cont 4 ces can often obtain palliation of symptoms with radiotherapy and/or a 5 en be appropriately deferred until symptoms or signs of progressive tu 6 or chemotherapy for palliation of symptoms from the primary tumor. Pa 7 ven to improve survival or control symptoms for a prolonged period of 8 enced by the presence of pulmonary symptoms, large tumor size (>3 cent 9 erm improvement in disease-related symptoms. In one study, symptomatic 10 val, and palliation of lung cancer symptoms [6]. With the paclitaxel p</p>				

¹⁸³ En Dik (1989: 101-3), el rol semántico *Force* se define como "the non-controlling entity instigating a Process". *Processed* es "the entity that undergoes a Process", y *Goal*, "the entity affected or effected by the operation of some controller (Agent/Positioner) or Force".

SYMPTOMATIC , symptomatically	Symptomatic relief 6	25	0,04
1	apy may be effective in palliating symptomatic local involvement with		
2	ily for palliative relief of local symptomatic tumor growth. 2 Che		
3	10]. In studies that have examined symptomatic response, improvement i		
4	s with good performance status and symptomatic recurrence can be offer		
5	al and peritoneal, represent major symptomatic problems for at least t		
6	can provide palliative relief from symptomatic effusions, discomfort c		
7	ND IV) Treatment options: 1 Symptomatic treatment to include dr		
8	n. Oral antihistamines may provide symptomatic relief in histamine-rel		
9	of delirium must take place while symptomatic and supportive therapie		
10	g in which care is being provided. Symptomatic treatment measures incl		

14.4.2.5. THERAPY

Relacionado también con el TRATAMIENTO, aparece entre los más frecuentes el lema THERAPY, que es muy productivo a la hora de formar UF que designan divesos tratamientos. Estas unidades terminológicas pueden especificar:

- el agente terapéutico (*radiation therapy, laser therapy*)
- el agente terapéutico y la parte del cuerpo donde se aplica (*endobronchial laser therapy*)
- lo que pretende combatir la terapia (*antineoplastic therapy, antibiotic therapy, cancer therapy, grief therapy*)
- el hecho de que se utiliza más de una terapia (*combined modality therapy, combined modality therapy*).

TABLA 139: Radio colocacional de *therap**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones / unidades fraseológicas			
THERAPY Therapy (185), therapeutic (16), therapies (14), therapist (3), therapeutically (1)	Radiation (54), patients (20), chemotherapy (17), cancer (17), grief (17), oral (15), antineoplastic (10)		219	0,31
	1 2 Antineoplastic therapy 9 Cancer therapy 9 Antibiotic therapy 7	Combined modality therapy 6 Laser therapy 6 Combination therapy 5 Endobronchial laser therapy 5		
1 RADIATION THERAPY	Chemotherapy (12), associated (6), patients (5)		43	0,06

1	ation chemotherapy with or without radiation therapy [15,16]. A small		
2	m 6% to 30% [1,3]. The addition of radiation therapy and/or chemothera		
3	improved survival [2]. The use of radiation therapy in pleural mesoth		
4	ion) with or without postoperative radiation therapy B) Extrapleural		
5	ural pneumonectomy C) Palliative radiation therapy Under clini		
6	d patients [2,3]. 3 Palliative radiation therapy [4,5]. 4 Sing		
7	nurse. Chemotherapy nurse. Radiation therapy nurse. Dietit		
8	cells respond to chemotherapy and radiation therapy. Some studies in		
9	f melatonin on chemotherapy and on radiation therapy may be beneficial		
10	d with dry desquamation related to radiation therapy, but should not b		
2	GRIEF THERAPY		13 0,02
1	ieving tasks. Rando [3].notes that grief therapy includes dealing with		
2	tization. Grief counseling and grief therapy are distinguished fro		
3	t on a time-limited basis (as with grief therapy). - helping the		
4	making referrals for professional grief therapy. Grief therapy		
5	professional grief therapy. Grief therapy is used with people w		
6	tion of this summary). The goal of grief therapy is to identify and re		
7	not present with cancer deaths). Grief therapy can be provided on an		
8	cal illness must be ruled out. Grief therapy requires talking abou		
9	f more ambivalent, angry feelings. Grief therapy may allow the individ		
10	s and vice versa. The focus of grief therapy is dependent upon an		

14.4.2.6. CHEMOTHERAPY

La quimioterapia es el tratamiento más mencionado en el corpus, ya sea mediante el lema CHEMOTHERAPY, o mediante otros relacionados como DOSES, DRUGS, COMBINATION y ADMINISTRATION, que se encuentran en los 55 primeros puestos de la lista lematizada¹⁸⁴. En el contexto del lema CHEMOTHERAPY encontramos referencias al paciente (*patients*), a otras terapias (*radiotherapy, radiation, treatment, etc.*) y a la interacción con estas (*adjuvant*), a la utilización de más de un medicamento (*combination, regimens*), a sus efectos secundarios (*antiemetic*), y a la acción de someterse a esa terapia (*receiving, treated*). Entre las UF más relevantes hemos señalado: *combination chemotherapy, chemotherapy regimen(s), adjuvant chemotherapy, emetogenic chemotherapy* y *cisplatin-based chemotherapy*. Esta última nos indica que el cisplatino es el medicamento más empleado en la quimioterapia contra el cáncer de pulmón.

¹⁸⁴ La lista lematizada contiene 4197 elementos.

TABLA 140: Radio colocacional de *chemotherap**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional		Frec	%
		Colocaciones / unidades fraseológicas			
CHEMOTHERAPY Chemotherapy (202) chemotherapeutic (12) chemoprevention (3)		Radiotherapy (35), patients (28), combination (25), radiation (20), treatment (16), adjuvant (15), surgery (15), regimens (14), therapy (14), receiving (14), emetogenic (14), alone (13), chest (13), irradiation (11), resection (11), treated (11)		217	0,3
CHEMOTHERAPY		Combination chemotherapy 20 Chemotherapy regimen(s) 10 Adjuvant chemotherapy 9 Emetogenic chemotherapy 8	Receiving chemotherapy 7 Cisplatin-based chemotherapy 6	202	0,28
<p>1 es of about 5% for those receiving chemotherapy and radiotherapy compa 2 [1]. With incorporation of current chemotherapy regimens into the trea 3 m surgery with or without adjuvant chemotherapy; these patients have a 4 ould initially receive combination chemotherapy regardless of the exte 5 n or surgical excision is added to chemotherapy alone for limited stag 6 include adding chest radiation to chemotherapy regimens, varying drug 7 ximately 40% of patients receiving chemotherapy for malignancies in ot 8 ients treated with higher doses of chemotherapy [1-3]. Even chemothera 9 r than others [1]. Cisplatin-based chemotherapy has been associated wi 10 doses of chemotherapy [1-3]. Even chemotherapy of the intensity used</p>					
1	COMBINATION CHEMOTHERAPY	Patients (5), treatment (3), chest (3), regimens (3) Combination chemotherapy regimens 3		20	0,03
<p>1 ically resected patients, adjuvant combination chemotherapy with cispl 2 uated either two to four cycles of combination chemotherapy or combina 3 ena cava syndrome are treated with combination chemotherapy with or wi 4 eatment options: Standard: 1 Combination chemotherapy with one o 5 ter define the role of these newer combination chemotherapy regimens i 6 t disseminated disease is present, combination chemotherapy is the cor 7 prove survival duration [6-9]. Combination chemotherapy plus chest 8 taining and carboplatin-containing combination chemotherapy regimens p 9 one study, symptomatic relief with combination chemotherapy was signif 10 ically resected patients, adjuvant combination chemotherapy with cispl</p>					
2	CHEMOTHERAPY REGIMEN(S)	Combination chemotherapy regimens 3		13	0,02
<p>1 [1]. With incorporation of current chemotherapy regimens into the trea 2 include adding chest radiation to chemotherapy regimens, varying drug 3 sted below. The use of alternating chemotherapy regimens has not prove 4 d those who have received multiple chemotherapy regimens rarely respon 5 val advantage. Currently no single chemotherapy regimen can be recomme 6 lowing surgery. Trials of adjuvant chemotherapy regimens have failed t 7 t have used modern cisplatin-based chemotherapy regimens [4-7]. A meta 8 carboplatin-containing combination chemotherapy regimens produce objec 9 he role of these newer combination chemotherapy regimens in the treatm 10 trials evaluating the role of new chemotherapy regimens. Refer to the</p>					
CHEMOTHERAPEUTIC		Chemotherapeutic agents 7		12	0,02

14.4.3. Unidades de significación de la lengua general *terminologizadas* en el subdominio de la oncología

En este apartado se presentan lemas de la lengua que adquieren un significado especial en los textos sobre oncología, en concreto, USE, STAGE y COMBINATION. Por otra parte, llamaremos la atención sobre la prosodia semántica de lemas con los que se expresan el efecto del tratamiento y la aparición del cáncer, de síntomas o de efectos secundarios.

14.4.3.1. USE, STAGE, COMBINATION-COMBINED

USE es el lexema genérico para designar la utilización de un determinado tratamiento contra el cáncer, por lo que podemos decir que su significado básico queda restringido. También hemos detectado una tendencia de este lema por coocurrir con lexemas relacionados con la quimioterapia (*chemotherapy, agents, drugs, medications*), como indica la siguiente casilla:

TABLA 141: Radio colocacional del lema USE

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones / unidades fraseológicas			
USE Use(150), used(128), Using(21), uses (2)	Patients (26), agents (13), treatment (12), therapy (11), drugs (11), management (10), medications (10), include (10), bowel (10)		301	0,42
	Commonly used 8 Used prophylactically 7	Used with caution 6		
1	tions. Staging procedures commonly used to document distant metastases			
2	nto clinical trials evaluating the use of adjuvant treatment with chem			
3	midplane of the known tumor volume using conventional fractionation. A			
4	Even chemotherapy of the intensity used in autologous bone marrow tran			
5	rent combinations of chemotherapy, using different schedules of chemot			
6	nt chest radiotherapy has now been used in multiple single institution			
7	hest radiation schedules have been used in regimens with etoposide and			
8	radiotherapy are listed below. The use of alternating chemotherapy reg			
9	stage disease. Doses and schedules used in current programs yield over			
10	extensively or are in less common use, including: cyclophosphamide			

Otro lexema de la lengua general cuyo significado queda restringido por el cotexto es *stage*, que en el campo de la oncología hace referencia a un sistema internacional de clasificación del cáncer que evalúa su extensión en el organismo. Este sistema llamado *TNM Staging System* fue creado en 1985 por tres organizaciones (*American Joint Committee on Cancer, Union Internationale Contre le Cancer* y *Japanese Cancer Committee*) y se fundamenta en tres aspectos: las características del tumor primario (T), la extensión a

ganglios linfáticos (N, por el término *lymph nodes*) y las metástasis producidas (M). Este sistema fue evaluado en 1997 para dar un mayor grado de especificidad al diagnóstico. Teniendo en cuenta que una evaluación acertada del estadio de la enfermedad condiciona el éxito del tratamiento, el lema STAGE es muy relevante en los textos para profesionales de la salud (0,23%).

La forma *stage* suele coocurrir con números romanos, solos o en combinación con las letras *A* y *B* y con los adjetivos *limited* y *extensive*. También le acompañan nomenclaturas de la enfermedad (*small cell lung cancer, non-small cell lung cancer, disease*) y alusiones al sistema *TNM*.

TABLA 142: Radio colocacional de *stag**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional		Frec	%
		Colocaciones / unidades fraseológicas			
STAGE Stage (120), staging(19), stages(18), staged(5),		Patients (50), cancer (39), lung (34), disease (33), cell (32), limited (24), M (21), T (20), I (20), small (17), treatment (12), extensive (12), non-small (12), NSCLC (6), IV (10), II (10)		162	0,23
		1 2 Stage small cell lung cancer 11 # stage disease 10 Stage information 6	Staging system 6 Pathologic stage 3 Carefully staged 3 Staging procedures 3		
1 s are designated as having limited stage disease, and most 2-year dise 2 on of this summary on treatment of stage IV disease). Many randomized 3 ted stage patients, results of the staging process have therapeutic im 4 r areas are said to have extensive stage disease and have a worse prog 5 ement. Therefore, the detailed TNM staging system developed for lung c 6 CELL LUNG CANCER Patients with stage IIIb non-small cell lung canc 7 from treatment. Regardless of stage, the current prognosis for pa 8 ical cure appears possible in most stage I patients. Careful diagnosis 9 mination. However, determining the stage of cancer by nonsurgical mean 10 ess have therapeutic implications. Staging procedures commonly used to					
1	LIMITED STAGE	Patients (11), small (11), cell (10) Limited stage small cell lung cancer 7 Limited stage disease 6 Patients with limited stage 6 Patients with limited stage small cell lung cancer 3		21	0,03

1	patients are designated as having limited stage disease, and most 2-y			
2	survivors come from this group. In limited stage disease, median survi			
3	small proportion of patients with limited stage disease may benefit f			
4	worse prognosis than patients with limited stage. Median survival of 6			
5	ormance status, female gender, and limited stage disease [3,9,10]. Pat			
6	is added to chemotherapy alone for limited stage patients, results of			
7	tandard treatment of patients with limited stage small cell lung cance			
8	E SMALL CELL LUNG CANCER As in limited stage small cell carcinoma			
9	included within and excluded from limited stage by various groups.			
10	is the cornerstone of treatment of limited stage small cell lung cance			
2	EXTENSIVE STAGE	Small (6), cell (6)	9	0,01
		Extensive stage small cell lung cancer	4	
		Extensive stage disease	3	
1	aclavicular areas are said to have extensive stage disease and have a			
2	stage by various groups. ----- Extensive stage Extensive stage s			
3	roups. ----- Extensive stage Extensive stage small cell lung can			
4	<References> ***** EXTENSIVE STAGE SMALL CELL LUNG CAN			
5	der to produce the best results in extensive stage disease. Doses and			
6	plete response rates of 20%-30% in extensive stage disease. Since over			
7	ompared with chemotherapy alone in extensive stage small cell lung can			
8	ction. Many more patients with extensive stage small cell carcinom			
9	s of active clinical evaluation in extensive stage small cell lung can			

La representación conceptual del estadiaje de la enfermedad en los textos es aún mayor si tenemos en cuenta otras formas que, si bien no contienen la palabra *stage*, designan estadios. Nos estamos refiriendo a las siguientes formas léxicas, que juntas constituyen un 0,08 % del corpus: *carcinoma in situ (Tis)*, *occult carcinoma* y tres lexemas que se anteponen a cualquier mención a la enfermedad, *advanced*, *localized* y *recurrent*.

(159)

1	Standard treatment for all but localized mesothelioma is generally			
2	<References> ***** LOCALIZED MALIGNANT MESOTHELIOMA			
3	t rated symptoms. Individuals with advanced disease and/or lung ca			
4	ally (T3-T4) or regionally (N2-N3) advanced lung cancer who have a			
5	n to be active in the treatment of advanced NSCLC. Chemoprevention			
6	clinical trials for patients with recurrent small cell lung cance			
7	LUNG CANCER Many patients with recurrent non-small cell lung c			
8	NANT MESOTHELIOMA Treatment of recurrent mesothelioma usually			
9	evidence of primary tumor. Tis: Carcinoma in situ. T1: A tumo			
10	lung cancer (NSCLC) is the same as carcinoma in situ of the lung.			
11	----- AJCC stage groupings Occult carcinoma: TX, N0, M0 /			

Por otra parte, comentamos cómo el cotexto restringe el significado de las formas léxicas *combination* y *combined*. Antes de analizar el corpus, se podría argumentar que estas dos formas léxicas son intercambiables cuando preceden al sustantivo *therapy* o *treatment*.

Sin embargo, las líneas de concordancia indican claras diferencias con respecto al uso y una cierta preferencia por la utilización de las formas *combination* y *combinations* cerca de términos relacionados con la quimioterapia.

TABLA 143: Radio colocacional de *combin**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional	Frec	%
	Colocaciones /unidades fraseológicas		
COMBINATION Combination (62), combined(27), combinations(16), combine(3), combining(2), combines(2)	Chemotherapy (37), drug* (18), therapy (18), modality (12), regimens (12), radiotherapy (11), antiemetic (10), patients (9), modalities (8), approaches (7), small cell (7), treatment (6), cisplatin (6)	112	0,16
1	laser therapy, brachytherapy, and combined modality approaches may le		
2	Herapy, chemotherapy, surgery, and combinations of these modalities. A		
3	These results are encouraging, and combined-modality therapy with neoa		
4	Apy [24]. 6 Clinical trials of combined modality therapy [22].		
5	Have been evaluated in single and combined modality studies. The most		
6	ions, though the efficacy of these combined approaches is untested. Th		
7	mg) can promote sleep and is often combined with other antidepressants		
8	ng aspirin. Cimetidine alone or in combination with aspirin has been u		
9	particular emetogenic drug or drug combination. Anticipatory symptoms		
10	ons. Because most patients receive combination chemotherapy, the emeto		
11	tes and mechanisms of action. Most combination regimens combine a dopa		
12	s been shown to be superior to the combination of intravenous metoclop		
13	ically resected patients, adjuvant combination chemotherapy with cispl		
14	chemotherapeutic regimen [153].The combination of a serotonin antagoni		

En la tabla 144 queda representado el radio colocacional de COMBINATION. En él predominan los lexemas relacionados con la quimioterapia (*chemotherapy, drug(s), regimens, cisplatin, dexamethasone*), seguidos de los que se refieren a otras terapias (*therapy, radiotherapy, treatment*), a sustancias que impiden el vómito (*antiemetics*) y a los pacientes. La UF *combination chemotherapy* aparece en veinte ocasiones. Cuando se especifican los medicamentos administrados al paciente, se emplean las expresiones:

- DRUG + *combinations*: *cisplatin-based combinations*
- *combination of* + DRUG: *the combination of cisplatin and paclitaxel*
- DRUG + *in combination with* + DRUG: *Both granisetron and ondansetron demonstrate greater antiemetic efficacy in combination with corticosteroids.*

TABLA 144: Radio colocacional de *combination*

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Unidades fraseológicas			
COMBINATION Combination 62, Combinations 16	Chemotherapy (26), drug* (13), regimens (11), antiemetic (10), patients (8), therapy (8), cisplatin (6), radiotherapy (6), dexamethasone (5), treatment (5)		78	
	combination chemotherapy 20	combination therapy 5		
	combination antiemetic regimens 4 combination chemotherapy regimens 3	used in combination 3		
1	patients should initially receive combination chemotherapy regardless			
2	- 24% [4,5]. Reports of paclitaxel combinations have shown relatively			
3	trials showed that cisplatin-based combinations plus radiotherapy resu			
4	gher 1 year survival rate than the combination of cisplatin and etopos			
5	trol tumors in the thorax. The combination of etoposide and cispla			
6	tiveness of many two- to four-drug combination programs appears simila			
7	sions for early change to standard combination therapy if there is no			
8	diotherapy or with radiotherapy in combination with other therapy moda			
9	therapy, chemotherapy, surgery, and combinations of these modalities. A			
10	ncer patients who may be receiving combinations of several treatments			

La mención a la interacción de terapias es mucho menos habitual que la coocurrencia de COMBINATION con realizaciones del concepto CHEMOTHERAPY. Encontramos la UF *combination therapy* y las expresiones *in combination with* y *combinations of*, tal y como ilustran las líneas de concordancia 8, 9 y 10.

COMBINED tiene en su radio colocacional el lexema *modality* con el que constituye las UF *combined modality therapy*, *combined modality approaches* y *combined modality treatment*. Cuando se precisa especificar las terapias que se combinan, se utiliza la expresión *combined with*, precedida y seguida por las terapias en cuestión, como ocurre en las líneas de concordancia número 6, 7, 8 y 9. En la línea 10 presentamos un uso de *combined* que nada tiene que ver con el área conceptual TRATAMIENTO.

TABLA 145: Radio colocacional de *combined*

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Unidades fraseológicas			
COMBINED	Combined modality	11	27	
	Combined modality therapy	6		
	Chemotherapy combined with	5		
	Chemotherapy combined with other modalities	4		
1	ive randomized trials suggest that combined modality therapy produces			
2	ce status should be considered for combined modality therapy. However,			
3	zers, radiolabeled antibodies, and combined modality approaches may le			
4	atients less than 65 years of age. Combined modality treatment is asso			
5	enic potential of all of the drugs combined and individual drug doses			
6	otherapy alone. 2 Chemotherapy combined with radiotherapy [1-3,8].			
7	ns to surgery). 3 Radiotherapy combined with curative surgery [8].			
8	y [8]. 4 Adjuvant chemotherapy combined with other modalities [4,6			
9	elected cases). 4 Chemotherapy combined with other modalities.			
10	carcinoma (small cell lung cancer combined with neoplastic squamous a			

14.4.3.2. PROSODIA SEMÁNTICA EN LA EXPRESIÓN DE LA CAUSA Y EL EFECTO Y DE LA APARICIÓN DE CÁNCER/ SÍNTOMAS / EFECTOS SECUNDARIOS

14.4.3.2.1. EFFECT, CAUSE, RESPONSE, PRODUCE, FACTOR

Entre las palabras de la lengua general más frecuentes en el texto son significativas las que expresan causalidad y, sobre todo, el efecto, positivo o negativo, que producen las terapias del cáncer. De los lemas que mencionamos en este apartado, EFFECT, CAUSE, RESPONSE y PRODUCE, nos hemos percatado de que la prosodia semántica de los dos primeros es negativa en la mayoría de los casos, mientras que RESPONSE y CAUSE indican un resultado neutro o positivo.

14.4.3.2.1.1. EFFECT

El lema EFFECT suele coocurrir con unidades léxicas que indican entidades, propiedades y eventos negativos, como indica la tabla 146, que recoge las unidades léxicas con las que forma colocación.

TABLA 146: Colocaciones de *effect** en el corpus de textos para especialistas

WORD	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
EFFECTS	106	6	2	2	1	2	1	0	98	0	0	1	0	1
EFFECT	67	3	0	1	1	1	0	0	64	0	0	0	0	0
SIDE	42	39	3	0	1	0	0	38	0	0	0	0	1	2
ADVERSE	19	19	0	0	0	0	1	18	0	0	0	0	0	0
TREATMENT	9	3	6	0	0	1	2	0	0	0	4	1	0	1
CANCER	9	4	5	2	0	2	0	0	0	0	2	0	2	1
INCLUDE	8	1	7	1	0	0	0	0	0	4	1	2	0	0
TUMOR	8	4	4	0	1	0	0	3	0	0	0	1	2	1
TOXIC	8	7	1	0	0	0	0	7	0	0	1	0	0	0
THERAPY	7	1	6	1	0	0	0	0	0	0	2	3	1	0
DRUGS	7	4	3	1	1	2	0	0	0	0	0	2	0	1
DAYS	7	3	4	1	0	1	1	0	0	0	0	1	1	2
PATIENTS	7	3	4	3	0	0	0	0	0	0	1	0	0	3
DOSES	6	2	4	0	2	0	0	0	0	0	0	2	2	0
HYPOCALCEMIC	6	6	0	0	1	0	0	5	0	0	0	0	0	0
ASSOCIATED	5	0	5	0	0	0	0	0	0	3	0	0	2	0

Esta prosodia semántica negativa también está presente en la mayoría de las UF que forma, *side effect(s)*, *toxic effect(s)* y *adverse effects*, como indica la tabla:

TABLA 147: UF formadas sobre la base *effect**

CLUSTER	FREQ.
side effects	31
adverse effects	14
side effect	9
toxic effects	8
CLUSTER	FREQ.
hypocalcemic effect	5
adverse effect	3
direct effect	3

Sin embargo, la forma léxica *effective* siempre indica resultados positivos, de ahí que en la ficha esté separada de *effects*.

TABLA 148: Radio colocacional de *effect**

PALABRA CLAVE y formas léxicas		Lexemas en su radio colocacional	Frec	%
		Colocaciones / unidades fraseológicas		
EFFECTS Effects (99), effective(69), effect(63), effectiveness(9), effectively(4)		Side (43), adverse (19), treatment (19), patients (13) / , therapy (11), agents (10), include (10), cancer (9)	244	0,34
EFFECTS Effects (99), effect (63)		Side (39), Adverse (19), treatment (9), cancer (9), include (8)	162	0,23
		1 2 Toxic effects 8, hypocalcemic effect 5, effects include 4,		
<p>1 more drugs are needed for maximal effect. Mature results of prosp 2 om pulmonary and hematologic toxic effects; proper administration requ 3 , those studies showing a positive effect for combined modality therap 4 0-200 cGy to avoid long-term toxic effects to normal brain tissue [13, 5 because of the high risk of toxic effects observed with such treatmen 6 acted out in destructive ways, the effects can be devastating. Likewis 7 tively on sleep patterns. Side effects of treatment that may affec 8 me anxiety, and serious withdrawal effects, such as seizures [3].The f 9 nzodiazepines have an anti-anxiety effect in low doses and a hypnotic 10 d that melatonin exerts a hypnotic effect through thermoregulatory mec</p>				
1	SIDE EFFECTS (31) Sside effect 9	Treatment (6), vomiting (4), sleep (3), extrapyramidal (3), cancer (3), nausea (3), pruritus (3)	40	0,06
<p>1 negatively on sleep patterns. Side effects of treatment that may 2 cs in reducing REM sleep; however, side effects prior to sleep inducti 3 therapy and biologic agent-induced side effects, antibiotic reactions, 4 d moist desquamation results. This side effect increases the risk of i 5 associated with a wide variety of side effects and toxic effects. Pru 6 toxic effects. Pruritus has been a side effect associated with several 7 .To date, reports of pruritus as a side effect of biologics are primar 8 ighed against the vasoconstrictive side effects. Topical steroids have 9 effective and does produce gastric side effects [17]. Aspirin seems 10 eated witnessing of treatments and side effects. For instance, in an i</p>				
2	ADVERSE EFFECTS Adverse effects (14), adverse effect (3)	Include (4)	17	0,02
<p>1 nts. Barbiturates have a number of adverse effects, including the deve 2 ize clinical response and minimize adverse effects. Some patients may 3 y doses; however, the incidence of adverse effects increases proportio 4 etween increments. The most common adverse effect at clinically useful 5 sing cayenne pepper candies due to adverse effects. Thus far, evid 6 thiazines are differences in their adverse effect profiles, which subs 7 een evaluated [34,45-50].The major adverse effects include headache (w 8 ntagonists is that they have fewer adverse effects. Granisetron ha 9 oclopramide, steroids may mitigate adverse effects, such as the freque 10 or up to 6 doses/day [116-120]. Adverse effects experienced along w</p>				
EFFECTIVE Effective (69), effectiveness (9), effectively (4)		Treatment (9), patients (6), agents (6), oral (6), found (4)	82	0,11
		More effective 7, most effective 5, minimally effective 4, effective management 3, effective in reducing 3, effective in treating 3		

1 ational treatment seems to be most effective for very early central tu
 2 chemotherapy has been shown to be effective in these clinical situati
 3 erapy regimens has not proven more effective than the consistent admin
 4 py) treatments and developing more effective systemic therapy. Several
 5 benefits. Radiotherapy may be effective in palliating symptomatic
 6 egimens that have been found to be effective by at least two independe
 7 etastases. 3 Identification of effective new agents is difficult i
 8 maximal benefit. The relative effectiveness of many two- to four-
 9 th T3 or N2 disease can be treated effectively with surgical resection
 10 rd fractionation) are necessary to effectively control tumors in the t

Lo significativo de estos resultados es que apuntan hacia diferencias entre la prosodia semántica de *effect/effects* en textos sobre oncología y la que tiene en un corpus de referencia en inglés. Para que la comparación no esté sesgada por factores diatópicos o por las diferencias que separan el lenguaje oral del escrito, se verá el cotexto de *effect/effects* en una porción de 9 millones de palabras en inglés norteamericano que contiene material escrito y transcripciones de programas de radio.

Esta porción ha sido extraída del *Bank of English*, un corpus auspiciado por la editorial HarperCollins y la Universidad de Birmingham, del que se pueden extraer líneas de concordancia y colocaciones mediante el servicio *CobuildDirect*, accesible en http://titania.cobuild.collins.co.uk/direct_info.html.

En el corpus de referencia predominan las instancias donde el lema EFFECT tiene un efecto neutro o positivo.

(160)

was due `just" to placebo effects and suggestion? The cure would be
 These essentially have the same effect as the traditional scrubs. Scrubs
 hallucination? And yet the total effect, as in any successful novel, is a
 there is some agreement on the effects. Children enrolled in Head Start
 conducted by skeptics showed no effect. Consider also the facts
 the point of the whole business." In effect, Dwight felt that arrest was an
 friends this has had a very positive effect. Exhibitions, parties, concerts and
 of the Association with effect from 1 July 1992. Substantial
 and ecstasy) have a stimulant effect, giving a rush of energy and making
 fellow and his views have a great effect on French". He hoped that Kitchener
 or correction of the negative effects. In Lilly's idiom, they can
 bends only 60°. The combined effect is that the `scope should be
 in the carer all have a positive effect on the outcome, and it is therefore
 it into a beehive. To get the full effect Mr. Dambar asked her to take the
 judging whether cause is related to effect. No part of math has a more direct
 even mediocre sex can have the same effect, not only because of the physical
 I was thrilled by the unifying effect of the scalloped edging of silver
 parents' behavior over time has the effect of changing the family system over
 New York University, 1985. For effect of Therapeutic Touch on anxiety,

within a week or so. The long-term effects of a cloud that slowly thins out but unprepared for the side effects of this Godlike interference with matter, however, the sexual effects of antidepressant drugs are a single answer as to the relative effects of heredity and environment. if scientific proof for the healing effects of prayer existed, surely it would see E. Keller and V. M. Bzkek, 'Effects of Therapeutic Touch on Tension Desert Storm. [p] Siegel: Also, the effect of the international embargo on pack. But others worried about the effect of his vice presidential choice. Instrument (SPI) or to evaluate the effects of treatment using the Stuttering This would have a far-reaching effect on the economy. But when the infant's birth may have a sizable effect on the marital relationship in the from smiling. Robin always had that effect on her. 'Oh, all right!' [p] She New York, the fire is also have an effect on securities trading in Chicago. since 1982, could have crippling effects on the nation's economy. NPR's a sensitive issue of racism and its effects on baseball; particularly concerns to have very concrete and long-term effects. Students who believe they can do [p] The attorney general says, in effect, that if there's enough evidence the razing of German cities would effect the destruction of Nazi industry Geneticists Identify Genetic Effects? The Nurture Side of the Equation and counter-claims as to just what effect the energy tax will have. [p] JOHN have some political psychological effect, though not much military effect. footswitch controllers that allow effects to be turned on or off within the be credited with negative traits. In effect, we have another instance of ` likewise had little or no cultural effect, whether because of the brief

Estos resultados también se confirman en el listado de adjetivos que preceden al sustantivo *effect(s)*, que hemos sacado de las 100 muestras ordenadas según el *t-score*¹⁸⁵ que proporciona la aplicación *CobuildDirect collocation sampler*. La prosodia semántica negativa que dominaba en los textos sobre oncología no se percibe aquí.

Greenhouse	631	161	12.641163
same	28385	120	8.483918
positive	4076	72	8.027287
immediate	2572	62	7.562573
negative	2175	58	7.343479
side	21609	89	7.250085
dramatic	2243	55	7.127835
desired	476	51	7.077879
beneficial	585	48	6.847697
opposite	2565	51	6.798980
profound	598	47	6.772489
devastating	673	47	6.762058
little	32454	92	6.365646
adverse	349	37	6.028059
overall	2787	41	5.988134
such	37145	86	5.454673
similar	6335	40	5.369543
maximum	2035	32	5.313864
cumulative	110	26	5.078451
detrimental	153	25	4.970825
lasting	1048	26	4.903060

¹⁸⁵ Al ordenar los datos conforme al *t-score*, se intenta relacionar la frecuencia absoluta de las colocaciones con la frecuencia de las colocaciones relativa a la frecuencia de la base y el colocador por separado.

significant	3428	29	4.778241
sound	7334	34	4.631746
net	3127	27	4.622382
psychological	1501	22	4.385302
great	28072	61	4.383355
greater	4722	27	4.329717
due	6713	28	4.081936
therapeutic	388	17	4.033383
combined	1886	19	3.946367
serious	7120	26	3.767690
doppler	21	14	3.736306
protective	638	14	3.579084
damaging	783	14	3.542136
healing	1136	14	3.452185
calming	178	12	3.415110
powerful	3875	18	3.371822

14.4.3.2.1.2. CAUSE

El lema CAUSE normalmente lleva como sujeto con rol semántico *Force*, bien un agente terapéutico o un término que designa el cáncer, y como objeto o sujeto paciente, términos relacionados con los efectos secundarios del tratamiento y los síntomas de la enfermedad, como se ve en las líneas de concordancia.

TABLA 149: Radio colocacional de *caus**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional	Frec	%
	Colocaciones /unidades fraseológicas		
CAUSE Cause(82), causes (45), caused(37), causing(13), causative(2)	Cancer (16), anxiety (15), fatigue (12), patients (10), factors (10), obstruction (8), tumor (8), pruritus (7), fever (7), bowel (7), constipation (7), death (7), hypercalcemia (7), disorder (6), increased (6), symptoms (6) Cause anxiety 5, cause insomnia 5	179	0,25
1	. The epithelial form occasionally causes confusion with peripheral an		
2	symptomatic effusions, discomfort caused by tumor burden, and pain ca		
3	t caused by tumor burden, and pain caused by invasive tumor. Operative		
4	ances. Hypnotic drugs can also cause insomnia in cancer patients.		
5	cohol, and thyroid preparations can cause insomnia. In addition, withdr		
6	caine, phencyclidine, opioids) may cause insomnia. Agents containing a		
7	Agents containing aspirin may also cause insomnia. The most commonly p		
8	mb movement disorder, and when the cause of insomnia is uncertain or w		
9	bance. When sleep disturbances are caused by symptoms of cancer or tre		
10	ythemia vera. Some conditions that cause iron deficiency, including ex		

14.4.3.2.1.3. RESPONSE

El lema RESPONSE es la forma neutra que marca el grado de respuesta del organismo a un determinado tratamiento. Esta respuesta, que cuando se produce es algo positivo, puede ser completa (*complete response*) o parcial (*partial response*). La efectividad de un tratamiento se mide con la expresión *response rates* seguida de un porcentaje.

TABLA 150: Radio colocacional de *respons**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones /unidades fraseológicas			
RESPONSE Response 79, responses 21	Rates (18), patients (15), complete (13), chemotherapy (11), biologic (8), treatment (7), modifiers (7), objective (7), survival (6), hypocalcemic (6), tumor (6)		124	0,17
	1 complete response(s) 10	hypocalcemic response 3		
	biologic response modifiers 7	partial responses 3		
	objective response 4	psychological responses 3		
1 of tumor that can be evaluated for response. In clinical situations wh 2 t programs yield overall objective response rates of 65%-90% and compl 3 onse rates of 65%-90% and complete response rates of 45%-75%. Because 4 CI given to patients with complete responses): The following regimens 5 hout PCI in patients with complete responses), especially in patients 6 rapy can produce objective partial responses and palliation of symptom 7 in current programs yield overall response rates of 70%-85% and compl 8 itability). Other common emotional responses associated with such trau 9 e is titrated to optimize clinical response and minimize adverse effec 10 with standard chemotherapy because response rates to agents, even of k				
1	RESPONSE RATES	Complete (5), patients (2)	15	0,02
1 t programs yield overall objective response rates of 65%-90% and compl 2 onse rates of 65%-90% and complete response rates of 45%-75%. Because 3 in current programs yield overall response rates of 70%-85% and compl 4 onse rates of 70%-85% and complete response rates of 20%-30% in extens 5 ctivity in stage IV patients, with response rates in the range of 21%- 6 therapy regimens produce objective response rates (including a few com 7 with standard chemotherapy because response rates to agents, even of k 8 nations have shown relatively high response rates, significant 1 year 9 clitaxel plus carboplatin regimen, response rates have been in the ran 10 have been reported to have higher response rates in small phase II tr				

14.4.3.2.1.4. PRODUCE

En lo que se refiere a PRODUCE, en nuestro corpus, su prosodia semántica es bastante abierta. El objeto directo de PRODUCE apunta hacia resultados positivos (*survival, responses, cure*), hacia resultados negativos, principalmente efectos secundarios y síntomas

dolorosos (*pain, pruritus*), o simplemente hacia las sustancias producidas por nuestro organismo (*hormone substances, melatonin*). Este significado neutro se actualiza principalmente mediante las formas léxicas *produced, production* y *products*. La forma *product(s)* también denomina productos cosméticos y de limpieza (*skin care products, antistatic products*).

TABLA 151: Radio colocacional de *produc**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional	Frec	%																		
	Colocaciones / unidades fraseológicas																				
PRODUCE Produce (38), produced(12), products(12), produces(11), production (10), producing (8), producers(4), product(1)	Survival (8), tumor (8), chemotherapy (8), agents (8), symptoms (6), use (6), substances (6), similar (6), regimens (6), pruritus (6), bulk (6), drugs (5), reactions (5), doses (5), responses (5), fatigue (5), pain (5), effects (5), blood (5)	96	0,13																		
PRODUCE Produced, produces, producing, producers	Survival (8), agents (8), chemotherapy (7), tumors (7), regimens (6), bulk (6), similar (6), producers (6), substances (5), pruritus (5), use (5), pain (5), doses (5), fatigue (5), drugs (5), effects (5), effect (4), outcomes (4), symptoms (4), toxicity (3), cisplatin (3), serum (3), cells (3), transient (3)	73	0,10																		
	<table border="0"> <tr> <td>Bulk producers</td> <td>6</td> <td>following regimens produce similar</td> </tr> <tr> <td>produce symptoms</td> <td>2</td> <td>survival outcomes</td> </tr> <tr> <td>produce fatigue</td> <td>2</td> <td>actch-producing tumors</td> </tr> <tr> <td>produce pruritus</td> <td>2</td> <td>produced modestly increased</td> </tr> <tr> <td>anxiety-producing drugs</td> <td>2</td> <td>disease-free survival</td> </tr> <tr> <td>chemotherapy produces</td> <td>2</td> <td>cyclophosphamide produced</td> </tr> <tr> <td>produce burning pain</td> <td>2</td> <td></td> </tr> </table>			Bulk producers	6	following regimens produce similar	produce symptoms	2	survival outcomes	produce fatigue	2	actch-producing tumors	produce pruritus	2	produced modestly increased	anxiety-producing drugs	2	disease-free survival	chemotherapy produces	2	cyclophosphamide produced
Bulk producers	6	following regimens produce similar																			
produce symptoms	2	survival outcomes																			
produce fatigue	2	actch-producing tumors																			
produce pruritus	2	produced modestly increased																			
anxiety-producing drugs	2	disease-free survival																			
chemotherapy produces	2	cyclophosphamide produced																			
produce burning pain	2																				
1 resection or radiotherapy, rarely produce long-term survival [1]. Wit 2 chemotherapy has been reported to produce short-term improvement in d 3 for this disease; radiotherapy can produce cure in a small minority an 4 lyses have shown that chemotherapy produces modest benefits in short-t 5 responses): The following regimens produce similar survival outcomes: 6 east moderate toxicity in order to produce the best results in extensi 7 e is not always effective and does produce gastric side effects [17]. 8 otherapeutic agents that typically produce mucosal toxicity are given 9 us malignant diseases are known to produce pruritus. Hodgkin's disease 10 asily removed. A thick layer could produce an undesired bolus effect w 11 their food intake. Tumors can also produce hormone substances, similar 12 Melatonin, a hormone produced by the pineal gland during																					
PRODUCT* Products, production, product	Blood (5), skin (3), reactions (3), fever (3), ingredients (3)	23	0,03																		
	Blood products 4 Steroid production 2 Product-associated fever 2																				

1 motherapy. Transfusions (blood products). Antibiotics. Gro
 2 syndromes associated with steroid production and by symptoms associat
 3 c syndromes with increased steroid production; symptoms associat
 4 gs known to alter normal melatonin production (e.g., beta-blockers and
 5 e tumors that diminish the natural production of the hormone [9].Howev
 6 the skin surface, encouraging the production of moisture in the layer
 7 pharmacologically stimulate saliva production from intact responsive s
 8 ge of the ingredients of skin care products is essential, since many i
 9 cause allergic skin reactions. Product selection and recommendatio
 10 as fabric softeners and antistatic products may aggravate pruritus. De

14.4.3.2.1.5. FACTOR

Por último, mencionamos un lema que lleva también implícito la idea de causalidad. Aparece con una gran variedad de lexemas que explicitan la naturaleza de aquello que puede causar o influir sobre algo: *risk factors*, *etiologic factors*, *growth factors*, *respiratory factors*, *biologic factors*, etc.

TABLA 152: Radio colocacional de *factor**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones / unidades fraseológicas			
FACTORS Factor (19)	Fatigue (19), risk (15), cancer (12), include (10), patients (9), psychological (8), associated (8), treatment (8), development (7), contribute (7)		147	0,21
	Risk factors	13	Factors that contribute	5
	Etiologic factors	5	Environmental factors	3
	Growth factor*	4	Psychological factors	3
	Respiratory factors	4	Social factors	3
	Biologic factors	3	Transforming growth factor	3
	Prognostic factors	3	Factor beta	3
	Situational factors	3	Factors associated	3
	Psychologic factors	3		
1	e. The pretreatment prognostic factors which consistently predict			
2	ing on a variety of tumor and host factors. Patients with resectable d			
3	he erbB-2 oncoprotein [1-6]. Other factors that have been identified a			
4	agnosis of cancer are contributory factors in sleep disturbances in th			
5	of environmental and psychological factors. Treatment of the malignanc			
6	th aggressive surgical approaches, factors associated with improved lo			
7	status is an important prognostic factor [3]. Median survival for mal			
8	red, and a variety of psychosocial factors, such needs range from low			
9	e dictated by medical and biologic factors, and also by demographics,			
10	red. Common nonmalignant etiologic factors include drug reactions, xer			

14.4.3.2.2. APARICIÓN DE SÍNTOMAS / EFECTOS SECUNDARIOS / ENFERMEDAD:

DEVELOPMENT, ASSOCIATED, OCCUR, EXPERIENCE, EFFECT,

Entre los lemas de la lengua general más frecuentes, abundan aquellas cuyo significado es el de APARICIÓN DE SÍNTOMAS / EFECTOS SECUNDARIOS / ENFERMEDAD. Entre estos destacamos DEVELOPMENT, ASSOCIATED, y en menor medida, OCCUR y EXPERIENCE, aunque aclaramos que este no es el único sentido que adquieren en el texto.

Comenzamos hablando del lema DEVELOPMENT, que hemos desglosado en dos en función de su sentido. Por una parte, las formas léxicas *development*, *develop(s)* y *developing* siempre tienen una prosodia semántica negativa. Son muy frecuentes en su cotexto los términos *cancer*, *symptoms*, *hypercalcemia*, *lymphedema*, *metastases* y *PTSD*. Destaca sobre todo la colocación *risk of developing*, seguida de un síntoma o del término *central nervous system metastases*¹⁸⁶. Por otra parte, con una prosodia semántica positiva en la mayoría de los casos encontramos la forma léxica *developed*, que se refiere a sistemas y técnicas desarrolladas para el diagnóstico y el tratamiento del cáncer, como se percibe en las líneas de concordancia.

TABLA 153: Radio colocacional de *develop**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%
	Colocaciones /unidades fraseológicas			
DEVELOPMENT Development (36), develop(34), developing(28), developed(18), Develops (8), developmental(6), developments(1)	Patients (15), cancer (11), symptoms (10), hypercalcemia (12), contribute (9), lymphedema (8), metastases (8), PTSD (8)		131	0,18
DEVELOPMENT Develop, developing, develops	1 Contribute to the development 6 Developing hypercalcemia 4 Developing lymphedema 4 Development of sleep apnea 3	Develop hypercalcemia 3 Prevent the development 3	106	0,15

¹⁸⁶ La UF *central nervous system* aparece en 25 ocasiones en el corpus, lo que la convierte en la tercera colocación de tres palabras plenas más frecuente.

1	n associated with a higher risk of developing clinical anxiety or depr		
2	l metastases [3,8,22]. The risk of developing central nervous system m		
3	brain have a 60% actuarial risk of developing central nervous system m		
4	reversals, can be early signs of a developing delirium. (A separate su		
5	and immunotherapy) treatments and developing more effective systemic		
6	n may be at increased risk for the development of delirium and organic		
7	recently has been applied to PTSD development in patients with cancer		
8	r mandibulectomy can result in the development of sleep apnea. All pat		
9	ts treated surgically subsequently develop regional or distant metasta		
10	toms or signs of progressive tumor develop. Treatment options:		
1	RISK OF DEVELOPING	Nervous (4), system (4), central (4), hypercalcemia (3), bowel obstruction (2)	12
		Risk of developing central nervous system metastases 4	
1	ide the brain have a 60% actuarial risk of developing central nervous		
2	r cranial metastases [3,8,22]. The risk of developing central nervous		
3	ide the brain have a 60% actuarial risk of developing central nervous		
4	ir cranial metastases [12-14]. The risk of developing central nervous		
5	has been associated with a higher risk of developing clinical anxiety		
6	edical conditions may increase the risk of developing lymphedema; ther		
7	ing hormonal therapy. An increased risk of developing hypercalcemia ha		
8	---- Prevention Individuals at risk of developing hypercalcemia ma		
9	ho have colostomies are at special risk of developing constipation. If		
10	in adequate hydration to avoid the risk of developing a bowel obstruct		
	DEVELOPED	Patients (4), specifically (3), cancer (2),	18 0,03
1	e, the detailed TNM staging system developed for lung cancer by the Am		
2	carcinoma. A simple 2-stage system developed by the Veterans Administr		
3	w-dose biweekly regimens have been developed for these patients, [19-23		
4	ly, and staff. Protocols should be developed for these transitions wit		
5	erized as an anxiety disorder that developed in response to a severe t		
6	re are few measures that have been developed specifically for use with		
7	emi-structured interview technique developed specifically for use with		
8	le no specific therapies have been developed for PTSD in the cancer se		
9	, oral assessment scales have been developed to grade the level of sto		
10	complexity. Once mucositis has developed, its severity and the pat		

ASSOCIATED es el lema que tiene en su radio colocacional un número mayor de síntomas y efectos secundarios. Cuando está acompañado por el nombre de una terapia, su sentido es similar al de *cause* o *produce*, es decir, indica los efectos secundarios producidos por la misma. En las líneas de concordancia hemos encontrado tres expresiones en las que su prosodia semántica sea positiva:

(161)

1 platin-based chemotherapy has been associated with short-term palliation o
 2 motherapy and chest irradiation is associated with improved survival compa
 3 ssive surgical approaches, factors associated with improved long-term surv
 4 otherapy and chest radiotherapy is associated with a better prognosis in t

TABLA 154: Radio colocacional de *associated*

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional	Frec	%																	
	Colocaciones /unidades fraseológicas																			
ASSOCIATED Associated (123)	Symptoms (15), fever* (15), pruritus (13), hypercalcemia (12), fatigue (11), factors (9), disease (9), chemotherapy (8), cancer (7), nausea (7), radiation (7), vomiting (6), increased (6), sleep (6),	130	0,18																	
	<table border="0"> <tr> <td style="padding-right: 20px;">Pruritus associated with</td> <td style="text-align: right;">9</td> <td style="padding-left: 20px;">Associated with ~</td> <td></td> </tr> <tr> <td>Symptoms associated</td> <td style="text-align: right;">7</td> <td>Hypercalcemia</td> <td style="text-align: right;">4</td> </tr> <tr> <td>Tumor-associated</td> <td style="text-align: right;">7</td> <td>Chemotherapy</td> <td style="text-align: right;">3</td> </tr> <tr> <td>Mecanisms of cancer-associated hypercalcemia</td> <td style="text-align: right;">3</td> <td>Fatigue</td> <td style="text-align: right;">3</td> </tr> <tr> <td></td> <td></td> <td>Radiation</td> <td style="text-align: right;">3</td> </tr> </table>			Pruritus associated with	9	Associated with ~		Symptoms associated	7	Hypercalcemia	4	Tumor-associated	7	Chemotherapy	3	Mecanisms of cancer-associated hypercalcemia	3	Fatigue	3	
Pruritus associated with	9	Associated with ~																		
Symptoms associated	7	Hypercalcemia	4																	
Tumor-associated	7	Chemotherapy	3																	
Mecanisms of cancer-associated hypercalcemia	3	Fatigue	3																	
		Radiation	3																	
1	ge. Combined modality treatment is associated with increased morbidity																			
2	ypnotic agent for sleep disruption associated with elevated temperatur																			
3	platin-based chemotherapy has been associated with short-term palliati																			
4	Invades the visceral pleura Associated with atelectasis or obst																			
5	tion, persistent insomnia has been associated with a higher risk of de																			
6	entions that diminish the distress associated with early insomnia and																			
7	ry of the patient. The agents most associated with hypersensitivities																			
8	eneral population [1 and are often associated with situational stress,																			
9	steroid production and by symptoms associated with tumor invasion, suc																			
10	ssive surgical approaches, factors associated with improved long-term																			
PRURITUS ASSOCIATED WITH + DISEASE		9	0,01																	
1	Kaposi's sarcoma. The incidence of pruritus associated with AIDS-relat																			
2	or weight loss ("B" symptoms) [3].Pruritus associated with Hodgkin's																			
3	ces of the upper extremities [1,3].Pruritus associated with malignant																			
4]. Hypothesized pathogenesis of pruritus associated with underlying																			
5	from white blood cells may trigger pruritus associated with lymphomas																			
6	** INTERVENTIONS Management of pruritus associated with neoplastic																			
7	een an acceptable intervention for pruritus associated with dry desqua																			
8	ents may be effective in relieving pruritus associated with renal or h																			
9	n used with some effectiveness for pruritus associated with Hodgkin's																			

La aparición de síntomas también se expresa con el lema EXPERIENCE, como indican las líneas de concordancia 1–4 de la tabla 155. No obstante, este apunta más bien hacia el proceso psicológico experimentado durante la enfermedad (líneas 5-7) y hacia la grata experiencia de la supervivencia (líneas 9 y 10).

TABLA 155: Radio colocacional de *experient**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional		Frec	%		
	Colocaciones / unidades fraseológicas					
EXPERIENCE Experience (73), experienced(25), experiences(17), experiencing(11)	Patients (31), cancer (28), grief (13), loss (11), nausea (11), vomiting (10), fatigue (10), symptoms (8), persons (8), patient (7), pain (6), anxiety (6), traumatic (6)		126	0,18		
	Cancer experience	9			Grief experienced	4
	Experience nausea	7			Patient experiences	3
	% Persons will experience nausea				Patients who experience	3
	and vomiting	5			Patients will experience	3
The cancer experience	5					
1	to moderate sleep disturbances may experience irritability and inability					
2	attment of malignancy are likely to experience pruritus. This same population					
3	ated fever: Cancer patients may experience fever as a response to contact					
4	tractility. People with cancer who experience dyspnea, another common					
5	ized. The most vulnerable of those experiencing cancer, such as those					
6	ble those seen in persons who have experienced events of extreme stress					
7	managed, patients and families can experience psychological growth and					
8	and autonomy constitute traumatic experiences for many cancer patients					
9	however, only the rare patient will experience long-term survival follow					
10	ative. Although some patients will experience long-term survival with					

En lo que se refiere al lema OCCUR, su prosodia semántica es en la mayoría de los casos negativa. De hecho, este verbo intransitivo lleva casi siempre como sujeto unidades léxicas o terminológicas que designan efectos secundarios (*vomiting*, *nausea*), síntomas (*hypercalcemia*) o incluso un fatal desenlace de la enfermedad (*death*). Asimismo, aunque más marginalmente, *occur* es el verbo intransitivo más utilizado al describir lo que sucede en el organismo bajo unas determinadas condiciones, algo que no está marcado axiológicamente y que incluso podría ser positivo.

- (162) For increasing numbers of individuals, challenging transitions also occur from "illness" to "health".
- (163) Changes in biopsychosocial status can occur rapidly [...].
- (164) The stages of sleep occur in a repeated pattern or cycle of NREM followed by REM [...].

Para ilustrar la prosodia semántica negativa de OCCUR de una forma más objetiva, incluimos una tabla con las palabras plenas que aparecen más de dos veces en su radio

Presentamos ahora la tabla que resume el comportamiento semántico de *occur**, donde se percibe de nuevo una tendencia a coocurrir con lexemas que designan experiencias negativas.

TABLA 157: Radio colocacional de *occur**

PALABRA CLAVE y formas léxicas	Lexemas en su radio colocacional	Frec	%
	Colocaciones /unidades fraseológicas		
OCCUR Occurs(26), occurrence(9), occurred(7), occurring(6),	Cancer (10), patients (9), therapy (8), vomiting (7), hypercalcemia (6), nausea (6), acute (6), death (6)	114	0,16
1	period during which most relapses occur. However, even these patients		
2	d. Like small cell lung cancer, it occurs primarily in cigarette smoke		
3	oms of burning and intense itching occurring on a localized skin area,		
4	tive symptoms has been reported to occur more frequently than objectiv		
5	uals, challenging transitions also occur from "illness" to "health." T		
6	daytime sedation and delirium can occur, especially in the elderly. T		
7	routine because many changes will occur over time in the patient's an		
8	n illness. These crisis points can occur at times of disease progressi		
9	elow). Other critical points could occur at the time of a psychosocial		
10	y alone [1,2]. Most of the benefit occurred in patients less than 65 y		

14.5. Observaciones en torno a las *palabras forma* más frecuentes del corpus

En Kennedy (1998: 97) se afirma que, si comparamos diversos corpóra, se hallan diferencias en el orden de las 50 palabras más frecuentes, las cuales reflejan la naturaleza de los mismos y son siempre *palabras forma*. Por eso, vamos estudiar las listas de frecuencia correspondientes a especialistas (*QD*) y a pacientes (*QP*) de los resúmenes sobre tratamiento y cuidados médicos del *PDQ*, tal y como se presentaban antes de la eliminación de palabras forma y de la lematización. Estudiaremos las 100 palabras más frecuentes para cada corpus.

Dada la procedencia de los textos objeto de análisis, hemos comparado dichas listas de frecuencia con las del *Brown Corpus*, un corpus de referencia de inglés americano que cuenta con un millón de palabras repartidas en 500 textos de aproximadamente 2000

palabras cada uno¹⁸⁷. También las hemos comparado con la sección de este corpus que recoge textos académicos escritos (*Learned*). Las ocho palabras más frecuentes de esta sección del *Brown*, que recibe la etiqueta *J*, son las mismas que aparecen en el *Brown corpus* completo (de ahora en adelante, *BR*). En el Apéndice VIII incluimos las 75 palabras más frecuentes de cada uno.

Un punto que llama la atención es que la *palabra forma* más frecuente de la parte para especialistas, la preposición *of*, ocupa el segundo lugar en el resto de las listas, incluso en la sección de textos académicos de *BR*. Podemos explicar esto teniendo en cuenta que en el lenguaje científico abundan UF de gran complejidad y que muchas de ellas se construyen con sintagmas preposicionales introducidos principalmente por *of*, *to*, *in* y *with*, que ocupan respectivamente la primera, cuarta, quinta y séptima posición. Estas preposiciones son más frecuentes en el corpus para especialistas que en el corpus para pacientes y que en el corpus *BR*, donde ocupan la segunda (*of*), cuarta (*to*), sexta (*in*) y décimo tercera¹⁸⁸ (*with*) posición.

La preposición *of* aparece en tercer lugar en el corpus destinado a pacientes, y no en segundo lugar como ocurre en el *BR* y en la sección *J*. Así se consiguen expresar conceptos sin sintagmas nominales demasiado complejos. El segundo lugar del corpus para pacientes lo ocupa la conjunción copulativa *and*, que ocupa el tercer lugar en el resto de córpora.

Otro rasgo destacado es la gran frecuencia de la conjunción disyuntiva *or* tanto en *QD* (sexta posición) como en *QP* (quinta posición) en comparación con *BR* (27) y *J* (31). Si reflexionamos sobre el hecho de que estos textos proponen distintas opciones de tratamiento y diagnóstico, distintas partes del cuerpo donde se pueden producir metástasis en un determinado estadio y distintas USE para designar un mismo concepto, no nos debe extrañar la frecuencia de esta conjunción. De esto hemos hallado numerosos ejemplos:

- (165) This group is treated with radiotherapy or with radiotherapy in combination with other therapy modalities. (QDT2)
- (166) Non-small cell lung cancer (NSCLC) is a heterogeneous aggregate of at least three distinct histologies of lung cancer including epidermoid or squamous carcinoma, adenocarcinoma, and large cell carcinoma. (QDT2)

¹⁸⁷ Cada texto acaba en el primer final de oración después de la palabra que ocupa el puesto 2000.

¹⁸⁸ En la parte académica del *Brown*, la preposición *with* ocupa el décimo quinto puesto.

- (167) Extensive stage: Cancer has spread outside of the lung where it began to other tissues in the chest or to other parts of the body. (QPT1)

También se percibe una clara diferencia en cuanto a los verbos modales más frecuentes tal y como indica la tabla 158, donde estos aparecen ordenados de mayor a menor frecuencia y acompañados de un número entre paréntesis que indica su orden en el corpus:

TABLA 158: Posición de los verbos modales de *PDQ* con respecto al corpus de referencia *Brown*.

Brown	%	Brown J	%	QD	%	QP	%
Would (38)	0,23	Would (39)	0,22	May (14)	0,77	May (8)	1,55
Will (47)	0,19	Can (60)	0,14	Can (18)	0,51	Can (17)	0,81
Can (61)	0,15	Will (63)	0,14	Should (23)	0,33	Should (34)	0,36
Could (67)	0,13	May (69)	0,12	Will (84)	0,13	Will (54)	0,26
May (70)	0,12	Could (72)	0,12				
Must (92)	0,08	Must (87)	0,10				
TOTAL	0,9		0,84		1,74		2,98

De esta tabla se infiere que entre las 100 palabras más frecuentes de nuestro corpus hay una concentración bastante elevada de modales, 1,7 por ciento en la parte para especialistas y un 2,98 en la parte para pacientes. En efecto, cuando se habla de cáncer predominan las expresiones de cautela, especialmente cuando los receptores del mensaje son pacientes. Asimismo, hay una cierta preferencia por modales que no son los más utilizados en *BR*, en concreto, *may* y *should*, que ni siquiera aparece entre las 100 palabras más frecuentes. En *QDT*, el modal *may* ocupa el décimo cuarto lugar, *can* el décimo octavo lugar y *should* el vigésimo tercero. Estos modales ocupan respectivamente la posición octava, décimo séptima y trigésimo cuarta en *QPT*.

Esta mayor presencia de modales con respecto al corpus de referencia es representativo de otros textos científicos, especialmente, los artículos experimentales. En estos, hay una serie de lexemas que la comunidad de discurso utiliza convencionalmente

para expresar cautela, modestia o las limitaciones del conocimiento científico¹⁸⁹ (Hyland 1998: 149). Este autor demuestra esta hipótesis al comparar la frecuencia de estos lexemas en distintos córpora:

- Un corpus de artículos experimentales elaborado por él (75 000 palabras)
- El corpus científico-técnico JDEST de la Universidad Jiao Dong de Shanghai (435850)
- Un corpus que compila las dos secciones académicas (J) del corpus *Brown* y *LOB*¹⁹⁰ (350 000)

Los resultados de su comparación dejan patente la mayor frecuencia de verbos modales en textos científicos, como se aprecia en la siguiente tabla:

TABLA 159: Lexemas para articular los matices de certeza más usuales: frecuencia por cada 10000 palabras (Hyland 1998: 149)

Forma	Artículos experimentales	JDEST	Brown / LOB (J)
Indicate	10,8	3,2	4,3
Would (not)	10,4	16,8	16,0
May (not)	9,2	8,0	4,4
Suggest	9,1	3,7	3,9
Could	6,4	3,2	0,4
About	4,0	-	-
Appear	4,0	2,7	3,7
Might (not)	3,6	2,4	4,0
Likely	2,8	2,5	2,7
Propose	2,8	1,6	0,9
Probably	2,7	1,6	2,8
Apparently	2,7	-	2,8
Should	2,4	0,8	0,4
Seem	2,3	4,0	7,7
Possible	2,3	1,5	1,3

¹⁸⁹ En inglés, la palabra *hedging* designa este concepto.

¹⁹⁰ Corpus de inglés británico de un millón de palabras diseñado según exactamente las mismas premisas del corpus *Brown*.

Por último, vamos a comentar la distribución de pronombres personales y posesivos porque también arroja resultados interesantes. Entre las 100 palabras más frecuentes del corpus *Brown* y de su sección académica, se encuentran *he, it*¹⁹¹, *his, I, they, you, her, she, their, we, him, its, them, me* y *your*. En *QD*, por el contrario, sólo encontramos *it* y *their*, para dar un mayor distanciamiento y para evitar ambigüedad y referencias erróneas. No ocurre lo mismo en la sección para pacientes, en la que la segunda persona del singular crea cercanía y una cierta complicidad con el lector. Así, en *QP*, el pronombre *you* ocupa el vigésimo octavo lugar, *it* aparece en el trigésimo, seguido a distancia por *your* (48), *they* (63) y *their* (75). No obstante, en ambas secciones de nuestro corpus hemos visto cómo muchas referencias a personas se hacen mediante el lema PATIENTS cuyo porcentaje en *QD* es de 1,20 y 1,41 en *QP*.

¹⁹¹ Aunque no lo hemos desglosado, suponemos que uno de los usos más frecuentes es el de *it* impersonal.

15. Cambios en la cohesión derivados de la traducción

Como ya se adelantó en 4.11., al traducir textos del inglés al español, se producen cambios con respecto a los recursos cohesivos utilizados y al grado de explicitación de la información que se pretende transmitir. Según Fernández Polo (1995, 1999) estas variaciones se deben a que las lenguas inglesa y española se estructuran de forma distinta desde el punto de vista retórico. En el capítulo 10, se hicieron evidentes algunas de estas diferencias retóricas, puesto que el texto analizado era una *traducción-adaptación* en forma de noticia de un artículo experimental en inglés.

En este capítulo se incluirán algunos ejemplos extraídos de nuestro corpus comparable y paralelo de textos en español que ilustran hasta qué punto los textos en español son menos explícitos desde el punto cohesivo tanto si estos son textos originales como si son traducciones del inglés, como afirma Fernández Polo.

15.1. La cohesión en textos originales en español

Los textos especializados, que Fernández Polo no incluye en su análisis, contienen un vocabulario consensuado y normalizado y, para evitar la ambigüedad, permiten la repetición léxica. Esta repetición es muy lícita desde el punto de vista retórico en los artículos científicos para especialistas, tal y como se percibe en siguiente *abstract* de la revista sobre oncología *Neoplasia* publicada por la editorial Doyma¹⁹². Con el mismo color se han aislado algunos conceptos importantes del texto con el fin de seguir sus referencias a lo largo del texto. Los conceptos seleccionados son los de TRATAMIENTO (azul), TIPO DE CÁNCER (rojo), PACIENTE (turquesa), TOXICIDAD (fucsia), DOSIS (verde azulado) y VOLUMEN (ocre).

¹⁹² Volumen 15, Número 6, Noviembre-Diciembre 1998

(168) **Efectos secundarios del tratamiento radioterápico en el cáncer de pulmón no microcítico**

R. Escó, J.E. Baquedano, C. Laínez y J. Valencia
 Servicio de Oncología Radioterápica. Hospital Clínico Universitario de Zaragoza.

Fundamentos: El **tratamiento radioterápico** en **pacientes** con **tumores en cuyo volumen esté incluido el pulmón** conlleva un riesgo de efectos adversos, según la **dosis administrada** y el **volumen** irradiado.

Pacientes y métodos: Presentamos un estudio retrospectivo de **tratamiento radioterápico** y la **toxicidad** aguda y tardía en 246 **pacientes** con **cáncer no microcítico de pulmón tratados** entre junio de 1978 y diciembre de 1989 en el Servicio de Oncología Radioterápica del Hospital Clínico Universitario de Zaragoza, según criterios del Radiation Therapy Oncology Group (RTOG), a partir de grado 2 de iatrogenia.

Resultados: El **tratamiento** fue radical en 165 **pacientes** (67%). La **dosis** media en **volumen** mediastínico de 44 Gy, y en **volumen** de **tumor** primitivo de 53 Gy. Se diagnosticó neumonitis en 27 **pacientes** (11%) y fibrosis en 33 (13%). Se valora la **toxicidad** en esófago, piel, pericardio y médula.

Conclusiones: En nuestro centro los porcentajes de **toxicidad** pulmonar coinciden con lo reflejado por otros **grupos**. Estos estudios proporcionan al radioterapeuta la posibilidad de **manejar dosis** umbrales seguras dependiendo de cada **tratamiento**.

Como se aprecia, hay una significativa repetición de las mismas unidades léxicas ("tratamiento", "pacientes", "toxicidad", "volumen", "dosis") para conseguir la univocidad. Y así, por ejemplo, de las 8 unidades léxicas que construyen la red de significado TRATAMIENTO, 6 se construyen sobre la raíz *trata-* y se repite 3 veces la USC "tratamiento radioterápico". De todas formas, se intuye una tendencia hacia la variedad terminológica en las formas "cáncer de pulmón no microcítico" y "cáncer no microcítico de pulmón", que ya se comentó en el capítulo 10 .

El segundo ejemplo se ha tomado de una noticia para especialistas publicada en Diario Médico (www.recoletos.es/dm)¹⁹³, un género con un menor grado de especialización y rigurosidad que el artículo experimental. En consecuencia, habrá una mayor adecuación a la norma retórica del español a la que aludía Montaña-Harmon

¹⁹³ Véase Apéndice II, texto 16

(véase p. 167) por la que se evita la repetición léxica. Se han destacado los conceptos TRATAMIENTO en azul y CANCER en rojo.

(169) **El tratamiento neoadyuvante incrementa la supervivencia en el cáncer de pulmón**

Carmen Fernández. Barcelona.

El Grupo Español de Cáncer de Pulmón, que engloba a 37 centros, tiene previsto estudiar los efectos de la **terapia neoadyuvante (preoperatoria)** con los nuevos **fármacos quimioterápicos** en la mejora de la supervivencia de los pacientes afectados por **tumores pulmonares** en estadios iniciales, según ha anunciado Rafael Rosell, presidente de la asociación y jefe del Servicio de Oncología Médica del Hospital Universitario Germans Trias, de Badalona, en Barcelona, que coordinará esta investigación.

"Creemos que puede ser mucho más efectivo que el **tratamiento postquirúrgico**", ha indicado Rosell, co-director científico, junto con Paul A. Bunn, de la División de Oncología Médica del Centro de Ciencias de la Salud de la Universidad de Colorado, en Denver, Estados Unidos, del Foro interactivo pan-europeo sobre nuevas perspectivas en el **manejo del cáncer de pulmón** celebrado en Sitges, Barcelona, y organizado por el laboratorio Bristol-Myers Squibb, que financiará parte del estudio que se propone realizar el grupo español de expertos.

En concreto se quiere reunir una muestra de 600 pacientes en dos años y seguirlos durante tres, con el objetivo de llegar a demostrar una reducción de la mortalidad del 20 por ciento a lo largo de ese período.

Rosell, junto con un equipo interdisciplinar de su propio centro y de los hospitales La Fe y General de Valencia, publicó en 1994 en la revista *The New England Journal of Medicine* que la **quimioterapia preoperatoria** con **fármacos** clásicos incrementaba la supervivencia media de los pacientes con **cáncer de pulmón** localmente avanzado, y el pasado mes de enero, a través de la revista *Annals of Oncology*, dieron a conocer un estudio que concluye que un 28 por ciento de los casos de **tumores** diagnosticados en estadios iniciales presentan micrometástasis - ADN **tumoral** circulante en la sangre- ya en el momento mismo de la detección de la **enfermedad**.

En este caso, el concepto TRATAMIENTO se actualiza mediante más formas que en el ejemplo anterior: "tratamiento neoadyuvante", "terapia neoadyuvante (preoperatoria)", "fármacos quimioterápicos", "tratamiento postquirúrgico", "manejo", "quimioterapia preoperatoria" y "fármacos". Por el carácter menos especializado de este tipo textual, el CÁNCER DE PULMÓN se designa con unidades de significación genéricas: "cáncer de pulmón", "tumor pulmonar", "tumor", "enfermedad".

15.2. Diferencias cohesivas entre textos originales en inglés y su traducción al español

Ahora es preciso comparar textos sobre oncología en inglés con sus traducciones al español para ver si las observaciones de Fernández Polo son válidas en el subdominio que nos ocupa. Se analizarán fragmentos de las traducciones al español que ofrece el *National Cancer Institute* a los visitantes del servicio *Physician Data Query (PDQ)*¹⁹⁴.

Se han alineado los textos en inglés con traducciones al español con la ayuda de *Wordsmith Tools* para que aparezcan en formato *bitexto*¹⁹⁵ (Apéndice II, texto 15). Las unidades léxicas objeto de análisis del texto original están color **fucsia** y la traducción que se propone en **verde azulado**. Se ha señalado la repetición variada léxica (véase p.192) mediante subrayado.

En el ejemplo (170), el traductor traslada el segmento repetido *tissue specimens* de dos formas diferentes, *muestras de tejido* y *especímenes de tejido* para que no se reiteren dos USE idénticas en un cotexto tan cercano. No obstante, la segunda forma resulta un tanto extraña y nada rigurosa desde el punto de vista médico. En (171), también se ofrecen dos alternativas de traducción para la expresión *surgical approaches*: “estrategias quirúrgicas” y “enfoques quirúrgicos”.

(170)

<!--L1, S 18-->It can be especially difficult to differentiate mesothelioma from carcinoma on small tissue **specimens**.

<!--L2, S 18-->Puede ser especialmente difícil diferenciar el mesotelioma de un carcinoma en **muestras** pequeñas de tejido.

<!--L1, S 19-->Thoracoscopy can be valuable in obtaining adequate tissue **specimens** for diagnostic purposes [1].

<!--L2, S 19-->La toracoscopia puede ser valiosa para la obtención de **especímenes** adecuados de tejido con fines de diagnóstico [1].

¹⁹⁴ Disponible en http://cancermet.nci.nih.gov/sp_menu.htm.

¹⁹⁵ La leyenda del texto alineado es la siguiente: S= Sentence; L1= Texto original; L2= Texto traducido.

(171)

<!--L1, S 6-->For patients treated with aggressive **surgical approaches**, factors associated with improved long-term survival include epithelial histology, negative lymph nodes, and negative surgical margins [3,4].

<!--L2, S 6-->Para pacientes tratados con **estrategias quirúrgicas** agresivas, los factores relacionados con una supervivencia mejor a largo plazo incluyen histología epitelial, ganglios linfáticos negativos y márgenes quirúrgicos negativos [3,4].

<!--L1, S 7-->For those patients treated with aggressive **surgical approaches**, nodal status is an important prognostic factor [3].

<!--L2, S 7-->Para aquellos pacientes tratados con **enfoces quirúrgicos** agresivos, el estado ganglionar es un factor pronóstico importante [3].

A continuación, se confirma que esta tendencia a la supresión de formas idénticas se mantiene incluso cuando la separación entre las oraciones es mayor. Buena prueba de ello lo constituye la traducción de tres ocurrencias del verbo *to show* con tres lemas diferentes: MOSTRAR, BRINDAR Y DEMOSTRAR. Por otra parte, en (173), *difficult to assess* aparece traducido como “difícil de evaluar” y “difícil de detectar”.

(172)

<!--L1, S 22-->Histologic appearance appears to be of prognostic value, with most clinical studies **showing** that epithelial mesotheliomas have a better prognosis than fibrous or sarcomatous mesotheliomas [2-4].

<!--L2, S 22-->La apariencia histológica parece tener valor pronóstico, al **mostrar** la mayoría de los estudios clínicos que los mesoteliomas epiteliales tienen un mejor pronóstico que los mesoteliomas fibrosos o sarcomatosos [2-4].

<!--L1, S 27-->A proposed staging system based upon thoracic surgery principles and clinical data is **shown** below [2].

<!--L2, S 27-->Un sistema de clasificación propuesto basado en principios quirúrgicos torácicos y datos clínicos **se brinda** a continuación [2].

<!--L1, S 45-->The use of radiation therapy in pleural mesothelioma has been **shown** to alleviate pain in the majority of patients treated.

<!--L2, S 45-->El uso de radioterapia en mesotelioma pleural **ha demostrado** alivio en el dolor en la mayoría de los pacientes tratados.

(173)

<!--L1, S 3-->Prognosis in this disease is difficult to **assess** consistently because there is great variability in the time before diagnosis and the rate of disease progression.

<!--L2, S 3-->El pronóstico de esta enfermedad es difícil de **evaluar** de manera consistente debido a la alta variabilidad de tiempo que existe antes del diagnóstico y la tasa de progresión de la enfermedad.

<!--L1, S 9-->In some instances the tumor grows through the diaphragm making site of origin difficult to **assess**.

<!--L2, S 9-->En algunos casos el tumor crece a través del diafragma haciendo así difícil de **detectar** el sitio de origen.

Si las USE repetidas están separadas por un número significativo de oraciones, en la traducción también se produce un cambio de término, aunque a veces se permite la repetición, como ocurre en las oraciones 42 y 51 del ejemplo (174). No obstante, cuando la unidad léxica en cuestión haya aparecido más de dos veces en la secuencia lineal del texto, es decir, ante la tercera repetición del término inglés, el traductor introduce la variación terminológica.

(174)

<!--L1, S 11-->**Effusions**, both pleural and peritoneal, represent major symptomatic problems for at least two thirds of the patients.

<!--L2, S 11-->Los **derrames**, tanto pleurales como peritoneales, representan problemas sintomáticos importantes en al menos dos tercios de los pacientes.

<!--L1, S 42-->Pleurectomy and decortication can provide palliative relief from symptomatic **effusions**, discomfort caused by tumor burden, and pain caused by invasive tumor.

<!--L2, S 42-->La pleurectomía y la decorticación pueden brindar alivio paliativo contra **efusiones** sintomáticas, malestar causado por el agobio tumoral y dolor causado por invasión tumoral.

<!--L1, S 51-->Recurrent pleural **effusions** may be treated with pleural sclerosing procedures; [...]

<!--L2, S 51-->Las **efusiones** pleurales recidivantes pueden tratarse con procedimientos esclerosantes pleurales; [...]

<!--L1, S 67-->1 Symptomatic treatment to include drainage of **effusions**, chest tube pleurodesis, or thoracoscopic pleurodesis [1].

<!--L2, S 67-->1 Tratamiento sintomático que incluya drenaje de **derrames**, pleurodesis del tubo torácico o pleurodesis toracoscópica [1].

<!--L1, S 74-->Intrapleural or intraperitoneal administration of chemotherapeutic agents (e.g, cisplatin, mitomycin, and cytarabine) has been reported to produce transient reduction in the size of tumor masses and temporary control of **effusions** in small clinical studies [14-16].

<!--L2, S 74-->Se ha informado que la administración intrapleural o intraperitoneal de agentes quimioterapéuticos (p. ej., cisplatino, mitomicina y citarabina) produce reducción transitoria en el tamaño de las masas tumorales y control temporal de **derrames** en estudios clínicos pequeños [14-16].

15.3. Recapitulación

La cohesión léxica contribuye a la textualidad y a la activación conceptual dentro de los textos. De ahí la necesidad de que el traductor tome consciencia de los mecanismos cohesivos presentes en el TO y el TT y de las cadenas léxicas del texto, al mismo tiempo que evalúa si se ha mantenido el valor discursivo de la cohesión.

Se han matizado las explicaciones de Blum-Kulka (1986) al fenómeno de la cohesión en textos traducidos para aplicarlas a la traducción inglés-español. Para este aspecto, han sido reveladoras las investigaciones de Fernández Polo (1995, 1999), que cuestionan la hipótesis de la explicitación cuando se traduce al español. En esta lengua hay una tendencia retórica a evitar que se repita una unidad léxica o terminológica en la secuencia textual más de dos veces, con lo cual, la relación de cohesión que suele ser explícita en inglés, pasa a oscurecerse en el TT.

A los motivos estilísticos ya mencionados, habría que añadir otros factores que también explican los cambios cohesivos detectables al comparar un texto inglés y su traducción al español. En primer lugar, la mediación lingüística lleva aparejada la eliminación de la repetición y la redundancia (Baker 1993: 243-4; Toury 1991a: 188), por lo que normalmente las traducciones al español presentan una mayor variedad terminológica (véase capítulo 10).

En un plano mucho menos idealizado, encontramos a veces traducciones que se han hecho con mucho apremio y probablemente poco rigor terminológico. Bajo estas condiciones, el traductor utiliza poco el término del que duda y en su lugar, se sirve de un sinónimo o un término superordinado del que sí tiene certeza, con lo que la variedad léxica es mayor en el original. Se puede detectar este fenómeno en la traducción española de *The Merck Manual of Diagnosis and Therapy* (1997)¹⁹⁶. En la versión española, se propone “carcinoma broncogénico” como traducción de *broncogenic carcinoma*. Si bien hemos constatado en los diccionarios *Dorland* y *DTCM* y mediante una búsqueda en Internet que esta forma se utiliza, los autores de artículos experimentales se decantan más por la USE “carcinoma broncogénico”¹⁹⁷. La

¹⁹⁶ Documento de Internet disponible en <http://www.merck.com/pubs/mmanual>.

¹⁹⁷ De hecho, la Sociedad Española de Neumología y Cirugía Torácica cuenta con una división denominada Grupo Cooperativo de Carcinoma Broncogénico (GCCB-S). Las revistas médicas de la editorial DOYMA, entre ellas, *Archivos de Bronconeumología*, también incluyen esta USE.

inseguridad acerca de si el término es el aceptado por los expertos junto con la mayor variedad terminológica permitida en español lleva al traductor a sustituir en muchas ocasiones “carcinoma broncígeno” por “cáncer de pulmón”, “carcinoma pulmonar”, “cáncer pulmonar” y “tumor”, cuando el TO contiene la expresión *broncogenic carcinoma*.

Por último, la gramática puede explicar también el hecho de que la traducción al español contenga un número mayor de formas distintas que el texto original¹⁹⁸. En efecto, el español es una lengua más flexiva que el inglés y cuenta con una mayor variedad de paradigmas verbales, con desinencias de género para los determinantes y para los adjetivos, con usos del pronombre SE que pueden sustituir a la pasiva con el auxiliar SER y con pronombres que pueden aparecer en forma enclítica.

De ahí inferimos que algunos datos estadísticos acerca de los textos como la *type-token ratio*, es decir, el cociente entre el número de formas distintas y el total de ocurrencias léxicas en el texto, no se puedan interpretar simplemente como un indicador de la variedad léxica y terminológica.

¹⁹⁸ En concreto, el original de donde se han sacado los ejemplos 170-174 cuenta con 420 formas léxicas distintas frente a las 469 de la traducción al español.

16. Conclusiones

El propósito de esta tesis ha sido el de investigar la relación entre dos elementos generadores de textualidad, el tipo textual y la cohesión léxica, para indagar los efectos que esta relación trae consigo a la traducción biomédica inglés-español.

Las tipologías textuales han ocupado una posición central en los estudios de traducción, donde el término *tipo textual* ha servido de cajón de sastre bajo el que agrupar instancias de variación textual según el tema, el tono, el medio, la función retórica y la función en el contexto cultural. Es decir, muchas veces ha sido utilizado como sinónimo de *registro* y *género*, en parte porque están muy relacionados. No se ha adoptado una aproximación interdisciplinar en la que confluyan los aspectos cognitivos, lingüísticos y sociales implicados en la producción y recepción de textos. Además, se ha prestado atención a dos facetas diferentes, por una parte, las tipologías textuales aplicadas al texto origen, y por otra, las tipologías de textos traducidos. Si a esta falta de consenso sobre el objeto de estudio, se añade el hecho de que los hablantes de una lengua están continuamente produciendo nuevos tipos textuales y la gran variedad de tipos de traducciones según el escopo, parece que la elaboración de un repertorio cerrado de tipos textuales es una tarea poco fructífera.

La imprecisión del término *tipo textual* ha hecho que revisemos un concepto con una mayor tradición, el concepto de *género*, que surge al constatarse la capacidad que tienen los hablantes de una lengua de reconocer una serie de características y funciones que están prototípicamente presentes en algunos textos y que están asociadas a situaciones sociales habituales en una determinada cultura.

Un determinado género es principalmente el resultado de dos elementos. En primer lugar, es el resultado de una situación comunicativa convencional que impone una organización retórica que está formada por *movimientos* o segmentos textuales (Swales 1990). Esta organización retórica responde a la función social que va a cumplir el género y ha sido consensuada tácitamente por los miembros de una comunidad de discurso o de una comunidad profesional, es decir, tiene un carácter social y convencional. En segundo lugar, los géneros son el resultado de una conjunción de

elementos lingüísticos que tejen una organización flexible en la que entran en juego la opcionalidad y la probabilidad.

En cualquier caso, existen regularidades en la organización de los géneros de carácter cognitivo que reflejan las estrategias que los miembros de una comunidad de discurso o de una comunidad profesional utilizan para comunicarse de un modo efectivo y lograr unas funciones específicas (Bhatia 1993: 19-20). Estas regularidades permiten la identificación de elementos discursivos, es decir, "divisiones de un texto en cuanto al modo en que sus partes se articulan para proporcionar diversas categorías de información" (Paltridge 1997: 90). La base lingüística de las cadenas léxicas cuyos nodos principales coinciden con dichos segmentos textuales nos hace pensar en que los segmentos textuales tienen un fundamento lingüístico, no sólo de contenido o de seguimiento de una convención social.

En este sentido, Hoey (1991) explica que la cohesión léxica es el fundamento de cualquier organización textual y que existe una correlación entre los patrones léxicos de un texto y el género al que pertenece el mismo. Asimismo, Hoey (1995: 90) encuentra en la cohesión una explicación al fenómeno de la intertextualidad y a la existencia de *géneros*, que podemos decir que sirven de pasaporte para poder pasar la frontera que separa a los miembros de un círculo profesional de los que no lo son:

If one bonds with what one has heard and read, and if everyone else in the field is doing the same, then certain expressions are going to become typical of that field and to distinguish the well-read from the novice.

En esta tesis se han revisado brevemente algunos de los géneros médicos más representativos. No se ha pretendido hacer una taxonomía contrastiva exhaustiva de géneros médicos porque, al comparar revistas especializadas médicas en lengua inglesa y española, los géneros que aparecen en el índice de contenidos son prácticamente los mismos, al igual que su estructura global. Esto se debe a que los investigadores y facultativos de la medicina leen los últimos avances de su especialidad en publicaciones internacionales que se escriben en inglés. De este modo van asimilando poco a poco la

organización retórica de estos géneros, que no plantea demasiados problemas en la translación de una lengua a otra, al no diferir demasiado de una lengua a otra.

Se ha indagado también sobre las diferencias entre los géneros típicos del subdominio de la biomedicina, que obedecen a los contenidos, el diseño de la investigación y los procedimientos utilizados, el formato y las secciones en que se dividen.

Las diferencias en cuanto a la cohesión léxica en distintos géneros no se debe tanto al género en sí, sino a la función comunicativa, a las dificultades comunicativas impuestas por el contexto situacional y por el canal de comunicación y a la relación y al conocimiento que comparten los participantes en la comunicación. Según la audiencia a la que va dirigido el texto y el grado de formalidad o tenor de la interacción, según el grado de tecnicidad de un texto (medicina general u oncología) y de normalización de su terminología, y por último, según el canal de comunicación, la cohesión léxica se fundamenta en un grado mayor o menor sobre la repetición de las mismas unidades léxicas y los recursos cohesivos se presentan de un modo más o menos explícito. No parece pues que el género estrictamente determine la cohesión sino que existe una correlación entre los recursos cohesivos utilizados y la interacción tenor-canal de comunicación. De ahí que en este estudio se adopte el término *tipo textual* en lugar del más restrictivo *género*. El tenor y el canal de comunicación son los aspectos que hemos tenido en cuenta al hablar de tipos textuales. Por eso, se ha comparado cómo cambia la cohesión en la comunicación escrita a través de Internet entre profesionales de la salud y entre el especialista y el paciente. Se han considerado los textos destinados a pacientes dentro de la divulgación porque transmiten los conocimientos científicos generados en oncología a un público distinto del originalmente previsto, los oncólogos y demás personal sanitario.

Como diferencias cohesivas entre los textos escritos para el *intercambio especializado* y los escritos para el *intercambio público* entre el médico y el paciente, podemos resaltar que los primeros se construyen en torno a la repetición de una serie limitada de unidades de significación clave, ampliamente consensuadas, y el uso de sinónimos como mecanismo de cohesión es poco habitual. Por el contrario, en los textos

para pacientes, la cohesión mediante sinónimos, términos superordinados, pronombres, etc. es muy frecuente.

De las cinco categorías cohesivas propuestas por Halliday y Hasan (1976), conjunción, referencia, sustitución, elipsis y cohesión léxica, nos hemos centrado en esta última por su frecuencia en los textos, por ser susceptible de estudio mediante métodos computacionales y por sus implicaciones terminológicas y traductológicas. Bajo el nombre de cohesión léxica se incluyen una variedad de relaciones semánticas que pueden existir entre lexemas y que se fundamentan en la *reiteración* y la *colocación*. Al revisar las distintas aproximaciones al fenómeno cohesivo, hemos puesto el énfasis principalmente en tres: Hoey (1991, 1995), Morris y Hirst (1991) y la *Teoría Sentido Texto* (Mel'čuk *et al.* 1984).

Hoey propone un método sistemático para estudiar el modo en que se relacionan unas oraciones con otras mediante recursos léxicos generadores de *redes de vínculos* que tejen los textos. También resalta la importante contribución de la repetición léxica a la cohesión.

Morris y Hirst sostienen que el contenido de un texto se crea a partir de *cadena léxica* que tratan sobre el mismo tema, y que están formadas por lexemas semánticamente relacionados en un texto. A diferencia de Hoey, su perspectiva se encuadra dentro de la lingüística computacional. En consecuencia, Morris y Hirst otorgan una menor importancia a la cohesión como fenómeno lingüístico interoracional y conceden la primacía al contenido semántico del texto.

Estas investigaciones han sido el soporte de dos tendencias de las que han surgido programas para el reconocimiento automático de la cohesión. La tendencia que se deriva de Hoey, ante la dificultad inherente al estudio de la cohesión, limita la cohesión a la repetición de formas léxicas, mientras que la que se deriva de Morris y Hirst además trata de reconocer la cohesión que se fundamenta en las relaciones conceptuales.

Esta segunda aproximación es la más interesante desde el punto de vista terminológico, ya que permite reconocer no sólo instancias de repetición léxica simple (*cancer – cancers*) y compleja (*treat – treatment*) sino también relaciones semánticas como la sinonimia, hiponimia o cohiponimia. Todos los elementos vinculados de este

modo formarían parte de la misma cadena léxica. El concepto de *cadena léxica* está presente en los estudios que consiguen la identificación automática de la cohesión mediante un ordenador dotado de tesoro (Green 1997 y Ellman 1998).

Por último, aunque la Teoría Sentido Texto trate la cohesión de forma tangencial, permite sistematizar las relaciones semánticas entre las unidades léxicas que crean cohesión mediante funciones léxicas abstractas. A partir de las funciones léxicas de la TST se han propuesto una serie de funciones léxicas cohesivas.

De la adaptación de estas tres fuentes deriva la metodología de esta tesis en la que se ha concebido la cohesión como un fenómeno interoracional e intraoracional que activa un universo léxico de relaciones paradigmáticas y sintagmáticas que se podrían representar en forma de marcos conceptuales. Este universo léxico activa un universo textual, crea textura. Asimismo, la cohesión léxica se fundamenta en la repetición de elementos léxicos que están relacionados no sólo respecto a la forma sino también en el plano sintagmático y paradigmático.

En consecuencia, hemos sugerido una clasificación de los recursos cohesivos que atiende a criterios formales, al grado de explicitación de la relación cohesiva y a las relaciones paradigmáticas y sintagmáticas. Desde el punto de vista formal, se proponen las categorías de repetición exacta, simple, variada léxica y variada sintáctica, que se pueden calificar como explícitas, semi-explícitas e implícitas. Asimismo, estas categorías se combinan con los parámetros de paradigmático y sintagmático y con las funciones léxicas cohesivas, de ahí que aparezcan los siguientes tipos de cohesión: repetición paradigmática simple, repetición paradigmática sintáctica (simple y variada léxica), repetición paradigmática semántica (hiponímica, sinonímica, opositiva, meronímica y asociativa), repetición sintagmática exacta, repetición sintagmática simple, repetición sintagmática variada sintáctica, repetición mixta de unidades fraseológicas y repetición mixta de unidad fraseológica y lexema.

Esta clasificación se ha tenido en cuenta en el estudio de textos muy breves como las noticias de avances científicos, al que hemos denominado *análisis intensivo*. Para textos completos y extensos, se ha utilizado un *análisis extensivo*, basado en la aplicación de recursos informáticos y textos en soporte electrónico. En este último aspecto, nos hemos servido de algunas nociones de la lingüística de corpus en relación a

la explotación de *córpore paralelos y comparables* en *sublenguajes* y de programas de análisis léxico. La referencia a datos estadísticos ha sido muy somera porque, aunque estos contribuyen al estudio empírico del lenguaje, algunas diferencias en cuanto a la relación entre formas léxicas y conceptos activados dependen del tamaño de los textos y de la frecuencia de determinados patrones sintácticos. Entre ellos, se puede citar la nominalización mediante *sintagmas nominales compuestos*, que en textos para especialistas es mucho más habitual que en los textos para pacientes.

Con esta metodología extensiva es posible vislumbrar la faceta formal y semántica de la cohesión, es decir, reconocer no sólo la cohesión que se deriva de la repetición de la misma raíz léxica (repetición léxica exacta y simple), sino también la derivada de relaciones semánticas (repetición variada léxica). Para tal objetivo nos hemos servido de un generador de cadenas léxicas llamado *Hesperus* (Ellman 1998). Este programa fue creado para optimizar la recuperación de información en Internet mediante búsquedas basadas no sólo en la forma de las palabras, sino también en su significado, de ahí que incorpore el tesoro *Roget*. Dado que el *Roget* fue creado por un médico, este tesoro ha sido útil en la detección automática de relaciones conceptuales en textos sobre oncología, a pesar de ser un tesoro de la lengua general.

Aparte de disponer de un tesoro, *Hesperus* proporciona datos relevantes sobre los textos, tales como el número de *palabras plenas*, el tipo de relación cohesiva y el porcentaje de activación de conceptos en cada texto analizado. No obstante, cuenta con limitaciones derivadas de la utilización de un tesoro de la lengua general no ajustado al subdominio de la oncología, que hacen que no se plasmen todas las relaciones cohesivas del texto. De entre estas limitaciones podemos citar que un ordenador:

- a) no es capaz de identificar fenómenos de polisemia, homonimia, ambigüedad y connotación.
- b) no detecta la interacción entre la maquetación del texto por una parte, y el género textual y la estructura informativa del texto, por otra. Por ejemplo, un lector sabe que el índice o las palabras en el título o en negrita apuntan hacia la estructura cohesiva del texto, algo que el ordenador no reconoce puesto que no tiene una base de conocimiento que relacione la ortotipografía con la cohesión.

- c) no percibe el solapamiento entre cadenas léxicas, que se debe a que las palabras no tienen límites claramente delimitados.
- d) no reconoce unidades poliléxicas

Por estas limitaciones ha sido preciso corregir los resultados obtenidos con *Hesperus* en numerosas ocasiones con la ayuda del programa de análisis léxico *Wordsmith Tools*, muy utilizado en lingüística de corpus, y de una hoja de cálculo.

Esta tesis ha servido de banco de pruebas de los esfuerzos que se están haciendo desde la ingeniería lingüística para reconocer automáticamente las relaciones conceptuales. Por tanto, supone una contribución a la gestión de terminología basada en el conocimiento (*knowledge based terminology management*).

En relación a las cadenas léxicas identificadas con esta metodología, se pueden inferir una serie de conclusiones. En primer lugar, existe un paralelismo entre el peso específico de las cadenas léxicas de un texto y las áreas conceptuales más activadas en el mismo. Además, se podría decir que existe un paralelismo entre las cadenas léxicas que elabora el investigador a partir de las palabras plenas del texto y la activación conceptual que se produce en la lectura. El lector recrea un texto coherente a partir de la cohesión de las unidades léxicas y terminológicas del texto.

En segundo lugar, los textos que comparten el mismo tema tienen unas cadenas léxicas parecidas. El tema queda reflejado en la cadena léxica de mayor importancia y normalmente en el título. Así, se ha comprobado que en textos que versan sobre el tratamiento del cáncer de pulmón, encabezados con el título *Treatment Summaries for Health Professionals* y *Treatment Summaries for Patients*, la cadena *TREATMENT* es la más relevante.

Sin embargo, como no están dirigidos a una misma audiencia, son perceptibles diferencias en los porcentajes de representación conceptual, el orden de importancia de las cadenas léxicas y la composición de las cadenas léxicas, a pesar de tratar sobre un tema parecido. Con esto se pone de manifiesto que el destinatario del texto es uno de los parámetros más importantes en la relación entre tipología textual y cohesión. Sintetizamos estas diferencias a continuación.

- a) En cuanto al peso de las cadenas que giran sobre el mismo concepto, hemos constatado que, por ejemplo, en los textos para pacientes, la segunda cadena más importante es la que agrupa conceptos anatómicos y le corresponde un 15,4%, mientras que en los textos para profesionales de la salud ocupa un quinto lugar y sólo un 8,3%. Estos cambios en el orden de las cadenas indican distintas perspectivas, intereses y prismas desde los que se percibe la misma la realidad. Y así, hay diferencias entre el peso conceptual otorgado a la cadena *OUTCOME OF TREATMENT* en textos para especialistas (10,15%) y para pacientes (2,9%). Los médicos tienen que decidir el tratamiento que aplicarán sobre el paciente y esta decisión está influenciada por la supervivencia y la respuesta a largo plazo del organismo al tratamiento. La perspectiva que se quiere dar al paciente es la de un resultado a corto plazo: su poder paliativo, el ataque a células cancerosas y la posibilidad de efectos secundarios. De esta forma el que ha redactado los textos evita crear falsas expectativas o desmoralizar a los pacientes.
- b) Dentro de las cadenas en torno al mismo concepto, el número de subcadenas en que se pueden desglosar es mayor en los textos para especialistas. En estos, los elementos de un sistema conceptual determinado tienen un alto grado de especificidad o granularidad. Si a partir de textos especializados se creara una ontología sobre terapias oncológicas, las representaciones de los conceptos tendrían un alto grado de granularidad y pertenecerían a diversas dimensiones, es decir, apuntarían hacia las diversas perspectivas desde las que se pueden considerar los conceptos. A esta multidimensionalidad contribuyen las UF construidas sobre los sustantivos más frecuentes en las cadenas léxicas, por ejemplo, los relacionados con la terapia del cáncer. En este sistema conceptual se pueden encontrar USE correspondientes a las dimensiones *LOCATION OF APPLICATION*, *AIM*, *TIME* y *THERAPEUTIC AGENT*. En los textos divulgativos también hay una especie de multidimensionalidad, aunque no se activa tanto mediante premodificación y unidades poliléxicas de etimología grecolatina sino mediante una progresión de vocablos que aparecen solos o enmarcados en patrones sintácticos y que están relacionados conceptualmente mediante sinonimia, hiponimia y meronimia principalmente. El redactor de textos divulgativos presenta el mismo concepto desde

distintas perspectivas de percepción y cognición con la esperanza de que, al menos, desde una de ellas, el lector lego llegue a comprender la compleja realidad científica.

- c) En los textos para pacientes, la proporción de unidades léxicas genéricas y de la lengua general es mayor, como cabría esperar, que en los textos para especialistas. Por ejemplo, se prefiere *cancer* a *tumor*; *radiation therapy* a *radiotherapy*. No obstante, en los textos especializados también existe una presencia elevada de lexemas de la lengua general que no es incompatible con la ya mencionada granularidad. Algunos de estos vocablos genéricos restringen su significado en este subdominio dado que el cotexto funciona como restrictor semántico, contribuye al proceso de terminologización de palabras de la lengua general y refleja la prosodia semántica de las unidades léxicas. Relacionado con estos vocablos de la lengua general, se ha encontrado gran parte del vocabulario semi-especializado común a los textos académicos y que hemos denominado *vocabulario organizador del discurso*. Este se ha incluido en una macrocadena denominada *RESEARCH / EXPERIMENTATION*, que se desglosa en subcadenas tales como *STUDY, INFORMATION, COGNITION, COMPARISON, CAUSE, MEASURE*. Lo curioso es que los lexemas de la lengua general y el vocabulario organizador del discurso aparecen en los dos tipos textuales analizados. Esto va en contra de la infundada creencia de que, en los textos especializados y en la traducción especializada, lo más importante son los términos. El traductor de textos biomédicos también tiene que asimilar las construcciones, colocaciones, giros idiomáticos que utilizan los expertos y conocer el sentido restringido que adquieren las unidades léxicas de la lengua general en textos especializados.
- d) En las cadenas, un mismo concepto puede adquirir un sentido distinto según el cotexto. Por ejemplo, para los especialistas, el concepto INCREASE está en el mismo cotexto que *survival* y *mortality* y aparece principalmente en secciones que evalúan los resultados de un determinado tratamiento. Por el contrario, en los textos para pacientes esta idea alude a la naturaleza expansiva de la enfermedad, y por eso los lemas más frecuentes asociados son SPREAD y GROW.

- e) El desarrollo de las cadenas, cuyos lemas hemos etiquetado con un sustantivo en la mayoría de los casos, deja patente la supremacía de los sustantivos sobre los verbos.
- f) En los textos para especialistas hay una tendencia a la variación terminográfica que responde al intento de presentar un mismo concepto desde distintas perspectivas mediante hipónimos, hiperónimos y comparaciones para facilitar la comprensión del lector lego. Este tipo de variación no debe confundirse con la de los textos para especialistas, que obedece a una mayor complejidad conceptual y a la multidimensionalidad.
- g) En definitiva, las similitudes y diferencias conceptuales entre tipos textuales obedecen obviamente a la idiosincrasia del tema tratado y también al factor *destinatario del texto*. Cuando el redactor piensa en su lector, focaliza los lexemas del texto hacia determinados aspectos del subdominio. Esto tiene consecuencias terminológicas relacionadas con la cohesión y la variedad léxica.

Además de estas diferencias en el peso de las cadenas y sus componentes, también hay divergencia entre estos dos tipos textuales en lo que se refiere a la presencia de recursos lingüísticos que hacen explícita la relación cohesiva entre lexemas. La presencia de estos recursos en los textos para pacientes sirve de puente entre el conocimiento del lector y el conocimiento especializado, produciéndose un movimiento desde la información asimilada por el lector hacia la información especializada.

Mientras que en los textos para especialistas, la cohesión léxica se fundamenta en gran medida en la repetición exacta y simple, en los destinados a pacientes se crea cohesión mediante la duplicidad *término especializado – lexema de la lengua general (término superordinado)*. La cohesión para hacer asequible el vocabulario de los especialistas al lector lego se consigue principalmente mediante repetición variada léxica y sintáctica y mediante la utilización de patrones sintácticos (paréntesis, aposiciones, oraciones de relativo, paráfrasis, etc.) que pueden resultar útiles al traductor y al terminógrafo en el análisis conceptual que llevan a cabo como parte de su tarea de documentación. Según Meyer y Mackintosh (1996: 21), estos recursos presentan un enorme potencial en la ingeniería del conocimiento y en terminología,

especialmente en la recuperación de la información y la búsqueda de sinónimos dentro de *córpora* informatizados.

En este trabajo, también se ha indagado sobre el papel que la frecuencia de determinadas unidades léxicas en un texto o en un corpus puede jugar en la configuración de cadenas léxicas, y en el reconocimiento de unidades fraseológicas.

Estudiar las unidades léxicas recurrentes en un corpus ayuda en la identificación de cadenas léxicas, ya que existe una correlación entre los términos más frecuentes y los conceptos que hilan las cadenas léxicas. Los términos más repetidos construyen gran parte de las cadenas léxicas más extensas, y representan en gran medida los marcos conceptuales más relevantes del texto y los focos de interés del profesional. No obstante, somos conscientes de que estas palabras no dan cuenta de todas las variadas relaciones cohesivas sobre las que se construyen las cadenas léxicas. Asimismo, el estudio de cómo las unidades léxicas más frecuentes se distribuyen a lo largo de distintos textos deja entrever la estructuración en secciones de contenido. Queda pues suficientemente comprobada la contribución de la cohesión léxica a la interacción entre el plano semántico y textual.

Al igual que con las palabras más frecuentes hemos etiquetado cadenas léxicas, pensamos que las palabras más frecuentes de un texto o un corpus constituyen el vocabulario básico sobre el que se construirían todas las definiciones de un glosario terminológico.

En definitiva, se puede aducir que la repetición de unidades de significación clave de un texto juega un papel primordial en su cohesión, y por tanto, en la consecución de una activación conceptual coherente. Si detectamos los lemas más frecuentes—bajo los que se agrupan variantes morfológicas que comparten la misma raíz y que se vinculan mediante repetición exacta, simple y variada sintáctica—percibiremos las áreas conceptuales más activadas en un texto e identificaremos posibles cadenas y subcadenas. Además, la frecuencia léxica es importante porque, según el modelo cognitivo de Bell (1991), a la hora de traducir, se ponen en juego en la memoria una base de datos con las palabras y estructuras sintácticas más frecuentes.

Como conclusión, diremos que las cadenas léxicas, que se fundamentan en la cohesión y en la repetición léxica, anclan el texto en las áreas conceptuales de un

determinado campo del saber. Las cadenas léxicas identificadas por el traductor, el terminólogo o el documentalista con la ayuda de medios informáticos, se convierten en un instrumento útil en la tarea indagadora de estos en el sentido de que son una puerta de entrada hacia la terminología que utilizan los especialistas. En consecuencia, el estudio de la cohesión tiene importantes implicaciones en terminología, documentación y traducción.

El intento de sistematizar y detectar estos mecanismos cohesivos en textos destinados a una audiencia distinta para facilitar la tarea terminológica y de documentación del traductor ha tomado prelación con respecto al estudio de las diferencias cohesivas entre textos originales y textos traducidos. Algunas de estas variaciones cohesivas derivadas del proceso traductor son las mismas que las que se basan en el conocimiento del lector del texto (especialistas o lectores lego) y las que se derivan de las convenciones retóricas de las lenguas inglesa y española. Al fin y al cabo, hay puntos de contacto entre la traducción y la divulgación, ya que esta última implica la existencia de dos textos, un texto fuente, el de especialidad, que es reformulado en un texto derivado, el propio de la divulgación. Al igual que el traductor, el divulgador es un intermediario entre los científicos y el público lego.

Al traducir textos del inglés al español, se producen cambios con respecto a los recursos cohesivos utilizados y al grado de explicitación de la información que se pretende transmitir. Según Fernández Polo (1995, 1999) estas variaciones se deben a que las lenguas inglesa y española se estructuran de forma distinta desde el punto de vista retórico. Los textos en español son menos explícitos desde el punto cohesivo tanto si estos son textos originales como si son traducciones del inglés.

En los capítulos 10 y 15 se ha constatado que en español hay una tendencia retórica a evitar que se repita una unidad léxica o terminológica en la secuencia textual más de dos veces seguidas, con lo cual, la relación de cohesión, que suele ser explícita en inglés, pasa a oscurecerse en el TT. Esta tendencia se atenúa cuanto mayor es el grado de especialización de los textos. Las publicaciones para especialistas contienen un vocabulario consensuado y normalizado, y para evitar la ambigüedad, permiten la repetición léxica desde el punto de vista retórico.

Hemos reflexionado sobre otros factores que explican los cambios cohesivos detectables al comparar un texto inglés y su traducción al español.

- a) La mediación lingüística lleva aparejada la eliminación de la repetición y la redundancia, por lo que normalmente las traducciones al español presentan una mayor variedad terminológica.
- b) La gramática del español, más flexiva que la del inglés, puede explicar también el hecho de que la traducción al español contenga un número mayor de formas distintas que el texto original.
- c) En un plano mucho menos idealizado, encontramos a veces traducciones realizadas con poco rigor terminológico. Bajo estas condiciones, el traductor utiliza poco el término del que duda y en su lugar, utiliza un sinónimo o un término superordinado del que sí tiene certeza, con lo que la variedad léxica es mayor en el original.

En definitiva, el interés de estudiar la cohesión léxica en relación a la traducción es evidente dado que la redacción de textos cohesivos facilita la transmisión de la función comunicativa de los mismos y su comprensión por parte del lector. En el proceso traslativo, el traductor detecta la conexión entre los elementos del texto origen y reformula esta unidad conceptual en el texto meta mediante recursos cohesivos.

En investigaciones posteriores, se podría aplicar la metodología de esta tesis al estudio cuantitativo de la cohesión en textos originales sobre el mismo tema tanto en inglés como en español y a un corpus más extenso de textos traducidos. Haciendo un seguimiento de las referencias al mismo concepto, se puede ver si el texto en español emplea un mayor número de formas que el correspondiente en inglés. Se tendrá en cuenta obviamente si el texto es original o traducido y las diferencias gramaticales entre una lengua y otra.

Asimismo, estimamos que esta tesis ha abierto el camino a nuevas aproximaciones pedagógicas a la traducción y a la gestión de la terminología para la traducción. Esto es así porque el reconocimiento de las oraciones más cohesionadas y de las cadenas léxicas tiene implicaciones importantes en la formación del futuro traductor. Al combinar en nuestro estudio las perspectivas de la terminología y la lingüística de corpus, hemos sugerido formas de aplicar la lingüística de corpus a la

terminología. Hemos explotado listas de frecuencia lematizadas de palabras plenas, líneas de concordancia y un detector de combinaciones léxicas reiterativas con el fin de identificar no sólo las USE más relevantes sino también las áreas conceptuales más evocadas y el vocabulario definicional que conformaría un glosario sobre oncología. Estos recursos también parecen útiles a la hora de indagar sobre los patrones sintácticos en los que encajan los términos.

Por último, nos gustaría concluir con las palabras que ponen punto final a Hoey (1991) e inician Hoey (1995). Con ellas, destacamos la importancia que tiene la cohesión no sólo en la transmisión y comprensión de la función comunicativa de los textos sino también en la construcción del conocimiento y la sociedad:

But books have to end somewhere, and this one will have served its purpose if it encourages readers to bond their own sentences to mine (whether in harmony or contradiction) and to add what they have to say to the endlessly expanding net that is the sum of all human discourse. (Hoey 1991: 245).

BIBLIOGRAFÍA

BIBLIOGRAFÍA

- Aarts, J. and W. Meijs (eds). 1984. *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi.
- Aarts, J. y W. Meijs (eds.). 1986. *Corpus Linguistics II: New studies in the analysis and exploitation of computer corpora*. Amsterdam: Rodopi.
- Abraham, Sunita. 1993. *Writing to Learn: Knowledge-telling and knowledge-transforming in undergraduate writing*. Tesis doctoral. University of Birmingham.
- Adams Smith, D. 1984. "Medical discourse: aspects of author's comment". *English for Specific Purposes* 3: 25-36.
- Ahmad, K. y H. Fulford. 1992. *Knowledge Processing 4: Semantic relations and their use in elaborating terminology*. Informe del Departamento de Informática nº CS-92-10. Guildford: University of Surrey.
- Aijmer, K. y B. Altenberg (eds). 1991. *English Corpus Linguistics*. Londres: Longman.
- Aitchison, J. 1987. *Words in the Mind. An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell.
- Aitchison, J. 1994. "Understanding words". En G. Brown *et al.* (eds). *Language and Understanding*. Oxford: Oxford University Press. 83-95.
- Alonso Ramos, M. y A. Tutin. 1996. "A classification and description of lexical functions for the analysis of their combinations". En L. Wanner (ed). *Lexical Functions in Lexicology and Natural Language Processing*. Amsterdam y Philadelphia: John Benjamins. 147-167.
- Alonso Ramos, M., A. Tutin y G. Lapalme. 1995. "Lexical functions of the *Explanatory Combinatorial Dictionary* for lexicalization in text generation". En P. Saint-Dizier y E. Viegas (eds). *Computational lexical semantics*. Cambridge: Cambridge University Press. 351-366.
- Arntz, R. y H. Picht. 1989. *Einführung in die Terminologearbeit*. Hildesheim: Georg Olms. [Trad. española de Amelia de Irazazábal *et al.* 1995. *Introducción a la terminología*. Madrid: Fundación Germán Sánchez Ruipérez].
- Atkins, S., J. Clear y N. Ostler. "Corpus design criteria". *Literary and Linguistic Computing* 7 (1): 1-16.
- Austin, J.L. 1962. *How to Do Things with Words*. Oxford: Clarendon Press.
- Aznar Más, L. *et al.* 1989. "Algunos problemas en la traducción al inglés del lenguaje científico". *Actas del XI del Congreso AEDEAN*: 25-31.

- Baddeley, A.D. 1990. *Human memory. Theory and Practice*. Londres: Lawrence Erlbaum Associates.
- Baddeley, A.D. 1992. "Working memory: the interface between memory and cognition". *Journal of Cognitive Neuroscience* 4(3): 281-288.
- Baker, M. 1992. *In Other Words: A coursebook on translation*. Londres/Nueva York: Routledge.
- Baker, M., G. Francis y E. Tognini-Bonelli (eds). 1993. *Text and Technology*. Amsterdam/Philadelphia: John Benjamins.
- Baker, Mona. 2000. "Using corpora for translation research and pedagogy". Conferencia ofrecida en los *IV Encuentros Internacionales de Traducción: Las nuevas tecnologías y el traductor* de la Universidad de Alcalá (Alcalá de Henares). 17-18 de febrero de 2000.
- Barlow, M. 1996. "Parallel texts in language teaching". En S. Botley *et al.* (eds). 1996. *Proceedings of Teaching and Language Corpora (UCREL Technical Papers, Volume 9)*. Lancaster: University of Lancaster. 45- 56.
- Barlow, M. 1994. "ParaConc: a concordancer for parallel texts". En A. Wilson y A. McEnery (eds). *Corpora in Language Education and Research: A Selection of Papers from Teaching and Language Corpora 94 (Unit for Computer Research on the English Language Technical Papers 4)*. Lancaster: University of Lancaster.
- Bazerman, C. 1988. *Shaping written knowledge*. Madison: University of Wisconsin Press.
- Bazerman, C. y J. Paradis. 1991. *Textual Dynamics of the Professions: Historical and contemporary studies of writing in professional communities*. Madison: University of Wisconsin.
- Beaugrande, R. de. 1980. *Text, Discourse and Process. Towards a multidisciplinary science of text*. Londres: Longman.
- Beaugrande, R. de y W. Dressler. 1981. *Introduction to Text Linguistics*. Londres: Longman.
- Bell, R.T. 1991. *Translation and Translating: Theory and practice*. Londres: Longman.
- Benbrahim, M. 1996. *Automatic Text Summarisation through Lexical Cohesion Analysis*. Tesis doctoral. Guilford: University of Surrey.
- Benbrahim, M. y K. Ahmad. 1994. "Computer-aided lexical cohesion analysis and text abridgment". *Technical Report Knowledge Processing 18*, Informe del Departamento de Informática nº CS-94-11, Guilford: University of Surrey.
- Berber Sardinha, A.P. 1994a. "Particle, wave, segmentation, and bonding". *DIRECT Papers* 18. Sao Paulo, CEPRIL, PUC-SP – AELSU, Liverpool University. [Disponible en http://www.liv.ac.uk/r_tony1/linguistics.html].

- Berber Sardinha, A.P. 1994b. "Lexis in annual reports: Text segmentation and lexical threads". *DIRECT Papers* 8. Sao Paulo, CEPRIL, PUC-SP – AELSU, Liverpool University. [Disponible en http://www.liv.ac.uk/r_tony1/linguistics.html].
- Berber Sardinha, A.P. 1995. "A preliminary study into patterns of lexis of business texts". En B. Warvik, S-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 157-165.
- Berber Sardinha, A.P. 1997a. *Automatic Identification of Segments in Written texts*. Tesis doctoral. University of Liverpool. [Disponible en <http://members.wbs.net/homepages/c/o/r/corpuslinguistics/homepage.html>]
- Berber Sardinha, A.P. 1997b. "Patterns of lexis in original and translated business reports". En K. Simms (ed). *Translating Sensitive Texts: Linguistic aspects*. Amsterdam: Rodopi. 147-153.
- Bhatia, V.K. 1993. *Analysing Genre. Language use in professional settings*. Londres: Longman.
- Biber, D. 1993. "Representativeness in corpus design". *Literary and Linguistic Computing* 8 (4): 243-257.
- Biber, D. et al. 1998. *Corpus Linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D. y E. Finegan. 1991. "On the exploitation of computerized corpora in variation studies". En K. Aijmer y B. Altenberg (eds). *English Corpus Linguistics*. Londres: Longman. 204-220.
- Bobrow, D.B. y A. Collins. 1975. *Representation and understanding. Studies and cognitive science*. Nueva York: Academic Press.
- Botley, S. et al. (eds). 1996. *Proceedings of Teaching and Language Corpora (UCREL Technical Papers, Volume 9)*. Lancaster: University of Lancaster.
- Bowker, L. e I. Meyer. 1993. "Beyond 'textbook' concept systems: handling multidimensionality in a new generation of term banks". Comunicación presentada en el 3rd *International Congress on Terminology and Knowledge Engineering*. Frankfurt: INDEKS Verlag.
- Brady, M. y R.C. Berwick (eds). 1983. *Computational Models of Discourse*. Cambridge, MA: MIT Press.
- Brown, G. y G. Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Brown, P. et al. 1993. "Machine-readable dictionaries and research in computational linguistics". *Computational Linguistics* 19(2).
- Budin, G. y S.E. Wright. 1997. *Handbook of Terminology Management*. Volumen 1 y 2. Amsterdam/Philadelphia: John Benjamins.

- Bühler, Karl. 1934. *Sprachtheorie*. Jena: Fischer.
- Busch-Lauer, I.A. 1995. "Textual organization in English and German medical abstracts". En B. Warvik, S-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 175-185.
- Butler, C. (ed). 1992. *Computers and Written Texts*. Oxford: Blackwell.
- Cabré Castellví, M.T. 1999a. *La terminología: representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré Castellví, M.T. 1999b. "Contexto y entorno del trabajo terminológico. La terminología, ¿para qué?". Conferencia ofrecida en el Curso *Gestión de Terminología y Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Cabré Castellví, M.T., R. Estopà y M. Lorente. 1996. "Terminología y fraseología". *Actas del V Simposio Iberoamericano de Terminología*, Ciudad de México. [Disponible en <http://www.unilat.org/dtil/MEXICO/indice.htm>].
- Carne, C. 1996. "Corpora, genre analysis and dissertation writing: an evaluation of the potential of corpus-based techniques in the study of academic writing". En S. Botley *et al.* (eds). 1996. *Proceedings of Teaching and Language Corpora (UCREL Technical Papers 9)*. Lancaster: University of Lancaster. 127- 137.
- Catford, J.C. 1965. *A Linguistic Theory of Translation*. Oxford: Oxford University Press.
- Church, K. y R. Mercer. 1993. "Introduction to the special issue on Computational Linguistics using large corpora". *Computational Linguistics* 19 (1): 1-24.
- Clear, J. 1993. "From Firth principles. Computational tools for the study of collocation". En M. Baker, G. Francis y E. Tognini-Bonelli (eds). *Text and Technology*. Amsterdam/Philadelphia: John Benjamins. 271-292.
- Clyne, M. 1981. "Culture and discourse structure". *Journal of Pragmatics* 5: 61-66.
- Clyne, M. 1987. "Cultural differences in the organization of academic texts". *Journal of Pragmatics* 11: 201-238.
- Clyne, M. 1991. "Zu kulturellen Unterschieden in der Produktion und Wahrnehmung englischer und deutscher wissenschaftlicher Texte". *Infor DaF* 18 (4): 376-383.
- Connor, U. 1995. "Methods of research in contrastive rhetoric: some issues". En B. Warvik, S-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 15-28.

- Connor, U. 1996. *Contrastive Rhetoric: Cross-cultural aspects of second-language writing*. Cambridge: Cambridge University Press.
- Connor, U. y R. Kaplan. 1987. *Writing across Languages: Analysis of L2 text*. Reading, MA: Addison Wesley.
- Corazzari, O. y E. Picchi. 1994. "A proposal for the construction of comparable multilingual corpora". Informe de *Multext Project*. Pisa: Instituto di Lingüística Computazionale, CNR.
- Corpas Pastor, G. (ed). 2000a. *Las lenguas en Europa: Estudios de fraseología, fraseografía y traducción*. Granada: Comares.
- Corpas Pastor, G. 2000b. "Acerca de la (in)traducibilidad de la fraseología". En G. Corpas (ed). *Las lenguas en Europa: Estudios de fraseología, fraseografía y traducción*. Granada: Comares. 483-522.
- Coseriu, E. 1977. *Principios de Semántica Estructural*. Madrid: Gredos.
- Cruse, D. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Díez Prados, M. 1999. "Frecuencia de los mecanismos de cohesión en textos escritos en lengua inglesa". Comunicación presentada en el *XVII Congreso Nacional de AESLA*. Alcalá de Henares.
- Dik, S.C. 1978. *Functional Grammar*. Amsterdam: North Holland.
- Dik, S.C. 1989. *The Theory of Functional Grammar. Part I: The structure of the clause*. Dordrecht: Foris Publications.
- Dixon, R.M.W. 1965. *What is language?: a new approach to linguistic description*. Londres: Longman.
- Dressler, W. (ed). 1978. *Current trends in text linguistics*. Berlín/Nueva York: De Gruyter.
- Dubuc, R. y A. Lauriston. 1997. "Terms and contexts". En G. Budin y S.E. Wright. *Handbook of Terminology Management*. Volumen 1. Amsterdam/Philadelphia: John Benjamins. 80-87.
- Eggins, S. y J.R. Martin. 1997. "Genres and registers of discourse". En T. A. van Dijk (ed). *Discourse as Structure and Process*. Londres: Sage Publications.
- Ellegard, A. 1978. *The Syntactic Structure of English Texts*. Goteborg: University of Goteborg.
- Ellman, Jeremy. 1999. "The generality of lexical links". [Disponible en <http://osiris.sunderland.ac.uk/~cs0jel/web-The.html>]
- Ellman, J. 1998. "Using the generic document profile to cluster similar texts". *Proceedings Computational Linguistics UK (CLUK 97)*. Sunderland: University of Sunderland. [Disponible en <http://osiris.sunderland.ac.uk/~cs0jel/publications.htm>].

- Ellman, J., I. Klincke y J. Tait. 1998. "SUSS: The Sunderland University Similarity System: Beneath the glass ceiling". *Proceedings SENSEVAL workshop*. Brighton: University of Brighton. [Disponible en <http://osiris.sunderland.ac.uk/~cs0jel/publications.htm>].
- Ellman, J. y J. Tait. 1997. "Using information density to navigate the web". *Colloquium on Intelligent World Wide Web Agents*. Sunderland: University of Sunderland. [Disponible en <http://osiris.sunderland.ac.uk/~cs0jel/publications.htm>].
- Escalier, M.C. y C. Fournier. 1996. "Towards a notional representation of meaning in the Meaning-Text Model: The case of the French SI". En L. Wanner (ed). *Lexical Functions in Lexicology and Natural Language Processing*. Amsterdam y Philadelphia: John Benjamins. 25-60.
- Estopà Bagot, R. 1999. *Extracció de terminologia: elements per a la construcció d'un SEACUSE (Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada)*. Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Faber, P. 1997. "Significado léxico y significado textual". Conferencia ofrecida en el Curso *Traducción y Pragmática* del Centro Mediterráneo de la Universidad de Granada.
- Faber, P. 1998a. "Conceptos neurolingüísticos y psicolingüísticos en la traducción". Conferencia ofrecida en el Curso *Aspectos Cognitivos de la Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Faber, P. 1998b. "Translation competence and language awareness". *Language Awareness* 7 (1): 9-21.
- Faber, P. 1999a. "Funciones léxicas y su aplicación en la traducción". Ponencia presentada en las *V Jornadas Internacionales sobre el estudio y la enseñanza del léxico*. Universidad de Granada.
- Faber, P. 1999b. "Conceptual analysis and knowledge acquisition in scientific translation". *Terminologie et Traduction* 2: 97-123.
- Faber, P. y R. Mairal. 1999. *Constructing a Lexicon of English Verbs*. Berlín: Mouton de Gruyter.
- Fairclough, N. 1992. *Discourse and Social Change*. Cambridge: Polity
- Fente Gómez, R. 1971. *Estilística del verbo en inglés y en español*. Madrid: SGEL.
- Fernández Polo, J. 1995. "Some discoursal aspects in the translation of popular science texts from English into Spanish". En B. Warvik, S-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 257-264

- Fernández Polo, F. J. 1999. *Traducción y retórica contrastiva. A propósito de la traducción de textos de divulgación científica del inglés al español*. Santiago de Compostela: Servicio de Publicaciones de la Universidad de Santiago de Compostela.
- Fillmore, Charles. 1977. "Scenes-and-frames semantics". En A. Zampolli (ed). *Linguistic Structures Processing*. Amsterdam: North Holland. 55-81.
- Fillmore, C. y B. Atkins. 1998. "FrameNet and lexicographic relevance". En A. Rubio *et al.* (eds). *Proceedings of the First International Conference of Language Resources and Evaluation*. Volumen 1. Paris: ELRA. 417-423.
- Firth, J.R. 1935. "The technique of semantics". *Transactions of the Philological Society*: 36-72.
- Firth, J.R. 1957a. "Modes of meaning". En J.R. Firth. *Papers in Linguistics: 1934-1951*. Londres: Oxford University Press.
- Firth, J.R. 1957b. "A synopsis of linguistic theory, 1930-1955". *Studies in Linguistic Analysis*. Oxford: Blackwell. 1-32.
- Firth, J.R. 1968. "Linguistics and translation". En F.R. Palmer (ed). *Selected Papers on J. R. Firth 1952-59*. Londres: Longman. 84-95.
- Fowler, R. 1991. *Language and the News*. Londres: Routledge.
- Fowler, R. 1996 (1986). *Linguistic Criticism*. Oxford: Oxford University Press.
- Fowler, R. *et al.* (eds). 1979. *Language and Control*. Londres: Routledge & Kegan Paul.
- Francis, W. y H. Kučera. 1982. *Frequency Analysis of English Usage*. Houghton Mifflin.
- Frawley, W. 1988. "Relational models and metascience". En M. Evens (ed). *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press. 335-372.
- Frawley, William J. y R. N. Smith. 1989. "Patterns of textual cohesion in genre-specific discourse". En S. Williams (ed). *Humans and machines: Fourth Delaware Symposium on Language Studies*. Norwood, NJ: Ablex. 145-154.
- Galtung, J. 1985. "Struktur, Kultu und intellektueller Stil". En A. Wierlacher (ed). *Das Fremde und das Eigene: Prolegomena zu einer interkulturellen Germanistik*. München: Iudicium Verlag. 151-193.
- García de Quesada, M. *Estructura definicional terminográfica en el subdominio de la oncología médica*. Tesis doctoral en curso. Universidad de Granada.
- Gerzymisch-Arbogast, H. 1986. "Zur Relevanz der Thema-Rhema-Gliederung für den Übersetzungsprozess". En M. Snell-Hornby (ed). *Übersetzungswissenschaft-eine Neuorientierung*. Tübinga: Francke. 160-183.

- Ghadessy, M. (ed). 1993. *Register Analysis. Theory and Practice*. Londres: Pinter Publishers.
- Givón, T. 1997. "Toward a neuro-cognitive interpretation of 'context' ". Informe presentado en la Universidad de Oregón.
- Givón, T. 1998a. "Lexical networks, spreading activation and contextual priming". Conferencia ofrecida en el Curso *Aspectos Cognitivos de la Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Givón, T. 1998b. "Lexical representation in the mind-brain". Conferencia ofrecida en el Curso *Aspectos Cognitivos de la Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Gledhill, C. 1996. "Science as collocation. Phraseology in cancer research articles". En S. Botley *et al.* (eds). 1996. *Proceedings of Teaching and Language Corpora (UCREL Technical Papers, Volume 9)*. Lancaster: University of Lancaster. 108-126.
- Gledhill, C. 1997. "Les collocations et la construction du savoir scientifique". *Asp la revue du GERAS* 15-18: 85-104.
- Göpferich, S. 1995. *Textsorten in Naturwissenschaften und Technik. Pragmatische Typologie-Kontrastierung-Translation*. Tubinga: Narr.
- Gopnik, M. 1972. *Linguistic Structures in Scientific Texts*. The Hague: Mouton.
- Graustein, G. y W. Thiele. 1987. *Properties of English Texts*. Leipzig: VEB Verlag Enzyklopädie.
- Green, Stephen J. 1997. *Automatically Generating Hypertext by Computing Semantic Similarity*. Tesis doctoral. Universidad de Toronto. [Disponible en <http://sjgreen.mpce.mq.edu.au/~sjgreen/Documents/sjgthesis.ps.gz>].
- Grice, H.P. 1975. "Logic and conversation". En P. Cole y J.L. Morgan (eds). *Syntax and Semantics 3, Speech acts*. Nueva York: Academic Press.
- Grimes, J. 1975. *The Thread of Discourse*. The Hague: Mouton.
- Gülich, E. y R. Raible (eds). 1972. *Textsorten*. Frankfurt: Athenäum.
- Gutiérrez Rodilla, B. M. 1998. *La ciencia empieza en la palabra. Análisis e historia del lenguaje científico*. Barcelona: Península.
- Halliday, M.A.K. 1973. *Explorations in the Functions of Language*. Londres: Longman.
- Halliday, M.A.K. 1978. *Language as Social Semiotic*. Londres: Arnold.
- Halliday, M.A.K. y R. Hasan. 1976. *Cohesion in English*. Londres: Longman.
- Halliday, M.A.K. y R. Hasan. 1985. *Language, Context and Text. Aspects of Language in a Social-semiotic Perspective*. Victoria: Deakin University Press.

- Halliday, M.A.K., A. McIntosh y P. Strevens. 1964. *The Linguistic Sciences and Language Teaching*. Londres: Longman.
- Hanks, Patrick. 2000. "Dictionaries of idioms and phraseology in English". En G. Corpas. (ed). *Las lenguas en Europa: Estudios de fraseología, fraseografía y traducción*. Granada: Comares. 303-320
- Hartmann, R.R.K. 1980. *Contrastive Textology. Comparative Discourse Analysis in Applied Linguistics*. Heidelberg: J. Groos.
- Hartmann, R.R.K. 1994. "The use of parallel text corpora in the generation of translation equivalents for bilingual lexicography". *EURALEX 1994 Proceedings*. Amsterdam: 291-297.
- Hartmann, R.R.K. 1996. "Contrastive textology and corpus linguistics: On the value of parallel texts". *Language Sciences* 3-4:947-957.
- Hasan, R. 1978. "Text in the systemic functional model". En W. Dressler, (ed). *Current Trends in Text Linguistics*. Berlín/Nueva York: De Gruyter.
- Hatim, Basil. 1997. *Communication Across Cultures. Translation Theory and Contrastive Text Linguistics*. Exeter: University of Exeter Press.
- Hatim, B. y I. Mason. 1990. *Discourse and the Translator*. Londres/Nueva York: Longman.
- Hatim, Basil y I. Mason. 1997. *The Translator as Communicator*. Londres: Routledge.
- Hinds, J. 1987. "Reader vs. writer responsibility: A new typology". En U. Connor y R. Kaplan (eds). *Writing across Languages: Analysis of L2 text*. Reading, MA: Addison Wesley. 141-152.
- Hockey, S. y D. Walker. 1993. "Developing effective resources for research on texts: collecting texts, tagging texts, cataloguing texts, using texts, and putting texts in context". *Literary and Linguistic Computing* 8(4): 235-242.
- Hoey, Michael. 1979. *Signalling in Discourse*. Discourse Analysis Monographs 6. Birmingham: English Language Research of the University of Birmingham.
- Hoey, M. 1983. *On the Surface of Discourse*. Londres: George Allen and Unwin.
- Hoey, M. 1988. "Writing to meet the reader's need: text patterning and reading strategies". *Trondheim Papers in Applied Linguistics* 4: 51-73.
- Hoey, M. 1991. *Patterns of Lexis in Text*. Oxford: Oxford University Press.
- Hoey, M. 1995. "The lexical nature of intertextuality: A preliminary study". En B. Warvik, S-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 73-94.

- Hoey, M. 1996. "The discourse's disappearing (and reappearing) subject: An exploration of the extent of intertextual interference in the production of texts". En K. Simms (ed). *Linguistic Representation of the Subject*. Berlín/La Haya: Mouton de Gruyter.
- Hoey, M. y E. Winter. 1986. "Clause relations and the writer's communicative task". En B. Couture (ed). *Functional Approaches to Writing Research Perspectives*. Londres: Frances Pinter. 120-141.
- Holland, D. y N. Quinn. 1987. *Cultural Models in Language and Thought*. Nueva York: Cambridge University Press.
- Hönig, H.G. y P. Kussmaul. 1982. *Strategie der Übersetzung. Ein Lehr-und Arbeitsbuch*. Tübinga: Narr.
- Horner, W.B. 1975. *Text Act Theory, A study of non-fiction texts*. Tesis doctoral sin publicar. Universidad de Michigan.
- House, J. 1977. *A Model for Translation Quality Assessment*. Tübinga: Günter Narr.
- House, J. 1986. "Acquiring Translational Competence in Interaction". En J. House y S. Blum-Kulka (eds). Tübinga: Günter Narr. 179-194.
- Huth, Edward J. 1987. *Medical Style and Format. An International Manual for Authors, Editors and Publishers*. Philadelphia: ISI press.
- Hyland, K. 1998. *Hedging in Scientific Research Articles*. Amsterdam/Philadelphia: John Benjamins.
- Hymes, D. 1972 (1964). "Toward ethnographics of communication: The analysis of communicative events". En P.P. Giglioli (ed). *Language and Social Context*. Harmondsworth: Penguin.
- Iser, W. 1976. *Der Akt des Lesens. Theorie ästhetischer Wirkung*. München: Fink.
- Jakobson, R. 1959. "On linguistic aspects of translation". En R.A. Brower (ed). *On Translation*. Cambridge, MA: Harvard University Press. 232-239.
- Jiménez Hurtado, C. 1994. *El componente pragmático en el lexicón verbal del español, alemán e inglés*. Tesis doctoral. Universidad de Granada.
- Johansson, S. 1985. "Some observations on word frequencies in three corpora of present-day English texts". *ITL Review of Applied Linguistics* 67-68: 117-126.
- Kennedy, C. y R. Bolitho. 1984. *English for Specific Purposes*. Londres: Macmillan.
- Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. Londres: Longman.
- Kintsch, W. 1977. *Memory and cognition*. Nueva York: Wiley & Sons.
- Kittredge, R. y J. Lehrberger (eds). 1982. *Sublanguage*. Berlín: Mouton de Gruyter.

- Klavans, J. y E. Tzoukermann. 1990. "The BI-CORD system". *Proceedings of the 15th International Conference on Computational Linguistics (COLING-90)*: 174-179.
- Kozima H. y T. Furugori. 1997. "Similarity between words computed by spreading activation on an English dictionary". [Disponible en <http://www-karc.clr.go.jp/kss/xkozima/work/paper/9304eacl.ps>].
- Kress, G. 1985. *Linguistic Processes in Sociocultural Practice*. Geelong: Deakin University Press.
- Kress, G. 1989. *Linguistic Processes in Sociocultural Practice*. Oxford: Oxford University Press.
- Kretzenbacher, H.L. 1990. *Rekapitulation. Textstrategien der Zusammenfassung von Zusammenfassung von wissenschaftlichen Fachtexten*. Tübingen: Narr.
- Kusmaul, Paul. 1997. "Text-type conventions and translating: Some methodological issues". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins. 67-83.
- Labov, W. 1972a. "The study of language in its social context". En J.B. Pride y J. Holmes (eds). *Sociolinguistics*. Harmondsworth: Penguin. 180-202.
- Labov, W. 1972b. *Sociolinguistics Patterns*. Oxford: Blackwell.
- Laffling, J. 1991. "Towards high-precision machine translation. Based on contrastive textology". *Distributed Language Translation 7*. Berlín: Foris. 43-57.
- Laffling, J. 1992. "On constructing a transfer dictionary for man and machine". *Target 4*(1): 17-31.
- Lankamp, R.E. 1989. *A Study of the Effect of Terminology on L2 Reading Comprehension*. Amsterdam: Rodopi.
- Latour, B. 1987. *Science in Action: How to follow scientists and engineers through society*. Milton Keynes: Open University Press.
- Lee, W. y M. Evens. 1996. "Generating cohesive text using lexical functions". En L. Wanner (ed). *Lexical Functions in Lexicology and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins. 299-306.
- Leech, G.N. 1969. *A Linguistic Guide to English Poetry*. Londres: Longman.
- Leech, G. N. y S. Fligelstone. 1992. "Computers and corpus analysis". En C. Butler (ed). *Computers and Written Texts*. Oxford: Blackwell. 115-140.
- Leech, G.N. y M. H. Short. 1981. *Style in fiction. A Linguistic Introduction to English Fictional Prose*. Londres: Longman.

- Lehnert, W.G. y M.H. Ringle (eds). 1982. *Strategies for Natural Language Processing*. Hillsdale, NJ: Erlbaum.
- Lévy-Leblond, J.M. 1996. "La langue tire la science". En R. Chartier y P. Corsi (eds). *Sciences et langues en Europe*. Paris: EHESS. 235-245.
- Lewis, K. David. 1969. *Convention. A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Longacre, R.E. 1983. *The Grammar of Discourse*. Nueva York: Plenum Press.
- López García, A. 1991. *Psicolingüística*. Madrid: Síntesis.
- López Rodríguez, C.I. (en prensa). "La traducción de valores culturales en textos sobre educación: un estudio de corpus". *Actas del XVI Congreso Nacional de AESLA*. Universidad de la Rioja (Logroño), abril 1998.
- López Rodríguez, C.I. 2000. "Detección automática de la cohesión léxica en textos sobre oncología: aplicaciones a la traducción". Comunicación presentada en los *IV Encuentros Internacionales de Traducción* de la Universidad de Alcalá (Alcalá de Henares). 17-18 de febrero de 2000.
- Lotfipour-Saedi, K. 1997. "Lexical cohesion and translation equivalence". *Meta* 42 (1): 185-192.
- Louw, B. 1993. "Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies". En M. Baker, G. Francis y E. Tognini-Bonelli (eds). *Text and Technology*. Amsterdam/Philadelphia: John Benjamins. 157-177
- Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Lyons, J. 1977. *Semantics*. 2 vols. Cambridge: Cambridge University Press
- Maher, J. 1986. "The development of English as an international language of medicine". *Applied Linguistics* 7 (2): 206-218
- Malinowski, B. 1923. "The problem of meaning in primitive languages". En C.K. Ogden e I.A. Richards. *The Meaning of Meaning*. Londres: Kegan Paul.
- Márquez Linares, C. 1998. *La polisemia en el campo léxico EL CUERPO HUMANO: un estudio contrastivo inglés-español*. Tesis doctoral. Universidad de Córdoba.
- Martín Mingorance, L. 1998. *El modelo lexemático-funcional*. Granada: Servicio de Publicaciones de la Universidad de Granada.
- Martin, J.R. 1984. "Language, register and genre". En F. Christie (ed). *Children's Writing: Studies Guide and Reader*. Geelong: Deakin University Press.

- Martin, J.R. 1987. *Factual writing: exploring and challenging social reality*. Melbourne: Deakin University Press.
- Mauranen, A. 1993. "Contrastive ESP rhetoric: Metatext in Finnish-English economics". *ESP* 12 (1): 3-22.
- Mayoral Asensio, R. 1997. *La traducción de la variación lingüística*. Tesis doctoral. Universidad de Granada.
- McCarthy, M. 1991. *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- McCarthy, M. y R. Carter. 1994. *Language as Discourse: Perspectives for Language Teaching*. Londres: Longman.
- McEnery, A. y A. Wilson. 1996. *Corpus Linguistics*. Edimburgo: Edinburgh University Press.
- McEnery, A. y A. Wilson (eds). 1994. "Corpora and translation: uses and future prospects". En M.A. Lorget (ed). *Atti della Fiera Internazionale della Traduzione II*. Bolonia: Editrice CLUEB Bolonia. 311-342.
- McKeown, R. 1992. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Texts*. Cambridge: Cambridge University Press
- Mel'cuk, Igor. 1996. "Lexical functions: A tool for the description of lexical relations in the lexicon". En L. Wanner (ed). *Lexical Functions in Lexicology and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins. 37-102.
- Mel'cuk, I. et al. 1984. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicos-sémantiques I*. Montreal: University of Montreal Press.
- Mel'cuk, I. y D. Zholkovskij. 1970. "Towards a functioning Meaning-Text Model of language". *Linguistics* 57: 10-47.
- Mel'cuk, I. y D. Zholkovskij. 1988. "The Explanatory Combinatorial Dictionary". En M. Evens (ed). *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge: Cambridge University Press. 41-74.
- Meyer, I. y K. Mackintosh. 1996. "Redefining the terminographer's concept-analysis methods: How can phraseology help?". *Terminology* 3 (1): 1-26.
- Minsky, M. 1975. "A framework for representing knowledge". En P.H. Winston (ed). *The Psychology of Computer Vision*. Nueva York: Mc Graw-Hill.
- Montaño-Harmon, M.R. 1991. "Discourse features of Mexican Spanish: Current research in contrastive rhetoric and its implications". *Hispania* 74: 417-425.
- Miller, G. et al. 1991. "Introduction to WordNet: An on-line lexical database". *Journal of Lexicography* 3: 235-244.

- Mirador, Josephine F. 1995. "The structure of scientific article conclusions in medicine". En P. Bruthiaux, T. Boswood y B. Du-Babcock (eds). *Explorations in English for Professional Communication*. Hong Kong: City University of Hong Kong. 85-97.
- Morris, A.J. 1988. "Lexical cohesion, the thesaurus, and the structure of text". *Technical Report 219*. Toronto: University of Toronto.
- Morris, A.J. y G. Hirst (1991). "Lexical cohesion computed by thesaural relations as an indicator of the structure of text". *Computational Linguistics* 17 (1): 21-48.
- Mosenthal, P.B. 1985. "Defining the expository discourse continuum: towards a taxonomy of expository text types". *Poetics* 14: 387-414.
- Muñoz Martín, R. 1995. *Lingüística para traducir*. Barcelona: Teide.
- Muñoz Martín, R. y E. Sánchez Trigo. 1994. "Parallel texts: a research project". En R. Martín-Gaitero (ed). *Actas de los V Encuentros Complutenses en torno a la traducción*. Madrid: Editorial Complutense. 169-174.
- Myers, Greg. 1990. *Writing biology: Text in the social construction of scientific knowledge*. Madison: University of Wisconsin Press.
- Myers, G. 1991. "Lexical cohesion and specialized knowledge in science and popular science texts". *Discourse Processes* 14 (1): 1-26.
- Myers, G. 1992. "Textbooks and the sociology of scientific knowledge". *English for Specific Purposes* 1: 3-17.
- Neubert, A. 1968. "Pragmatische Aspekte der Übersetzung". En A. Neubert y O. Kade (eds). *Grundfragen der Übersetzungswissenschaft*. Leipzig: Enzyklopädie Verl.
- Neubert, A. 1981. "Translation, interpreting and text linguistics". *Studia Linguistica* 35 (1-2): 130-145.
- Neubert, A. 1983. "Discourse analysis of translation". *Zeitschrift-fur-Anglistik-und-Amerikanistik* 31 (2): 160-162.
- Neubert, A. 1984. "Text-bound translation teaching". En W. Wilss y Thome (eds). *Die Theorie des Übersetzens und ihr Aufschlusswert für die Übersetzungs-und Dometschdidaktik*. Tübinga: Narr. 61-70.
- Neubert, A. 1985. *Text and Translation*. Leipzig: Verlag Enzyklopädie.
- Neubert, A. 1991. "Models of translation". En S. Tirkonnen-Condit (ed). *Empirical Research in Translation and Intercultural Studies*. Tübinga: Günter Narr.
- Neubert, A. 1997. "Postulates for a theory of translatio". En J.H. Danks *et al.* (eds). *Cognitive Processes in Translation and Interpreting*. Londres: Sage Publications. 1-24.

- Newmark, P. 1981. *Approaches to Translation*. Oxford: Pergamon.
- Newmark, P. 1988. *A Textbook of Translation*. Londres/Nueva York: Prentice Hall.
- Nida, E.A. 1964. *Toward a Science of Translating*. Leiden: Brill.
- Nord, Christiane. 1987. "Übersetzungsprobleme-Übersetzungsschwierigkeiten. Was in den Köpfen von Übersetzern vorgehen sollte...". *Mitteilungsblatt für Dolmetscher und Übersetzer* 2: 5-8.
- Nord, C. 1991a. "Scopos, loyalty and translational conventions". *Target* 3(1): 91-109.
- Nord, C. 1991b. *Text Analysis in Translation Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Amsterdam: Rodopi B.V.
- Nord, C. 1996a. "Revisiting the classics: Text type and translation method. An objective approach to translation criticism". *The Translator* 2 (1): 81-88.
- Nord, C. 1996b. *Translating as a Purposeful Activity. Functionalist Approaches Explained*. Manchester: St. Jerome.
- Nord, C. 1997. "A functional typology of translations". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins. 43-66.
- Nwogu, K. N. 1991. "Structure of science popularizations: A genre-analysis approach to the schema of popularized medical texts". *English for Specific Purposes* 10: 111-122.
- Nwogu, K. N. 1997. "The medical research paper: structure and functions". *English for Specific Purposes* 16 (2): 119-138.
- Paltridge, B. 1994. "Genre analysis and the identification of textual boundaries". *Applied Linguistics* 15 (3): 288-299.
- Paltridge, B. 1997. *Genre, Frames and Writing in Research Settings*. Amsterdam/Philadelphia: John Benjamins.
- Pérez Guerra, J. 1998. *Análisis computerizado de textos. Una introducción al TACT*. Vigo: Servicio de publicaciones de la Universidad de Vigo.
- Pérez Hernández, M.C. 2000. *Explotación de los corpórea textuales informatizados para la creación de bases terminológicas basadas en el conocimiento*. Tesis doctoral. Universidad de Málaga.
- Peters, C. y E. Picchi. 1997. "Reference corpora and lexicon for translators and translation studies". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins. 247-274.
- Peters, C., E. Picchi, y L. Biagini. 1996. "Parallel and comparable bilingual corpora in language teaching and learning". En S. Botley et al. (eds). *Proceedings of Teaching and*

- Language Corpora (UCREL Technical Papers, Volume 9)*. Lancaster: University of Lancaster. 68-80.
- Petöfi, J.S. 1979. *Text vs. Sentence. Basic Questions of Text Linguistics*. Papiere zur Textlinguistik 20. 2 vols. Hamburg: Buske.
- Petöfi, J.S. y H. Rieser (eds). 1973. *Studies in Text Grammar*. Dordrech: Reidel.
- Phillips, M. 1985. *Aspects of Text Structure: An investigation of the lexical organisation of text*. Amsterdam: North-Holland.
- Pike, K.L. y E.G. Pike. 1977. *Grammatical Analysis*. Summer Institute of Linguistics Publications in Linguistics No 53. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington.
- Pilegaard, Morten. 1997. "Translation of medical research articles". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins. 159-184.
- Puerta López-Cózar, J.L. y A. Mauri Más. 1995. *Manual para la redacción, traducción y publicación de textos médicos*. Barcelona: Mason.
- Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Quirk, R. et al. 1972. *A Grammar of Contemporary English*. Londres: Longman.
- Rabadán, Rosa. 1991. *Equivalencia y traducción: problemática de la equivalencia transléctica inglés-español*. León: Universidad de León.
- Reiss, Katharina. 1968. "Überlegungen zu einer Theorie der Übersetzungskritik". *Linguistica Antverpiensia* 2: 369-383.
- Reiss, K. 1976. *Texttyp und Übersetzungsmethode. Der operative Text*. Kronberg/Ts: Scriptor.
- Reiss, K. 1979. "Types de textes et types de traductions". *Le langage et l'Homme* 40: 48-54.
- Reiss, K. y J. Vermeer. 1984. *Grundlegung einer Allgemeinen Translationstheorie*. Tubinga: Niemeyer.
- Rector, A.L. et al. 1994. "The GALEN CORE Model Schemata for Anatomy: Towards a Re-usable Application-Independent Model of Medical Concepts". *Proceedings of Medical Informatics Europe (MIE)*, Lisboa: 229-233.
- Renouf, A. 1984. "Corpus development at Birmingham University". En J. Aarts y W. Meijs (eds). *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi. 3-39
- Renouf, A. y A. Collier. 1995. "A system of automatic abridgment". Informe sin publicar del *Research and Development Unit for English Studies*. Liverpool: University of Liverpool.

- Rodríguez Reina, M.P. 2000. "Cómo *abordan* la fraseología los diccionarios bilingües italiano-español del siglo XX desde perspectivas marinas". En G. Corpas. (ed). *Las lenguas en Europa: Estudios de fraseología, fraseografía y traducción*. Granada: Comares. 321-346.
- Rosch, E. 1977. "Classification of real world objects: origins and representations in cognition". En P.N. Johnson-Laird y P.C. Wagon (eds). *Thinking: readings in cognitive science*. Cambridge: Cambridge University Press. 212-222.
- Rosch, E. y C.B. Mervis. 1975. "Family resemblances: studies in the internal structure of categories". En E. Hunt (ed). *Cognitive Psychology*. Nueva York/Londres: Academic Press. 573-603.
- Saeed, J. I. 1997. *Semantics*. Londres: Blackwell.
- Sager, J.C. 1997. "Text types and translation". En A. Trosborg (ed). *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins. 25-41.
- Sager, J.C. 1998. "What distinguishes major types of translation?". *The Translator* 4 (1): 69-89.
- Saint-Dizier, P. y E. Viegas (eds). 1995. *Computational Lexical Semantics*. Cambridge: Cambridge University Press.
- Salager-Meyer, F. 1991. "Hedging in medical discourse: 1980-1990". *Interface* 6(1): 33-54.
- Salager-Meyer, F. 1994. "Hedges and textual communicative function in medical English written discourse". *English for Specific Purposes* 13(2): 149-170.
- Salager-Meyer, F. *et al.* 1989. "Communicative function and grammatical variations in medical English scholarly papers: A genre analysis study". En C. Lauren y M. Nordman (eds). *Special Language: From humans thinking to thinking machines*: 151-160.
- Saussure, F. de. 1983 (1916). *Curso de Lingüística General*. [Trad. de A. Alonso. Ed. Crítica de T. de Mauro]. Madrid: Alianza Editorial.
- Schank, R. y R. Abelson. 1977. *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Erlbaum.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press
- Searle, J. 1979. *Expression and Meaning. Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press
- Sinclair, John M. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J.M. 1993. "Written discourse structure". En J.M. Sinclair, M. Hoey y G. Fox (eds). *Techniques of Description: Spoken and written discourse*. Londres: Routledge. 6-31.
- Snell-Hornby, M. 1988. *Translation Studies. An Integrated Approach* Amsterdam/Philadelphia: John Benjamins.

- Snell-Hornby, M. 1997. "Translation as a cross-cultural event". Conferencia ofrecida en el Curso *Traducción y Pragmática* del Centro Mediterráneo de la Universidad de Granada.
- Somers, H. 1996. *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins.
- Sonneveld, H.B. y K.L. Loening. 1993. *Terminology. Applications in Interdisciplinary Communication*. Amsterdam/Philadelphia: John Benjamins.
- Stubbs, Michael. 1987 (1983). *Análisis del discurso. Análisis sociolingüístico del lenguaje natural*. [Trad. española de Celina González]. Madrid: Alianza Editorial.
- Stubbs, M. 1996. *Text and Corpus Analysis: Computer-assisted studies of language and culture*. Londres: Blackwell Publishers.
- Suau Jiménez, F.A. (en prensa). "La repetición léxica como instrumento pragmático de cohesión y coherencia". *Actas del XVI Congreso Nacional de AESLA*, Universidad de la Rioja (Logroño), abril 1998.
- Svartvik, J. (ed). 1992. *Directions in Corpus Linguistics*. Berlín: Mouton de Gruyter.
- Swales, John M. 1981. *Aspects of Article Introductions*. Birmingham: Aston University.
- Swales, J.M. 1985. "A genre-based approach to language across the curriculum". Ponencia presentada en *RELC Conference*, Singapore.
- Swales, J. M. 1990. *Genre Analysis. English in Academic an Research Settings*. Cambridge: Cambridge University Press
- Swales, J. M. y C.B. Feak. 1994. *Academic Writing for Graduate Students*. The University of Michigan Press.
- Swales, J.M. 1981. *Aspects of Article Introductions*. Birmingham: Aston University.
- Sweetser, E.E. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Tanskanen, S.-K. 1995. "Lexical cohesion meets communicative conditions". En B. Warvik, S.-K. Tanskanen y R. Hiltunen (eds). *Organization in Discourse. Proceedings from the Turku Conference. Anglicana Turkuensia 14*. Turku: University of Turku. 531-538.
- Taylor, J.R. 1995 (1989). *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.
- Tercedor Sánchez, M.I. 1999. *La fraseología en el lenguaje biomédico: análisis desde las necesidades del traductor*. Tesis doctoral. Universidad de Granada. [Disponible en <http://elies.rediris.es/elies6/>].

- Teubert, W. 1996. "Comparable or parallel corpora". *International Journal of Lexicography* 9 (3): 238-264.
- Thomas, S. y T.-P. Hawes. 1994. "Reporting verbs in medical journal articles". *English for Specific Purposes* 13 (2): 129-148.
- Tomaszczyk, J. 1989. "L1-L2 technical translation and dictionaries". *Translation and Lexicography*. Innsbruck: John Benjamins.
- Toury, G. 1980. *In Search of a Theory of Translation*. Tel Aviv: Porter Institute for Poetics and Semiotics.
- Trosborg, Anna (ed). 1997. *Text Typology and Translation*. Amsterdam/Philadelphia: John Benjamins.
- Tyler, A. 1992. "Discourse structure and specification of relationship: A crosslinguistic analysis". *Text* 12(1): 1-18.
- Van Dijk, Teun A. 1972. *Some Aspects of Text Grammars*. The Hague: Mouton.
- Van Dijk, T.A. 1977a. *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. Londres: Longman.
- Van Dijk, T.A. 1977b. "Semantic macro-structures and knowledge frames in discourse comprehension". En M. Just y P. Carpenter (eds). *Cognitive Processes in Comprehension*. Hillsdale, NJ: Erlbaum. 3-32.
- Van Dijk, T.A. 1981. "Discourse studies in education". *Applied Linguistics* 2 (1).
- Van Dijk, T.A. 1980. *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van Dijk, T.A. y W. Kintsch. 1978. "Towards a model of discourse comprehension and production". *Psychological review* 85: 363-394.
- Van Rinjn-Van Tongeren, G.W. 1997. *Metaphors in Medical texts*. Amsterdam: Rodopi B.V.
- Ventola, E. 1987. *The Structure of Social Interaction: A systemic approach to the semiotics of service encounters*. Londres: Frances Pinter.
- Vermeer, H.J. 1978. "Ein rahmen für eine allgemeine translationstheorie". En H. J. Vermeer (ed). 1983. *Aufsätze zur Translationstheorie*. Heidelberg: Vermeer. 48-61.
- Vihla, M. 1998. "Medicor: A corpus of contemporary American medical texts". *ICAME* 22: 73-80.
- Walker, D. y R.A. Amsler. 1991. "The ecology of language". *Proceedings of the International Workshop on Electronic Dictionaries*. Tokio: Japan Electronic Dictionary Research Institute. 10-22.

- Wanner, L. (eds). 1996. *Lexical Functions in Lexicology and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins.
- Webber, Pauline. 1994 "The function of questions in different medical journal genres". *English for Specific Purposes* 13 (3): 257-268.
- Werlich, E. 1976. *A Text Grammar of English*. Heidelberg: Quelle & Meyer.
- Wertheimer, M. 1959 (1912). *Productive Thinking*. Chicago: Chicago University Press.
- Williams, I.A. 1996. "A translator's reference needs: dictionaries or parallel texts?". *Target* 8(2): 275-299.
- Wilson, A. y A. McEnery (eds). 1994. *Corpora in Language Education and Research: A selection of papers from Teaching And Language Corpora 94 (Unit for Computer Research on the English Language Technical Papers 4)*. Lancaster: University of Lancaster.
- Wilss, W. 1995. *Knowledge and Skills in Translator Behaviour*. Amsterdam: John Benjamins.
- Winter, E.O. 1974. *Replacement as a Function of Repetition: a study of some of its principal features in the clause relations of contemporary English*. Tesis doctoral. University of London.
- Winter, E.O. 1979. "Replacement as a fundamental function of the sentence in context". *Forum Linguisticum* 4(2): 95-133.
- Winter, E.O. y M. Hoey. 1986. "Clause relations and the writer's communicative task". En B. Couture (ed). *Functional Approaches to Writing Research Perspectives*. Londres: Frances Pinter. 120-141.
- Wotjak, G. 1998. "El significado léxico como transmisor/puente entre el conocimiento enciclopédico y la comunicación lingüística". Conferencia ofrecida en el Curso *Aspectos Cognitivos de la Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Wright, S.E. 1999. "Basic principles of terminology management". Conferencia ofrecida en el Curso *Gestión de Terminología y Traducción* del Centro Mediterráneo de la Universidad de Granada.
- Wunderlich, D. 1972. "Zur Konventionalität von Sprechhandlungen". En D. Wunderlich (ed). *Linguistisch Pragmatik*. Frankfurt: Athenäum. 11-58.
- Yang, H. 1985. "The JDEST computer corpus of texts in English for Science and Technology". *ICAME News* 9: 24-25.
- Zanettin, F. 1994. "Parallel words: designing a bilingual database for translation activities". En A. Wilson y A. McEnery (eds). *Corpora in Language Education and Research: A Selection of Papers from TALC 94*. Lancaster: Lancaster University. 99-111.

- Zanettin, F. 1996. "Using newspapers and computers in the classroom". En S. Botley *et al.* (eds). 1996. *Proceedings of Teaching and Language Corpora (UCREL Technical Papers 9)*. Lancaster: University of Lancaster.
- Zhu, Q. 1989. "A quantitative look at the Guangzhou Petroleum English Corpus". *ICAME Journal* 13: 28-38.
- Zydatiss, W. 1983. "Text typologies and translation". *The Incorporated Linguist* 22(4): 212-221.

OBRAS LEXICOGRÁFICAS Y TERMINOGRÁFICAS UTILIZADAS

- Altman, R. y M.J. Sarg. 1991. *The Cancer Dictionary*. Nueva York: Facts on file.
- Collins Cobuild English Language Dictionary*. 1987. Londres y Glasgow: Harper Collins.
- CLAVE Diccionario de Uso del Español Actual*. 1996. Madrid: S.M.
- Concise Oxford Dictionary*. 1993. Oxford: Oxford University Press.
- Diccionario Enciclopédico Ilustrado de Medicina Dorland*. 1985. Buenos Aires: Internamericana, 26 edic.
- DTCM: Diccionario Terminológico de Ciencias Médicas*. 1992. Barcelona: Masson, 13 edic.
- EURODICAUTOM*. <http://www2.echo.lu/edic/>
- Frenae, A.C. y R.M. Mahoney (eds). 1997. *Understanding Medical Terminology*. Boston: Mc-Graw Hill. 1997.
- MEDLINE*. <http://www4.ncbi.nlm.nih.gov/PubMed>
- Medicine on-line Cancer glossary*. [Disponible en <http://www.meds.com/glossary.html>].
- National Cancer Institute's Cancer glossary*. [Disponible en <http://cancernet.nci.nih.gov/dictionary.html>].
- Oxford Advanced Learner's Dictionary*. 1995. Oxford: Oxford University Press.
- RAE*. 1992. *Diccionario de la lengua española*. Madrid: Espasa-Calpé, 21 edic. añola. 1992.
- Roget's Thesaurus of English Words and Phrases*. 1982. Harmondsworth: Penguin books.

PUBLICACIONES SOBRE EL CÁNCER

Publicaciones especializadas

En inglés:

British Medical Journal 1998. Vol. 317: 245-252.

British Medical Journal 1998. Vol. 317: 771-775.

British Medical Journal 1998. Vol. 317: 346.

British Medical Journal 1998. Vol. 317: 771-775.

British Medical Journal 1998. Vol. 315: 980-988.

British Medical Journal 1997. Vol. 314: 1860.

British Medical Journal 1997. Vol. 314: 974.

CA. A Cancer Journal for Clinicians 1998. Vol. 48:167-176

CA. A Cancer Journal for Clinicians 1998. Vol. 48:164-166

CA. A Cancer Journal for Clinicians 1998. Vol. 48:167-176

CA. A Cancer Journal for Clinicians 1997. Vol. 47:161-170

CA. A Cancer Journal for Clinicians 1996. Vol. 46:185-192

CA. A Cancer Journal for Clinicians 1996. Vol. 46:303-320

Cancer (American Cancer Society) 1998. Vol. 83:1335-43.

Cancer (American Cancer Society) 1998. Vol. 82:301-8.

Cancer (American Cancer Society) 1997. Vol. 79:1527-32.

Cancer (American Cancer Society) 1997. Vol. 79:1329-37.

Cancer (American Cancer Society) 1997. Vol. 79:468-73.

Cancer (American Cancer Society) 1997. Vol. 79:35-44.

Cancer (American Cancer Society) 1997. Vol. 80:1134-40.

Lancet 1998. Vol. 352, nº 9124

Lancet 1998. Vol. 351, nº 9111

Lancet 1997. Vol. 350, nº 9081

Medline

The New England Journal of Medicine. Vol. 339, nº 1. 2 de julio de 1998

The New England Journal of Medicine. Vol. 337, nº 17 23 de octubre de 1997

The New England Journal of Medicine. Vol. 330, nº 24, 16 de junio de 1994.

The New England Journal of Medicine. Vol. 328, nº 19, 13 de mayo de 1993.

En español:

Archivos Bronconeumológicos 1998. Vol. 34: 99-101.

Asociación española de endoscopia respiratoria

Medicina Clínica 1997. Vol. 109: 629-631.

Medicina Clínica 1997. Vol. 108: 201-206.

Medicina Clínica 1997. Vol. 108: 121-127.

Medicina Clínica 1997. Vol. 108: 81-86.

Neoplasia. Nº 1. Enero - Febrero 1998.

Neoplasia. Nº 3. Mayo-Junio 1998.

Neoplasia. Nº 4. Julio-Agosto 1998.

Neoplasia. Nº 5. Septiembre-October 1998

Neoplasia. Nº 6. Noviembre-Diciembre 1998

Revisiones en Cancer 1998. Vol. 12, nº 1: 3-35.

Revisiones en Cancer 1998. Vol. 12, nº 1: 36-41.

Bases de datos bibliográficas

CANCERLIT

Medline

Manuales

En inglés:

DeVita, V. T., S. Hellman y S. A. Rosenberg, (eds). 1997. *Cancer: Principles and Practice of Oncology*. Lippincott Williams & Wilkins, 5ª edic.

The Merck Manual of Diagnosis and Therapy. 1997. [Disponible en <http://www.merck.com/pubs/mmanual>].

En español:

DeVita, V. T., S. Hellman y S. A. Rosenberg, (eds). 2000 (1997). *Cancer. Principios y Práctica de Oncología* [Trad de la obra original *Cancer: Principles and Practice of Oncology*]. Madrid: Ediciones Arán y Editorial Médica Panamericana.

Díaz-Rubio, E. y V. Guillén Porta (eds). 1998. *Oncología Médica. Guía de Oncología Médica*. Madrid: Rhône Poulenc Rorer Oncology, 2ª edic.

Manual Merck en español. [Disponible en <http://www.msd.es/mmerck>].

Rodes, J. 1997. "Neoplasias pulmonares". *Medicina Interna*. Tomo II. Barcelona: Masson: 1160-75.

Publicaciones de divulgación especializada y general sobre el cáncer

En inglés

Daily Telegraph. [Disponible en <http://www.telegraph.co.uk>].

Oncolink (CNN Health and Food). [Disponible en <http://www.cnn.com/HEALTH>].

Reuters Health

Scientific American, número monográfico sobre el cáncer, septiembre de 1996. [Disponible en <http://www.sciam.com/0996issue>].

TIME, número monográfico sobre el cáncer, Vol. 151, nº 19. 18 de mayo de 1998.

En español:

Blanco y negro, 14 de junio de 1998.

El Mundo (CD-ROM de 1996 y versión *on-line* en <http://www2.el-mundo.es/salud>)

El Semanal, 25 de octubre de 1998.

Ideal (Suplemento Campus), 4 de diciembre de 1997.

Investigación y Ciencia, monografía sobre el cáncer, noviembre de 1996.

Muy interesante, marzo de 1999.

QUO, monografía sobre el cáncer, enero de 1997.

Enciclopedias en soporte electrónico

Enciclopedia Microsoft Encarta 97

Listas de distribución sobre el cáncer de pulmón

LUNG-ONC

LUNG-ONC@LISTSERV.ACOR.ORG

Sitios sobre cáncer consultados en la red

En inglés

Alcase

<http://www.alcase.org>

American Cancer Society

<http://www3.cancer.org>

American Lung Association

<http://www.lungusa.org/diseases/index2.html>

CancerBacup

<http://www.cancerbacup.org.uk>

CancerHelp UK

<http://medweb.bham.ac.uk/cancerhelp/>

Imperial Cancer Research Fund (ICRF)

<http://www.icnet.uk/>

National Cancer Institute

Cancernet <http://cancernet.nci.nih.gov>

PDQ

En inglés:

<http://cancernet.nci.nih.gov/clinpdq/facts.sp.html>

En español:

<http://cancernet.nci.nih.gov/clinpdq/soa.sp.html>

Website for Patients, Public and Mass Media

http://rex.nci.nih.gov/INTRFCE_GIFS/MASSMED_INTR_DOC.htm

Oncolink: The University of Pennsylvania Cancer Center Resource

<http://oncolink.upenn.edu>

Patrick Connelly-M.D.-Lung Cancer Page

<http://www.erinet.com/fnadoc/lung.htm>

PSL

<http://www.pslgroup.com>

Virtual Hospital

<http://www.vh.org>

En español

Atheneum / Jano, Medicina y Humanidades

<http://www.atheneum.doyma.es>

Diario Médico

<http://www.diariomedico.com/oncologia>

Instituto Madrileño de Oncología

<http://www.ctv.es/imo/primera.htm>

Sarenet

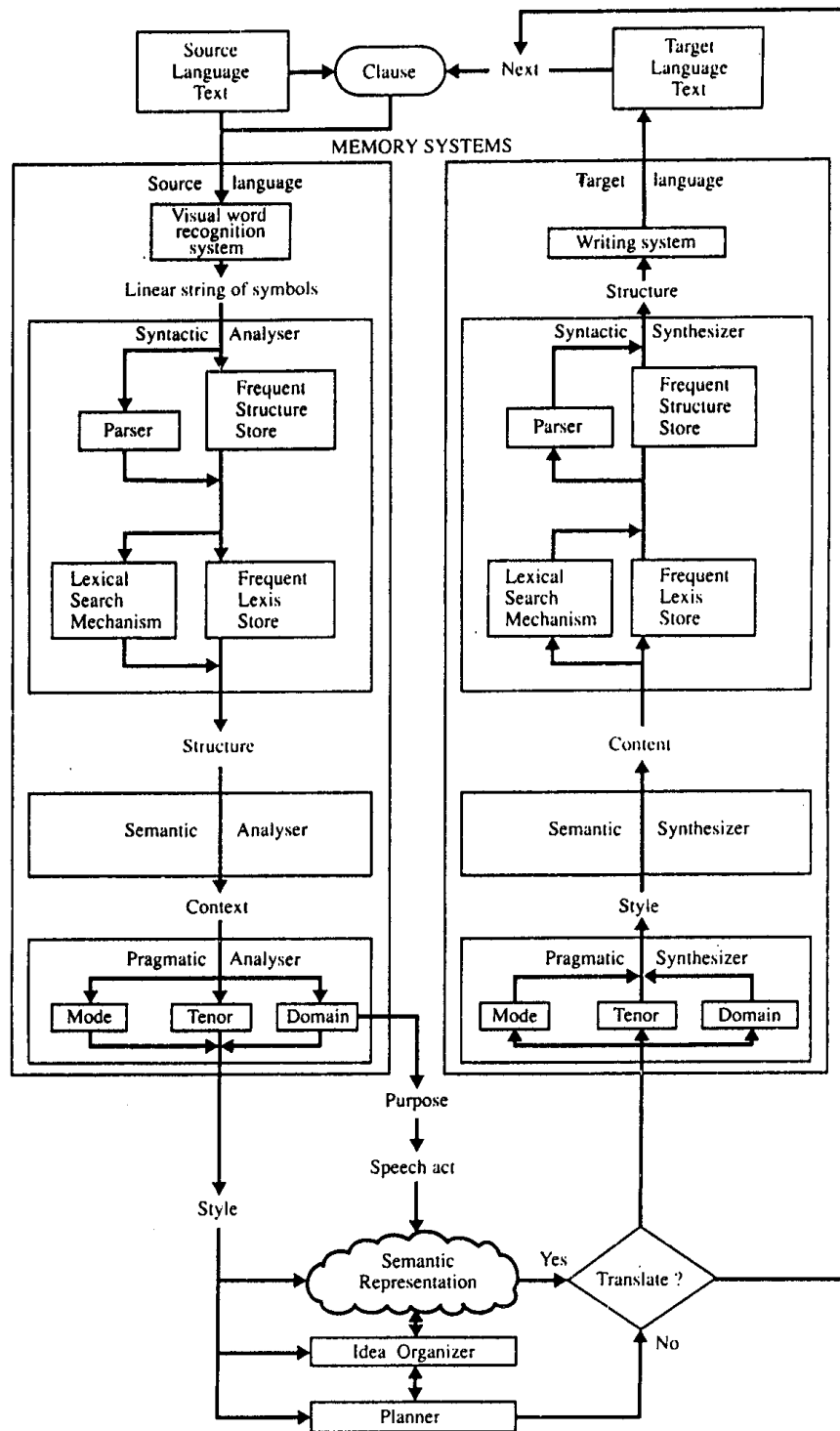
<http://www.ecomedic.com/em>

Sociedad Iberoamericana de Información Científica

<http://www.siicsalud.com>

APÉNDICES

APÉNDICE I: El proceso cognitivo de la traducción (Bell 1991: 59)



APÉNDICE II: Textos analizados

CAPÍTULOS 4 y 6

- Texto 1: Altered Retinoblastoma Protein Expression in Nonsmall Cell Lung Cancer (p.623)
Texto 2: Prevenir el cáncer de cuello de útero (p.628)
Texto 3: Fórmulas prostáticas (p.629)
Texto 4: Role of vindesine in induction chemotherapy in locally-advanced non-small-cell lung cancer (p.630)
Texto 5: El triunfo del taxol (p. 631)
Texto 6: ¿Qué es el cáncer? (p. 632)

CAPÍTULOS 7 y 10

- Texto 7: Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary Chromosome 9 Aberrations (p. 634)
Texto 8: Alteraciones del cromosoma 9 (p. 641)

CAPÍTULOS 11, 12, 13 y 14:

SERVICIO *PDQ* DEL *NATIONAL CANCER INSTITUTE*

Treatment Summaries for Health Professionals

- Texto 9: Malignant mesothelioma (QDT1) (p. 644)
Texto 10: Non-small cell lung cancer (QDT2) (p. 647)
Texto 11: Small cell lung cancer (QDT3) (p. 667)

Treatment Summaries for Patients

- Texto 12: Malignant mesothelioma (QPT1) (p. 674)
Texto 13: Non-small cell lung cancer (QPT2) (p. 678)
Texto 14: Small cell lung cancer (QPT3) (p. 682)

CAPÍTULO 15

SERVICIO *PDQ* DEL *NATIONAL CANCER INSTITUTE*

Tratamiento para pacientes

- Texto 15: Malignant mesothelioma-Mesotelioma maligno (versión bitexto de QDT1)(p. 685)

NOTICIA PARA ESPECIALISTAS PUBLICADA EN *DIARIO MÉDICO*

- Texto 16: El tratamiento neoadyuvante incrementa la supervivencia en el cáncer de pulmón (p. 691)

TEXTO 1

Cancer 79:1329-37, 1997

Copyright © 1997 American Cancer Society. All rights reserved.

Published by John Wiley & Sons, Inc.

Altered Retinoblastoma Protein Expression in Nonsmall Cell Lung Cancer

Its Synergistic Effects with Altered *ras* and p53 Protein Status on Prognosis

Hirotohi Dosaka-Akita, M.D.¹, Shi-Xue Hu, M.D.³, Michihiro Fujino, M.D.¹, Masao Harada, M.D.¹, Ichiro Kinoshita, M.D.¹, Hong-Ji Xu, M.D.³, Noboru Kuzumaki, M.D.², Yoshikazu Kawakami, M.D.¹, William F. Benedict, M.D.³

¹ First Department of Medicine, Hokkaido University School of Medicine, Sapporo, Japan. ² Laboratory of Gene Regulation, Cancer Institute, Hokkaido University School of Medicine, Sapporo, Japan. ³ University of Texas M. D. Anderson Cancer Center, The Woodlands, Texas.

Received August 12, 1996; revision received December 2, 1996; accepted December 2, 1996.

ABSTRACT

BACKGROUND. Inactivation of the retinoblastoma (Rb) gene has been documented in various types of cancer, including lung cancer. Alterations of the p53 and *ras* genes are also common features in the molecular biology of lung carcinoma, and the authors of this article have reported previously on the prognostic significance of both of them. In the present study, the authors evaluated the prognostic significance of the loss of Rb protein expression alone, then performed a combined analysis of Rb protein and *ras* p21 status (Rb/*ras*) as well as an analysis of Rb and p53 protein status (Rb/p53) in patients with nonsmall cell lung cancer (NSCLC).

METHODS. Ninety-one patients with NSCLC underwent potentially curative resection between 1977 and 1988, 65 of whom received postoperative combination chemotherapy. Tumor specimens were analyzed for Rb protein expression by immunohistochemistry. Univariate and multivariate analyses were performed to assess the association between Rb protein expression and survival.

RESULTS. Nineteen (21%) of the 91 NSCLCs showed negative Rb protein expression. Positive or negative Rb protein expression (Rb+ or Rb-) as an individual factor was not statistically correlated with survival or prognosis in this cohort of NSCLC patients, although a tendency among Rb- patients to do worse was observed. The authors then combined the Rb protein status with previously studied results of *ras* p21 and p53 protein expression in the same tumor specimens, and compared the prognosis between the individuals with theoretically the best pattern of gene expression in their tumors and those with theoretically the worst pattern of expression, i.e., Rb+/*ras*- versus Rb-/*ras*+ and Rb+/p53- versus Rb-/p53+. In patients with adenocarcinoma, those with Rb-/*ras*+ tumors survived for a significantly shorter period after surgery (13% 5-year survival) than those with Rb+/*ras*- tumors (82% 5-year survival) ($P = 0.01$). Similarly, patients with Rb-/p53+ tumors survived for a significantly shorter period (20% 5-year survival) compared with those who had Rb+/p53- tumors (73% 5-year survival) ($P = 0.008$). Rb/*ras* status was a significant prognostic factor ($P = 0.02$ by univariate analysis, $P = 0.048$ by multivariate analysis), and Rb/p53 status tended to be significant as a prognostic factor ($P = 0.04$ by univariate analysis, $P = 0.08$ by multivariate analysis). In patients with squamous cell carcinoma, neither Rb/*ras* nor Rb/p53 status was a significant prognostic factor in this cohort.

CONCLUSIONS. These results suggest that combined immunohistochemical analyses of Rb and *ras* p21 proteins and of Rb and p53 proteins may indicate their potentially synergistic effects on survival and prognosis. These analyses may also be useful for stratifying patients with adenocarcinoma of the lung into different prognostic groups and identifying populations with different risks of recurrence. Larger prospective studies with Stage I NSCLC patients are necessary to confirm the current findings.

KEYWORDS

retinoblastoma protein, *ras* p21 protein, p53 protein, nonsmall cell lung cancer, prognosis.

Lung cancer is one of the leading causes of cancer-related death throughout the world. Despite major advances in cancer treatment in the past two decades, the prognosis of patients with lung cancer has improved only minimally.¹ A new understanding of the molecular pathogenesis and multistep carcinogenesis of lung cancer has emerged from research advances in the field of molecular biology.² Increased knowledge of the biologic role of genetic changes and other aberrations in tumorigenesis provokes an intriguing search for clinical applications of these alterations.³⁻⁵ It may be possible to predict the responses of individual patients with lung cancer to treatment and to identify patients with a good or poor prognosis using molecular biologic alterations as clinical biomarkers.⁶⁻⁸

Inactivation of the retinoblastoma (Rb) gene has been documented in various types of cancers, including lung cancer.⁹⁻¹⁴ It has also been associated with aggressive tumor behavior and poor clinical outcome in specific types of cancer, including sarcomas as well as bladder and lung cancers.¹⁵⁻¹⁸ Loss of Rb protein expression was shown to occur frequently in primary nonsmall cell lung cancer (NSCLC) by both Western blot and immunohistochemical analyses.¹³ Alterations of the p53 gene have been shown to be among the most common molecular biologic changes in lung cancer.¹⁹⁻²¹ Several studies have yielded conflicting results regarding the prognostic relevance of alterations in the p53 gene itself or in p53 protein expression.²²⁻²⁷ Abnormalities in the *ras* family of oncogenes, including point mutations and overexpression, are also common features in the molecular biology of lung cancer.^{28,29}

In previous studies,^{26,29} we performed immunohistochemical analysis of *ras* p21 and p53 protein expression in the 96 NSCLC tumor specimens of all stages, including Stages IIIb and IV. In these studies, we found that patients with *ras* p21 positive (*ras*+) or p53 positive (p53+) tumors had significantly shorter survival than those with *ras* p21 negative (*ras*-) or p53 negative (p53-) tumors, respectively, that both *ras* p21 and p53 protein status were significant and independent prognostic factors for patients with surgically treated NSCLC, and that combined analysis of *ras* p21 and p53 protein expression could stratify these NSCLC patients into more accurate prognostic groups. In the present study, we extended the number of tumor specimens that were taken from patients who had undergone potentially curative surgery and had been analyzed for *ras* p21 and p53 protein expression to 91, and we included Rb protein expression as an additional potential prognostic factor. We combined the current results of Rb protein expression with those of *ras* p21 and p53 protein expression in the same tumor specimens, and we analyzed the survival and prognosis of NSCLC patients based on Rb, *ras* p21, and p53 protein expression.

In the present study, loss of Rb protein expression was not a statistically significant prognostic variable as an individual factor in the cohort examined. However, our data suggested that the combined analysis of Rb and *ras* p21 or of Rb and p53 protein status could potentially identify adenocarcinoma patients with significantly better or worse prognoses, although this association was not found for patients with squamous cell carcinoma of the lung.

MATERIALS AND METHODS

Tumor Specimens and Survival Data

Tumor specimens from 91 patients with NSCLC who had undergone potentially curative resection at the Hokkaido University Medical Hospital between 1977 and 1988 were analyzed. The patients consisted of 61 men and 30 women (average age at diagnosis, 61.8 years). The histologic classification of the tumor specimens was based on World Health Organization criteria,³⁰ and the specimens included 44 adenocarcinomas, 41 squamous cell carcinomas, four large cell carcinomas, and 2 adenosquamous cell carcinomas. They represented 49 Stage I, 10 Stage II, and 32 Stage IIIa tumors. The postsurgical pathologic TNM stage (pTNM) was determined according to the guidelines of the American Joint Committee on Cancer.³¹ Survival was analyzed for the 89 patients who met the following criteria: (1) survived for more than 3 months after surgery; (2) did not die of causes other than lung cancer within 5 years after surgery; and (3) were followed for more than 3 years after surgery (for patients who remained alive). Two of 91 patients who did not meet the above criteria (because they died within 3 months after surgery) were excluded from the survival analysis. Because all the patients enrolled in the current study were coded, they could not be individually identified. Sixty-five patients received combination chemotherapy as postsurgical treatment. These chemotherapy-treated patients had 34 of the 44 adenocarcinomas, 28 of the 41 squamous cell carcinomas, and 3 of the 6 adenosquamous and large cell carcinomas. Among these 65 were also 41 of the 59 Stage I and II patients and 24 of the 32 Stage III patients. Radiation therapy was not performed before or after surgery for any patients. We had previously reported on 63 of the 91 patients for the prognostic significance of *ras* p21 and p53 protein expression in NSCLC.^{26,29}

Immunohistochemical Analysis

The methods for staining of Rb nuclear protein in paraffin embedded sections have been described previously.^{13,17} Briefly, the highly specific affinity-purified polyclonal anti-Rb antibody, Rb-WL-1,^{32,33} was used. It recognizes both the phosphorylated and unphosphorylated Rb protein and can be completely blocked in the presence of an excess of the immunizing Rb peptide.³² Four investigators (H.D.A., I.K., H.-J.X., and W.F.B.) separately evaluated the Rb nuclear protein staining pattern. At the time of review, none of these investigators were aware of the clinical outcome of the patients, since all the slides had been coded and the clinical outcome was disclosed only after all four investigators were in agreement on the Rb protein status for a given specimen. In all cases, adequate nuclear staining

was obtained in the adjacent normal lung tissues, which represented an internal positive control for the Rb protein staining of each section. A tumor was considered to be Rb positive (Rb+) if the Rb protein staining was heterogeneous, with a portion of the tumor cells showing typical Rb nuclear protein staining. Tumors were scored as Rb negative (Rb-) only if all malignant cells showed no Rb nuclear protein staining and normal stromal cells were Rb+ as an internal control. Two tumors had large areas of malignant cells with no Rb nuclear protein staining that were surrounded by Rb+ stroma cells, whereas in other small tumor areas the Rb staining was heterogeneously positive. These two tumors were considered to have altered Rb protein expression and were included as Rb- tumors.

For the *ras* p21 and p53 protein staining of 63 of 91 tumor specimens, the slides and results that had been previously reported^{26,29} were used for the current study. The remaining 28 tumor specimens were newly analyzed for p53 and *ras* p21 protein expression by immunohistochemical staining; paraffin embedded sections were used, as previously described.^{26,29} Both *ras* p21 and p53 protein staining were scored as positive (+) and negative (-) for the current study, as in one of the previous studies.²⁶ Therefore, the positive staining of *ras* p21 protein in the current study included moderate and strong staining in our first *ras* p21 report.²⁹ Typical staining patterns for Rb, p53, and *ras* p21 proteins are shown in [Figure 1](#).

Statistical Analysis

The correlation of various characteristics of patients and tumors with Rb protein expression was analyzed by the chi square test. Survival curves were estimated using the Kaplan-Meier method,³⁴ and differences in survival distributions were evaluated by the generalized Wilcoxon test.³⁵ For simultaneous adjustment of the effects of other variables on the association between Rb, Rb/*ras*, and Rb/p53 protein status and survival outcome, the Cox proportional hazards general linear model was used.³⁶ All *P* values reported are two-sided. These computations were performed with the SAS program package (SAS Institute, Cary, NC)³⁶ at the Hokkaido University Computer Center.

RESULTS

Immunohistochemical Detection of Rb Protein in NSCLC

Nineteen (21%) of the 91 NSCLCs analyzed showed negative Rb protein expression ([Fig. 1](#)). Rb protein status was analyzed in relation to various clinical and clinicopathologic characteristics ([Table 1](#)). No characteristics except gender ($P = 0.04$) correlated with Rb protein status. In patients with squamous cell carcinoma, no correlation was observed between Rb protein status and the various characteristics examined (data not shown). However, when patients with adenocarcinoma were analyzed, loss of Rb protein expression correlated with lymph node metastasis, or pN status ($P = 0.02$ for pN0 vs. pN1-2) ([Table 1](#)).

Prognostic Significance of Rb Protein Expression in NSCLC

The relationship between Rb protein expression and postsurgical survival was analyzed for 89 patients. When all of the NSCLC patients were evaluated together, patients with Rb+ and Rb- tumors did not show a significant difference in survival (5-year survival rates, 50% and 34%, respectively; $P = 1.0$) ([Fig. 2A](#)). When only patients with squamous cell carcinoma were analyzed, Rb protein status still did not show a significant effect on survival (5-year survival rates, 47% for Rb+ patients and 69% for Rb- patients, respectively; $P = 0.2$). Among adenocarcinoma patients, those with Rb- tumors survived shorter periods than those with Rb+ tumors, although this difference in survival did not reach statistical significance (5-year survival rates, 23% and 59%, respectively; $P = 0.07$) ([Fig. 2B](#)). Univariate analysis of various potential prognostic factors ([Table 2](#)) indicated that Rb protein expression was not a significant prognostic factor for NSCLC patients overall ($P = 0.8$) or for adenocarcinoma ($P = 0.07$) or squamous cell carcinoma patients ($P = 0.2$) individually.

Combined Analysis of Rb and *ras* p21 and of Rb and p53 Protein Expression in NSCLC

The significance of *ras* p21 and p53 protein status as potential prognostic markers was analyzed to confirm our previous studies of patients with all stages of NSCLC, which included, unlike our current cohort, Stages IIIb and IV.^{26,29} Both *ras* p21 and p53 protein status were significant prognostic factors ($P < 0.01$ and $P = 0.04$, respectively, by univariate analysis) for the cohort overall ([Table 2](#)). Multivariate analysis was performed to adjust for differences and influences of the other prognostic factors, pStage and chemotherapy, on the results. Again, both *ras* p21 and p53 protein status were independent prognostic factors for these patients overall ($P < 0.01$ and $P = 0.045$, respectively; Models 1 and 2, [Table 3](#)), confirming results from previous studies.^{26,29} When examined by histologic type for squamous cell carcinoma, *ras* p21 protein status was a significant prognostic factor by both univariate and multivariate analyses ($P < 0.01$), although p53 protein status was not a significant prognostic factor ($P = 0.2$ by univariate analysis). In contrast, when patients with adenocarcinoma were studied, neither *ras* p21 nor p53 protein status was a significant predictor for prognosis ($P = 0.1$ and $P = 0.2$, respectively, by univariate analysis) or for survival ($P = 0.1$ and $P = 0.09$, respectively, by the generalized Wilcoxon test) ([Figs. 3A and 4A](#)).

The results of Rb protein expression was then combined with *ras* p21 or p53 protein status, and the combinations were analyzed as prognostic markers in patients with theoretically the best or worst pattern of expression, i.e., Rb+/*ras*- versus Rb-/*ras*+ or Rb+/p53- versus Rb-/p53+, respectively. For the overall cohort, patients with Rb-/*ras*+ tumors had a significantly lower 5-year survival rate than those with Rb+/*ras*- tumors (21% and 75%, respectively, $P = 0.03$, data not shown). Patients with Rb-/p53+ tumors also showed a lower 5-year survival rate than those with Rb+/p53- tumors (28% and 63%, respectively), although this did not reach statistical significance ($P = 0.1$, data not shown). However, when only patients with adenocarcinoma were analyzed, individuals with Rb-/*ras*+ or Rb-/p53+ tumors had significantly lower 5-year survival rates than those with Rb+/*ras*- or Rb+/p53- tumors, respectively (13% and 82%, $P = 0.01$ for Rb/*ras*, Fig. 3B; 20% and 73%, $P = 0.008$ for Rb/p53, Fig. 4B). In contrast, there were no statistical differences in 5-year survival rates for patients with squamous cell carcinoma using the same comparisons (50% and 75%, $P = 0.5$ for Rb/*ras*; 38% and 56%, $P = 0.8$ for Rb/p53), although the same trends were observed.

Finally, univariate and multivariate analyses using the Cox proportional hazards general linear model were performed to assess the prognostic significance of Rb/*ras* (Rb+/*ras*- vs. Rb-/*ras*+) and Rb/p53 (Rb+/p53- vs. Rb-/p53+) status before and after adjusting for differences in other variables (Tables 2 and 3). For the overall cohort, Rb/*ras* status tended to be a significant prognostic factor ($P = 0.03$ by univariate analysis; $P = 0.07$ by multivariate analysis in Model 3, Table 3), although Rb/p53 status did not reach statistical significance as a prognostic factor by univariate analysis ($P = 0.1$). When patients with adenocarcinoma were analyzed, Rb/*ras* status was a significant prognostic factor ($P = 0.02$ by univariate analysis; $P = 0.048$ by multivariate analysis in Model 1, Table 3), and Rb/p53 status tended to be significant as a prognostic factor ($P = 0.04$ by univariate analysis; $P = 0.08$ by multivariate analysis in Model 2, Table 3). Wide confidence intervals for Rb/*ras* and Rb/p53 status for patients with adenocarcinoma were observed; these might be due to the smaller numbers of patients in each group. For patients with squamous cell carcinoma, neither Rb/*ras* nor Rb/p53 status was a significant prognostic factor ($P = 0.6$ and $P = 1.0$, respectively, by univariate analysis).

DISCUSSION

Contrary to the results of our previous study,¹⁸ Rb protein status was not significant as an individual factor for survival and prognosis in the current NSCLC cohort, although this study included Stage IIIa patients, whereas the other¹⁸ did not. Nevertheless, when both Rb and *ras* p21 protein expression in the same tumors were considered, patients whose tumors showed both Rb and *ras* p21 protein alterations had a markedly shorter survival than those having tumors without measurable changes in these two proteins; this was the case for the overall NSCLC cohort and particularly for patients with adenocarcinoma. Similarly, patients whose tumors exhibited both Rb and p53 protein alterations had shorter survival than those having tumors without such changes; this was the case for the overall NSCLC cohort, and again especially for those with adenocarcinoma.

Both *ras* p21 and p53 protein status were significant and independent prognostic factors for the overall NSCLC cohort, confirming the results of previous studies.^{26,29} Postsurgical chemotherapy was also a significant prognostic factor in the univariate analysis and was an independent prognostic factor in one of three multivariate analyses. However, it was not a significant prognostic factor by univariate analysis either for patients with adenocarcinoma or for those with squamous cell carcinoma when the two groups were analyzed separately. These data suggest that postsurgical chemotherapy may have only limited importance for the prognosis of NSCLC patients. Both *ras* p21 and p53 protein status were independent as prognostic factors from postsurgical chemotherapy and pStage by multivariate analysis, indicating that the effects on prognosis of altered *ras* p21 and p53 proteins were not dependent on the chemotherapy administered or on the stage of the disease.

In the cohort of patients with adenocarcinoma, however, Rb, *ras* p21, and p53 protein status, when examined individually, were not significant predictors for survival or prognosis, indicating their limited value as prognostic markers. However, Rb studied simultaneously with *ras* p21 and Rb studied simultaneously with p53 protein status became potential predictors of shortened survival for patients with adenocarcinoma. Therefore, alterations of Rb and *ras* p21 or of Rb and p53 proteins may have synergistic effects on survival and prognosis for lung carcinoma patients, especially for those with adenocarcinoma. However, the multivariate analysis showed borderline significance for Rb/*ras* ($P = 0.048$) and Rb/p53 status ($P = 0.08$) as independent prognostic factors for adenocarcinoma patients. These observations might have resulted from the smaller numbers of patients in each group. Larger studies that include patients with more homogeneous stages of disease are required to confirm whether the combinations of Rb and *ras* p21 or of Rb and p53 protein status can predict clinical outcome. It is noteworthy that similar synergistic effects on prognosis were found when Rb and p53 protein alterations were analyzed together in two other cohorts of Stage I and II NSCLC.^{18,37}

Combined analyses of biologically important factors in addition to Rb protein have also been reported to be useful as prognostic markers in NSCLC in several studies, including ours.^{26,38,39} We reported previously that the combined analysis of *ras* p21 and p53 protein expression could divide NSCLC patients into more accurate prognostic groups in a cohort that was basically the same in composition as the current cohort, but smaller.²⁶ We also reported the importance of a combination of *ras* p21 protein status and nuclear DNA ploidy as a biologic marker for the survival of resected NSCLC patients.³⁸ Kern et al. also reported additive effects of *c-erb* B2 protein expression and *K-ras* mutations on the survival of patients with adenocarcinoma of the lung.³⁹ They found that patients with

adenocarcinoma of the lung that expressed *c-erb* B2 protein and harbored an activating *K-ras* point mutation had significantly shorter survival than patients whose tumor had none or only one of these two alterations.

The alterations of Rb protein expression were correlated with pN classification in the patients with adenocarcinoma. This finding is also consistent with our previous study,¹³ which showed a significant difference in the occurrence of altered Rb protein expression between patients with Stage I/II disease and Stage III/IV disease, since pN classification is the most important determinant of the stage of NSCLC.³¹ A recently cloned tumor suppressor gene, *CDKN2/MTS1*,^{40,41} the gene product of which (p16^{INK4}) is known to function as a cell-cycle inhibitor through a Rb protein-dependent pathway,⁴² was also shown to be frequently inactivated in NSCLC.^{43,44} It is noteworthy that the inactivation of *CDKN2/MTS1* was more frequently observed in Stage III/IV disease than in Stage I/II disease,⁴⁴ like the alterations of Rb protein. From these studies of Rb protein and *CDKN2/MTS1* in NSCLC, including the current study, inactivation of the p16/Rb pathway may be involved in the clinical progression of NSCLC, especially in lymph node metastasis. Lymph node metastasis is accepted as the most important determinant of the stage of surgically treated NSCLC³¹ and as a significant prognostic factor in such cases.⁷ Taken together, the inactivation of the p16/Rb pathway may make the clinical outcome worse by having a role in lymph node metastasis and clinical progression. It would therefore be interesting to see the relationship between p16^{INK4} inactivation and prognosis in NSCLC patients.

In conclusion, it may be possible to identify patients with lung cancer, especially adenocarcinoma, with better or worse prognosis by using alterations of Rb protein expression in combination with *ras* p21 or p53 protein changes as combined clinical markers. If the current findings are confirmed in larger prospective studies, including especially those with Stage I NSCLC, then combined immunohistochemical analysis of Rb and *ras* p21 proteins or of Rb and p53 proteins can be used to stratify patients with lung cancer into different prognostic groups and to identify populations with different risks of recurrence.

ONCOLOGIA

Prevenir el cáncer de cuello de útero

Experimentan un vacuna contra este tumor

EL MUNDO

Un tratamiento experimental podría llegar a salvar a miles de mujeres de tener que tratarse de un cáncer de cuello de útero. Investigadores de Baltimore acaban de usar en ratones una vacuna contra este tipo de tumor y la experiencia ha resultado exitosa.

El estudio, publicado en la revista *Cancer Research*, revela que los científicos utilizaron una proteína, llamada E7, que juega un papel fundamental en la progresión del cáncer de cuello de útero. Esta proteína se encuentra en la superficie del papiloma virus humano, que causa el 90% de este tipo de tumores.

Los autores del trabajo utilizaron la vacuna antes y después de haber inyectado células cancerosas en los ratones. El 80% de los roedores en los que se usó la vacuna de forma preventiva, tres meses después de inocular las células malignas se encontraron libres de tumores. Mientras que los expertos comprobaron que aquéllos que fueron tratados una semana después de recibir la inyección de células malignas se curó de los tumores pequeños que habían desarrollado. El siguiente paso será probar la efectividad de esta vacuna en humanos.

UROLOGIA**Fórmulas prostáticas**

Sigue la polémica sobre la terapia del cáncer de la glándula

EL MUNDO

La polémica continúa. La publicación en el último número del JAMA de un trabajo de la Clínica Mayo (EEUU), en el que se aconseja qué es lo que hay que hacer con los cánceres de próstata, añade más leña al fuego de la discusión que hace años existe sobre la terapia del tumor maligno más frecuente en el hombre.

Especialistas de la prestigiosa Institución estadounidense, revisaron las últimas 337 prostatectomías radicales que habían realizado en otros tantos enfermos afectados de cáncer de la glándula para contrastar la edad de los pacientes, el tamaño del tumor, la malignidad del mismo y el índice de crecimiento anual que habían tenido. De los resultados de este examen, los autores del trabajo concluyen que ante un cáncer de próstata se puede utilizar una fórmula para saber cómo tratarlo.

Los parámetros que -según los expertos de Rochester- habría que tener

siempre en cuenta antes de tomar decisiones prostáticas demasiado radicales serían: la edad del paciente (menor de 70 años), el volumen tumoral y las posibilidades de crecimiento del mismo.

De hecho, cuando los especialistas revisaron su casuística concluyeron que en el 49% de las operaciones realizadas, el resultado del tratamiento hubiera sido el mismo si los cirujanos no hubieran actuado. Los investigadores recomiendan prudencia basándose en la biología tan particular que tienen los tumores prostáticos. Aunque el 40% de los varones mayores de 50 años tiene cáncer de próstata, sólo el 8% de todos los tumores acabará produciendo un problema al paciente. Por eso es necesario encontrar métodos y fórmulas que concreten mucho más el qué hacer con los enfermos. Dado que la prostatectomía radical es una operación con bastantes problemas secundarios, hay que llevarla a cabo sólo cuando vaya a beneficiar claramente al enfermo.

TEXTO 4

CANCERLIT Topic Searches

CITATIONS AND ABSTRACTS FROM THE NCI'S CANCERLIT DATABASE

CHEMOTHERAPY FOR NONSMALL CELL LUNG CANCER - April 1998

Role of vindesine in induction chemotherapy in locally-advanced non-small-cell lung cancer.

Unique Identifier: 98158593

Author: Fischer JR; Drings P

Source: J Cancer Res Clin Oncol 1998;124(1):4-9

Address: Department of Medical Oncology, Thoraxklinik der LVA Baden, Heidelberg, Germany.

Abstract:

Locally advanced non-small-cell lung cancer (NSCLC) in most cases is not curable at the present time. Owing to the local extent of the tumor, the rate of complete resections is low and, therefore, survival in these patients is poor. For this reason, induction chemotherapy is being *investigated* in patients expected to have a poor prognosis after standard surgery and radiotherapy. The rationale for induction chemotherapy is to increase the rate of complete resections and achieve early elimination of micrometastases. Clinical *investigations* have reported an improvement of survival in stage III NSCLC after induction chemotherapy by using different combinations of cytotoxic drugs. Vindesine ranks among the most active single agents in this disease and has been part of a number of combination regimens in induction chemotherapy. The combination of mitomycin, vindesine or vinblastine and cisplatin has produced encouraging results in several *studies*, indicating a possible improvement of survival in stage III NSCLC, although its superiority to other combinations yet has to be demonstrated. (32 Refs)

ONCOLOGIA

El triunfo del taxol

El fármaco pasa a primera línea en el tratamiento del cáncer ovárico

JOSE LUIS DE LA SERNA

Ha dejado de ser un segundón. El taxol (paclitaxel en su nombre genérico), el avance quimioterápico frente al cáncer más importante de la última década, ha adquirido hoy la categoría de producto oncológico de primera línea. Con la publicación en el número de este jueves del New England Journal of Medicine de un trabajo en el que se demuestra la superioridad de una combinación de taxol y cisplatino sobre la quimioterapia estándar del cáncer de ovario, los oncólogos tendrán que pensar en él desde el principio cuando se enfrenten a un tumor maligno de la gónada.

En el trabajo, realizado con el patrocinio de los Institutos Nacionales de la Salud de EEUU, se comprobó el resultado de tratar inicialmente a 386 mujeres con cáncer del ovario con dos tipos de quimioterapia diferentes. La mitad de ellas -elegidas de forma aleatoria- recibió cisplatino y ciclofosfamida y la otra mitad cisplatino y taxol. Los resultados al cabo de cuatro años de seguimiento clínico fueron significativamente mejores -tanto en supervivencia como en número de recidivas- para el grupo que recibió inicialmente el taxol que para las féminas tratadas de forma estándar.

Toxicidad

Sin embargo, la toxicidad secundaria al tratamiento fue superior en las mujeres tratadas con taxol y cisplatino, que en

aquellas que formaban parte del grupo control. No obstante, los investigadores aseguran que los efectos secundarios, aunque serios, no fueron excesivamente peligrosos. En el editorial que acompaña al artículo, escrito por el doctor Jan Neijt, de la Universidad de Utrech en Holanda, se pone énfasis en que el taxol debería ser considerado de entrada en el tratamiento de todos los cánceres de ovario.

Este tipo de tumor, sin ser excesivamente frecuente, es el más común de todos los cánceres ginecológicos. Tiene, por desgracia, una mortalidad elevada, aunque su tratamiento ha mejorado mucho en los últimos años. Con cirugía y quimioterapia, buena parte de las afectadas responden bien a la medicación. Sin embargo, el tumor suele recidivar con gran frecuencia por lo que su mortalidad a los 10 años de haberse presentado es todavía muy elevada, en torno al 80%. La posibilidad de aumentar la supervivencia de este tipo de cáncer usando taxol es sin duda una buena noticia con la que los oncólogos empiezan el año.

Por otra parte, el paclitaxel ha ampliado sus acciones y éstas no sólo estarán limitadas al tratamiento del tumor del ovario sino que serán beneficiosas también en el de mama y el de pulmón. En uno de ellos -el de mama- se utiliza como producto de segunda línea con buenos resultados y en el otro -el de pulmón- su uso es aún experimental pero todos los datos disponibles indican que prometedor.



Servicio producido por **Sarenet** Información mantenida por **Pulso100**. Patrocinado por **Sanitas**

¿QUE ES EL CANCER?

DESCRIPCION

El cáncer no es una sola enfermedad sino un grupo de más de 200 enfermedades distintas en las que se produce un crecimiento anormal de las células, hasta convertirse en masas de tejidos llamados tumores. Hay dos tipos de tumores: benignos o no cancerosos y malignos o cancerosos.

Los tumores benignos tienen seis características principales:

- Sólo crecen hasta un determinado tamaño.
- Normalmente no crecen muy rápido.
- No destruyen células normales.
- No se propagan al tejido que les rodea.
- Normalmente no producen efectos secundarios graves.
- Por lo general crecen de una manera ordenada.

Los tumores malignos se conocen por su capacidad para invadir y destruir tejidos y órganos, tanto los que están cerca como los que están lejos del tumor original. La muerte se produce cuando la propagación del cáncer daña los tejidos y los órganos vitales, de tal manera que no pueden funcionar.

Las células del cáncer atacan el tejido sano y nunca dejan de multiplicarse. El cáncer tiene un comportamiento distinto en cada persona, según su tipo. Puede darse a cualquier edad, pero es más probable que afecte a personas de edad avanzada; por lo general a partir de los 55 años. El cáncer también puede presentarse en niños, y de hecho, es la segunda causa principal de muerte en niños de edades comprendidas entre 1 y 15 años.

El cáncer puede ser causado por causas externas al cuerpo, o por causas internas.

Los factores externos que pueden causar el cáncer incluyen el estar expuesto a determinados productos químicos como el benceno y el asbesto (amiante). Los pintores, los fabricantes de neumáticos, los destiladores y los fabricantes de zapatos, están expuestos a menudo al benceno. Los trabajadores de minas, aislamiento y astilleros están expuestos a menudo al asbesto. Otros factores ambientales que causan cáncer incluyen:

- Exposición a agentes contaminantes ambientales, como los gases del escape del automóvil.
- Exposición a las radiaciones del sol.
- Exposición a niveles altos de rayos X.
- Exposición a radiaciones Electromagnéticas
- Dieta con gran cantidad de grasas y poca fibra.
- Consumo de tabaco.
- Abuso de las bebidas alcohólicas o de determinadas drogas.

Los factores internos que pueden causar cáncer incluyen la obesidad, las infecciones causadas por virus, como la hepatitis B crónica y un historial familiar en el que exista el cáncer.

Las posibilidades de supervivencia al cáncer dependen del lugar del cuerpo en que se encuentre el cáncer y de las clases de tratamiento utilizadas. Hay cinco formas principales de tratar el cáncer:

- Cirugía
- Radioterapia
- Quimioterapia
- Terapia hormonal
- Terapia biológica.

Para el tratamiento del cáncer, el médico puede utilizar un solo método una combinación de varios.

VER TAMBIEN:

- [Cáncer de Mama, Cirugía](#)
- [Cáncer de Testículos](#)
- [Tumores de Páncreas](#)
- [Cáncer colorectal](#)
- [Cáncer de cuello de uterino](#)
- [Cancer de mama](#)
- [Melanoma](#)
- [Cáncer de pulmón de células grandes](#)
- [Cáncer de pulmón de células pequeñas](#)
- [Cáncer laringe](#)
- [Dolor en el cáncer](#)
- [Linfoma de Hodgkin en adultos](#)

Ultima actualización : Enero 98

<http://www.ecomedic.com/em>

TEXTO 7

Cancer 79:1527-32, 1997

Copyright © 1997 [American Cancer Society](#). All rights reserved.

Published by [John Wiley & Sons, Inc.](#)

Lung Carcinoma Patients with a Family History of Cancer and Lymphocyte Primary Chromosome 9 Aberrations

Xifeng Wu, M.D., Ph.D.^{1,2}, Bhavana J. Dave, Ph.D.³, Hong Jiang, M.S.¹, Sen Pathak, Ph.D.³, Margaret R. Spitz, M.D., M.P.H.¹

¹ Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

² Southwestern Center for Occupational and Environmental Health, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas. ³ Department of Cell Biology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Received September 12, 1996; revision received December 4, 1996; accepted December 26, 1996.

ABSTRACT

BACKGROUND. Deletion of chromosome 9p has been reported in numerous tumor types. The authors demonstrated in an earlier study that spontaneous chromosome aberrations on chromosome 9 in peripheral blood lymphocytes (PBLs) were a significant risk predictor for lung carcinoma.

METHODS. The current study evaluated the relationship between self-reported family history of cancer and spontaneous chromosome aberrations in the PBLs of 97 previously untreated lung carcinoma patients. The authors' hypothesis was that individuals exhibiting specific chromosome aberrations might have inherited genetic instability and thus be more likely to report a family history of cancer. For each individual, a personal interview was conducted to construct a detailed family history, and 100 metaphases from PBLs were analyzed for spontaneous aberrations by G-banding.

RESULTS. The patients reported having 829 first-degree relatives, including 74 (8.9%) with cancer. A significantly elevated odds ratio (OR) of 2.7 was noted in smokers for having chromosome 9 aberrations and a first-degree relative with cancer. When the family history of cancer was dichotomized into lung carcinoma or other cancers, the OR associated with chromosomal aberrations was 8.5 for lung carcinoma but only 2.3 for other cancers. In addition to chromosome 9 aberrations, other spontaneous chromosome aberrations and family history of cancer were also evaluated, but no associations were found. There were no associations between age, gender, ethnicity, or smoking status and the chromosome 9 aberration profile.

CONCLUSIONS. The findings of this study suggest that chromosome 9 aberrations may be a marker of cancer susceptibility and may be associated with familial aggregation of cancer.

KEYWORDS

chromosome 9, lymphocytes, family history of cancer, lung carcinoma.

Carcinoma of the lung is an excellent model for studying genetic-environmental interactions. The incidence of lung carcinoma is dependent not only on the dose, duration, form, and type of tobacco smoked but also host susceptibility factors. In this molecular epidemiologic case-control study of lung carcinoma in two minority populations, the authors have integrated epidemiologic data with a panel of susceptibility markers. One of the study objectives was to examine primary chromosome changes in peripheral blood lymphocytes (PBLs) as a predictor of lung carcinoma risk.

Screening of PBLs for cytogenetic defects has proven useful in studying other cancers.¹⁻⁹ The nonrandom chromosome 13 anomalies in retinoblastoma and chromosome 11 anomalies in Wilms' tumor were first identified in the lymphocyte cultures of children with morphological birth defects and mental retardation.¹⁻⁴ Later, it became apparent that similar specific cytogenetic anomalies were present in the tumor cells of children with retinoblastoma and Wilms' tumor who were otherwise phenotypically normal. It is well established that abnormalities of chromosome 3p are a primary defect in human renal cell carcinoma,⁶ and they have also been reported in the lymphocyte cultures of certain renal cell carcinoma patients.⁷ Similar abnormalities have been observed in the lymphocyte cultures of patients with adenomatous polyps and colorectal carcinomas, which have specific chromosome 5 defects.⁸ Studies of the lymphocyte cultures of patients with breast carcinoma (both familial and sporadic) and their asymptomatic relatives have revealed chromosome 1q rearrangement, a defect previously reported in a number of breast carcinoma cell lines and freshly harvested pleural effusions.⁹

Previously, the authors demonstrated by univariate analysis that the odds ratio (OR) was significantly higher for spontaneous chromosome 7, 9, 12, and 21 aberrations in the PBLs of lung carcinoma cases than of controls.¹⁰ However, only aberrations in chromosomes 9 and 21 were associated with significantly higher risk in lung carcinoma cases than in controls after adjustment by age, gender, ethnicity, and pack-years smoked. Of 97 cases, 43 (44.3%) exhibited chromosome 9 aberrations, compared with 17 (23.0%) of 74 controls. The hot spots on chromosome 9 were located at 9p13-9q13. The multiple tumor suppressor MSTI (p16) and MST2 (p15) genes are located at 9p21,¹¹ and 9p deletion is frequent in lung carcinomas.¹²⁻¹⁴

Based on this information, it was hypothesized that spontaneous chromosome aberrations reflect genetic instability and that individuals with such aberrations would be more likely to report a family history of cancer than individuals without such aberrations. Therefore, in this study, the authors evaluated the relationship between self-reported family history of cancer and spontaneous chromosome aberrations in PBLs of lung carcinoma patients. Family history of cancer was defined as any cancer reported in the patient's first-degree relatives.

MATERIALS AND METHODS

Study Subjects

The patients included in this report are a subset from a molecular epidemiologic study of lung carcinoma in minority populations (African Americans and Mexican Americans). The 97 patients studied had newly diagnosed, histologically confirmed lung carcinoma, had not been previously treated with radiotherapy or chemotherapy, and reported themselves as being of African-American (n = 73) or Mexican-American (n = 24) ancestry. The patients were recruited from The University of Texas M. D. Anderson Cancer Center; from county, community, and Veterans Administration hospitals in the Houston metropolitan area; and from community hospitals in San Antonio, Texas. There were no age, histologic, or stage restrictions.

Data Collection

The epidemiologic data were collected by personal interviews.¹⁵ After informed consent was obtained, a structured interview approximately 45 minutes long was conducted in English or Spanish by trained interviewer-phlebotomists. Data were collected on sociodemographic characteristics, recent and past tobacco use, and other lifestyle habits using standardized questions. A detailed family history of the patient's first-degree relatives was collected and included the number of siblings and offspring and the dates of birth, death, and cancer diagnoses in first-degree relatives.

Lymphocyte Cultures and Chromosome Analysis

From each subject, 10 mL of peripheral blood was drawn into heparinized tubes. The lymphocyte culture and chromosome preparation techniques used were those of Hsu et al.¹⁶ All G-banded slides were coded before analysis. One hundred metaphase spreads from each individual were examined with a Genetiscan (PSI, Houston, TX). For each subject, abnormal metaphases and three normal metaphases were karyotyped.

Statistical Analysis

Chromosome 9 aberration was defined as translocation, deletion, or inversion of chromosome 9 or the presence of a dicentric chromosome 9. A smoker was defined as someone who had smoked at least 100 cigarettes in his or her lifetime. To test for significant associations between family history of cancer and specific chromosomal aberrations, ORs with 95% confidence limits were computed as estimates of the relative risk. Logistic regression was conducted to estimate risks associated with chromosome 9 aberrations, which were adjusted for multiple factors using STATA statistical software.¹⁷

RESULTS

Host Characteristics

The study population was described previously.¹⁰ There were 68 men and 29 women, 24 of whom were Mexican American and 73 African American. The mean age was 60.3 years. As reported previously,¹⁰ overall, 43 of the patients (44.3%) exhibited spontaneous chromosome 9 aberrations, (compared with 23% of the controls). For 9 p aberrations specifically, the comparable percentages were 22% and 13%, respectively, in patients and controls.

The patients reported having 829 first-degree relatives, including 74 (8.9%) with cancer. [Table 1](#) summarizes the proportion of relatives with any reported cancer by chromosome 9 aberration status. The proportions of parents, siblings, and all first-degree relatives with cancer in individuals with chromosome 9 aberrations were higher than in individuals without chromosome 9 aberrations (26.0% vs. 15.0%, 10.1% vs. 8.0%, and 10.6% vs. 7.6%, respectively). None of these differences were statistically significant.

The associations between age, gender, ethnicity, smoking status, and chromosome 9 aberration profile were also evaluated. There were no statistically significant associations between chromosome 9 aberration status and current or former smoking status, gender, ethnicity, age, pack-years smoked, or family size ([Table 2](#)). There were also no statistically significant differences in the first-degree relatives' mean age at cancer diagnosis or in the distribution of cancer sites (data not shown).

Table 3 shows the risk estimates among smokers for chromosome 9 aberrations by family history of any cancer, of lung carcinoma, and of other cancers. A significantly elevated OR of 2.7 was noted for chromosome 9 aberrations and any reported cancer in a first-degree relative. When the data were dichotomized into lung carcinoma or other cancers, the OR was 8.5 for having a family history of lung carcinoma, but only 2.1 for having a family history of other cancers.

In the control population selected for the previous study¹⁰, 11 of those with chromosome 9 aberrations (64.7%) reported a positive family history, compared with 28 of the controls without such aberrations (49.1%). Although this pattern was consistent with that for the cases, these differences were not statistically significant. In addition to chromosome 9 aberrations, the relationship between other spontaneous chromosome aberrations and family history of cancer were also evaluated in both the cases and controls, but no associations were noted. There was no trend for increasing risk with increased frequency of aberrations. However, of four cases with six or more different chromosomal aberrations, three reported a history of cancer in a first-degree relative.

DISCUSSION

Lung carcinoma is the epitome of an environmentally induced disease. Approximately 90% of lung carcinomas are attributed to tobacco exposure. However, only a fraction of smokers develop neoplastic lesions and genetically determined modulation of environmental exposures is an attractive mechanism to explain the variation in host susceptibility. Epidemiologic studies of familial aggregation of lung carcinoma provide indirect evidence for the role of genetic predisposition to this disease. Tokuhata and Lilienfeld¹⁸ and Ooi et al.¹⁹ found increased risks for lung carcinoma for both smoking and nonsmoking relatives of lung carcinoma patients. Sellers et al.²⁰ also documented an increased familial risk for lung carcinoma after controlling for the effects of age, gender, occupation, and smoking. They subsequently performed segregation analysis and reported results compatible with Mendelian codominant inheritance of a rare major autosomal gene for lung carcinoma predisposition.²¹ Twin studies have also confirmed that monozygotic twins discordant for smoking behavior have different risks for lung carcinoma.²²

In vitro chromosomal analyses have been frequently used to study individual sensitivity to genotoxicity and cancer risk. Using a molecular epidemiologic approach, the authors and other investigators have found that lymphocyte analysis can be used to identify individuals at a higher risk of developing cancer.^{15,23-28}

In their previous lung carcinoma study, the authors demonstrated that spontaneous chromosome 9 aberrations detected by G-banding in PBLs were a significant risk predictor for lung carcinoma with an OR of 3.6, after adjusting for age, gender, ethnicity, cigarette smoking, and occupational exposure. In this study, the authors found that these chromosome 9 aberrations were also significantly associated with family history of cancer. Furthermore, the current data showed that there was no significant association between chromosome 9 aberration and smoking status, pack-years smoked, gender or ethnicity.

A number of cytogenetic studies in lung carcinoma specimens have revealed consistent deletions involving chromosome 9p.^{11,29,30} Center et al.³⁰ reported nonrandom abnormalities of 9p in most primary samples from patients with nonsmall cell lung carcinoma (NSCLC). Several recent molecular studies complemented these cytogenetic observations by showing loss of heterozygosity (LOH) in lung carcinoma tumors and cell lines.^{12-14,31-35} Merlo et al.³² have reported frequent LOH on chromosome 9p, especially at 9p,²¹⁻²² and identified novel tumor suppressor gene loci on 9q in small cell lung carcinoma α and NSCLC.

A recent report suggested that the putative tumor suppressor gene MTS1 is very near the interferon gene cluster, at chromosome 9p21.¹¹ Xiao et al.¹² concluded that p15 and p16 are deleted or mutated in most primary NSCLCs and also provided evidence that at least one additional tumor suppressor gene on chromosome 9 is involved in NSCLC. Kishimoto et al.¹⁴ tested for deletions and LOH of 9p loci in preneoplastic and neoplastic loci in the lungs of patients with NSCLC by polymerase chain reaction-based assays for polymorphism in dinucleotide repeats (microsatellite markers) in the interferon and D9S171 loci on 9p. They concluded that LOH at 9p loci occurs at the earliest stage in the pathogenesis of lung carcinoma and involves all regions of the respiratory tract.

To determine whether the same chromosomal changes were present in the normal and target, diseased tissue, Dave et al.³⁶ analyzed the concordance of chromosomal alterations in primary lung tumors and PBLs from 10 lung carcinoma patients. Using molecular cytogenetic techniques, they reported that each paired sample (PBL and lung tumor tissue) had at least three chromosomes and two chromosomal regions with the same structural rearrangements. Genomic instability at the chromosomal level in PBLs thus appears to reflect the genetic changes in tumors and it is possible that cytogenetic analysis of primary chromosome changes in PBLs could provide a clue to the primary role of those changes in tumorigenesis.

The current hypothesis-generating study is limited by the small sample size and thus the results require cautious interpretation. Furthermore, the authors relied on the probands' self-reported family history of cancer without validation by medical documentation. However, Bondy et al.³⁷ found that self-reported cancer information regarding first-degree relatives was 88% accurate when compared with medical documentation.

The complex pattern of karyotypic changes in tumor cells emphasizes the need for more investigations of normal tissue (PBLs) and premalignant bronchial lesions to identify the primary genetic changes important for the identification of high risk populations, early detection, and intervention in this aggressive neoplasm. Chromosome 9 aberrations may be a marker of lung carcinoma susceptibility and may be associated with familial aggregation of cancer. These chromosome aberrations could be used as biomarkers to identify high risk populations. The authors are currently using molecular cytogenetic techniques to confirm these findings in a larger study of lung carcinoma patients.

REFERENCES

1. Yunis JJ, Ramsay N. Retinoblastoma and subband deletion of chromosome 13. *Am J Diseased Child* 1978;132:161-3.
2. Yunis JJ, Ramsay N. Familial occurrence of the aniridia Wilm's tumor syndrome with deletion 11p13-14.1. *J Pediatr* 1980;96:1027-30. [Medline](#)
3. Lee WH, Bookstein R, Hong F, Young LH, Shew JY, Lee EY. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science* 1987;235:1394-9. [Medline](#)
4. Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, et al. Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* 1990;60:509-20. [Medline](#)
5. Pathak S. Cytogenetics of epithelial malignancies. *Cancer* 1992;70:1660-70. [Medline](#)
6. Pathak S, Dhaliwal MK. Cytogenetic lesions in renal cell carcinomas. *Cancer Bull* 1989;41:324-9.
7. Pathak S, Strong LC, Ferrell RE, Trindade A. Familial renal cell carcinoma with a 3;11 chromosome translocation limited to tumor cells. *Science* 1982;217:939-41. [Medline](#)
8. Pathak S, Hopwood VL, Hughes JI, Jackson GL. Identification of colon cancer-predisposed individuals: a cytogenetic analysis. *Am J Gastroenterol* 1991;86:679-84. [Medline](#)

9. Pathak S, Siciliano MJ, Cailleau R, Wisemen CL, Hsu TC. A human breast adenocarcinoma with chromosome and isoenzyme markers similar to those of the HeLa line. *J Natl Cancer Inst* 1979;62:263-71. [Medline](#)
10. Dave BJ, Hopwood VL, King TM, Jiang H, Spitz MR, Pathak S. Genetic susceptibility to lung cancer as determined by lymphocytic chromosome analysis. *Cancer Epidemiol Biomarkers Prev* 1995;4:743-9. [Medline](#)
11. Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavitigian SV, et al. A cell cycle regulator potentially involved in genesis of many tumor types. *Science* 1994;264:436-40. [Medline](#)
12. Xiao S, Li D, Corson JM, Vijg J, Fletcher JA. Codeletion of p15 and p16 genes in primary non-small cell lung carcinoma. *Cancer Res* 1995;55:2968-71. [Medline](#)
13. Wiest J, Franklin W, Otstot J, Anderson M. Identification of a novel region of homozygous deletion on chromosome 9p in squamous cell carcinoma of the lung, [abstract 4067]. *Proc Am Assoc Cancer Res* 1996;37:593.
14. Kishimoto Y, Sugio K, Hung JY, Virmani AK, McIntire DD, Minna JD, et al. Allele-specific loss in chromosome 9p loci in preneoplastic lesions accompanying non-small-cell lung cancers. *J Natl Cancer Inst* 1995;87:1224-9. [Medline](#)
15. Spitz MR, Hsu TC, Wu XF, Fueger JJ, Amos CI, Roth JA. Mutagen sensitivity as a biologic marker of lung cancer risk in African-Americans. *Cancer Epidemiol Biomarkers Prev* 1995;4:99-103. [Medline](#)
16. Hsu TC, Johnston DA, Cherry LM, Ramkisson D, Schantz SP, Jessup JM, et al. Sensitivity to geneotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. *Int J Cancer* 1989;43:403-9. [Medline](#)
17. STATA Release 3.1 Reference Manual. Edition 6. College Station, TX: STATA Corporation, 1993.
18. Tokuhata GK, Lilienfeld AM. Familial aggregation of lung cancer in humans. *J Natl Cancer Inst* 1963;30:289-312.
19. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. Increased familial risk for lung cancer. *J Natl Cancer Inst* 1986;76:217-22. [Medline](#)
20. Sellers TA, Elston RC, Rothschild H. Familial risk of cancer among randomly selected cancer probands. *Genet Epidemiol* 1988;5:381-92. [Medline](#)
21. Sellers TA, Bailey-Wilson JE, Elston RC, Wilson AF, Elston GZ, Ooi WL, et al. Evidence for mendelian inheritance in the pathogenesis of lung cancer. *J Natl Cancer Inst* 1990; 82:1272-9. [Medline](#)
22. Carmelli D, Swan GE, Robiner D, Fabsitz R. Genetic influence on smoking--a study of male twins. *N Engl J Med* 1992; 327:829-33. [Medline](#)
23. Strom SS, Wu XF, Sigurdson AJ, Hsu TC, Fueger JJ, Lopez J, et al. Lung cancer, smoking patterns, and mutagen sensitivity in Mexican-Americans. *Monogr Natl Cancer Inst* 1995;18:29-33. [Medline](#)
24. Wu XF, Hsu TC, Annegers JF, Amos CI, Fueger JJ, Spitz MR. A case-control study of nonrandom distribution of bleomycin-induced chromatid breaks in lymphocytes of lung cancer patients. *Cancer Res* 1995;55:557-61.
25. Schantz SP, Spitz MR, Hsu TC. Mutagen sensitivity in patients with head and neck cancers: a biologic marker for risk of multiple primary malignancies. *J Natl Cancer Inst* 1990;82:1773-5. [Medline](#)
26. Wu XF, Hsu TC, Spitz MR. Mutagen sensitivity exhibits a dose-response relationship in case-control studies. *Cancer Epidemiol Biomarkers Prev* 1996;5:577-8.
27. Spitz MR, Hoque A, Trizna Z, Schantz SP, Amos CI, King TM, et al. Mutagen sensitivity as a risk factor for second malignant tumors following upper aerodigestive tract malignancies. *J Natl Cancer Inst* 1994;86:1681-4. [Medline](#)
28. Hagmar L, Brogger A, Hansteen I, Heim S, Hogstedt B, Knudsen L, et al. Cancer risk in humans predicted by increased levels of chromosome aberrations in lymphocytes: Nordic study group on the health risk of chromosome damage. *Cancer Res* 1994;54:2919-22. [Medline](#)

29. Testa JR, Siegfried JM. Chromosome abnormalities in human non-small cell lung cancer. *Cancer Res* 1992;52(Suppl):2702s-6s.
30. Center R, Lukeis R, Dietzsch E, Gillespie M, Garson OM. Molecular deletion of 9p sequences in non-small cell lung cancer and malignant mesothelioma. *Genes Chromosom Cancer* 1993;7:47-53. [Medline](#)
31. Merlo A, Gabrielson E, Askin F, Sidransky D. Frequent loss of chromosome 9 in human primary non-small cell lung cancer. *Cancer Res* 1994;54:640-2. [Medline](#)
32. Merlo A, Gabrielson E, Mabry M, Vollmer R, Baylin SB, Sidransky D. Homozygous deletion on chromosome 9p and loss of heterozygosity on 9q, 6p, and 6q in primary human small cell lung cancer. *Cancer Res* 1994;54:2322-6. [Medline](#)
33. Neville EM, Stewart M, Myskow M, Donnelly RJ, Field JK. Loss of heterozygosity at 9p23 defines a novel locus in non-small cell lung cancer. *Oncogene* 1995;11:581-5. [Medline](#)
34. Tarmin L, Yin J, Zhou X, Suzuki H, Jiang HY, Rhyu MG, et al. Frequent loss of heterozygosity on chromosome 9 in adenocarcinoma and squamous cell carcinoma of the esophagus. *Cancer Res* 1994;54:6094-6. [Medline](#)
35. Mead LJ, Gillespie MT, Irving LB, Campbell LJ. Homozygous and hemizygous deletions of 9p centromeric to the interferon genes in lung cancer. *Cancer Res* 1994;54:2307-9. [Medline](#)
36. Dave BJ, Hopwood VL, Spitz MR, Pathak S. Shared cytogenetic abnormalities in lung tumors and corresponding peripheral blood lymphocytes. *Int J Oncol* 1995;7:1297-305.
37. Bondy ML, Strom SS, Colopy MW, Brown BW, Strong LC. Comparison of accuracy of cancer information in a family study of childhood soft tissue sarcoma. *J Clin Epidemiol* 1994;47:89-96. [Medline](#)

FOOTNOTES AND ACKNOWLEDGMENTS

Supported by Grant RO1 CA 55769 from the National Cancer Institute (to M. R. S.) and R03 CA 70191 (to X. W.) and by the Richard C. Devereaux Award from the Cancer Research Foundation of America (to X. W.).

The authors thank Dr. Maureen Goode of the Department of Scientific Publications for editorial assistance with the article.

Address for reprints: Xifeng Wu, M.D., Ph.D, Department of Epidemiology, Box 189, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030.

TEXTO 8



16/04/1997

El cáncer de pulmón se asocia a alteraciones del cromosoma 9

Jano Diario, Barcelona

Investigadores de la Universidad de Texas (Estados Unidos) han descubierto que ciertas alteraciones en el cromosoma 9 parecen ser marcadores del riesgo de cáncer pulmonar y de la agregación familiar de este tipo de tumor, que es precisamente el que parece ser en mayor medida inducido por factores ambientales.

Investigaciones previas del equipo que dirige el Dr. Xifeng Wu mostraban que las anomalías del cromosoma 9 en los linfocitos periféricos constituían un predictor de riesgo significativo de carcinoma pulmonar. En el nuevo estudio, publicado en el último número de "Cancer", trataron de determinar si esta asociación se debía a una inestabilidad genética heredada. Identificaron a 97 pacientes con cáncer de pulmón y alteraciones en el cromosoma 9, y a continuación evaluaron a sus familiares de primer grado.

Los resultados indican que estos familiares presentan una probabilidad 8,5 veces superior de desarrollar carcinoma pulmonar si tienen mutaciones del cromosoma 9 respecto a los que no las tienen. Aunque se trata de resultados preliminares, el Dr. Wu sugiere que el estudio de estas alteraciones cromosómicas podrían utilizarse como marcadores para identificar a poblaciones de alto riesgo.

<http://www.atheneum.doyma.es/socios/noticias/main.htm>

What information do you need?

NEWS

▪ **Types of Cancer**

- Breast, Colon, Leukemia...
- Statistical Data

▪ **Treatment Options**

- Treatment Information
- Chemotherapy, Radiation Therapy,...
- Alternative/Complementary Medicine

▪ **Clinical Trials**

- cancerTrials
- Finding Clinical Trials
- Understanding Clinical Trials
- Clinical Trial Resources

▪ **Genetics, Causes, Risk Factors, Prevention**

- Prevention Information
- Genetics, Genetic Testing,...
- Search for Genetics Professionals

▪ **Testing for Cancer**

- Detection/Screening Information
- Mammograms, PSA Test, Colonoscopy

▪ **Coping with Cancer**

- Side Effects of Treatment
- Complications Caused by Cancer
- Emotional and Other Concerns

▪ **Support and Resources**

- Support Groups
- Hospice, Home Care
- Resources for Patients/Survivors, Caregivers, Health Professionals

▪ **Cancer Literature**

- CANCERLIT
- Journal of the National Cancer Institute

▪ **PDQ®**

- What is PDQ?
- Current Cancer Information Summaries by Cancer Experts
- Searching Clinical Trials
- Submitting Trials to PDQ

TEXTOS EXTRAIDOS DEL SERVICIO *PDQ* DEL *NATIONAL CANCER INSTITUTE*

Treatment – Health Professionals

Malignant Mesothelioma (PDQ®)	TEXTO 9
Non-Small Cell Lung Cancer (PDQ®)	TEXTO 10
Small Cell Lung Cancer (PDQ®)	TEXTO 11

Treatment – Patients

Malignant Mesothelioma (PDQ®)	TEXTO 12
Non-Small Cell Lung Cancer (PDQ®)	TEXTO 13
Small Cell Lung Cancer (PDQ®)	TEXTO 14

PDQ Tratamiento Profesionales de la Salud

Malignant Mesothelioma – Mesotelioma maligno (versión bitexto de QDT1)	TEXTO 15
---	----------

TEXTO 9: Treatment summaries for health professionals (QDT1)

----- MALIGNANT MESOTHELIOMA -----

***** GENERAL INFORMATION

Prognosis in this disease is difficult to assess consistently because there is great variability in the time before diagnosis and the rate of disease progression. Various surgical procedures may be possible in selected patients, providing long-term survival without cure. In large retrospective series of pleural mesothelioma patients, important prognostic factors were found to be stage, age, performance status, and histology [1,2]. For patients treated with aggressive surgical approaches, factors associated with improved long-term survival include epithelial histology, negative lymph nodes, and negative surgical margins [3,4]. For those patients treated with aggressive surgical approaches, nodal status is an important prognostic factor [3]. Median survival for malignant local pleural disease has been reported as 16 months, and extensive disease as 5 months. In some instances the tumor grows through the diaphragm making site of origin difficult to assess. Cautious interpretation of treatment results in this disease is imperative because of the selection differences among series. Effusions, both pleural and peritoneal, represent major symptomatic problems for at least two thirds of the patients. A history of asbestos exposure is reported in about 70%-80% of all cases of mesothelioma [1,5,6].

<References>

***** CELLULAR CLASSIFICATION

Histologically, these tumors are composed of fibrous or epithelial elements or both. The epithelial form occasionally causes confusion with peripheral anaplastic lung carcinomas or metastatic carcinomas. Attempts at diagnosis by cytology or needle biopsy of the pleura are often noncontributory. It can be especially difficult to differentiate mesothelioma from carcinoma on small tissue specimens. Thoracoscopy can be valuable in obtaining adequate tissue specimens for diagnostic purposes [1]. Examination of the gross tumor at surgery and use of special stains or electron microscopy can often help. The special stains reported to be most useful include periodic acid-Schiff diastase, hyaluronic acid, mucicarmine, CEA, and Leu M1 [2]. Histologic appearance appears to be of prognostic value, with most clinical studies showing that epithelial mesotheliomas have a better prognosis than fibrous or sarcomatous mesotheliomas [2-4].

<References>

***** STAGE INFORMATION

Patients with stage I disease have a significantly better prognosis than those with more advanced stages. However, because of the relative rarity of this disease, exact survival information based upon stage is limited [1]. A proposed staging system based upon thoracic surgery principles and clinical data is shown below [2]. It is a modification of the older system proposed by Butchart et al [3]. Other staging systems that have been employed, including a proposed new international TNM staging system, are summarized by the International Mesothelioma Interest Group [4].

Stage I: Disease confined within the capsule of the parietal pleura: ipsilateral pleura, lung, pericardium, and diaphragm.

Stage II: All of stage I with positive intrathoracic (N1 or N2) lymph nodes.

Stage III: Local extension of disease into the following: chest wall or mediastinum; heart or through the diaphragm, peritoneum; with or without extrathoracic or contralateral (N3) lymph node involvement.

Stage IV: Distant metastatic disease.

Localized malignant mesothelioma: See description of stage I above.

Advanced malignant mesothelioma: See descriptions of stages II, III, and IV above.

For the purposes of the discussion of treatment in this statement, the disease is categorized as either localized or advanced.

<References>

***** TREATMENT OPTION OVERVIEW

Standard treatment for all but localized mesothelioma is generally not curative. Although some patients will experience long-term survival with aggressive treatment approaches, it remains unclear if overall survival has been significantly altered by the different treatment modalities or by combinations of modalities. Extrapleural pneumonectomy in selected patients with early stage disease may improve recurrence-free survival, but its impact on overall survival is unknown [1]. Pleurectomy and decortication can provide palliative relief from symptomatic effusions, discomfort caused by tumor burden, and pain caused by invasive tumor. Operative mortality from pleurectomy/decortication is less than 2%, [2]. while mortality from extrapleural pneumonectomy has ranged from 6% to 30% [1,3]. The addition of radiation therapy and/or chemotherapy following surgical intervention has not demonstrated improved survival [2]. The use of radiation therapy in pleural mesothelioma has been shown to alleviate pain in the majority of patients treated. However, the duration of symptom control is short-lived [4,5]. Single agent and combination chemotherapy have been evaluated in single and combined modality studies. The most studied agent is doxorubicin, which has produced partial responses in approximately 15%-20% of patients studied [6]. Some combination chemotherapy regimens have been reported to have higher response rates in small phase II trials. However the toxicity reported is also higher and there is no evidence that combination regimens result in longer survival or longer control of symptoms [6,7]. Recurrent pleural effusions may be treated with pleural sclerosing procedures; however, failure rates are high secondary to the restrictive nature of the tumor.

The designations in PDQ that treatments are "standard" or "under clinical evaluation" are not to be used as a basis for reimbursement determinations.

<References>

***** LOCALIZED MALIGNANT MESOTHELIOMA (STAGE I)

Treatment options:[1].

Standard:

1 Solitary mesotheliomas: Surgical resection en bloc including contiguous structures to ensure wide disease-free margins. Sessile polypoid lesions should be treated with surgical resection to ensure maximal potential for cure [2].

2 Intracavitary mesothelioma:

A) Palliative surgery (pleurectomy and decortication) with or without postoperative radiation therapy

- B) Extrapleural pneumonectomy
- C) Palliative radiation therapy.

Under clinical evaluation:

- 1 Intracavitary chemotherapy following resection [3,4].
- 2 Multimodality therapy [4-6].
- 3 Other clinical trials.

<References>

***** ADVANCED MALIGNANT MESOTHELIOMA (STAGES II, III, AND IV)

Treatment options:

- 1 Symptomatic treatment to include drainage of effusions, chest tube pleurodesis, or thoracoscopic pleurodesis [1].
- 2 Palliative surgical resection in selected patients [2,3].
- 3 Palliative radiation therapy [4,5].
- 4 Single-agent chemotherapy. Partial responses have been reported with doxorubicin, epirubicin, mitomycin, cyclophosphamide, cisplatin, carboplatin, and ifosfamide [6-8].
- 5 Multimodality clinical trials [9-13].
- 6 Intracavitary therapy. Intrapleural or intraperitoneal administration of chemotherapeutic agents (e.g, cisplatin, mitomycin, and cytarabine) has been reported to produce transient reduction in the size of tumor masses and temporary control of effusions in small clinical studies [14-16]. Additional studies are needed to define the role of intracavitary therapy.

<References>

***** RECURRENT MALIGNANT MESOTHELIOMA

Treatment of recurrent mesothelioma usually utilizes procedures and/or agents not previously employed in the initial treatment attempt. No standard treatment approaches have been proven to improve survival or control symptoms for a prolonged period of time. These patients should be considered candidates for phase I and II clinical trials evaluating new biologicals, chemotherapeutic agents, or physical approaches [1-5]. Consult the PDQ protocol file for a current listing of active clinical trials.

<References>

TEXTO 10: Treatment summaries for health professionals (QDT2)

PDQ® Treatment Health Professionals

Important: This information is intended for use by doctors and other health care professionals. If you are a cancer patient, your doctor can explain how it applies to you, or you can call the Cancer Information Service at 1-800-422-6237.

Non-small cell lung cancer

Table of Contents

GENERAL INFORMATION

CELLULAR CLASSIFICATION

STAGE INFORMATION

The Revised International Staging System for Lung Cancer

TNM definitions

AJCC stage groupings

Occult carcinoma

Stage 0

Stage IA

Stage IB

Stage IIA

Stage IIB

Stage IIIA

Stage IIIB

Stage IV

TREATMENT OPTION OVERVIEW

OCCULT NON-SMALL CELL LUNG CANCER

TX, N0, M0

STAGE 0 NON-SMALL CELL LUNG CANCER

Tis, N0, M0

STAGE I NON-SMALL CELL LUNG CANCER

T1, N0, M0 or T2, N0, M0

STAGE II NON-SMALL CELL LUNG CANCER

T1, N1, M0 or T2, N1, M0 or T3, N0, M0

STAGE IIIA NON-SMALL CELL LUNG CANCER

T1, N2, M0 or T2, N2, M0 or T3, N1, M0 or T3, N2, M0

Superior sulcus tumor (T3, N0 or N1, M0)

Chest wall tumor (T3, N0 or N1, M0)

STAGE IIIB NON-SMALL CELL LUNG CANCER

T4 or N3, M0

STAGE IV NON-SMALL CELL LUNG CANCER

Any T, any N, M1

RECURRENT NON-SMALL CELL LUNG CANCER

GENERAL INFORMATION

Non-small cell lung cancer (NSCLC) is a heterogeneous aggregate of at least three distinct histologies of lung cancer including epidermoid or squamous carcinoma, adenocarcinoma, and large cell carcinoma. These histologies are often classified together because, when localized, all have the potential for cure with surgical resection. Systemic chemotherapy can produce objective partial responses and palliation of symptoms for short durations. Local control can be achieved with radiation in a large number of patients with unresectable disease, but cure is seen only in a small minority of patients.

At diagnosis, patients with NSCLC can be divided into three groups that reflect the extent of disease and treatment approach. The first group of patients has tumors that are surgically resectable (generally stages I and II). This is the group with the best prognosis, depending on a variety of tumor and host factors. Patients with resectable disease who have medical contraindications to surgery can be considered for curative radiotherapy. The second group includes patients with either locally (T3-T4) or regionally (N2-N3) advanced lung cancer who have a diverse natural history. This group is treated with radiotherapy or with radiotherapy in combination with other therapy modalities. Selected patients with T3 or N2 disease can be treated effectively with surgical resection alone. The final group of patients have distant metastases (M1) found at the time of diagnosis. This group can be treated with radiotherapy or chemotherapy for palliation of symptoms from the primary tumor. Patients with good performance status, women, and patients with distant metastases confined to a single site appear to live longer than others.[1] Cisplatin-based chemotherapy has been associated with short-term palliation of symptoms and a small survival advantage. Currently no single chemotherapy regimen can be recommended for routine use.

For operable patients, prognosis is adversely influenced by the presence of pulmonary symptoms, large tumor size (>3 centimeters), and presence of the erbB-2 oncoprotein.[1-6] Other factors that have been identified as adverse prognostic factors in some series of patients with resectable non-small cell lung cancer include mutation of the K-ras gene, vascular invasion, and increased numbers of blood vessels in the tumor specimen.[3,7,8]

Since treatment is not satisfactory for almost all patients with NSCLC, with the possible exception of a subset of pathologic stage I (T1, N0, M0) patients treated surgically, eligible patients should be considered for clinical trials.

References:

1. Albain KS, Crowley JJ, LeBlanc M, et al.: Survival determinants in extensive-stage non-small-cell lung cancer: the Southwest Oncology Group experience. *Journal of Clinical Oncology* 9(9): 1618-1626, 1991.
2. Macchiarini P, Fontanini G, Hardin MJ, et al.: Blood vessel invasion by tumor cells predicts recurrence in completely resected T1 N0 M0 non-small-cell lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 106(1): 80-89, 1993.
3. Harpole DH, Herndon JE, Wolfe WG, et al.: A prognostic model of recurrence and death in stage I non-small cell lung cancer utilizing presentation, histopathology, and oncoprotein expression. *Cancer Research* 55(1): 51-56, 1995.
4. Ichinose Y, Yano T, Asoh H, et al.: Prognostic factors obtained by a pathologic examination in completely resected non-small-cell lung cancer: an analysis in each pathologic stage. *Journal of Thoracic and Cardiovascular Surgery* 110(3): 601-605, 1995.
5. Martini N, Bains MS, Burt ME, et al.: Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 109(1): 120-129, 1995.
6. Strauss GM, Kwiatkowski DJ, Harpole DH, et al.: Molecular and pathologic markers in stage I non-small-cell carcinoma of the lung. *Journal of Clinical Oncology* 13(5): 1265-1279, 1995.
7. Slebos RJ, Kibbelaar RE, Dalesio O, et al.: K-RAS oncogene activation as a prognostic marker in adenocarcinoma of the lung. *New England Journal of Medicine* 323(9): 561-565, 1990.
8. Fontanini G, Bigini D, Vignati S, et al.: Microvessel count predicts metastatic disease and survival in non-small cell lung cancer. *Journal of Pathology* 177: 57-63, 1995.

CELLULAR CLASSIFICATION

Prior to initiating treatment of any patient with lung cancer, a review of pathologic material by an experienced lung cancer pathologist is critical since some cases of small cell lung cancer (which responds well to chemotherapy) can be confused on microscopic examination with non-small cell carcinoma.[1] Nonsquamous cell cancers may be more likely to recur after surgical resection of early stage I tumors than other types of non-small cell lung cancers.[2] Bronchoalveolar carcinoma represents 10%-25% of adenocarcinomas and sometimes has a distinct presentation and biologic behavior.[3-5] Bronchoalveolar cancer may present as a more diffuse lesion than other types of cancer; 30%-40% of patients undergoing an attempt at surgical resection present with an infiltrate on their chest radiograph. Bronchoalveolar

cancer is more common in women and in patients who do not smoke cigarettes than other histologic types of lung cancer.

Histologic classification of non-small cell lung cancer:

- squamous cell (epidermoid) carcinoma
- spindle cell variant
- adenocarcinoma
- acinar
- papillary
- bronchoalveolar [4,5]
- solid tumor with mucin
- large cell carcinoma
- giant cell
- clear cell
- adenosquamous carcinoma
- undifferentiated carcinoma

References:

1. Kreyberg L, Liebow AA, Uehlinger EA: International Histologic Classification of Tumours: No. 1. Histological Typing of Lung Tumours. Geneva: World Health Organization, 2nd ed., 1981.
2. Thomas P, Rubinstein L: Cancer recurrence after resection: T1 N0 non-small cell lung cancer: Lung Cancer Study Group. *Annals of Thoracic Surgery* 49(2): 242-246, 1990.
3. Harpole DH, Bigelow C, Young WG, et al.: Alveolar cell carcinoma of the lung: a retrospective analysis of 205 patients. *Annals of Thoracic Surgery* 46: 502-507, 1988.
4. Grover FL, Piantadosi S: Recurrence and survival following resection of bronchioloalveolar carcinoma of the lung: the Lung Cancer Study Group experience. *Annals of Surgery* 209(6): 779-790, 1989.
5. Daly RC, Trastek VF, Pairolero PC, et al.: Bronchoalveolar carcinoma: factors affecting survival. *Annals of Thoracic Surgery* 51(3): 368-377, 1991.

STAGE INFORMATION

Since determination of stage has important therapeutic and prognostic implications, careful initial diagnostic evaluation to define location and extent of primary and metastatic tumor involvement is critical for the appropriate care of patients.

Stage has a critical role in the selection of therapy. The stage of disease is based on a combination of clinical (physical examination, radiologic, and laboratory studies) and pathologic (biopsy of lymph nodes, bronchoscopy, mediastinoscopy, or anterior mediastinotomy).[1] The distinction between clinical stage and pathologic stage should be considered when evaluating reports of survival outcome. Surgical staging of the mediastinum is considered standard if accurate evaluation of the nodal status is needed to determine therapy. The Radiology Diagnostic Oncology Group reported that the sensitivity and specificity of computerized tomography (CT) scanning is only 52% and 69%, respectively.[2] Magnetic resonance imaging (MRI) does not appear to improve the accuracy of staging.[2] Early evaluation of the role of positron emission tomography (PET) suggests that the combination of CT and PET may have greater sensitivity and specificity than CT alone.[3] A report evaluating the staging of 1,400 patients undergoing tumor resection found that clinical staging by radiologic studies accurately assessed the T stage in 78% of patients and the N stage in only 47% of patients. Errors in clinical staging were equally divided between overstaging and understaging.[4]

The Revised International Staging System for Lung Cancer

The Revised International System for Staging Lung Cancer was adopted in 1997 by the American Joint Committee on Cancer and the Union Internationale Contre le Cancer.[5] These revisions were made to provide greater specificity for patient groups. Stage I is divided into two categories by the size of the tumor; IA, T1N0M0 and IB, T2N0M0. Stage II is divided into two categories by the size of the tumor and by the nodal status; IIA, T1N1M0 and IIB, T2N1M0. T3N0 has been moved from stage IIIA in the 1986

version of the staging system to stage IIB. The other change has been to clarify the classification of multiple tumor nodules. Satellite tumor nodules in the same lobe as the primary lesion that are not lymph nodes should be classified as T4 lesions. Intrapulmonary ipsilateral metastasis in a lobe other than the lobe containing the primary lesions should be classified as an M1 lesion (stage IV). The American Joint Committee on Cancer (AJCC) has designated staging by TNM classification.[6]

TNM definitions

Primary tumor (T)

TX: Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy

T0: No evidence of primary tumor

Tis: Carcinoma in situ

T1: A tumor that is 3 cm or less in greatest dimension, surrounded by lung or visceral pleura, and without bronchoscopic evidence of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus)*

T2: A tumor with any of the following features of size or extent:

More than 3 cm in greatest dimension

Involves the main bronchus, 2 cm or more distal to the carina

Invades the visceral pleura

Associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the entire lung

T3: A tumor of any size that directly invades any of the following: chest

wall (including superior sulcus tumors), diaphragm, mediastinal pleura,

parietal pericardium; or tumor in the main bronchus less than 2 cm

distal to the carina but without involvement of the carina; or

associated atelectasis or obstructive pneumonitis of the entire lung

T4: A tumor of any size that invades any of the following: mediastinum,

heart, great vessels, trachea, esophagus, vertebral body, carina; or

separate tumor nodules in the same lobe; or tumor with a malignant

pleural effusion **

*Note: The uncommon superficial tumor of any size with its invasive component limited to the bronchial wall, which may extend proximal to the main bronchus, is also classified as T1.

**Note: Most pleural effusions associated with lung cancer are due to tumor. However, there are a few patients in whom multiple cytopathologic examinations of pleural fluid are negative for tumor. In these cases, fluid is non-bloody and is not an exudate. When these elements and clinical judgement dictate that the effusion is not related to the tumor, the effusion should be excluded as a staging element and the patient should be staged as T1, T2, or T3.

Regional lymph nodes (N)

NX: Regional lymph nodes cannot be assessed

N0: No regional lymph node metastasis

N1: Metastasis to ipsilateral peribronchial and/or ipsilateral hilar lymph nodes, and intrapulmonary nodes including involvement by direct extension of the primary tumor

N2: Metastasis to ipsilateral mediastinal and/or subcarinal lymph node(s)

N3: Metastasis to contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s)

Distant metastasis (M)

MX: Distant metastasis cannot be assessed

M0: No distant metastasis

M1: Distant metastasis present

Note: M1 includes separate tumor nodule(s) in a different lobe (ipsilateral or contralateral).

Specify sites according to the following notations:

BRA = brain	EYE = eye	HEP = hepatic
LYM = lymph nodes	MAR = bone marrow	OSS = osseous
OTH = other	OVR = ovary	PER = peritoneal
PLE = pleura	PUL = pulmonary	SKI = skin

AJCC stage groupings

Occult carcinoma

TX, N0, M0

Stage 0

Tis, N0, M0

Stage IA

T1, N0, M0

Stage IB

T2, N0, M0

Stage IIA

T1, N1, M0

Stage IIB

T2, N1, M0

T3, N0, M0

Stage IIIA

T1, N2, M0

T2, N2, M0

T3, N1, M0

T3, N2, M0

Stage IIIB

Any T, N3, M0

T4, Any N, M0

Stage IV

Any T, Any N, M1

References:

1. Ginsberg RJ: Invasive and noninvasive techniques of staging in potentially operable lung cancer. *Seminars in Surgical Oncology* 6(5): 244-247, 1990.
 2. Webb WR, Gatsonis C, Zerhouni EA, et al.: CT and MR imaging in staging non-small cell bronchogenic carcinoma: report of the Radiologic Diagnostic Oncology Group. *Radiology* 178(3): 705-713, 1991.
 3. Vansteenkiste JF, Stroobants SG, De Leyn PR, et al.: Lymph node staging in non-small-cell lung cancer with FDG-PET scan: a prospective study on 690 lymph node stations from 68 patients. *Journal of Clinical Oncology* 16(6): 2142-2149, 1998.
 4. Bulzebruck H, Bopp R, Drings P, et al.: New aspects in the staging of lung cancer: prospective validation of the International Union Against Cancer TNM classification. *Cancer* 70(5): 1102-1110, 1992.
 5. Mountain CF: Revisions in the International System for Staging Lung Cancer. *Chest* 111(6): 1710-1717, 1997.
 6. Lung. In: American Joint Committee on Cancer: *AJCC Cancer Staging Manual*. Philadelphia: Lippincott-Raven Publishers, 5th ed., 1997, pp 127-137.
-

TREATMENT OPTION OVERVIEW

In non-small cell lung cancer (NSCLC), results of standard treatment are poor in all but the most localized cancers. All newly diagnosed patients with NSCLC are potential candidates for studies evaluating new forms of treatment. Surgery is the major potentially curative therapeutic option for this disease; radiotherapy can produce cure in a small minority and palliation in the majority of patients. In advanced-stage disease, chemotherapy offers modest improvements in median survival although overall survival is poor.^[1,2] Where studied, chemotherapy has been reported to produce short-term improvement in disease-related symptoms. In one study, symptomatic relief with combination chemotherapy was significant but independent of objective response.^[3,4] The impact of chemotherapy on quality of life requires more study.

Current areas under evaluation include combining local (surgery), regional (radiotherapy), and systemic (chemotherapy and immunotherapy) treatments and developing more effective systemic therapy. Several new agents, including paclitaxel (Taxol), docetaxel (Taxotere), topotecan, irinotecan, vinorelbine, and gemcitabine have been shown to be active in the treatment of advanced NSCLC. Chemoprevention of second primary cancers of the upper aerodigestive tract is also under active investigation in early-stage lung cancer.^[5]

The designations in PDQ that treatments are "standard" or "under clinical evaluation" are not to be used as a basis for reimbursement determinations.

References:

1. Souquet PJ, Chauvin F, Boissel JP, et al.: Polychemotherapy in advanced non small cell lung cancer: a meta-analysis. *Lancet* 342(8862): 19-21, 1993.
2. Non-small Cell Lung Cancer Collaborative Group: Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *British Medical Journal* 311(7010): 899-909, 1995.
3. Hardy JR, Noble T, Smith IE: Symptom relief with moderate dose chemotherapy (mitomycin-C, vinblastine and cisplatin) in advanced non-small cell lung cancer. *British Journal of Cancer* 60(5): 764-766, 1989.
4. Ellis PA, Smith IE, Hardy JR, et al.: Symptom relief with MVP (mitomycin C, vinblastine and cisplatin) chemotherapy in advanced non-small-cell lung cancer. *British Journal of Cancer* 71(2): 366-370, 1995.
5. Pastorino U, Infante M, Maioli M, et al.: Adjuvant treatment of stage I lung cancer with high-dose vitamin A. *Journal of Clinical Oncology* 11(7): 1216-1222, 1993.

OCCULT NON-SMALL CELL LUNG CANCER

TX, N0, M0

In occult lung cancer, a diagnostic evaluation often includes chest x-ray and selective bronchoscopy with close follow-up (e.g., computed tomography scan), when needed, to define the site and nature of the primary tumor; tumors discovered in this fashion are generally early stage and curable by surgery. After discovery of the primary tumor, treatment is determined by establishing the stage of the patient's tumor. Therapy is identical to that recommended for other non-small cell lung cancer patients with similar stage disease.

STAGE 0 NON-SMALL CELL LUNG CANCER

Tis, N0, M0

Stage 0 non-small cell lung cancer (NSCLC) is the same as carcinoma in situ of the lung. Because these tumors are by definition noninvasive and incapable of metastasizing, they should be curable with surgical resection; however, there is a high incidence of second primary cancers, many of which are unresectable. Endoscopic phototherapy with a hematoporphyrin derivative has been described as an alternative to surgical resection in carefully selected patients.[1-3] This investigational treatment seems to be most effective for very early central tumors that extend less than 1 centimeter within the bronchus.[2] Efficacy of this treatment modality in the management of early NSCLC remains to be proven.

Treatment options:

1. Surgical resection using the least extensive technique possible (segmentectomy or wedge resection) to preserve maximum normal pulmonary tissue since these patients are at high risk for second lung cancers.
2. Endoscopic photodynamic therapy.[2,3]

References:

1. Woolner LB, Fontana RS, Cortese DA, et al.: Roentgenographically occult lung cancer: pathologic findings and frequency of multicentricity during a 10-year period. *Mayo Clinic Proceedings* 59(7): 453-466, 1984.
2. Furuse K, Fukuoka M, Kato H, et al.: A prospective phase II study on photodynamic therapy with photofrin II for centrally located early-stage lung cancer. *Journal of Clinical Oncology* 11(10): 1852-1857, 1993.
3. Edell ES, Cortese DA: Photodynamic therapy in the management of early superficial squamous cell carcinoma as an alternative to surgical resection. *Chest* 102(5): 1319-1322, 1992.

STAGE I NON-SMALL CELL LUNG CANCER

T1, N0, M0 or T2, N0, M0

Surgery is the treatment of choice for patients with stage I non-small cell lung cancer (NSCLC). Careful preoperative assessment of the patient's overall medical condition, especially the patient's pulmonary reserve, is critical in considering the benefits of surgery. The immediate postoperative mortality rate is age-related, but 3%-5% with lobectomy can be expected.[1] Patients with impaired pulmonary function may be considered for segmental or wedge resection of the primary tumor; the Lung Cancer Study Group has conducted a randomized study (LCSG-821) to compare lobectomy with limited resection for patients with stage I cancer of the lung. The results of this study show a reduction in local recurrence for patients

treated with lobectomy compared with those treated with limited excision but no significant difference in overall survival.[2] Similar results have been reported from a nonrandomized comparison of anatomic segmentectomy and lobectomy.[3] A survival advantage was noted with lobectomy for patients with tumors greater than 3 centimeters, but not for those with tumors smaller than 3 centimeters. However, the rate of local/regional recurrence was significantly less after lobectomy, regardless of primary tumor size. Another study of stage I patients showed that those treated with wedge or segment resections had a local recurrence rate of 50% (31 of 62) despite having undergone complete resections.[4] Exercise testing may aid in the selection of patients with impaired pulmonary function who can tolerate lung resection.[5] The availability of video-assisted thoracoscopic wedge resection permits limited resections in patients with poor pulmonary function who are not usually considered candidates for lobectomy.[6]

Inoperable patients with stage I disease and with sufficient pulmonary reserve may be considered for radiotherapy with curative intent. In one report of patients older than 70 years of age who had resectable lesions smaller than 4 centimeters but who were medically inoperable or who refused surgery, survival at 5 years following radiotherapy with curative intent was comparable to a historical control group of patients of similar age resected with curative intent.[7] In the two largest retrospective radiotherapy series, inoperable patients treated with definitive radiotherapy achieved 5-year survival rates of 10% and 27%. Both series found that patients with T1, N0 tumors had better outcomes, with 5-year survival rates of 60% and 32% in this subgroup.[8,9]

Primary radiotherapy should consist of approximately 6,000 cGy delivered with megavoltage equipment to the midplane of the known tumor volume using conventional fractionation. A boost to the cone-down field of the primary tumor is frequently used to further enhance local control. Careful treatment planning with precise definition of target volume and avoidance of critical normal structures to the extent possible is needed for optimal results and requires the use of a simulator.

Many patients treated surgically subsequently develop regional or distant metastases.[10] Therefore, patients should be considered for entry into clinical trials evaluating adjuvant treatment with chemotherapy or radiotherapy following surgery. Trials of adjuvant chemotherapy regimens have failed to demonstrate a consistent benefit. Smokers who undergo complete resection of stage I NSCLC are also at risk for second malignant tumors. In the Lung Cancer Study Group trial of 907 stage T1, N0 resected patients, the rate of nonpulmonary second cancers was 1.8% per year and 1.6% per year for new lung cancers.[11] Others have reported even higher risks of second tumors in long-term survivors, including rates of 10% for second lung cancers and 20% for all second cancers.[4] A randomized trial of vitamin A versus observation in resected stage I patients showed a trend toward decreased second primary cancers in the vitamin A arm with no difference in overall survival rates.[12] An ongoing intergroup clinical trial will evaluate the role of isotretinoin in the chemoprevention of second cancers in patients resected for stage I NSCLC.[13]

Treatment options:

1. Lobectomy or segmental, wedge, or sleeve resection as appropriate.
2. Radiotherapy with curative intent (for potentially resectable patients who have medical contraindications to surgery).
3. Clinical trials of adjuvant chemotherapy following resection.[14,15]
4. Adjuvant chemoprevention trials.[12,13]
5. Endoscopic photodynamic therapy (under clinical evaluation in highly selected T1, N0, M0 patients).[16]

References:

1. Ginsberg RJ, Hill LD, Eagan RT, et al.: Modern thirty-day operative mortality for surgical resections in lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 86(5): 654-658, 1983.
2. Ginsberg RJ, Rubinstein LV: Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. *Annals of Thoracic Surgery* 60(5): 615-623, 1995.
3. Warren WH, Faber LP: Segmentectomy versus lobectomy in patients with stage I pulmonary carcinoma. *Journal of Thoracic and Cardiovascular Surgery* 107(4): 1087-1094, 1994.
4. Martini N, Bains MS, Burt ME, et al.: Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 109(1): 120-129, 1995.
5. Morice RC, Peters EJ, Ryan MB, et al.: Exercise testing in the evaluation of patients at high risk for complications from lung resection. *Chest* 101(2): 356-361, 1992.
6. Shennib HA, Landreneau R, Mulder DS, et al.: Video-assisted thoracoscopic wedge resection of T1 lung cancer in high-risk patients. *Annals of Surgery* 218(4): 555-560, 1993.
7. Noordijk EM, Clement EP, Hermans J, et al.: Radiotherapy as an alternative to surgery in elderly patients with resectable lung cancer. *Radiotherapy and Oncology* 13(2): 83-89, 1988.
8. Dosoretz DE, Katin MJ, Blitzer PH, et al.: Radiation therapy in the management of medically inoperable carcinoma of the lung: results and implications for future treatment strategies. *International Journal of Radiation Oncology, Biology, Physics* 24(1): 3-9, 1992.
9. Gauden S, Ramsay J, Tripcony L: The curative treatment by radiotherapy alone of stage I non-small cell carcinoma of the lung. *Chest* 108(5): 1278-1282, 1995.
10. Martini N, Bains MS, Burt ME, et al.: Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 109(1): 120-129, 1995.
11. Thomas P, Rubinstein L: Cancer recurrence after resection: T1 N0 non-small cell lung cancer: Lung Cancer Study Group. *Annals of Thoracic Surgery* 49(2): 242-246, 1990.
12. Pastorino U, Infante M, Maioli M, et al.: Adjuvant treatment of stage I lung cancer with high-dose vitamin A. *Journal of Clinical Oncology* 11(7): 1216-1222, 1993.
13. Lippman SM, University of Texas - MD Anderson Cancer Center: Phase III, Double-Blind, Randomized Trial of 13-CRA vs Placebo to Prevent Second Primary Tumors in Patients with Totally Resected Stage I non-Small Cell Lung Cancer (Summary Last Modified 07/97), MDA-ID-91025, clinical trial, closed, 04/09/1997.
14. Feld R, Rubinstein L, Thomas PA, et al.: Adjuvant chemotherapy with cyclophosphamide, doxorubicin, and cisplatin in patients with completely resected stage I non-small-cell lung cancer. *Journal of the National Cancer Institute* 85(4): 299-306, 1993.
15. Niiranen A, Niitamo-Korhonen S, Kouri M, et al.: Adjuvant chemotherapy after radical surgery for non-small-cell lung cancer: a randomized study. *Journal of Clinical Oncology* 10(12): 1927-1932, 1992.
16. Furuse K, Fukuoka M, Kato H, et al.: A prospective phase II study on photodynamic therapy with photofrin II for centrally located early-stage lung cancer. *Journal of Clinical Oncology* 11(10): 1852-1857, 1993.

STAGE II NON-SMALL CELL LUNG CANCER

T1, N1, M0 or T2, N1, M0 or T3, N0, M0

Surgery is the treatment of choice for patients with stage II non-small cell lung cancer (NSCLC). Careful preoperative assessment of the patient's overall medical condition, especially the patient's pulmonary reserve, is critical in considering the benefits of surgery. The immediate postoperative mortality rate is age-related, but up to 5%-8% with pneumonectomy or 3%-5% with lobectomy can be expected.

Inoperable patients with stage II disease and with sufficient pulmonary reserve may be considered for radiotherapy with curative intent.[1] Among patients with excellent performance status, up to a 20% 3-year survival rate may be expected if a course of radiotherapy with curative intent can be completed. In the largest retrospective series reported to date, 152 patients with medically inoperable NSCLC treated with definitive radiotherapy achieved a 5-year overall survival rate of 10%; however, the 44 patients with T1 tumors achieved an actuarial disease-free survival rate of 60%. This retrospective study also suggested that improved disease-free survival was obtained with radiotherapy doses greater than 6,000 cGy.[2] Primary radiotherapy should consist of approximately 6,000 cGy delivered with megavoltage equipment to the midplane of the volume of known tumor using conventional fractionation. A boost to the cone-down field of the primary tumor is frequently used to further enhance local control. Careful treatment planning with precise definition of target volume and avoidance of critical normal structures to the extent possible is needed for optimal results and requires the use of a simulator.

Many patients treated surgically subsequently develop regional or distant metastases.[3] Therefore, patients should be considered for entry into clinical trials evaluating the use of adjuvant treatment with chemotherapy or radiotherapy following surgery.[4] One controlled trial has failed to demonstrate an overall survival benefit for patients with carefully staged squamous cell carcinoma receiving postoperative irradiation, although local recurrences were significantly reduced.[5] In two controlled trials in carefully staged, surgically resected patients, adjuvant combination chemotherapy with cisplatin, doxorubicin, and cyclophosphamide produced modestly increased disease-free survival and a trend toward improved overall survival, especially in the first year after surgery.[6,7] Based on these data, participation in clinical trials evaluating adjuvant chemotherapy after surgical resection should be encouraged.

Treatment options:

1. Lobectomy, pneumonectomy, or segmental, wedge, or sleeve resection as appropriate.
2. Radiotherapy with curative intent (for potentially operable patients who have medical contraindications to surgery).
3. Radiotherapy combined with curative surgery.[8]
4. Adjuvant chemotherapy combined with other modalities.[4,6-8]

References:

1. Komaki R, Cox JD, Hartz AJ, et al.: Characteristics of long-term survivors after treatment for inoperable carcinoma of the lung. *American Journal of Clinical Oncology* 8(5): 362-370, 1985.
2. Dosoretz DE, Katin MJ, Blitzer PH, et al.: Radiation therapy in the management of medically inoperable carcinoma of the lung: results and implications for future treatment strategies. *International Journal of Radiation Oncology, Biology, Physics* 24(1): 3-9, 1992.
3. Martini N, Bains MS, Burt ME, et al.: Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *Journal of Thoracic and Cardiovascular Surgery* 109(1): 120-129, 1995.

4. Keller SM, Eastern Cooperative Oncology Group: NCI HIGH PRIORITY CLINICAL TRIAL -- - Phase III Comparison of Thoracic Radiotherapy Alone vs Combined with CDDP/VP-16 Chemotherapy as Postoperative Adjuvant Therapy for Stage II/IIIA Non-small Cell Lung Cancer (Summary Last Modified 06/97), EST-3590, clinical trial, closed, 02/04/1997.
5. Weisenburger TH, Holmes EC, Gail M, et al.: Effects of postoperative mediastinal radiation on completely resected stage II and stage III epidermoid cancer of the lung. *New England Journal of Medicine* 315(22): 1377-1381, 1986.
6. Holmes AC: Surgical adjuvant therapy for stage II and stage III adenocarcinoma and large cell undifferentiated carcinoma. *Chest* 106(6 Suppl): 293S-296S, 1994.
7. Lad T, Rubinstein L, Sadeghi A: The benefit of adjuvant treatment for resected locally advanced non-small-cell lung cancer. *Journal of Clinical Oncology* 6(1): 9-17, 1988.
8. Holmes EC: Adjuvant treatment in resected lung cancer. *Seminars in Surgical Oncology* 6(5): 263-267, 1990.

STAGE IIIA NON-SMALL CELL LUNG CANCER

Some citations in the text of this section are followed by a level of evidence. The PDQ editorial boards use a formal ranking system to help the reader judge the strength of evidence linked to the reported results of a therapeutic strategy. Refer to the PDQ levels of evidence summary for more information.

T1, N2, M0 or T2, N2, M0 or T3, N1, M0 or T3, N2, M0

Depending on clinical circumstances, the principal forms of treatment that are considered for patients with stage III non-small cell lung cancer (NSCLC) are radiotherapy, chemotherapy, surgery, and combinations of these modalities. Although the majority of these patients do not achieve a complete response to radiotherapy, there is a reproducible long-term survival benefit in 5%-10% of patients treated with standard fractionation to 6,000 cGy, and significant palliation often results. Patients with excellent performance status and those who require a thoracotomy to prove that surgically unresectable tumor is present are most likely to benefit from radiotherapy.[1] Because of the poor long-term results, these patients should be considered for clinical trials. Trials examining fractionation schedules, endobronchial laser therapy, brachytherapy, and combined modality approaches may lead to improvement in the control of this regional disease.[2] One prospective randomized clinical study showed that radiotherapy given as three daily fractions improved overall survival compared to radiotherapy given as one daily fraction.[3][Level of evidence: 1iiA]

The addition of chemotherapy to radiotherapy has been reported to improve survival in prospective clinical studies that have used modern cisplatin-based chemotherapy regimens.[4-7] A meta-analysis of patient data from 11 randomized clinical trials showed that cisplatin-based combinations plus radiotherapy resulted in 10% reduction in the risk of death compared with radiotherapy alone.[8] The optimal sequencing of modalities and schedule of drug administration remains to be determined and is under study in ongoing clinical trials.[9]

Patients with N2 disease apparent on chest radiograph and documented by biopsy or discovered by prethoracotomy exploration have a 5-year survival rate of only about 2%. The use of preoperative (neoadjuvant) chemotherapy has been shown to be effective in these clinical situations in two small randomized studies of a total of 120 patients with stage IIIa NSCLC.[10,11] The 58 patients randomized to three cycles of cisplatin-based chemotherapy followed by surgery had a median survival more than three times as long as patients treated with surgery but no chemotherapy in both these studies. Two additional single-arm studies have evaluated either two to four cycles of combination chemotherapy or combination chemotherapy plus chest irradiation for 211 patients with histologically confirmed N2 stage IIIa NSCLC.[12] Sixty-five percent to 75% of patients were able to have a resection of their cancer, and 27%-28% were alive at 3 years. These results are encouraging, and combined-modality therapy with neoadjuvant chemotherapy with surgery and/or chest radiotherapy should be considered for patients with good performance status who have stage IIIa NSCLC.

Although most retrospective studies suggest that postoperative radiotherapy can improve local control for node-positive patients whose tumors were resected, it remains controversial whether it can improve survival.[13,14] One controlled trial in patients with completely resected stage II or III squamous cell lung cancer failed to demonstrate a survival benefit for patients who received postoperative irradiation,

although local recurrences were significantly reduced.[15] In two controlled trials with carefully staged surgically resected patients, adjuvant combination chemotherapy with cisplatin, doxorubicin, and cyclophosphamide produced modestly increased disease-free survival and a trend toward improved survival, especially in the first year after surgery.[16-18] Based on these data, participation in clinical trials evaluating adjuvant chemotherapy after surgical resection should be encouraged.

No consistent benefit from any form of immunotherapy has been demonstrated thus far in the treatment of NSCLC.

Treatment options:

1. Surgery alone in highly selected cases.[19-21]
 2. Chemotherapy combined with other modalities.[4-6,12,16-18]
 3. Surgery and postoperative radiotherapy.[13,15,22]
 4. Radiotherapy alone.[1,2]
-

Superior sulcus tumor (T3, N0 or N1, M0)

Another category that merits a special approach is that of superior sulcus tumors, a locally invasive problem usually with a reduced tendency for distant metastases. Consequently, local therapy has curative potential, especially for T3, N0 disease. Radiotherapy alone, radiotherapy preceded or followed by surgery, or surgery alone (in highly selected cases) may be curative in some patients, with a 5-year survival rate of 20% or more in some studies.[23] Patients with more invasive tumors of this area, or true Pancoast tumors, have a worse prognosis and generally do not benefit from primary surgical management. Follow-up surgery may be used to verify complete response in the radiotherapy field and to resect necrotic tissue.

Treatment options:

1. Radiotherapy and surgery.
 2. Radiotherapy alone.
 3. Surgery alone (selected cases).
 4. Chemotherapy combined with other modalities.
 5. Brachytherapy.[24]
 6. Clinical trials of combined modality therapy.[22]
-

Chest wall tumor (T3, N0 or N1, M0)

Selected patients with bulky primary tumors that directly invade the chest wall can obtain long-term survival with surgical management provided that their tumor is completely resected.

Treatment options:

1. Surgery.[\[21,25\]](#)
2. Surgery and radiotherapy.
3. Radiotherapy alone.
4. Chemotherapy combined with other modalities.

References:

1. Komaki R, Cox JD, Hartz AJ, et al.: Characteristics of long-term survivors after treatment for inoperable carcinoma of the lung. *American Journal of Clinical Oncology* 8(5): 362-370, 1985.
2. Johnson DH, Einhorn LH, Bartolucci A, et al.: Thoracic radiotherapy does not prolong survival in patients with locally advanced, unresectable non-small cell lung cancer. *Annals of Internal Medicine* 113(1): 33-38, 1990.
3. Saunders M, Dische S, Barrett A, et al.: Continuous hyperfractionated accelerated radiotherapy (CHART) versus conventional radiotherapy in non-small-cell lung cancer: a randomised multicentre trial. *Lancet* 350(9072): 161-165, 1997.
4. Dillman RO, Seagren SL, Propert KJ, et al.: A randomized trial of induction chemotherapy plus high-dose radiation versus radiation alone in stage III non-small-cell lung cancer. *New England Journal of Medicine* 323(14): 940-945, 1990.
5. LeChevalier T, Arriagada R, Quoix E, et al.: Radiotherapy alone versus combined chemotherapy and radiotherapy in nonresectable non-small-cell lung cancer: first analysis of a randomized trial in 353 patients. *Journal of the National Cancer Institute* 83(6): 417-423, 1991.
6. Schaake-Koning C, Van dan Bogaert W, Dalesio O, et al.: Effects of concomitant cisplatin and radiotherapy on inoperable non-small-cell lung cancer. *New England Journal of Medicine* 326(8): 524-530, 1992.
7. Sause WT, Scott C, Taylor S, et al.: Radiation Therapy Oncology Group (RTOG) 88-08 and Eastern Cooperative Oncology Group (ECOG) 4588: preliminary results of a phase III trial in regionally advanced, unresectable non-small-cell lung cancer. *Journal of the National Cancer Institute* 87(3): 198-205, 1995.
8. Non-small Cell Lung Cancer Collaborative Group: Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *British Medical Journal* 311(7010): 899-909, 1995.
9. Curran WJ, Radiation Therapy Oncology Group: Phase III Randomized Study of Standard Thoracic Irradiation Following VBL/CDDP vs Standard Thoracic Irradiation and Concurrent VBL/CDDP vs Hyperfractionated Thoracic Irradiation and Concurrent VP-16/CDDP for Locally Advanced, Unresectable, non-Small Cell Lung Cancer (Summary Last Modified 09/98), RTOG-9410, clinical trial, closed, 07/31/1998.
10. Rosell R, Gomez-Codina J, Camps C, et al.: A randomized trial comparing preoperative chemotherapy plus surgery with surgery alone in patients with non-small-cell lung cancer. *New England Journal of Medicine* 330(3): 153-158, 1994.

11. Roth JA, Fossella F, Komaki R, et al.: A randomized trial comparing perioperative chemotherapy and surgery with surgery alone in resectable stage IIIA non-small-cell lung cancer. *Journal of the National Cancer Institute* 86(9): 673-680, 1994.
12. Albain KS, Rusch VW, Crowley JJ, et al.: Concurrent cisplatin/etoposide plus chest radiotherapy followed by surgery for stages IIIA(N2) and IIIB non-small-cell lung cancer: mature results of Southwest Oncology Group phase II study 8805. *Journal of Clinical Oncology* 13(8): 1880-1892, 1995.
13. Emami B, Kaiser L, Simpson J, et al.: Postoperative radiation therapy in non-small cell lung cancer. *American Journal of Clinical Oncology* 20(5): 441-448, 1997.
14. Sawyer TE, Bonner JA, Gould PM, et al.: Effectiveness of postoperative irradiation in stage IIIA non-small cell lung cancer according to regression tree analyses of recurrence risks. *Annals of Thoracic Surgery* 64(5): 1402-1408, 1997.
15. Weisenburger TH, Holmes EC, Gail M, et al.: Effects of postoperative mediastinal radiation on completely resected stage II and stage III epidermoid cancer of the lung. *New England Journal of Medicine* 315(22): 1377-1381, 1986.
16. Niiranen A, Niitamo-Korhonen S, Kouri M, et al.: Adjuvant chemotherapy after radical surgery for non-small-cell lung cancer: a randomized study. *Journal of Clinical Oncology* 10(12): 1927-1932, 1992.
17. Holmes EC: Adjuvant treatment in resected lung cancer. *Seminars in Surgical Oncology* 6(5): 263-267, 1990.
18. Lad T, Rubinstein L, Sadeghi A: The benefit of adjuvant treatment for resected locally advanced non-small-cell lung cancer. *Journal of Clinical Oncology* 6(1): 9-17, 1988.
19. Shields TW: The significance of ipsilateral mediastinal lymph node metastasis (N2 disease) in non-small cell carcinoma of the lung: a commentary. *Journal of Thoracic and Cardiovascular Surgery* 99(1): 48-53, 1990.
20. Mountain CF: The biological operability of stage III non-small cell lung cancer. *Annals of Thoracic Surgery* 40(1): 60-64, 1985.
21. Van Raemdonck DE, Schneider A, Ginsberg RJ: Surgical treatment for higher stage non-small cell lung cancer. *Annals of Thoracic Surgery* 54(5): 999-1013, 1992.
22. Keller SM, Eastern Cooperative Oncology Group: NCI HIGH PRIORITY CLINICAL TRIAL -- - Phase III Comparison of Thoracic Radiotherapy Alone vs Combined with CDDP/VP-16 Chemotherapy as Postoperative Adjuvant Therapy for Stage II/IIIA Non-small Cell Lung Cancer (Summary Last Modified 06/97), EST-3590, clinical trial, closed, 02/04/1997.
23. Komaki R, Mountain CF, Holbert JM, et al.: Superior sulcus tumors: treatment selection and results for 85 patients without metastasis (M0) at presentation. *International Journal of Radiation Oncology, Biology, Physics* 19(1): 31-36, 1990.
24. Miller JJ, Phillips TW: Neodymium:YAG laser and brachytherapy in the management of inoperable bronchogenic carcinoma. *Annals of Thoracic Surgery* 50(2): 190-196, 1990.
25. McCaughan BC, Martini N, Bains MS, et al.: Chest wall invasion in carcinoma of the lung: therapeutic and prognostic implications. *Journal of Thoracic and Cardiovascular Surgery* 89(6): 836-841, 1985.

STAGE IIIB NON-SMALL CELL LUNG CANCER

Patients with stage IIIB non-small cell lung cancer (NSCLC) do not benefit from surgery alone and are best managed by initial chemotherapy, chemotherapy plus radiotherapy, or radiotherapy alone, depending

on sites of tumor involvement and performance status. Most patients with excellent performance status should be considered for combined modality therapy. However, patients with malignant pleural effusion are rarely candidates for radiotherapy, and should generally be treated similarly to stage IV patients (see separate section of this summary on treatment of stage IV disease). Many randomized studies of unresectable patients with stage III NSCLC show that treatment with neoadjuvant or concurrent cisplatin-based chemotherapy and chest irradiation is associated with improved survival compared to treatment with radiotherapy alone.[1-5] A meta-analysis of patient data from 11 randomized clinical trials showed that cisplatin-based combinations plus radiotherapy resulted in 10% reduction in the risk of death compared with radiotherapy alone.[6]

Patients with stage IIIb disease with poor performance status are candidates for chest irradiation to palliate pulmonary symptoms (e.g., cough, shortness of breath, or local chest pain). No consistent benefit from any form of immunotherapy has been demonstrated thus far.

T4 or N3, M0

An occasional patient with supraclavicular node involvement who is otherwise a good candidate for irradiation with curative intent will survive 3 years. Although the majority of these patients do not achieve a complete response to radiotherapy, significant palliation often results. Patients with excellent performance status and those who are found to have advanced-stage disease at the time of resection are most likely to benefit from radiotherapy.[7] Adjuvant systemic chemotherapy with radiotherapy has been tested in randomized trials for patients with inoperable or unresectable locoregional NSCLC.[1-3,8] Some patients have shown a modest survival advantage with adjuvant chemotherapy. The addition of chemotherapy to radiotherapy has been reported to improve long-term survival in some,[1,3,4] but not all,[9] prospective clinical studies. A meta-analysis of patient data from 54 randomized clinical trials showed an absolute survival benefit of 4% at 2 years with the addition of cisplatin-based chemotherapy to radiotherapy.[10] One study showed improvement in local control and survival when cisplatin was given daily with concurrent radiotherapy but not when it was given weekly.[4] It is not yet clear whether the schedule of drug administration is responsible for the variation in results; the optimal sequencing of modalities remains to be determined and is under study in ongoing clinical trials.[11]

Because of the poor overall results, these patients should be considered for clinical trials; trials examining fractionation schedules, radiosensitizers, radiolabeled antibodies, and combined modality approaches may lead to improvement in the control of regional disease.

Patients with NSCLC can present with superior vena cava syndrome. Refer to the PDQ supportive care summary on superior vena cava syndrome for more information. Regardless of stage, this problem should generally be managed with radiotherapy with or without chemotherapy.

Treatment options:

1. Radiotherapy alone.
2. Chemotherapy combined with radiotherapy.[1-3,8]
3. Chemotherapy and concurrent radiotherapy followed by resection.[12,13]
4. Chemotherapy alone.

References:

1. LeChevalier T, Arriagada R, Quoi E, et al.: Radiotherapy alone versus combined chemotherapy and radiotherapy in nonresectable non-small-cell lung cancer: first analysis of a randomized trial in 353 patients. *Journal of the National Cancer Institute* 83(6): 417-423, 1991.
2. Morton RF, Jett JR, McGinnis WL, et al.: Thoracic radiation therapy alone compared with combined chemoradiotherapy for locally unresectable non-small cell lung cancer. *Annals of Internal Medicine* 115(9): 681-686, 1991.

3. Dillman RO, Seagren SL, Propert KJ, et al.: A randomized trial of induction chemotherapy plus high-dose radiation versus radiation alone in stage III non-small-cell lung cancer. *New England Journal of Medicine* 323(14): 940-945, 1990.
4. Schaake-Koning C, Van dan Bogaert W, Dalesio O, et al.: Effects of concomitant cisplatin and radiotherapy on inoperable non-small-cell lung cancer. *New England Journal of Medicine* 326(8): 524-530, 1992.
5. Sause WT, Scott C, Taylor S, et al.: Radiation Therapy Oncology Group (RTOG) 88-08 and Eastern Cooperative Oncology Group (ECOG) 4588: preliminary results of a phase III trial in regionally advanced, unresectable non-small-cell lung cancer. *Journal of the National Cancer Institute* 87(3): 198-205, 1995.
6. Non-small Cell Lung Cancer Collaborative Group: Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *British Medical Journal* 311(7010): 899-909, 1995.
7. Komaki R, Cox JD, Hartz AJ, et al.: Characteristics of long-term survivors after treatment for inoperable carcinoma of the lung. *American Journal of Clinical Oncology* 8(5): 362-370, 1985.
8. Ihde DC: Chemotherapy combined with chest irradiation for locally advanced non-small cell lung cancer. *Annals of Internal Medicine* 115(9): 737-739, 1991.
9. Blanke C, Ansari R, Mantravadi R, et al.: Phase III trial of thoracic irradiation with or without cisplatin for locally advanced unresectable non-small-cell lung cancer: a Hoosier Oncology Group protocol. *Journal of Clinical Oncology* 13(6): 1425-1429, 1995.
10. Pignon JP, Stewart LA, Souhami RL, et al.: A meta-analysis using individual patient data from randomised clinical trials (RCTS) of chemotherapy (CT) in non-small cell lung cancer (NSCLC): (2) survival in the locally advanced (LA) setting. *Proceedings of the American Society of Clinical Oncology* 13: A-1109, 334, 1994.
11. Curran WJ, Radiation Therapy Oncology Group: Phase III Randomized Study of Standard Thoracic Irradiation Following VBL/CDDP vs Standard Thoracic Irradiation and Concurrent VBL/CDDP vs Hyperfractionated Thoracic Irradiation and Concurrent VP-16/CDDP for Locally Advanced, Unresectable, non-Small Cell Lung Cancer (Summary Last Modified 09/98), RTOG-9410, clinical trial, closed, 07/31/1998.
12. Rusch VW, Albain KS, Crowley JJ, et al.: Surgical resection of stage IIIA and stage IIIB non-small-cell lung cancer after concurrent induction chemoradiotherapy: a Southwest Oncology Group trial. *Journal of Thoracic and Cardiovascular Surgery* 105(1): 97-106, 1993.
13. Albain KS, Rusch VW, Crowley JJ, et al.: Concurrent cisplatin/etoposide plus chest radiotherapy followed by surgery for stages IIIA(N2) and IIIB non-small-cell lung cancer: mature results of Southwest Oncology Group phase II study 8805. *Journal of Clinical Oncology* 13(8): 1880-1892, 1995.

STAGE IV NON-SMALL CELL LUNG CANCER

Any T, any N, M1

Cisplatin-containing and carboplatin-containing combination chemotherapy regimens produce objective response rates (including a few complete responses) that are higher than those achieved with single-agent chemotherapy. Although toxic effects may vary, outcome is similar with most cisplatin-containing regimens; a randomized trial comparing five cisplatin-containing regimens showed no significant difference in response, duration of response, or survival.[1] Patients with good performance status and a limited number of sites of distant metastases have superior response and survival when given chemotherapy when compared to other patients.[2] A prospective randomized comparison of vinorelbine

plus cisplatin versus vindesine plus cisplatin versus single agent vinorelbine has reported improved response rate (30%) and median survival (40 weeks) with the vinorelbine plus cisplatin regimen.[3] Two small phase II studies reported that paclitaxel (Taxol) has single-agent activity in stage IV patients, with response rates in the range of 21%- 24%.[4,5] Reports of paclitaxel combinations have shown relatively high response rates, significant 1 year survival, and palliation of lung cancer symptoms.[6] With the paclitaxel plus carboplatin regimen, response rates have been in the range of 27%-53% with 1-year survival rates of 32%-54%.[6,7] The combination of cisplatin and paclitaxel was shown to have a higher response rate and higher 1 year survival rate than the combination of cisplatin and etoposide.[8] Additional clinical studies should better define the role of these newer combination chemotherapy regimens in the treatment of advanced non- small cell lung cancer.[8] Meta-analyses have shown that chemotherapy produces modest benefits in short-term survival compared to supportive care alone in patients with inoperable stages IIIb and IV disease.[9-11]

Although these results support further evaluation of chemotherapeutic approaches for both metastatic and locally advanced non-small cell lung cancer (NSCLC), efficacy of current programs is such that no specific regimen can be regarded as standard therapy. Appropriate patients should be encouraged to participate in clinical trials. Outside of a clinical trial setting, chemotherapy should be given only to patients with good performance status and evaluable tumor lesions who desire such treatment after being fully informed of its anticipated risks and limited benefits.

Radiotherapy may be effective in palliating symptomatic local involvement with NSCLC such as tracheal, esophageal, or bronchial compression, bone or brain metastases, pain, vocal cord paralysis, hemoptysis, or superior vena cava syndrome. In some cases, endobronchial laser therapy and/or brachytherapy has been used to alleviate proximal obstructing lesions.[12] Such therapeutic intervention may be critical in the prolongation of an acceptable lifestyle in an otherwise functional patient. In the rare patient with synchronous presentation of a resectable primary tumor in the lung and a single brain metastasis, surgical resection of the solitary brain lesion is indicated with resection of the primary tumor and appropriate postoperative chemotherapy and/or irradiation of the primary tumor site and with postoperative whole-brain irradiation delivered in daily fractions of 180-200 cGy to avoid long-term toxic effects to normal brain tissue.[13,14]

In asymptomatic patients kept under close observation, treatment may often be appropriately deferred until symptoms or signs of progressive tumor develop.

Treatment options:

1. External-beam radiotherapy, primarily for palliative relief of local symptomatic tumor growth.
2. Chemotherapy. The following regimens produce similar survival outcomes: cisplatin plus vinblastine.[1] cisplatin plus vinblastine plus mitomycin.[15] cisplatin plus vinorelbine.[3] cisplatin plus vindesine.[3] cisplatin plus paclitaxel.[8] carboplatin plus paclitaxel.[6,7]
3. Clinical trials evaluating the role of new chemotherapy regimens. Refer to the clinical trials section of PDQ for a list of clinical trials. The clinical trials in PDQ are also available on CancerNet (<http://cancernet.nci.nih.gov/>).
4. Endobronchial laser therapy and/or brachytherapy for obstructing lesions.[12]

References:

1. Weick JK, Crowley J, Natale RB, et al.: A randomized trial of five cisplatin-containing treatments in patients with metastatic non-small-cell lung cancer: a Southwest Oncology Group study. *Journal of Clinical Oncology* 9(7): 1157-1162, 1991.
2. O'Connell JP, Kris MG, Gralla RJ, et al.: Frequency and prognostic importance of pretreatment clinical characteristics in patients with advanced non-small-cell lung cancer treated with combination chemotherapy. *Journal of Clinical Oncology* 4(11): 1604-1614, 1986.

3. Le Chevalier T, Brisgand D, Douillard JY, et al.: Randomized study of vinorelbine and cisplatin versus vindesine and cisplatin versus vinorelbine alone in advanced non-small-cell lung cancer: results of a European multicenter trial including 612 patients. *Journal of Clinical Oncology* 12(2): 360-367, 1994.
4. Chang AY, Kim K, Glick J, et al.: Phase II study of taxol, merbarone, and piroxantrone in stage IV non-small-cell lung cancer: the Eastern Cooperative Oncology Group results. *Journal of the National Cancer Institute* 85(5): 388-394, 1993.
5. Murphy WK, Fossella FV, Winn RJ, et al.: Phase II study of taxol in patients with untreated advanced non-small-cell lung cancer. *Journal of the National Cancer Institute* 85(5): 384-388, 1993.
6. Johnson DH, Paul DM, Hande KR, et al.: Paclitaxel plus carboplatin in advanced non-small-cell lung cancer: a phase II trial. *Journal of Clinical Oncology* 14(7): 2054-2060, 1996.
7. Langer CJ, Leighton JC, Comis RL, et al.: Paclitaxel and carboplatin in combination in the treatment of advanced non-small-cell lung cancer: a phase II toxicity, response, and survival analysis. *Journal of Clinical Oncology* 13(8): 1860-1870, 1995.
8. Bonomi P, Kim K, Chang A, et al.: Phase III trial comparing etoposide (E) cisplatin (C) versus taxol (T) with cisplatin-G-CSF(G) versus taxol-cisplatin in advanced non-small cell lung cancer. An Eastern Cooperative Oncology Group (ECOG) trial. *Proceedings of the American Society of Clinical Oncology* 15: A-1145, 382, 1996.
9. Souquet PJ, Chauvin F, Boissel JP, et al.: Polychemotherapy in advanced non small cell lung cancer: a meta-analysis. *Lancet* 342(8862): 19-21, 1993.
10. Grilli R, Oxman AD, Julian JA: Chemotherapy for advanced non-small-cell lung cancer: how much benefit is enough? *Journal of Clinical Oncology* 11(10): 1866-1872, 1993.
11. Marino P, Pampallona S, Preatoni A, et al.: Chemotherapy vs supportive care in advanced non-small-cell lung cancer: results of a meta-analysis of the literature. *Chest* 106(3): 861-865, 1994.
12. Miller JI, Phillips TW: Neodymium:YAG laser and brachytherapy in the management of inoperable bronchogenic carcinoma. *Annals of Thoracic Surgery* 50(2): 190-196, 1990.
13. Mandell L, Hilaris B, Sullivan M, et al.: The treatment of single brain metastasis from non-oat cell lung carcinoma: surgery and radiation versus radiation therapy alone. *Cancer* 58(3): 641-649, 1986.
14. DeAngelis LM, Mandell LR, Thaler HT, et al.: The role of postoperative radiotherapy after resection of single brain metastases. *Neurosurgery* 24(6): 798-805, 1989.
15. Veeder MH, Jett JR, Su JQ, et al.: A phase III trial of mitomycin C alone versus mitomycin C, vinblastine, and cisplatin for metastatic squamous cell lung carcinoma. *Cancer* 70(9): 2281-2287, 1992.

RECURRENT NON-SMALL CELL LUNG CANCER

Many patients with recurrent non-small cell lung cancer (NSCLC) are eligible for clinical trials. Radiotherapy may provide excellent palliation of symptoms from a localized tumor mass. Patients who present with a solitary cerebral metastasis after resection of a primary NSCLC lesion and who have no evidence of extracranial tumor can achieve prolonged disease-free survival with surgical excision of the brain metastasis and postoperative whole-brain irradiation.^[1,2] Unresectable brain metastases in this setting may be treated radiosurgically.^[3] Because of the small potential for long-term survival, radiotherapy should be delivered by conventional methods in daily doses of 180-200 cGy, while higher daily doses over a shorter period of time (hypofractionated schemes) should be avoided because of

the high risk of toxic effects observed with such treatments.[4] Most patients not suitable for surgical resection should receive conventional whole-brain radiotherapy. Selected patients with good performance status and small metastases can be considered for stereotactic radiosurgery.[5]

Approximately one half of patients treated with resection and postoperative radiotherapy will develop recurrence in the brain; some of these patients will be suitable for additional treatment.[6] In those selected patients with good performance status and without progressive metastases outside of the brain, treatment options include reoperation or stereotactic radiosurgery.[3,6] For most patients, conventional radiotherapy can be considered; however, the palliative benefit of this treatment is limited.[7]

A solitary pulmonary metastasis from an initially resected bronchogenic carcinoma is unusual. The lung is frequently the site of second primary malignancies in patients with primary lung cancers. Determining whether the new lesion is a new primary cancer or a metastasis may be difficult. Studies have indicated that in the majority of patients the new lesion is a second primary tumor, and following resection some patients may achieve long-term survival. Thus, if the first primary tumor has been controlled, the second primary tumor should be resected if possible.[8,9]

The use of chemotherapy has produced objective responses and small improvement in survival for patients with metastatic disease.[10] In studies that have examined symptomatic response, improvement in subjective symptoms has been reported to occur more frequently than objective response.[11,12] Informed patients with good performance status and symptomatic recurrence can be offered treatment with a cisplatin-based chemotherapy regimen for palliation of symptoms.

Treatment options:

1. Palliative radiotherapy.
2. Chemotherapy alone.
3. Surgical resection of isolated cerebral metastasis (highly selected patients).[6]
4. Laser therapy or interstitial radiotherapy for endobronchial lesions.[13]
5. Stereotactic radiosurgery (highly selected patients).[3,5,14]

References:

1. Patchell RA, Tibbs PA, Walsh JW, et al.: A randomized trial of surgery in the treatment of single metastases to the brain. *New England Journal of Medicine* 322(8): 494-500, 1990.
2. Mandell L, Hilaris B, Sullivan M, et al.: The treatment of single brain metastasis from non-oat cell lung carcinoma: surgery and radiation versus radiation therapy alone. *Cancer* 58(3): 641-649, 1986.
3. Loeffler JS, Kooy HM, Wen PY, et al.: The treatment of recurrent brain metastases with stereotactic radiosurgery. *Journal of Clinical Oncology* 8(4): 576-582, 1990.
4. DeAngelis LM, Mandell LR, Thaler HT, et al.: The role of postoperative radiotherapy after resection of single brain metastases. *Neurosurgery* 24(6): 798-805, 1989.
5. Alexander E, Moriarty TM, Davis RB, et al.: Stereotactic radiosurgery for the definitive, noninvasive treatment of brain metastases. *Journal of the National Cancer Institute* 87(1): 34-40, 1995.
6. Arbit E, Wronski M, Burt M, et al.: The treatment of patients with recurrent brain metastases: a retrospective analysis of 109 patients with nonsmall cell lung cancer. *Cancer* 76(5): 765-773, 1995.
7. Hazuka MB, Kinzie JJ: Brain metastases: results and effects of re-irradiation. *International Journal of Radiation Oncology, Biology, Physics* 15(2): 433-437, 1988.

8. Salerno TA, Munro DD, Blundell PE, et al.: Second primary bronchogenic carcinoma: life-table analysis of surgical treatment. *Annals of Thoracic Surgery* 27(1): 3-6, 1979.
9. Yellin A, Hill LR, Benfield JR: Bronchogenic carcinoma associated with upper aerodigestive cancer. *Journal of Thoracic and Cardiovascular Surgery* 91(5): 674-683, 1986.
10. Souquet PJ, Chauvin F, Boissel JP, et al.: Polychemotherapy in advanced non small cell lung cancer: a meta-analysis. *Lancet* 342(8862): 19-21, 1993.
11. Ellis PA, Smith IE, Hardy JR, et al.: Symptom relief with MVP (mitomycin C, vinblastine and cisplatin) chemotherapy in advanced non-small-cell lung cancer. *British Journal of Cancer* 71(2): 366-370, 1995.
12. Medical Research Council Lung Cancer Working Party: Randomized trial of etoposide cyclophosphamide methotrexate and vincristine versus etoposide and vincristine in the palliative treatment of patients with small-cell lung cancer and poor prognosis. *British Journal of Cancer* 67(Suppl 20): A-4;2, 14, 1993.
13. Miller JI, Phillips TW: Neodymium:YAG laser and brachytherapy in the management of inoperable bronchogenic carcinoma. *Annals of Thoracic Surgery* 50(2): 190-196, 1990.
14. Loeffler JS, Kooy HM, Wen PY, et al.: The treatment of recurrent brain metastases with stereotactic radiosurgery. *Journal of Clinical Oncology* 8(4): 576-582, 1990.

Date Last Modified: 12/98

TEXTO 11: Treatment summaries for health professionals (QDT3)

----- SMALL CELL LUNG CANCER -----

***** GENERAL INFORMATION

Without treatment, small cell carcinoma of the lung has the most aggressive clinical course of any type of pulmonary tumor, with median survival from diagnosis of only 2-4 months. Compared with other cell types of lung cancer, small cell carcinoma has a greater tendency to be widely disseminated by the time of diagnosis, but is much more responsive to chemotherapy and irradiation.

Because of its propensity for distant metastases, localized forms of treatment, such as surgical resection or radiotherapy, rarely produce long-term survival [1]. With incorporation of current chemotherapy regimens into the treatment program, however, survival is unequivocally prolonged, with at least a 4- to 5-fold improvement in median survival compared with patients who are given no therapy. Furthermore, about 10% of the total population of patients remain free of disease over two years from the start of therapy, the time period during which most relapses occur. However, even these patients are at risk of dying from lung cancer (both small and non-small cell types) [2]. The overall survival at 5 years is 5%-10% [2-4].

At the time of diagnosis, approximately 40% of patients with small cell carcinoma will have tumor confined to the hemithorax of origin, the mediastinum, or the supraclavicular lymph nodes. These patients are designated as having limited stage disease, and most 2-year disease-free survivors come from this group. In limited stage disease, median survival of 16-24 months with current forms of treatment can reasonably be expected [5-8]. A small proportion of patients with limited stage disease may benefit from surgery with or without adjuvant chemotherapy; these patients have an even better prognosis. Patients with tumor that has spread beyond the supraclavicular areas are said to have extensive stage disease and have a worse prognosis than patients with limited stage. Median survival of 6-12 months is reported with currently available therapy, but long-term disease-free survival is rare.

The pretreatment prognostic factors which consistently predict for prolonged survival include good performance status, female gender, and limited stage disease [3,9,10]. Patients with involvement of the central nervous system or liver at the time of diagnosis have a significantly worse outcome [3,9-11]. In general, patients who are confined to bed tolerate aggressive forms of treatment poorly, have increased morbidity, and rarely attain 2-year disease-free survival. However, patients with poor performance status can often derive significant palliative benefit and prolongation of survival from treatment.

Regardless of stage, the current prognosis for patients with small cell lung cancer is unsatisfactory even though considerable improvements in diagnosis and therapy have been made over the past 10-15 years. Therefore, all patients with this type of cancer may appropriately be considered for inclusion in clinical trials at the time of diagnosis.

<References>

***** CELLULAR CLASSIFICATION

Review of pathologic material by an experienced lung cancer pathologist is important prior to initiating treatment of any patient with small cell lung cancer. The intermediate subtype of small cell carcinoma and the more readily recognized lymphocyte-like or "oat cell" subtype are equally responsive to treatment.

The current classification of subtypes of small cell lung cancer are:[1]

- small cell carcinoma;
- mixed small cell/large cell carcinoma;

combined small cell carcinoma (small cell lung cancer combined with neoplastic squamous and/or glandular components).

There is increasing evidence that light microscopy has some limitations as a means of classifying bronchogenic carcinomas, particularly small cell carcinomas. Electron microscopy, which can detect neuroendocrine granules, may help to differentiate between small cell and non-small cell cancers [2]. Neuroendocrine carcinomas of the lung represent a spectrum of disease. At one extreme is small cell lung cancer, which has a poor prognosis. At the other extreme are bronchial carcinoids, with an excellent prognosis after surgical excision [3]. Between these extremes is an unusual entity called well-differentiated neuroendocrine carcinoma of the lung [4]. It has been referred to as malignant carcinoid, metastasizing bronchial adenoma, pleomorphic carcinoid, nonbenign carcinoid tumor, and atypical carcinoid. Like small cell lung cancer, it occurs primarily in cigarette smokers, but it metastasizes less frequently. The 5-year survival rate is greater than 50% in some series, and surgical cure appears possible in most stage I patients. Careful diagnosis is important, however, since the differential pathologic diagnosis from small cell lung cancer may be difficult.

<References>

***** STAGE INFORMATION

Staging procedures do not currently have a major impact on treatment of small cell lung cancer since patients should initially receive combination chemotherapy regardless of the extent of tumor dissemination. However, determining the stage of cancer by nonsurgical means allows a better assessment of prognosis and identifies sites of tumor that can be evaluated for response. In clinical situations where the choice of treatment is affected by stage, particularly when chest irradiation or surgical excision is added to chemotherapy alone for limited stage patients, results of the staging process have therapeutic implications. Staging procedures commonly used to document distant metastases include bone marrow examination, computed tomographic or magnetic resonance imaging scans of the brain, computerized tomographic scans of the chest and the abdomen, and radionuclide bone scans.

Because occult or overt metastatic disease is present at diagnosis in most patients, survival is usually not affected by small differences in the amount of locoregional tumor involvement. Therefore, the detailed TNM staging system developed for lung cancer by the American Joint Committee on Cancer (AJCC) is not commonly employed in patients with small cell carcinoma. A simple 2-stage system developed by the Veterans Administration Lung Cancer Study Group is most commonly used.

The international staging system, which is outlined in detail in the PDQ summary for non-small cell lung cancer, may also be used, particularly for the small minority of patients who may be candidates for surgical resection.

---- Limited stage

Limited stage small cell lung cancer means tumor confined to the hemithorax of origin, the mediastinum, and the supraclavicular nodes, which can be encompassed within a "tolerable" radiotherapy port. There is no universally accepted definition of this term, and patients with pleural effusion, massive pulmonary tumor, and contralateral supraclavicular nodes have been both included within and excluded from limited stage by various groups.

---- Extensive stage

Extensive stage small cell lung cancer means tumor that is too widespread to be included within the definition of limited stage disease [1,2].

<References>

***** TREATMENT OPTION OVERVIEW

In small cell lung cancer, the majority of patients die of their tumor despite state-of-the-art treatment. Most of the improvements in survival in small cell lung cancer are attributable to clinical trials which have attempted to improve on the best available, accepted therapy. Patient entry into such studies is highly desirable.

Methods under clinical evaluation in small cell lung cancer include adding chest radiation to chemotherapy regimens, varying drug doses in current regimens, alternating different combinations of chemotherapy, using different schedules of chemotherapeutic agents, and using new drug regimens composed of standard and new agents.

Prospective randomized trials have not demonstrated a consistent survival advantage for patients treated with higher doses of chemotherapy [1-3]. Even chemotherapy of the intensity used in autologous bone marrow transplant regimens has not clearly been shown to improve survival in patients with small cell lung cancer [4,5].

The designations in PDQ that treatments are "standard" or "under clinical evaluation" are not to be used as a basis for reimbursement determinations.

<References>

***** LIMITED STAGE SMALL CELL LUNG CANCER

In patients with small cell lung cancer, combination chemotherapy produces results that are clearly superior to single-agent treatment, and moderately intensive doses of drugs are superior to doses that produce only minimal or mild hematologic toxicity. Current programs yield overall objective response rates of 65%-90% and complete response rates of 45%-75%. Because of the frequent presence of occult metastatic disease, chemotherapy is the cornerstone of treatment of limited stage small cell lung cancer. Combinations containing two or more drugs are needed for maximal effect.

Mature results of prospective randomized trials suggest that combined modality therapy produces a modest but significant improvement in survival compared with chemotherapy alone. Two meta-analyses showed an improvement in 3-year survival rates of about 5% for those receiving chemotherapy and radiotherapy compared to those receiving chemotherapy alone [1,2]. Most of the benefit occurred in patients less than 65 years of age. Combined modality treatment is associated with increased morbidity and, in some trials, increased treatment-related mortality from pulmonary and hematologic toxic effects; proper administration requires close collaboration between medical and radiation oncologists [3]. In general, those studies showing a positive effect for combined modality therapy employed thoracic irradiation early in the course of treatment, concurrently with chemotherapy [3-6].

Studies strongly suggest that minimal tumor doses in the range of 4,000 to 4,500 cGy or more (standard fractionation) are necessary to effectively control tumors in the thorax.

The combination of etoposide and cisplatin chemotherapy with concurrent chest radiotherapy has now been used in multiple single institutional studies and in cooperative group studies. These studies have consistently achieved median survivals of 18-24 months and 40%-50% 2-year survival with less than 3% treatment-related mortality [3-9]. Once daily and twice daily chest radiation schedules have been used in regimens with etoposide and cisplatin, and a randomized study has not shown a significant survival advantage for patients treated with twice daily chest radiotherapy compared to once daily [9]. The current standard treatment of patients with limited stage small cell lung cancer should be a combination containing etoposide and cisplatin plus chest radiotherapy administered during the first or second cycle of chemotherapy administration.

The relative effectiveness of 2-5 drug regimens and different schedules of chest radiotherapy appear to be similar. A representative selection of regimens incorporating chemotherapy plus chest radiotherapy are listed below. The use of alternating chemotherapy regimens has not proven more effective than the consistent administration of a single regimen [3,6,8-12]. The optimal duration of chemotherapy for patients with limited stage small cell lung cancer is not clearly defined but there is no improvement in survival after the duration of drug administration exceeds 3-6 months [3,8,9,13]. There is no evidence from randomized trials that maintenance chemotherapy prolongs survival for patients with limited stage small cell lung cancer [10,14].

Patients presenting with superior vena cava syndrome are treated with combination chemotherapy with or without radiation therapy [15,16]. A small minority of limited stage patients with adequate pulmonary function and with tumor pathologically confined to the lung of origin, or the lung and ipsilateral hilar lymph nodes, may possibly benefit from surgical resection with or without adjuvant chemotherapy [17-20].

Patients with small cell lung cancer treated with chemotherapy with or without chest irradiation who have achieved a complete remission can be considered for administration of prophylactic cranial irradiation (PCI). Patients whose cancer can be controlled outside the brain have a 60% actuarial risk of developing central nervous system metastases within 2-3 years after starting treatment [21,22]. The majority of these patients relapse only in their brain and nearly all of those who relapse in their central nervous system die of their cranial metastases [3,8,22]. The risk of developing central nervous system metastases can be reduced by more than 50% by the administration of PCI in doses of 2400 cGy [22]. Retrospective studies have shown that long-term survivors of small cell lung cancer (>2 years from the start of treatment) have a high incidence of central nervous system impairment [23-25]. However, prospective studies have shown that patients treated with PCI do not have detectably different neuropsychological function than patients not treated [22]. In addition, the majority of patients with small cell lung cancer have neuropsychological abnormalities present before the start of cranial irradiation and have no detectable decline in their neurological status up to 2 years after the start of their cranial irradiation [26]. Patients treated for small cell lung cancer continue to have declining neuropsychologic function after 2 years from the start of treatment [23-25]. Therefore, additional neuropsychologic testing of patients beyond 2 years from the start of treatment will be needed before concluding that PCI does not contribute to the decline in intellectual function.

Treatment options:

Standard:

1 Combination chemotherapy with one of the following regimens and chest irradiation (with or without PCI given to patients with complete responses): The following regimens produce similar survival outcomes: EC: etoposide + cisplatin + 4000-4500 cGy chest radiotherapy [3,7-9]. ECV: etoposide + cisplatin + vincristine + 4500 cGy chest radiotherapy [5].

2 Combination chemotherapy (with or without PCI in patients with complete responses), especially in patients with impaired pulmonary function or poor performance status.

3 Surgical resection followed by chemotherapy or chemotherapy plus chest radiotherapy (with or without PCI in patients with complete responses) for patients in highly selected cases [17-20].

Under clinical evaluation:

Areas of active clinical evaluation in limited stage small cell lung cancer include new drug regimens, variation of drug doses in current regimens, surgical resection of the primary tumor, new radiotherapy schedules and techniques (e.g, 3-dimensional treatment planning), and timing of thoracic radiation [27-29].

<References>

***** EXTENSIVE STAGE SMALL CELL LUNG CANCER

As in limited stage small cell carcinoma, chemotherapy should be given as multiple agents in doses associated with at least moderate toxicity in order to produce the best results in extensive stage disease. Doses and schedules used in current programs yield overall response rates of 70%-85% and complete response rates of 20%-30% in extensive stage disease. Since overt disseminated disease is present, combination chemotherapy is the cornerstone of treatment of this stage of small cell lung cancer. Combinations containing two or more drugs are needed for maximal benefit.

The relative effectiveness of many two- to four-drug combination programs appears similar, and there are a large number of potential combinations. Therefore, a representative selection of regimens that have been found to be effective by at least two independent groups has been provided. Some physicians have administered two of these or other regimens in alternating sequences, but there is no proof that this strategy yields substantial survival improvement [1-3]. Optimal duration of chemotherapy is not clearly defined, but there is no obvious improvement in survival when the duration of drug administration exceeds 6 months [4,5]. There is no clear evidence from reported data that maintenance chemotherapy will improve survival duration [6-9].

Combination chemotherapy plus chest irradiation does not appear to improve survival compared with chemotherapy alone in extensive stage small cell lung cancer. However, radiotherapy plays an extremely important role in palliation of symptoms of the primary tumor and of metastatic disease, particularly brain, epidural, and bone metastases.

Chest irradiation is sometimes given for superior vena cava syndrome, but chemotherapy alone (with irradiation reserved for nonresponding patients) is appropriate initial treatment. Brain metastases are appropriately treated with whole-brain radiotherapy. However, intracranial metastases from small cell carcinoma may respond to chemotherapy as readily as metastases in other organs [9,10].

Patients with small cell lung cancer treated with chemotherapy with or without chest irradiation who have achieved a complete remission can be considered for administration of prophylactic cranial irradiation (PCI). Patients whose cancer can be controlled outside the brain have a 60% actuarial risk of developing central nervous system metastases within 2-3 years after starting treatment [11,12]. The majority of these patients relapse only in their brain and nearly all of those who relapse in their central nervous system die of their cranial metastases [12-14]. The risk of developing central nervous system metastases can be reduced by more than 50% by the administration of PCI in doses of 2400 cGy [12]. Retrospective studies have shown that long-term survivors of small cell lung cancer (>2 years from the start of treatment) have a high incidence of central nervous system impairment [15-17]. However, prospective studies have shown that patients treated with PCI do not have detectably different neuropsychological function than patients not treated [12]. In addition, the majority of patients with small cell lung cancer have neuropsychological abnormalities present before the start of cranial irradiation and have no detectable decline in their neurological status up to 2 years after the start of their cranial irradiation [18]. Patients treated for small cell lung cancer continue to have declining neuropsychologic function after 2 years from the start of treatment [15-17]. Therefore, additional neuropsychologic testing of patients beyond 2 years from the start of treatment will be needed before concluding that PCI does not contribute to the decline in intellectual function.

Many more patients with extensive stage small cell carcinoma have greatly impaired performance status at the time of diagnosis when compared to patients with limited stage disease. Such patients have a poor prognosis and tolerate aggressive chemotherapy or combined modality therapy poorly. Single-agent intravenous, oral, and low-dose biweekly regimens have been developed for these patients [19-23]. However, prospective randomized studies have shown that patients with a poor prognosis who are treated with conventional regimens live longer than those treated with the single-agent or low-dose regimens [22-24].

Treatment options:

Standard:

1 Combination chemotherapy with one of the following regimens with or without PCI given to patients with complete responses: The following regimens produce similar survival outcomes: CAV: cyclophosphamide + doxorubicin + vincristine [25,26]. CAE: cyclophosphamide + doxorubicin + etoposide [27]. EP or EC: etoposide + cisplatin or carboplatin [28,29]. ICE: ifosfamide + carboplatin + etoposide [30].

Other regimens appear to produce similar survival outcomes but have been studied less extensively or are in less common use, including:

- cyclophosphamide + methotrexate + lomustine [31].
- cyclophosphamide + methotrexate + lomustine + vincristine [32].
- cyclophosphamide + doxorubicin + etoposide + vincristine [33].
- CEV: cyclophosphamide + etoposide + vincristine [34].
- single-agent etoposide or teniposide [19,20].

2 Radiotherapy to sites of metastatic disease unlikely to be immediately palliated by chemotherapy, especially brain, epidural, and bone metastases.

3 Identification of effective new agents is difficult in patients who have previously been treated with standard chemotherapy because response rates to agents, even of known efficacy, are known to be lower than in previously untreated patients. This situation led to the suggestion that patients with extensive disease who are medically stable be treated with new agents under evaluation, with provisions for early change to standard combination therapy if there is no response [35]. Such a strategy has been shown to be feasible, with survival comparable to survival with initial standard therapy, as long as the patients with extensive disease are carefully chosen [36-38]. A variety of other strategies have been proposed, depending on the activity of the new agent in other tumors, in preclinical small cell lung cancer models, or the activity of drug analogs [39]. Active single agents undergoing further evaluation include teniposide, paclitaxel, and topotecan [19,40,41].

Under clinical evaluation:

Areas of active clinical evaluation in extensive stage small cell lung cancer include evaluation of new drug regimens, variation of drug doses, alternative drug schedules, and bone marrow transplants [42,43].

<References>

***** RECURRENT SMALL CELL LUNG CANCER

The prognosis for small cell lung carcinoma that has progressed despite chemotherapy is exceedingly poor regardless of stage. Expected median survival is 2-3 months. These patients should be considered for palliative therapy or clinical trials. Patients who are primarily resistant to chemotherapy and those who have received multiple chemotherapy regimens rarely respond to additional treatment. However, patients who have initially responded and relapsed more than 6 months following initial treatment are more likely to respond to additional chemotherapy [1-8].

Some patients with intrinsic endobronchial obstructing lesions or extrinsic compression due to tumor have achieved successful palliation with endobronchial laser therapy (for endobronchial lesions only) and/or brachytherapy [9]. Expandable metal stents can be safely inserted under local anesthesia via the bronchoscope, resulting in improved symptoms and pulmonary function in patients with malignant airways obstruction [10]. Patients with progressive intrathoracic tumor after failing initial chemotherapy can achieve significant tumor responses, palliation of symptoms, and short-term local control with external-beam

radiotherapy. However, only the rare patient will experience long-term survival following "salvage" radiotherapy [11].

Patients with central nervous system recurrences can often obtain palliation of symptoms with radiotherapy and/or additional chemotherapy. The majority of patients treated with radiotherapy obtain objective responses and improvement following radiotherapy [12]. A retrospective review showed that 43% of patients treated with additional chemotherapy at the time of CNS relapse respond to second-line chemotherapy [13].

Treatment options:

1 Palliative radiotherapy [11].

2 Salvage chemotherapy can provide some palliative benefit for patients previously sensitive to standard chemotherapy [2-5,8].

3 Local palliation with endobronchial laser therapy, endobronchial stents, and/or brachytherapy [9,10].

4 Clinical trials of phase I or phase II drugs. Refer to the PDQ protocol directory for active clinical trials for patients with recurrent small cell lung cancer.

<References>

TEXTO 12: Treatment summaries for patients (QPT1)

----- MALIGNANT MESOTHELIOMA -----

***** OVERVIEW OF PDQ

---- What is PDQ?

PDQ is a computer system that gives up-to-date information on cancer and its prevention, detection, treatment, and supportive care. It is a service of the National Cancer Institute (NCI) for people with cancer and their families and for doctors, nurses, and other health care professionals.

To ensure that it remains current, the information in PDQ is reviewed and updated each month by experts in the fields of cancer treatment, prevention, screening, and supportive care. PDQ also provides information about research on new treatments (clinical trials), doctors who treat cancer, and hospitals with cancer programs. The treatment information in this summary is based on information in the PDQ summary for health professionals on this cancer.

---- How to use PDQ

PDQ can be used to learn more about current treatment of different kinds of cancer. You may find it helpful to discuss this information with your doctor, who knows you and has the facts about your disease. PDQ can also provide the names of additional health care professionals who specialize in treating patients with cancer.

Before you start treatment, you also may want to think about taking part in a clinical trial. PDQ can be used to learn more about these trials. A clinical trial is a research study that attempts to improve current treatments or finds information on new treatments for patients with cancer. Clinical trials are based on past studies and information discovered in the laboratory. Each trial answers certain scientific questions in order to find new and better ways to help patients with cancer. Information is collected about new treatments, their risks, and how well they do or do not work. When clinical trials show that a new treatment is better than the treatment currently used as "standard" treatment, the new treatment may become the standard treatment. Listings of current clinical trials are available on PDQ. Many cancer doctors who take part in clinical trials are listed in PDQ.

To learn more about cancer and how it is treated, or to learn more about clinical trials for your kind of cancer, call the National Cancer Institute's Cancer Information Service. The number is 1-800-4-CANCER (1-800-422-6237); TTY at 1-800-332-8615. The call is free and a trained information specialist will be available to answer cancer-related questions.

PDQ is updated whenever there is new information. Check with the Cancer Information Service to be sure that you have the most up-to-date information-----.

***** DESCRIPTION

---- What is malignant mesothelioma?

Malignant mesothelioma, a rare form of cancer, is a disease in which cancer (malignant) cells are found in the sac lining the chest (the pleura) or abdomen (the peritoneum). Most people with malignant mesothelioma have worked on jobs where they breathed asbestos.

A doctor should be seen if a person has shortness of breath, pain in the chest, or pain or swelling in the abdomen. If there are symptoms, the doctor may order an x-ray of the chest or abdomen.

The doctor may look inside the chest cavity with a special instrument called a thoracoscope. A cut will be made through the chest wall and the thoracoscope will be put into the chest between two ribs. This test, called thoracoscopy, is usually done in the hospital. Before the test, the patient will be given a local anesthetic (a drug that causes a loss of feeling for a short period of time). Some pressure may be felt, but usually there is no pain.

The doctor may also look inside the abdomen (peritoneoscopy) with a special tool called a peritoneoscope. The peritoneoscope is put into an opening made in the abdomen. This test is also usually done in the hospital. Before the test is done, a local anesthetic will be given.

If tissue that is not normal is found, the doctor will need to cut out a small piece and have it looked at under a microscope to see if there are any cancer cells. This is called a biopsy. Biopsies are usually done during the thoracoscopy or peritoneoscopy.

The chance of recovery (prognosis) depends on the size of the cancer, where the cancer is, how far the cancer has spread, how the cancer cells look under the microscope, how the cancer responds to treatment, and the patient's age.

***** STAGE EXPLANATION

----- Stages of malignant mesothelioma

Once malignant mesothelioma is found, more tests will be done to find out if cancer cells have spread to other parts of the body. This is called staging. A doctor needs to know the stage of the cancer to plan treatment. The following stages are used for malignant mesothelioma.

----- Localized malignant mesothelioma

Stage I: The cancer is found in the lining of the chest cavity near the lung and heart or in the diaphragm or the lung.

----- Advanced malignant mesothelioma

Stage II: The cancer has spread beyond the lining of the chest to lymph nodes in the chest.

Stage III: Cancer has spread into the chest wall, center of the chest, heart, through the diaphragm, or abdominal lining, and in some cases into nearby lymph nodes.

Stage IV: Cancer has spread to distant organs or tissues.

----- Recurrent malignant mesothelioma

Recurrent disease means that the cancer has come back (recurred) after it has been treated. It may come back in the lining of the chest or abdomen or in another part of the body.

***** TREATMENT OPTION OVERVIEW

----- How malignant mesothelioma is treated

There are treatments for all patients with malignant mesothelioma. Three kinds of treatment are used:

surgery (taking out the cancer) /
radiation therapy (using high-dose x-rays or other high-energy rays to kill cancer cells) /
chemotherapy (using drugs to fight the cancer).

Surgery is a common treatment of malignant mesothelioma. The doctor may remove part of the lining of the chest or abdomen and some of the tissue around it. Depending on how far the cancer has spread, a lung also may be removed in an operation called a pneumonectomy. Sometimes part of the diaphragm, the muscle below the lungs that helps with breathing, is also removed.

Radiation therapy uses high-energy x-rays to kill cancer cells and shrink tumors. Radiation may come from a machine outside the body (external radiation therapy) or from putting materials that produce radiation (radioisotopes) through thin plastic tubes in the area where the cancer cells are found (internal radiation therapy).

If fluid has collected in the chest or abdomen, the doctor may drain the fluid out of the body by putting a needle into the chest or abdomen and using gentle suction to remove the fluid. If fluid is removed from the chest, this is called thoracentesis. If fluid is removed from the abdomen, this is called paracentesis. The doctor may also put drugs through a tube into the chest to prevent more fluid from accumulating.

Chemotherapy uses drugs to kill cancer cells. Chemotherapy may be taken by pill, or it may be put into the body by a needle in the vein or muscle. Chemotherapy is called a systemic treatment because the drug enters the bloodstream, travels through the body, and can kill cancer cells throughout the body. In mesothelioma, chemotherapy may be put directly into the chest (intrapleural chemotherapy).

Intraoperative photodynamic therapy is a new type of treatment that uses special drugs and light to kill cancer cells during surgery. A drug that makes cancer cells more sensitive to light is injected into a vein several days before surgery. During surgery to remove as much of the cancer as possible, a special light is used to shine on the pleura. This treatment is being studied for early stages of mesothelioma in the chest.

----- Treatment by stage

Treatment depends on where the cancer is, how far it has spread, and the patient's age and general health.

Standard treatment may be considered because of its effectiveness in patients in past studies, or participation in a clinical trial may be considered. Not all patients are cured with standard therapy and some standard treatments may have more side effects than are desired. For these reasons, clinical trials are designed to find better ways to treat cancer patients and are based on the most up-to-date information. Clinical trials are ongoing in many parts of the country for many patients with malignant mesothelioma. To learn more about clinical trials, call the Cancer Information Service at 1-800-4-CANCER (1-800-422-6237); TTY at 1-800-332-8615.

***** LOCALIZED MALIGNANT MESOTHELIOMA (STAGE I)

If the cancer is only in one place in the chest or abdomen, treatment will probably be surgery to remove part of the pleura and some of the tissue around it.

If the cancer is found in a larger part of the pleura, treatment may be one of the following:

1 Surgery to remove the pleura and the tissue near it to relieve symptoms, with or without radiation therapy after surgery.

2 Surgery to remove sections of the pleura, the lung, part of the diaphragm, and part of the lining around the heart.

3 External beam radiation therapy to relieve symptoms.

4 A clinical trial of surgery followed by chemotherapy given inside the chest.

5 A clinical trial of surgery, radiation therapy, and/or chemotherapy.

***** ADVANCED MALIGNANT MESOTHELIOMA (STAGES II, III, AND IV)

Treatment may be one of the following:

1 Draining of fluid in the chest or abdomen (thoracentesis or paracentesis) to reduce discomfort. Drugs also may be put into the chest or abdomen to prevent further collection of fluid.

2 Surgery to relieve symptoms.

3 Radiation therapy to relieve symptoms.

4 Chemotherapy.

5 A clinical trial of surgery, radiation therapy, and chemotherapy.

6 Chemotherapy given in the chest or abdomen.

***** RECURRENT MALIGNANT MESOTHELIOMA

Treatment depends on many factors, including where the cancer came back and what treatment the patient received before. Clinical trials are testing new treatments.

<TO LEARN MORE>

TEXT0 13: Treatment summaries for patients (QPT2)

----- NON-SMALL CELL LUNG CANCER -----

***** DESCRIPTION

----- What is non-small cell lung cancer?

Lung cancers can be divided into two types: small cell lung cancer and non-small cell lung cancer. The cancer cells of each type grow and spread in different ways, and they are treated differently. Non-small cell lung cancer is usually associated with prior smoking, passive smoking, or radon exposure. (A separate patient information summary on small cell lung cancer is also available in PDQ).

The main kinds of non-small cell lung cancer are named for the type of cells found in the cancer: squamous cell carcinoma (also called epidermoid carcinoma), adenocarcinoma, large cell carcinoma, adenosquamous carcinoma, and undifferentiated carcinoma.

Non-small cell lung cancer is a common disease. It is usually treated by surgery (taking out the cancer in an operation) or radiation therapy (using high-dose x-rays to kill cancer cells). However, chemotherapy may be used in some patients.

The prognosis (chance of recovery) and choice of treatment depend on the stage of the cancer (whether it is just in the lung or has spread to other places), tumor size, the type of lung cancer, whether there are symptoms, and the patient's general health.

***** STAGE EXPLANATION

----- Stages of non-small cell lung cancer

Once lung cancer has been found (diagnosis), more tests will be done to find out if the cancer has spread from the lung to other parts of the body (staging). A doctor needs to know the stage to plan treatment. The following stages are used for non-small cell lung cancer:

----- Occult stage

Cancer cells are found in sputum, but no tumor can be found in the lung.

----- Stage 0

Cancer is only found in a local area and only in a few layers of cells. It has not grown through the top lining of the lung. Another term for this type of lung cancer is carcinoma in situ.

----- Stage I

The cancer is only in the lung, and normal tissue is around it.

----- Stage II

Cancer has spread to nearby lymph nodes.

----- Stage III

Cancer has spread to the chest wall or diaphragm near the lung; or the cancer has spread to the lymph nodes in the area that separates the two lungs (mediastinum); or to the lymph nodes on the other side of the chest or in the neck. Stage III is further divided into stage IIIA (usually can be operated on) and stage IIIB (usually cannot be operated on).

----- Stage IV

Cancer has spread to other parts of the body.

----- Recurrent

Cancer has come back (recurred) after previous treatment.

***** TREATMENT OPTION OVERVIEW

----- How non-small cell lung cancer is treated

Chemotherapy uses drugs to kill cancer cells. Chemotherapy may be taken by pill, or it may be put into the body by a needle in the vein or muscle. Chemotherapy is called a systemic treatment because the drug enters the bloodstream, travels through the body, and can kill cancer cells outside the lungs.

Chemoprevention uses drugs to prevent a second cancer from occurring.

Radiation therapy uses high-energy x-rays to kill cancer cells and shrink tumors. Radiation may come from a machine outside the body (external radiation therapy) or from putting materials that produce radiation (radioisotopes) through thin plastic tubes in the area where the cancer cells are found (internal radiation therapy).

One new type of radiation therapy is called radiosurgery. In radiosurgery, radiation is directly focused on the tumor, and involves as little normal tissue as possible. Radiosurgery is usually used as treatment of tumors that involve the brain.

Cryosurgery freezes the tumor and kills it. Photodynamic therapy uses a certain type of light and a special chemical to kill cancer cells. Laser therapy uses a narrow beam of light to kill cancer cells. Cryosurgery and photodynamic therapy are usually used in clinical trials.

Surgery, radiation therapy, and chemotherapy are used to treat non-small cell lung cancer. However, these treatments often do not cure the disease.

If lung cancer is found, a patient may want to think about taking part in one of the many clinical trials being done to improve treatment. Clinical trials are ongoing in most parts of the country for all stages of non-small cell lung cancer. Treatment choices can be discussed with a doctor.

Patients with non-small cell lung cancer can be divided into three groups, depending on the stage of the cancer and the treatment that is planned. The first group (stages 0, I, and II) includes patients whose cancers can be taken out by surgery. The operation that takes out only a small part of the lung is called a wedge resection. When a whole section (lobe) of the lung is taken out, the operation is called a lobectomy. When one whole lung is taken out, it is called a pneumonectomy.

Radiation therapy may be used to treat patients in this group who cannot have surgery because they have other medical problems. Like surgery, radiation therapy is called local treatment because it works only on the cells in the area being treated.

The second group of patients has lung cancer that has spread to nearby tissue or to lymph nodes. These patients can be treated with radiation therapy alone or with surgery and radiation, chemotherapy and radiation, or chemotherapy alone.

The third group of patients has lung cancer that has spread to other parts of the body. Radiation therapy may be used to shrink the cancer and to relieve pain. Chemotherapy may be used to treat some patients in this group.

***** OCCULT NON-SMALL CELL LUNG CANCER

Tests are done to find the main tumor (cancer). Lung cancer that is found at this early stage can be cured by surgery.

***** STAGE 0 NON-SMALL CELL LUNG CANCER

Treatment may be one of the following:

1 Surgery to cure these very early cancers. However, these patients may get a second lung cancer that may not be able to be taken out by surgery.

2 Photodynamic therapy used internally.

***** STAGE I NON-SMALL CELL LUNG CANCER

Treatment may be one of the following:

1 Surgery.

2 Radiation therapy (for patients who cannot be operated on).

3 Clinical trials of chemotherapy following surgery.

4 Clinical trials of chemoprevention following other therapy.

5 Clinical trials of photodynamic therapy used internally.

***** STAGE II NON-SMALL CELL LUNG CANCER

Treatment may be one of the following:

1 Surgery to take out the tumor and lymph nodes.

2 Radiation therapy (for patients who cannot be operated on).

3 Surgery and/or radiation therapy with or without chemotherapy.

***** STAGE III NON-SMALL CELL LUNG CANCER

---- Stage IIIA non-small cell lung cancer

Treatment may be one of the following:

1 Surgery alone.

2 Chemotherapy with other treatments.

3 Surgery and radiation therapy.

4 Radiation therapy alone.

5 Laser therapy and/or internal radiation therapy.

----- Stage IIIB non-small cell lung cancer

Treatment may be one of the following:

- 1 Radiation therapy alone.
- 2 Chemotherapy plus radiation therapy.
- 3 Chemotherapy plus radiation therapy followed by surgery.
- 4 Chemotherapy alone.
- 5 Cryotherapy plus radiation therapy.

***** STAGE IV NON-SMALL CELL LUNG CANCER

Treatment may be one of the following:

- 1 Radiation therapy.
- 2 Chemotherapy.
- 3 Chemotherapy and radiation therapy.
- 4 Laser therapy and/or internal radiation therapy.

***** RECURRENT NON-SMALL CELL LUNG CANCER

Treatment may be one of the following:

- 1 Radiation therapy to control symptoms.
- 2 Chemotherapy.
- 3 Chemotherapy with radiotherapy.
- 4 For some patients who have a very small amount of tumor that has spread to the brain, surgery may be used to remove the tumor.
- 5 Laser therapy or internal radiation therapy.
- 6 Radiosurgery (for certain patients who cannot be operated on).

<TO LEARN MORE>

TEXTO 14: Treatment summaries for patients (QPT3)

----- SMALL CELL LUNG CANCER -----

***** DESCRIPTION

----- What is small cell lung cancer?

Small cell lung cancer is a disease in which cancer (malignant) cells are found in the tissues of the lungs. The lungs are a pair of cone-shaped organs that take up much of the room inside the chest. The lungs bring oxygen into the body and take out carbon dioxide, which is a waste product of the body's cells. Tubes called bronchi make up the inside of the lungs.

There are two kinds of lung cancer based on how the cells look under a microscope: small cell and non-small cell. If a patient has non-small cell lung cancer, see the PDQ patient information summary on non-small cell lung cancer.

Small cell lung cancer is usually found in people who smoke or who used to smoke cigarettes. A doctor should be seen if there are any of the following symptoms: a cough or chest pain that doesn't go away, a wheezing sound when breathing, shortness of breath, coughing up blood, hoarseness, or swelling in the face and neck.

If there are symptoms, a doctor may want to look into the bronchi through a special instrument, called a bronchoscope, that slides down the throat and into the bronchi. This test, called bronchoscopy, is usually done in the hospital. Before the test, the patient will be given a local anesthetic (a drug that causes a loss of feeling for a short period of time) in the back of the throat. Some pressure may be felt, usually with no pain. The doctor can take cells from the walls of the bronchi tubes or cut small pieces of tissue to look at under the microscope to see if there are any cancer cells. This is called a biopsy.

The doctor may also use a needle to remove tissue from a place in the lung that may be hard to reach with the bronchoscope. A cut will be made in the skin and the needle will be put in between the ribs. This is called a needle aspiration biopsy. The doctor will look at the tissue under the microscope to see if there are any cancer cells. Before the test, a local anesthetic will be given to keep the patient from feeling pain.

The chance of recovery (prognosis) and choice of treatment depend on the stage of the cancer (whether it is just in the lung or has spread to other places), and the patient's gender and general state of health.

***** STAGE EXPLANATION

----- Stages of small cell lung cancer

Once small cell lung cancer has been found, more tests will be done to find out if cancer cells have spread from one or both lungs to other parts of the body (staging). A doctor needs to know the stage of the disease to plan treatment. The following stages are used for small cell lung cancer:

----- Limited stage

Cancer is found only in one lung and in nearby lymph nodes. (Lymph nodes are small, bean-shaped structures that are found throughout the body. They produce and store infection-fighting cells.)

----- Extensive stage

Cancer has spread outside of the lung where it began to other tissues in the chest or to other parts of the body.

----- Recurrent stage

Recurrent disease means that the cancer has come back (recurred) after it has been treated. It may come back in the lungs or in another part of the body.

***** TREATMENT OPTION OVERVIEW

----- How small cell lung cancer is treated

There are treatments for all patients with small cell lung cancer. Three kinds of treatment are used:

surgery (taking out the cancer) /

radiation therapy (using high-dose x-rays or other high-energy rays to kill cancer cells) /

chemotherapy (using drugs to kill cancer cells).

Additionally, clinical trials are testing the effect of new therapies on the treatment of small cell lung cancer. Surgery may be used if the cancer is found only in one lung and in nearby lymph nodes. Because this type of lung cancer is usually not found in only one lung, surgery alone is not often used. Occasionally, surgery may be used to help determine exactly which type of lung cancer the patient has. If a patient does have surgery, the doctor may take out the cancer in one of the following operations:

Wedge resection removes only a small part of the lung.

Lobectomy removes an entire section (lobe) of the lung.

Pneumonectomy removes the entire lung.

During surgery, the doctor will also take out lymph nodes to see if they contain cancer.

Radiation therapy uses x-rays or other high-energy rays to kill cancer cells and shrink tumors. Radiation therapy for small cell lung cancer usually comes from a machine outside the body (external beam radiation therapy). It may be used to kill cancer cells in the lungs or in other parts of the body where the cancer has spread. Radiation therapy may also be used to prevent the cancer from growing in the brain. This is called prophylactic cranial irradiation (PCI). Because PCI may affect brain function, the doctor will help the patient decide whether to have this kind of radiation therapy. Radiation therapy can be used alone or in addition to surgery and/or chemotherapy.

Chemotherapy is the most common treatment of all stages of small cell lung cancer. Chemotherapy may be taken by pill, or it may be put into the body by a needle in the vein or muscle. Chemotherapy is called a systemic treatment because the drug enters the bloodstream, travels through the body, and can kill cancer cells outside the lungs, including cancer cells that have spread to the brain.

----- Treatment by stage

Treatment of small cell lung cancer depends on the stage of the disease, and the patient's age and overall condition.

Standard treatment may be considered because of its effectiveness in patients in past studies, or participation in a clinical trial may be considered. Most patients are not cured with standard therapy and some standard treatments may have more side effects than are desired. For these reasons, clinical trials

are designed to find better ways to treat cancer patients and are based on the most up-to-date information. Clinical trials are ongoing in most parts of the country for most stages of small cell lung cancer. To learn more about clinical trials, call the Cancer Information Service at 1-800-4-CANCER (1-800-422-6237); TTY at 1-800-332-8615.

***** LIMITED STAGE SMALL CELL LUNG CANCER

Treatment may be one of the following:

1 Chemotherapy and radiation therapy to the chest with or without radiation therapy to the brain to prevent spread of the cancer (prophylactic cranial irradiation).

2 Chemotherapy with or without prophylactic cranial irradiation.

3 Surgery followed by chemotherapy with or without prophylactic cranial irradiation.

Clinical trials are testing new drugs and new ways of giving all of the above treatments.

***** EXTENSIVE STAGE SMALL CELL LUNG CANCER

----- Treatment may be one of the following:

1 Chemotherapy with or without radiation therapy to the brain to prevent spread of the cancer (prophylactic cranial irradiation).

2 Radiation therapy to places in the body where the cancer has spread, such as the brain, bone, or spine to relieve symptoms.

Clinical trials are testing new drugs and new ways of giving all of the above treatments.

***** RECURRENT SMALL CELL LUNG CANCER

----- Treatment may be one of the following:

1 Radiation therapy to reduce discomfort.

2 Chemotherapy to reduce discomfort.

3 Laser therapy, radiation therapy, and/or surgical implantation of devices to keep the airways open to relieve discomfort.

4 A clinical trial testing new drugs.

<TO LEARN MORE>

TEXTO 15: Versión bitexto (inglés-español) de QDT1

164 Sentences and headings

<!--L1, S 1-->----- MALIGNANT MESOTHELIOMA -----

<!--L2, S 1-->Mesotelioma maligno

<!--L1, S 2-->***** GENERAL INFORMATION

<!--L2, S 2-->***** INFORMACION GENERAL

<!--L1, S 3-->Prognosis in this disease is difficult to assess consistently because there is great variability in the time before diagnosis and the rate of disease progression.

<!--L2, S 3-->El pronóstico de esta enfermedad es difícil de evaluar de manera consistente debido a la alta variabilidad de tiempo que existe antes del diagnóstico y la tasa de progresión de la enfermedad.

<!--L1, S 4-->Various surgical procedures may be possible in selected patients, providing long-term survival without cure.

<!--L2, S 4-->Es posible efectuar varios procedimientos quirúrgicos en pacientes seleccionados, proporcionando supervivencia sin curación a largo plazo.

<!--L1, S 5-->In large retrospective series of pleural mesothelioma patients, important prognostic factors were found to be stage, age, performance status, and histology [1,2].

<!--L2, S 5-->En series retrospectivas extensas de pacientes con mesotelioma pleural, se encontró que los factores de diagnóstico importantes fueron etapa, edad, nivel de funcionamiento e histología [1,2].

<!--L1, S 6-->For patients treated with aggressive surgical approaches, factors associated with improved long-term survival include epithelial histology, negative lymph nodes, and negative surgical margins [3,4].

<!--L2, S 6-->Para pacientes tratados con estrategias quirúrgicas agresivas, los factores relacionados con una supervivencia mejor a largo plazo incluyen histología epitelial, ganglios linfáticos negativos y márgenes quirúrgicos negativos [3,4].

<!--L1, S 7-->For those patients treated with aggressive surgical approaches, nodal status is an important prognostic factor [3].

<!--L2, S 7-->Para aquellos pacientes tratados con enfoques quirúrgicos agresivos, el estado ganglionar es un factor pronóstico importante [3].

<!--L1, S 8-->Median survival for malignant local pleural disease has been reported as 16 months, and extensive disease as 5 months.

<!--L2, S 8-->Se ha reportado que la supervivencia promedio de enfermedad pleural maligna local es de alrededor 16 meses, y en el caso de enfermedad extensa, de 5 meses.

<!--L1, S 9-->In some instances the tumor grows through the diaphragm making site of origin difficult to assess.

<!--L2, S 9-->En algunos casos el tumor crece a través del diafragma haciendo así difícil de detectar el sitio de origen.

<!--L1, S 10-->Cautious interpretation of treatment results in this disease is imperative because of the selection differences among series.

<!--L2, S 10-->La interpretación cautelosa de los resultados del tratamiento en esta enfermedad es indispensable debido a las diferencias de selección entre las series.

<!--L1, S 11-->Effusions, both pleural and peritoneal, represent major symptomatic problems for at least two thirds of the patients.

<!--L2, S 11-->Los derrames, tanto pleurales como peritoneales, representan problemas sintomáticos importantes en al menos dos tercios de los pacientes.

<!--L1, S 12-->A history of asbestos exposure is reported in about 70%-80% of all cases of mesothelioma [1,5,6].

<!--L2, S 12-->Se reporta una historia de exposición a asbestos en cerca del 70%-80% de todos los casos de mesotelioma [1,5,6].

<!--L1, S 13--><References>

<!--L2, S 13--><Bibliografía>

<!--L1, S 14-->***** CELLULAR CLASSIFICATION

<!--L2, S 14-->***** CLASIFICACION CELULAR

<!--L1, S 15-->Histologically, these tumors are composed of fibrous or epithelial elements or both.

<!--L2, S 15-->Histológicamente estos tumores están compuestos por elementos fibrosos o epiteliales o ambos.

<!--L1, S 16-->The epithelial form occasionally causes confusion with peripheral anaplastic lung carcinomas or metastatic carcinomas.

<!--L2, S 16-->La forma epitelial ocasionalmente causa confusión con carcinomas anaplásicos periféricos de pulmón o carcinomas metastáticos.

<!--L1, S 17-->Attempts at diagnosis by cytology or needle biopsy of the pleura are often noncontributory.

<!--L2, S 17-->Los intentos diagnósticos por citología o biopsia con aguja de la pleura a menudo ayudan.

<!--L1, S 18-->It can be especially difficult to differentiate mesothelioma from carcinoma on small tissue specimens.

<!--L2, S 18-->Puede ser especialmente difícil diferenciar el mesotelioma de un carcinoma en muestras pequeñas de tejido.

<!--L1, S 19-->Thoracoscopy can be valuable in obtaining adequate tissue specimens for diagnostic purposes [1].

<!--L2, S 19-->La toracoscopia puede ser valiosa para la obtención de especímenes adecuados de tejido con fines de diagnóstico [1].

<!--L1, S 20-->Examination of the gross tumor at surgery and use of special stains or electron microscopy can often help.

<!--L2, S 20-->El examen del tumor macroscópico durante la cirugía y el uso de colorantes especiales o microscopia electrónica pueden a menudo ayudar.

<!--L1, S 21-->The special stains reported to be most useful include periodic acid-Schiff diastase, hyaluronic acid, mucicarmine, CEA, and Leu M1 [2].

<!--L2, S 21-->Se ha reportado que los siguientes colorantes especiales son los más útiles: diastasa con ácido peryódico de Schiff, ácido hialurónico, mucicarmín, CEA y Leu M1 [2].

<!--L1, S 22-->Histologic appearance appears to be of prognostic value, with most clinical studies showing that epithelial mesotheliomas have a better prognosis than fibrous or sarcomatous mesotheliomas [2-4].

<!--L2, S 22-->La apariencia histológica parece tener valor pronóstico, al mostrar la mayoría de los estudios clínicos que los mesoteliomas epiteliales tienen un mejor pronóstico que los mesoteliomas fibrosos o sarcomatosos [2-4].

<!--L1, S 23--><References>

<!--L2, S 23--><Bibliografía>

<!--L1, S 24-->***** STAGE INFORMATION

<!--L2, S 24-->***** INFORMACION DE LAS ETAPAS

<!--L1, S 25-->Patients with stage I disease have a significantly better prognosis than those with more advanced stages.

<!--L2, S 25-->Los pacientes con enfermedad en etapa I tienen un pronóstico significativamente mejor que aquéllos con etapas más avanzadas.

<!--L1, S 26-->However, because of the relative rarity of this disease, exact survival information based upon stage is limited [1].

<!--L2, S 26-->Sin embargo, debido a que esta enfermedad es poco común, la información exacta de supervivencia basada en etapas es limitada [1].

<!--L1, S 27-->A proposed staging system based upon thoracic surgery principles and clinical data is shown below [2].

<!--L2, S 27-->Un sistema de clasificación propuesto basado en principios quirúrgicos torácicos y datos clínicos se brinda a continuación [2].

<!--L1, S 28-->It is a modification of the older system proposed by Butchart et al [3].

<!--L2, S 28-->Es una modificación del sistema más antiguo propuesto por Butchart [3].

<!--L1, S 29-->Other staging systems that have been employed, including a proposed new international TNM staging system, are summarized by the International Mesothelioma Interest Group [4].

<!--L2, S 29-->Otros sistemas de clasificación que se han empleado, incluyendo un nuevo sistema internacional de clasificación de TNM propuesto, han sido resumidos por el International Mesothelioma Interest Group [4].

<!--L1, S 30-->Stage I: Disease confined within the capsule of the parietal pleura: ipsilateral pleura, lung, pericardium, and diaphragm.

<!--L2, S 30-->Etapa I: Enfermedad confinada dentro de la cápsula de la pleura parietal: pleura ipsilateral, pulmón, pericardio y diafragma.

<!--L1, S 31-->Stage II: All of stage I with positive intrathoracic (N1 or N2) lymph nodes.

<!--L2, S 31-->Etapa II: Todo lo de la etapa I con ganglios linfáticos intratorácicos positivos (N1 o N2)

<!--L1, S 32-->Stage III: Local extension of disease into the following: chest wall or mediastinum; heart or through the diaphragm, peritoneum; with or without extrathoracic or contralateral (N3) lymph node involvement.

<!--L2, S 32-->Etapa III: Extensión local de la enfermedad a: pared torácica o mediastino; corazón o a través del diafragma, peritoneo; con o sin complicación extratorácica o contralateral (N3) de los ganglios linfáticos.

<!--L1, S 33-->Stage IV: Distant metastatic disease.

<!--L2, S 33-->Etapa IV: Enfermedad metastática distante

<!--L1, S 34-->----- Localized malignant mesothelioma: See description of stage I above.

<!--L2, S 34-->----- Mesotelioma maligno localizado: Ver descripción de etapa I arriba.

<!--L1, S 35-->----- Advanced malignant mesothelioma: See descriptions of stages II, III, and IV above.

<!--L2, S 35-->----- Mesotelioma maligno avanzado: Ver descripción de etapas II, III y IV arriba.

<!--L1, S 36-->For the purposes of the discussion of treatment in this statement, the disease is categorized as either localized or advanced.

<!--L2, S 36-->Para fines de discusión del tratamiento en este documento, la enfermedad se categoriza ya sea como localizada o avanzada.

<!--L1, S 37--><References>

<!--L2, S 37--><Bibliografía>

<!--L1, S 38-->***** TREATMENT OPTION OVERVIEW

<!--L2, S 38-->***** ASPECTOS DE LAS OPCIONES DE TRATAMIENTO

<!--L1, S 39-->Standard treatment for all but localized mesothelioma is generally not curative.

<!--L2, S 39-->El tratamiento estándar para todos los casos de mesotelioma con excepción del mesotelioma localizado es generalmente no curativo.

<!--L1, S 40-->Although some patients will experience long-term survival with aggressive treatment approaches, it remains unclear if overall survival has been significantly altered by the different treatment modalities or by combinations of modalities.

<!--L2, S 40-->Aunque algunos pacientes experimentarán supervivencia a largo plazo con estrategias agresivas de tratamiento, aún no está claro si la supervivencia en general ha sido alterada significativamente por las diferentes modalidades de tratamiento o por las combinaciones de modalidades.

<!--L1, S 41-->Extrapleural pneumonectomy in selected patients with early stage disease may improve recurrence-free survival, but its impact on overall survival is unknown [1].

<!--L2, S 41-->Una neumonectomía extrapleural en pacientes seleccionados con enfermedad en etapa inicial puede obtener una supervivencia libre de recidiva, pero su impacto en la supervivencia general es desconocido [1].

<!--L1, S 42-->Pleurectomy and decortication can provide palliative relief from symptomatic effusions, discomfort caused by tumor burden, and pain caused by invasive tumor.

<!--L2, S 42-->Pleurectomía y decorticación pueden brindar alivio paliativo contra efusiones sintomáticas, malestar causado por el agobio tumoral y dolor causado por invasión tumoral.

<!--L1, S 43-->Operative mortality from pleurectomy/decortication is less than 2%, [2]. while mortality from extrapleural pneumonectomy has ranged from 6% to 30% [1,3].

<!--L2, S 43-->La mortalidad quirúrgica por pleurectomía/decorticación es menos del 2%, [2] mientras que la mortalidad por neumonectomía extrapleural ha oscilado entre 6% y 30% [1,3].

<!--L1, S 44-->The addition of radiation therapy and/or chemotherapy following surgical intervention has not demonstrated improved survival [2].

<!--L2, S 44-->La adición de radioterapia y/o quimioterapia después de intervención quirúrgica no ha demostrado mejoras en la supervivencia [2].

<!--L1, S 45-->The use of radiation therapy in pleural mesothelioma has been shown to alleviate pain in the majority of patients treated.

<!--L2, S 45-->El uso de radioterapia en mesotelioma pleural ha demostrado alivio en el dolor en la mayoría de los pacientes tratados.

<!--L1, S 46-->However, the duration of symptom control is short-lived [4,5].

<!--L2, S 46-->Sin embargo, la duración del control de los síntomas es corta [4,5].

<!--L1, S 47-->Single agent and combination chemotherapy have been evaluated in single and combined modality studies.

<!--L2, S 47-->La quimioterapia de agente único y la quimioterapia de combinación han sido evaluadas en estudios de modalidad simple y combinada.

<!--L1, S 48-->The most studied agent is doxorubicin, which has produced partial responses in approximately 15%-20% of patients studied [6].

<!--L2, S 48-->El agente más estudiado es doxorrubicina, el cual ha producido respuestas parciales en aproximadamente el 15%-20% de los pacientes estudiados [6].

<!--L1, S 49-->Some combination chemotherapy regimens have been reported to have higher response rates in small phase II trials.

<!--L2, S 49-->Algunos regímenes de quimioterapia de combinación han reportado tener tasas de respuestas más altas en ensayos de fase II no muy extensos.

<!--L1, S 50-->However the toxicity reported is also higher and there is no evidence that combination regimens result in longer survival or longer control of symptoms [6,7].

<!--L2, S 50-->Sin embargo, la toxicidad reportada también es más alta y no hay evidencia de que los regímenes de combinación resulten en una supervivencia más larga o en control de síntomas más extenso [6,7].

<!--L1, S 51-->Recurrent pleural effusions may be treated with pleural sclerosing procedures; however, failure rates are high secondary to the restrictive nature of the tumor.

<!--L2, S 51-->Las efusiones pleurales recidivantes pueden tratarse con procedimientos esclerosantes pleurales; sin embargo, las tasas de fracaso son altas debido a la naturaleza restrictiva del tumor.

<!--L1, S 52-->The designations in PDQ that treatments are "standard" or "under clinical evaluation" are not to be used as a basis for reimbursement determinations.

<!--L2, S 52-->Las designaciones en PDQ que indican que un tratamiento es "estándar" o que está "en evaluación clínica" no deben emplearse como base para la determinación de reembolsos.

<!--L1, S 53--><References>

<!--L2, S 53--><Bibliografía>

<!--L1, S 54-->***** LOCALIZED MALIGNANT MESOTHELIOMA (STAGE I)

<!--L2, S 54-->***** MESOTELIOMA MALIGNO LOCALIZADO (ETAPA I)

<!--L1, S 55-->Treatment options:[1].

<!--L2, S 55-->Opciones de tratamiento:[1].

<!--L1, S 56-->Standard:

<!--L2, S 56-->Estándar:

<!--L1, S 57-->1 Solitary mesotheliomas: Surgical resection en bloc including contiguous structures to ensure wide disease-free margins.

<!--L2, S 57-->1 Mesoteliomas solitarios: Resección quirúrgica en bloque incluyendo estructuras contiguas para asegurar márgenes libres de enfermedad amplios.

<!--L1, S 58-->Sessile polypoid lesions should be treated with surgical resection to ensure maximal potential for cure [2].

<!--L2, S 58-->Las lesiones polipoides sésiles deben tratarse con resección quirúrgica para asegurar el potencial curativo máximo [2].

<!--L1, S 59-->2 Intracavitary mesothelioma: A) Palliative surgery (pleurectomy and decortication) with or without postoperative radiation therapy B) Extrapleural pneumonectomy C) Palliative radiation therapy.

<!--L2, S 59-->2 Mesotelioma intracavitario: A) Cirugía paliativa (pleurectomía y decorticación) con o sin radioterapia postoperatoria B) Neumonectomía extrapleural C) Radioterapia paliativa.

<!--L1, S 60-->Under clinical evaluation:

<!--L2, S 60-->En evaluación clínica:

<!--L1, S 61-->1 Intracavitary chemotherapy following resection [3,4].
 <!--L2, S 61-->1 Quimioterapia intracavitaria después de resección [3,4].
 <!--L1, S 62-->2 Multimodality therapy [4-6].
 <!--L2, S 62-->2 Terapia de modalidades múltiples [4-6].
 <!--L1, S 63-->3 Other clinical trials.
 <!--L2, S 63-->3 Otras pruebas clínicas.
 <!--L1, S 64--><References>
 <!--L2, S 64--><Bibliografía>
 <!--L1, S 65-->***** ADVANCED MALIGNANT MESOTHELIOMA (STAGES II, III, AND IV)
 <!--L2, S 65-->***** MESOTELIOMA MALIGNO AVANZADO (ETAPAS II, III Y IV)
 <!--L1, S 66-->Treatment options:
 <!--L2, S 66-->Opciones de tratamiento:
 <!--L1, S 67-->1 Symptomatic treatment to include drainage of effusions, chest tube pleurodesis, or thoracoscopic pleurodesis [1].
 <!--L2, S 67-->1 Tratamiento sintomático que incluya drenaje de derrames, pleurodesis del tubo torácico o pleurodesis toracoscópica [1].
 <!--L1, S 68-->2 Palliative surgical resection in selected patients [2,3].
 <!--L2, S 68-->2 Resecciones quirúrgicas paliativas en pacientes seleccionados [2,3].
 <!--L1, S 69-->3 Palliative radiation therapy [4,5].
 <!--L2, S 69-->3 Radioterapia paliativa [4,5].
 <!--L1, S 70-->4 Single-agent chemotherapy.
 <!--L2, S 70-->4 Quimioterapia con un solo agente.
 <!--L1, S 71-->Partial responses have been reported with doxorubicin, epirubicin, mitomycin, cyclophosphamide, cisplatin, carboplatin, and ifosfamide [6-8].
 <!--L2, S 71-->Se ha informado de respuestas parciales con doxorubicina, epirubicina, mitomicina, ciclofosfamida, cisplatino, carboplatino e ifosfamida [6-8].
 <!--L1, S 72-->5 Multimodality clinical trials [9-13].
 <!--L2, S 72-->5 Pruebas clínicas de modalidades múltiples [9-13].
 <!--L1, S 73-->6 Intracavitary therapy.
 <!--L2, S 73-->6 Terapia intracavitaria.
 <!--L1, S 74-->Intrapleural or intraperitoneal administration of chemotherapeutic agents (e.g, cisplatin, mitomycin, and cytarabine) has been reported to produce transient reduction in the size of tumor masses and temporary control of effusions in small clinical studies [14-16].
 <!--L2, S 74-->Se ha informado que la administración intrapleural o intraperitoneal de agentes quimioterapéuticos (p. ej., cisplatino, mitomicina y citarabina) produce reducción transitoria en el tamaño de las masas tumorales y control temporal de derrames en estudios clínicos pequeños [14-16].
 <!--L1, S 75-->Additional studies are needed to define the role of intracavitary therapy.
 <!--L2, S 75-->Se requieren estudios adicionales para definir la función de terapia intracavitaria.
 <!--L1, S 76--><References>
 <!--L2, S 76--><Bibliografía>
 <!--L1, S 77-->***** RECURRENT MALIGNANT MESOTHELIOMA
 <!--L2, S 77-->***** MESOTELIOMA MALIGNO RECURRENTE
 <!--L1, S 78-->Treatment of recurrent mesothelioma usually utilizes procedures and/or agents not previously employed in the initial treatment attempt.
 <!--L2, S 78-->El tratamiento de mesotelioma recidivante generalmente utiliza procedimientos y/o agentes que no han sido previamente empleados en el tratamiento tentativo inicial.
 <!--L1, S 79-->No standard treatment approaches have been proven to improve survival or control symptoms for a prolonged period of time.
 <!--L2, S 79-->Ningún enfoque de tratamiento ha demostrado mejorar la supervivencia o el control de síntomas por un período de tiempo prolongado.
 <!--L1, S 80-->These patients should be considered candidates for phase I and II clinical trials evaluating new biologicals, chemotherapeutic agents, or physical approaches [1-5].
 <!--L2, S 80-->Estos pacientes deben considerarse como candidatos para pruebas clínicas de fase I y II donde se evalúen productos biológicos nuevos, agentes quimioterapéuticos o enfoques físicos [1-5].
 <!--L1, S 81-->Consult the PDQ protocol file for a current listing of active clinical trials.

<!--L2, S 81-->Consulte el PDQ para obtener una lista actual de pruebas clínicas activas.
<!--L1, S 82--><References>
<!--L2, S 82--><Bibliografía>

TEXTO 16: Noticia para profesionales de la salud publicada en *Diario Médico*

17/03/98: El tratamiento neoadyuvante incrementa la supervivencia en el cáncer de pulmón

<http://www.diariomedico.com/oncologia/noticias.html>



El tratamiento neoadyuvante incrementa la supervivencia en el cáncer de pulmón

El Grupo Español de Cáncer de Pulmón, que engloba a 37 centros, tiene previsto estudiar los efectos de la terapia neoadyuvante (preoperatoria) con los nuevos fármacos quimioterápicos en la mejora de la supervivencia de los pacientes afectados por tumores pulmonares en estadios iniciales, según ha anunciado Rafael Rosell, presidente de la asociación y jefe del Servicio de Oncología Médica del Hospital Universitario Germans Trias, de Badalona, en Barcelona, que coordinará esta investigación.

"Creemos que puede ser mucho más efectivo que el tratamiento postquirúrgico", ha indicado Rosell, co-director científico, junto con Paul A. Bunn, de la División de Oncología Médica del Centro de Ciencias de la Salud de la Universidad de Colorado, en Denver, Estados Unidos, del Foro interactivo pan-europeo sobre nuevas perspectivas en el manejo del cáncer de pulmón celebrado en Sitges, Barcelona, y organizado por el laboratorio Bristol-Myers Squibb, que financiará parte del estudio que se propone realizar el grupo español de expertos.

En concreto se quiere reunir una muestra de 600 pacientes en dos años y seguirlos durante tres, con el objetivo de llegar a demostrar una reducción de la mortalidad del 20 por ciento a lo largo de ese período.

Rosell, junto con un equipo interdisciplinar de su propio centro y de los hospitales La Fe y General de Valencia, publicó en 1994 en la revista *The New England Journal of Medicine* que la quimioterapia preoperatoria con fármacos clásicos incrementaba la supervivencia media de los pacientes con cáncer de pulmón localmente avanzado, y el pasado mes de enero, a través de la revista *Annals of Oncology*, dieron a conocer un estudio que concluye que un 28 por ciento de los casos de tumores diagnosticados en estadios iniciales presentan micrometástasis -ADN tumoral circulante en la sangre- ya en el momento mismo de la detección de la enfermedad.

Evidencias

Estas evidencias avalan la hipótesis de la utilidad de la quimioterapia preoperatoria incluso en pacientes con tumores muy iniciales y ADN tumoral circulante.

"Lo convencional hoy día es operar y aplicar la quimioterapia posteriormente", ha precisado. Los estudios realizados indican que de esta manera el 50 por ciento de los pacientes presentan recidiva a distancia al año y medio después de ser tratados.

Por otro lado, los centros integrados dentro del Grupo Español de Cáncer de Pulmón publicaron en 1994 su experiencia con el quimioterápico taxol en cáncer de pulmón

avanzado no quirúrgico -estadios IIIb y IV-, que es el mismo fármaco que ahora se probará en los estadios iniciales preoperatoriamente y que se ha demostrado que por sí solo tiene una actividad similar a la de las combinaciones farmacológicas con cisplatino. "El taxol y otras moléculas como el taxotere y la gemcitabina han cambiado el panorama de la quimioterapia en el cáncer de pulmón, que hasta hace poco tiempo se basaba en el cisplatino, que es eficaz pero presenta efectos secundarios. Estos fármacos han permitido individualizar más el tratamiento, en función de los estadios y tratar incluso tumores iniciales", ha explicado Rosell. El Hospital Clínico de Madrid, el Clínico de Valencia y el Hospital Germans Trias de Badalona estudiaron los efectos del taxol en 60 pacientes con cáncer diseminado y vieron que inhibe la angiogénesis del tumor, consiguiendo un 30 por ciento de respuestas objetivas y un 50 por ciento de enfermedad estable. "La respuesta radiográfica estable durante más de un año es una buena respuesta", ha precisado Rosell.

Por otro lado, el experto ha destacado la importancia del trabajo coordinado que llevan a cabo los centros investigadores del cáncer de pulmón en España y del que se está desarrollando conjuntamente con otros países europeos. "Esto permite homogeneidad en la calidad del tratamiento en el ámbito europeo y respuestas más rápidas".

Comparación

En Sitges se han dado a conocer, por otra parte, resultados preliminares de un estudio multicéntrico internacional -16 países, entre ellos España- sobre el taxol combinado con carboplatino versus taxol mezclado con cisplatino en enfermos con cáncer de pulmón no microcítico.

Los datos confirman la hipótesis de que la primera opción es tan eficaz como la segunda y con menores efectos secundarios. Las conclusiones se presentarán en mayo durante el Congreso Americano de Oncología Médica que se celebrará en Los Angeles, en

California.

Carmen Fernández. Barcelona.

APÉNDICE III: Tipos de repetición

TABLA 1. Detección de posibles instancias de cohesión léxica

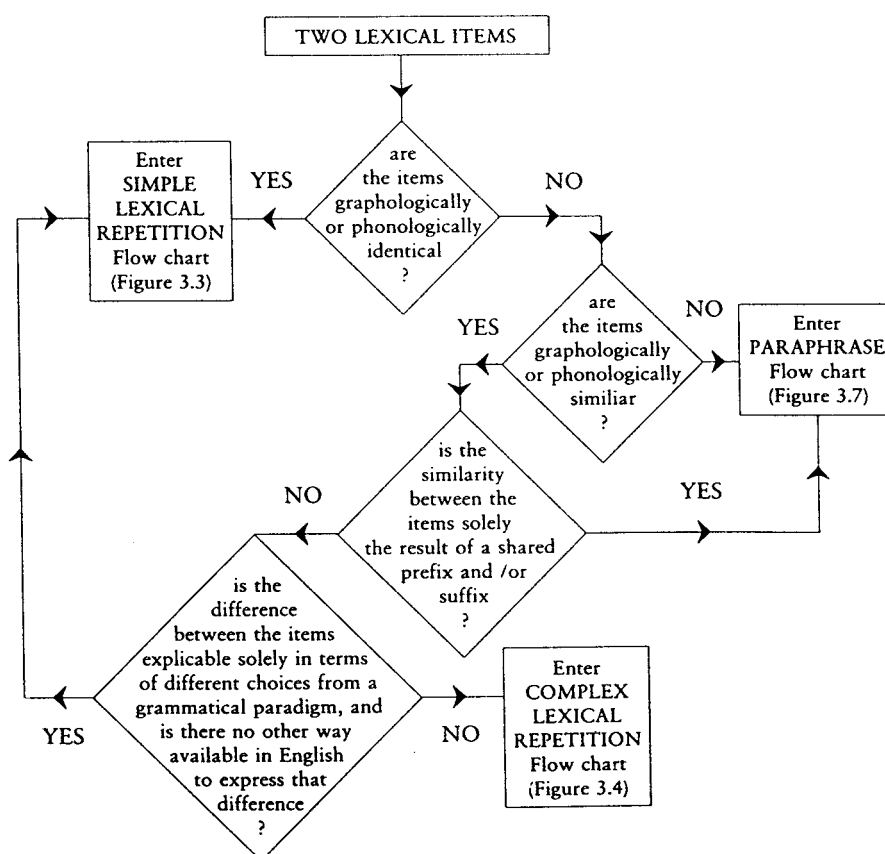


TABLA 2: Repetición léxica simple

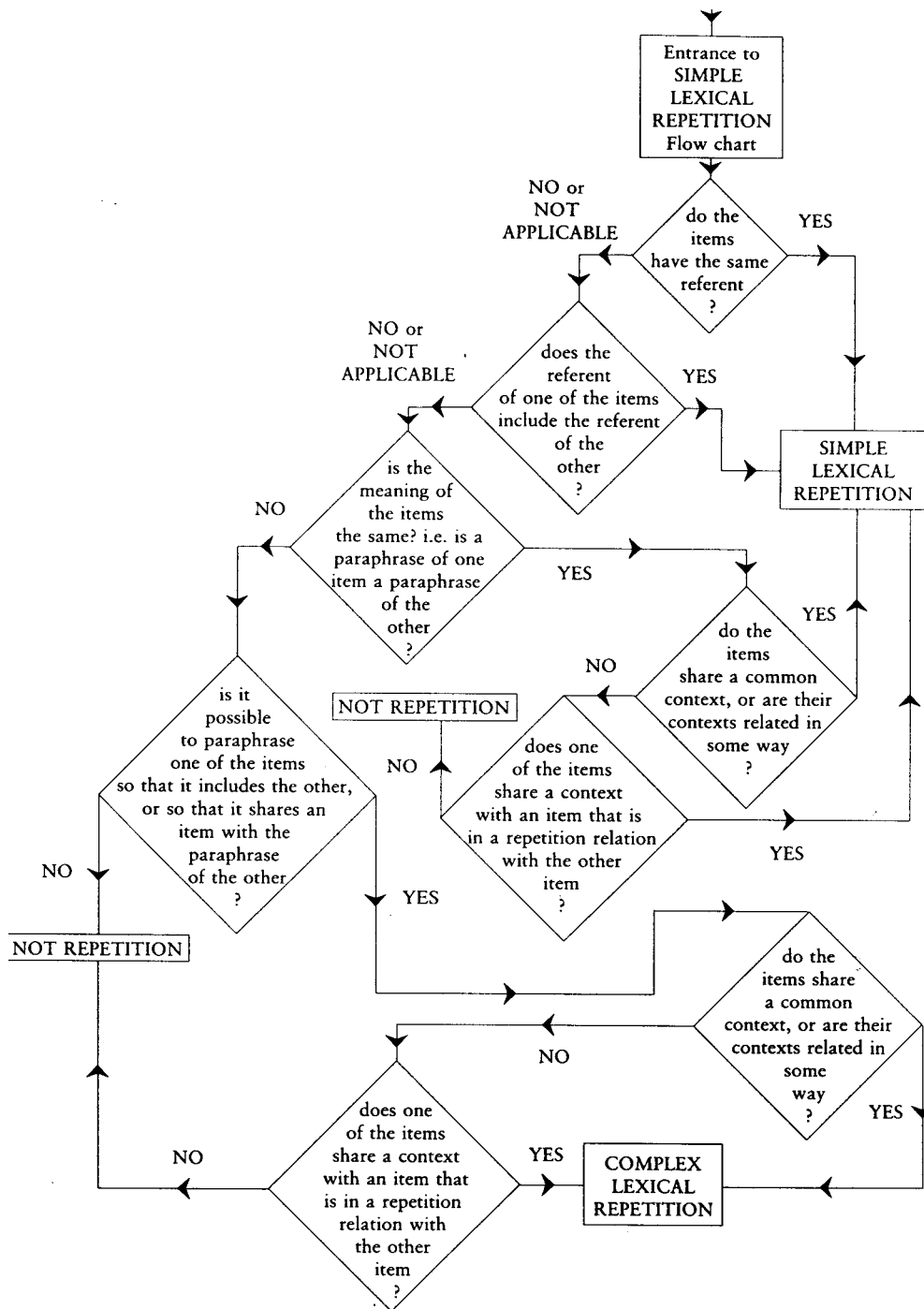


TABLA 3: Repetición léxica compleja

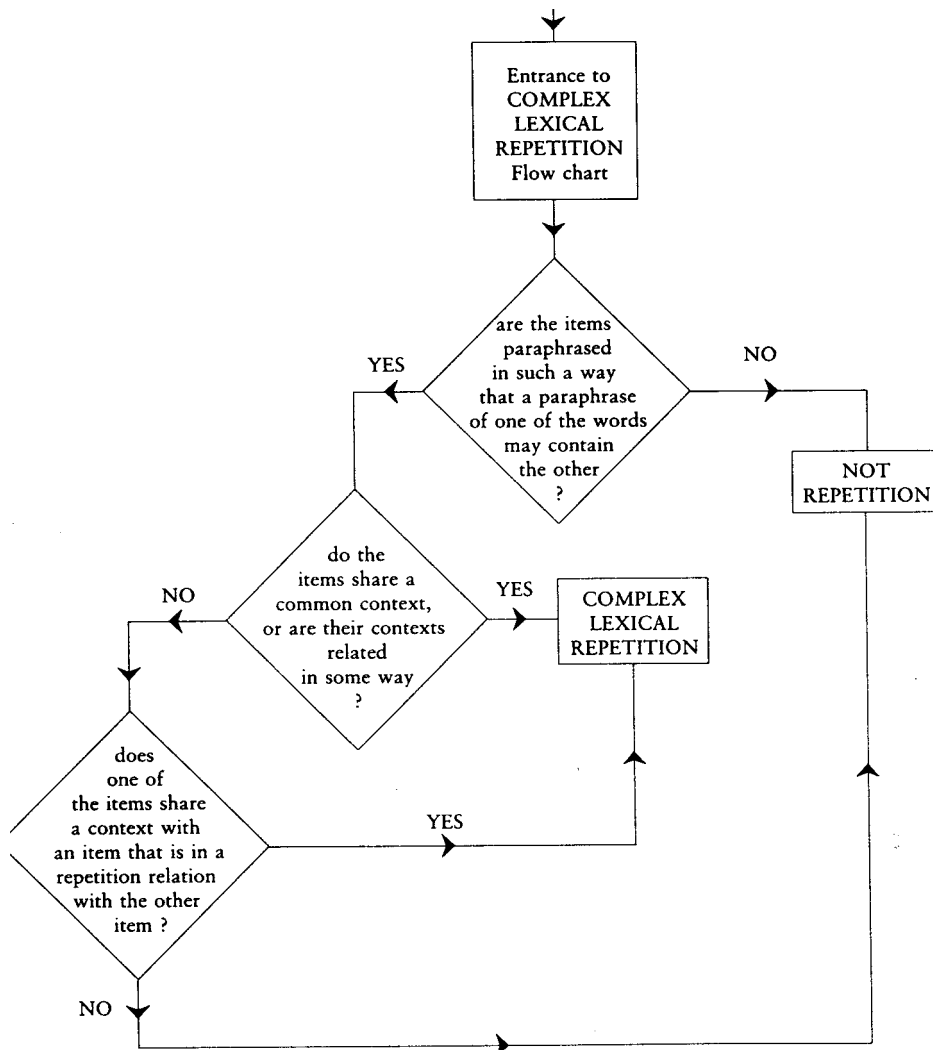
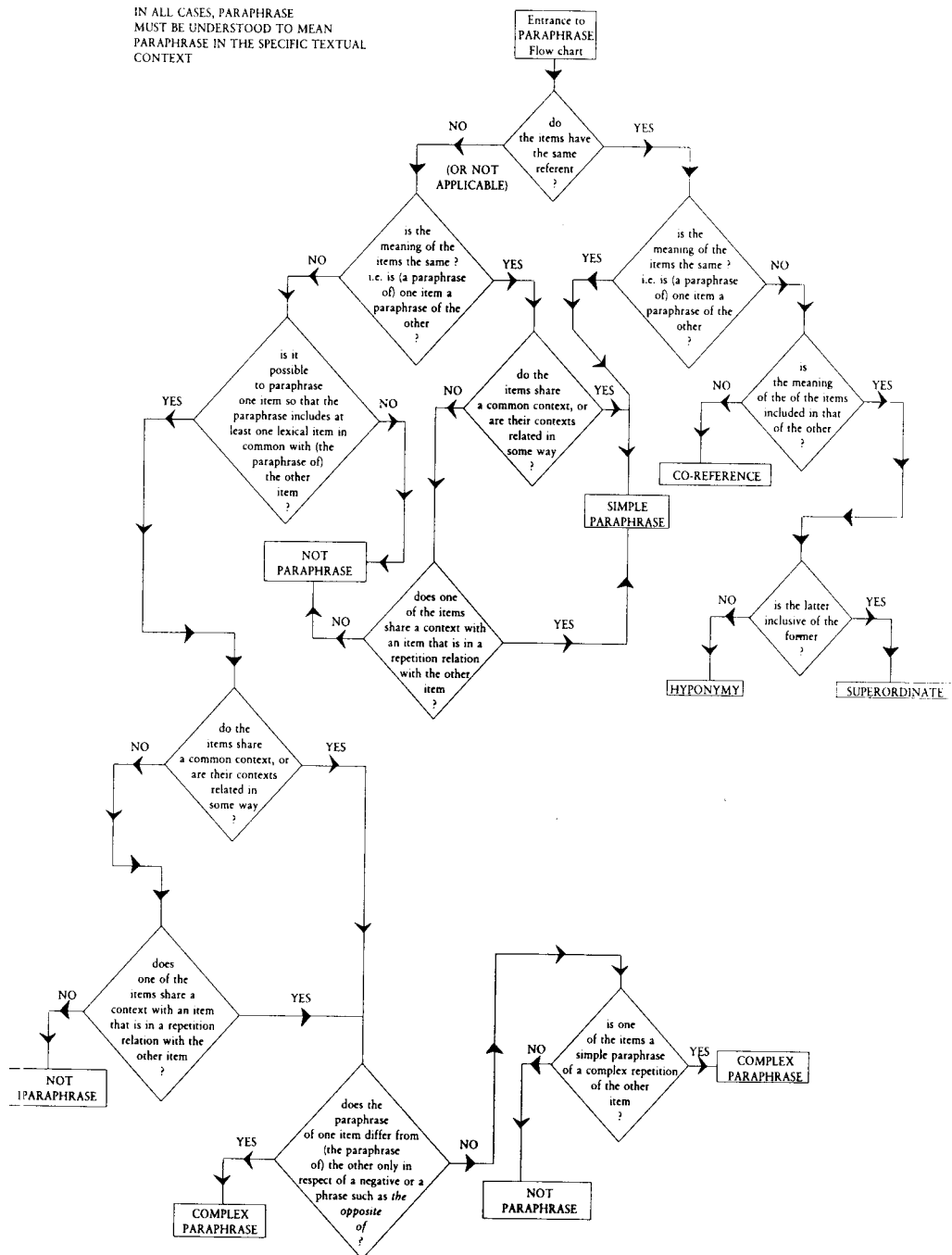


TABLE 4: Paráfrasis



APÉNDICE IV: Fragmento inicial de *Masters of Political Thought* (en Hoey 1991: 246-247)

1 What is attempted in the following volume is to present to the reader a series of actual excerpts from the writings of the greatest political theorists of the past; selected and arranged so as to show the mutual coherence of various parts of an author's thought and his historical relation to his predecessors or successors; and accompanied by introductory notes and intervening comments designed to assist the understanding of the meaning and importance of the doctrine quoted. 2 The book does not purport to be a history of political theory, with quotations interspersed to illustrate the history. 3 It is rather a collection of texts, to which I have endeavoured to supply a commentary. 4 I have tried rather to render the work of Aristotle, Augustine, and the rest accessible to the student, than to write a book about them; and the main object of this work will have been achieved if it serves not as a substitute for a further study of the actual works of these authors, but as an incentive to undertake it.

5 Nor does the commentary make any pretension of being exhaustive. 6 Very often after a long passage has been quoted a single point only has been selected for comment; and sometimes this point has been selected not because it was the most important, but because it was one on which I had something to say. 7 I have not tried to cover all the ground, and shall have done my part if the reader is stimulated, by the samples which I have offered, to complete a commentary of his own.

8 The selection has been confined to a few authors, for reasons not only of space, or of limitations of my own knowledge (though either of these reasons would have been sufficient), but because it is part of the plan of the book to concentrate attention upon the most important works. 9 A knowledge of Plato's *Republic*, of Aristotle's *Politics*, of parts of Augustine's *City of God*, belongs to a general education. 10 The works of lesser writers, or the lesser works of these writers, are doubtless worth reading; but a man who is not a specialist may ignore them without reproach.

11 If the commentary is secondary to the text, still more so must be any introductory remarks which I make here. 12 In commending the writings which follow to the reader's attention, I will indeed stake my credit on the assertion that the study of them will correct the judgment and enlighten the understanding upon matters in which it is important to be enlightened and correct. 13 But if a proof of this assertion is demanded, there is no proof except that of asking the inquirer to make the experiment. 14 The introducer may suggest lines of reasoning, he may try to convey certain lights which he has himself derived from the study, but in doing this he must be tentative and not dogmatic, and in the last resort he must say to the reader, "Go and read for yourself, and try whether this is confirmed by your experience." 15 In this respect his position is like that of the critic of a work of art. 16 However useful the critic's remarks may be in preparing an approach to the work, they can never dispense the reader from the necessity of studying the work itself, nor deprive him of the right, on the basis of this study, of turning critic himself and standing in judgment on the reasonings by which he was led to it in the first place.

APÉNDICE V: Matriz de repetición del texto del Apéndice IV

TABLA 1. Desarrollo de la primera línea de la matriz de repetición en Tabla 2 (Hoey 1991: 88-89)

	(1)	attempted	following	volume	present	reader	actual	writings	political	theorists	selected	author's	historical	introductory	notes	intervening	comments	importance	quoted	excerpt
(2)	sr cr cr smp spp			book					political theory			history				interspersed				quotations
(3)	cr spp	endeavoured																	commentary	
(4)	sr sr smp spp spp	tried		book			actual works					author								
(5)	cr																		commentary	
(6)	sr cr smp																		comment	quoted passage
(7)	sr cr smp spp	tried			offered	reader														commentary
(8)	sr cr cr smp spp			book				words			selection	author								important
(9)	cr					reading		writers	'politics'											
(10)	cr cr																			
(11)	sr* cr spp*														introductory	remarks			commentary	
(12)	sr sr cr cr*				follow	reader		writings												important
(13)																				
(14)	cr cr* spp	try					read								introducer					
(15)																				
(16)	sr						reader													

APÉNDICE VI: Matrices de repetición en forma de pares de coordenadas para las oraciones 3-12 del texto 6 del apéndice II.

COHESIÓN CON RESPECTO A LA TERCERA ORACIÓN (TABLA A)

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 3)	Véase (3, 1)
(2, 3)	Véase (3, 2)
(4, 3)	55. Spontaneous chromosome aberrations 56. Peripheral blood lymphocytes (PLBs) / PLBs 57. Lung carcinoma / lung carcinoma patients
(5, 3)	58. Authors 59. Spontaneous chromosome aberrations / chromosome aberrations 60. Lung carcinoma / cancer
(6, 3)	61. Spontaneous chromosome aberrations / spontaneous aberrations
(7, 3)	62. Lung carcinoma / cancer
(8, 3)	63. Spontaneous chromosome aberrations / chromosome 9 aberrations Significant / significantly 64. Lung carcinoma / cancer
(9, 3)	65. Lung carcinoma 66. Spontaneous chromosome aberrations / chromosome 9 aberrations
(10, 3)	67. Spontaneous chromosome aberrations 68. Lung carcinoma / cancer
(11, 3)	69. Spontaneous chromosome aberrations / chromosome 9 aberration 70. Lung carcinoma / smoking
(12, 3)	71. Spontaneous chromosome aberrations / chromosome 9 aberrations 72. Lung carcinoma / cancer

COHESIÓN CON RESPECTO A LA CUARTA ORACIÓN (TABLA A)

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 4)	Véase (4, 1)
(2, 4)	Véase (4, 2)
(3, 4)	Véase (4, 3)
(5, 4)	73. Spontaneous chromosome aberrations / chromosome aberrations 74. Patients / individuals 75. self-reported / report 76. Family history of cancer
(6, 4)	77. Patients / individual 78. Family history of cancer / family history 79. Spontaneous chromosome aberrations / spontaneous aberrations
(7, 4)	80. Lung carcinoma patients / patients 81. Cancer
(8, 4)	82. Spontaneous chromosome aberrations / chromosome 9 aberrations 83. Cancer
(9, 4)	84. Family history of cancer 85. Lung carcinoma
(10, 4)	86. Spontaneous chromosome aberrations 87. Family history of cancer 88. Evaluated
(11, 4)	89. Spontaneous chromosome aberrations / chromosome 9 aberrations profile 90. Cancer / smoking
(12, 4)	91. Spontaneous chromosome aberrations / chromosome 9 aberrations 92. Family history of cancer / familial aggregation of cancer 93. Study

**COHESIÓN CON RESPECTO A LA
QUINTA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 5)	Véase (5, 1)
(2, 5)	Véase (5, 2)
(3, 5)	Véase (5, 3)
(4, 5)	Véase (5, 4)
(6, 5)	94. Individuals / individual 95. Family history of cancer / family history 96. Chromosome aberrations / spontaneous aberrations
(7, 5)	97. Individuals / patients 98. Family / relatives 99. Cancer

(8, 5)	100. Chromosome aberrations / Chromosome 9 aberrations 101. Family history of cancer / first-degree relative with cancer
(9, 5)	102. Family history of cancer 103. Chromosome aberrations / chromosomal aberrations
(10, 5)	104. Chromosome aberrations / spontaneous chromosome aberrations 105. Family history of cancer
(11, 5)	106. Chromosome aberrations / Chromosome 9 aberration (profile) 107. Cancer / smoking
(12, 5)	108. Chromosome aberrations / Chromosome 9 aberrations 109. Family history of cancer / familial aggregation of cancer

**COHESIÓN CON RESPECTO A LA
SEXTA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 6)	Véase (6, 1)
(2, 6)	Véase (6, 2)
(3, 6)	Véase (6, 3)
(4, 6)	Véase (6, 4)
(5, 6)	Véase (6, 5)
(7, 6)	110. Individual / patients 111. Family / relatives
(8, 6)	112. Spontaneous aberrations/ Chromosome 9 aberrations 113. Family / relative

(9, 6)	114. Family history / family history of cancer 115. Spontaneous aberrations/ Chromosomal aberrations
(10, 6)	116. Spontaneous aberrations/ Spontaneous chromosome aberrations 117. Family history / family history of cancer
(11, 6)	118. Spontaneous aberrations/ chromosome 9 aberration (profile)
(12, 6)	119. Spontaneous aberrations/ chromosome 9 aberrations 120. Family history / familial aggregation of cancer

**COHESIÓN CON RESPECTO A LA
SÉPTIMA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 7)	Véase (7, 1)
(2, 7)	Véase (7, 2)
(3, 7)	Véase (7, 3)
(4, 7)	Véase (7, 4)
(5, 7)	Véase (7, 5)
(6, 7)	Véase (7, 6)
(8, 7)	121. First-degree relatives with cancer / first degree relative with cancer
(9, 7)	122. First-degree relatives with cancer / family history of cancer
(10, 7)	123. First-degree relatives with cancer / family history of cancer
(11, 7)	124. Cancer / smoking
(12, 7)	125. First-degree relatives with cancer / familial aggregation of cancer

**COHESIÓN CON RESPECTO A LA
OCTAVA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 8)	Véase (8, 1)
(2, 8)	Véase (8, 2)
(3, 8)	Véase (8, 3)
(4, 8)	Véase (8, 4)
(5, 8)	Véase (8, 5)
(6, 8)	Véase (8, 6)
(7, 8)	Véase (8, 7)
(9, 8)	126. Relative / family 127. Chromosome 9 aberrations chromosomal aberrations 128. Cancer / cancers
(10, 8)	129. Chromosome 9 aberrations 130. First-degree relative with cancer / Family history of cancer
(11, 8)	131. Chromosome 9 aberrations / Chromosome 9 aberration (profile) 132. Cancer / smoking
(12, 8)	133. Chromosome 9 aberrations 134. First-degree relative with cancer / Familial aggregation of cancer

**COHESIÓN CON RESPECTO A LA
NOVENA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 9)	Véase (9, 1)
(2, 9)	Véase (9, 2)
(3, 9)	Véase (9, 3)
(4, 9)	Véase (9, 4)
(5, 9)	Véase (9, 5)
(6, 9)	Véase (9, 6)
(7, 9)	Véase (9, 7)
(8, 9)	Véase (9, 8)

(10, 9)	135. Family history of cancer 136. Chromosomal aberrations / Chromosome 9 aberrations 137. Associated / associations
(11, 9)	138. Cancers / smoking 139. Associated / associations 140. Chromosomal aberrations / Chromosome 9 aberration (profile)
(12, 9)	141. Family history of cancer / familial aggregation of cancer 142. Chromosomal aberrations / Chromosome 9 aberrations 143. Associated 144. Cancers / cancer (Adj)

**COHESIÓN CON RESPECTO A LA
DÉCIMA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 10)	Véase (10, 1)
(2, 10)	Véase (10, 2)
(3, 10)	Véase (10, 3)
(4, 10)	Véase (10, 4)
(5, 10)	Véase (10, 5)
(6, 10)	Véase (10, 6)
(7, 10)	Véase (10, 7)
(8, 10)	Véase (10, 8)
(9, 10)	Véase (10, 9)
(11, 10)	145. Chromosome 9 aberrations/ chromosome 9 aberration (profile) 146. Associations 147. Cancer smoking
(12, 10)	148. Chromosome 9 aberrations 149. Family history of cancer / familial aggregation of cancer 150. Associations / associated

**COHESIÓN CON RESPECTO A LA
UNDÉCIMA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 11)	Véase (11, 1)
(2, 11)	Véase (11, 2)
(3, 11)	Véase (11, 3)
(4, 11)	Véase (11, 4)
(5, 11)	Véase (11, 5)
(6, 11)	Véase (11, 6)

(7, 11)	Véase (11, 7)
(8, 11)	Véase (11, 8)
(9, 11)	Véase (11, 9)
(10, 11)	Véase (11, 10)
(12, 11)	151. Chromosome 9 aberration (profile) / Chromosome 9 aberrations 152. Associations /associated 153. Smoking / cancer

**COHESIÓN CON RESPECTO A LA
DUODÉCIMA ORACIÓN (TABLA A)**

PAR COHESIVO	UNIDAD LÉXICA EN ORAC. PRECEDENTE / EN ORACIÓN POSTERIOR
(1, 12)	Véase casillas correspondientes
(2, 12)	
(3, 12)	
(4, 12)	
(5, 12)	
(6, 12)	
(7, 12)	
(8, 12)	
(9, 12)	
(10, 12)	
(11, 12)	

APÉNDICE VII: Resultados de *Hesperus*

VII a: Perspectiva general de los 36 textos del *PDQ* (*Cancer Treatment Information summaries, Supportive Care and Advocacy Issues*)

VII b: Perfil del texto QDT2 (TEXTO 10)

VII c: Versión de QDT2 con hipervínculos hacia las cadenas léxicas

VII d: Representación de la cadena 1 del texto QDT2

VII e: Representación tabular de las cadenas reconocidas por *Hesperus*.

APÉNDICE VII a: Perspectiva general de los 36 textos del *PDQ*

QDT: Cancer Treatment Information Summaries for Health Professionals

QPT: Cancer Treatment Information Summaries for Patients

QDC: Supportive Care and Advocacy Issues for Health Professionals

QPC: Supportive Care and Advocacy Issues for Patients

Case: 35

[QDT1.txt](#) 1.00 (505 words, 10536 total)
([Chains](#), [Profile](#))

Case: 34

[QDCOra.txt](#) 0.21 (2180 words, 40659 total)
([Chains](#), [Profile](#))

Case: 33

[QDT2.txt](#) 0.20 (2470 words, 46217 total)
([Chains](#), [Profile](#))

Case: 32

[QDCPos.txt](#) 0.19 (1824 words, 23812 total)
([Chains](#), [Profile](#))

Case: 31

[QDCCons.txt](#) 0.19 (2159 words, 58349 total)
([Chains](#), [Profile](#))

Case: 30

[QDCNut.txt](#) 0.19 (2272 words, 55327 total)
([Chains](#), [Profile](#))

Case: 29

[QPCLos.txt](#) 0.18 (2039 words, 46469 total)
([Chains](#), [Profile](#))

Case: 28

[QDCPru.txt](#) 0.18 (1758 words, 42276 total)
([Chains](#), [Profile](#))

Case: 27

[QDT3.txt](#) 0.17 (1648 words, 30223 total)
([Chains](#), [Profile](#))

Case: 26

[QDCHyp.txt](#) 0.17 (3025 words, 63480 total)
([Chains](#), [Profile](#))

Case: 25

[QDCAnx.txt](#) 0.17 (1489 words, 35427 total)
([Chains](#), [Profile](#))

Case: 24
[QDCTra.txt](#) 0.16 (1412 words, 30701 total)
([Chains](#), [Profile](#))

Case: 23
[QDCNau.txt](#) 0.16 (2333 words, 38513 total)
([Chains](#), [Profile](#))

Case: 22
[QPCNau.txt](#) 0.15 (707 words, 17882 total)
([Chains](#), [Profile](#))

Case: 21
[QPCFat.txt](#) 0.15 (1198 words, 26372 total)
([Chains](#), [Profile](#))

Case: 20
[QDCSle.txt](#) 0.15 (1554 words, 38774 total)
([Chains](#), [Profile](#))

Case: 19
[QPCOra.txt](#) 0.15 (1415 words, 29522 total)
([Chains](#), [Profile](#))

Case: 18
[QDCLos.txt](#) 0.14 (2871 words, 63084 total)
([Chains](#), [Profile](#))

Case: 17
[QPCAnx.txt](#) 0.14 (931 words, 24276 total)
([Chains](#), [Profile](#))

Case: 16
[QDCFat.txt](#) 0.14 (1987 words, 44515 total)
([Chains](#), [Profile](#))

Case: 15
[QPT1.txt](#) 0.14 (704 words, 17351 total)
([Chains](#), [Profile](#))

Case: 14
[QPCHyp.txt](#) 0.13 (990 words, 20224 total)
([Chains](#), [Profile](#))

Case: 13
[QPT3.txt](#) 0.12 (562 words, 15327 total)
([Chains](#), [Profile](#))

Case: 12
[QDCLym.txt](#) 0.11 (1717 words, 33784 total)
([Chains](#), [Profile](#))

Case: 11
[QPCLym.txt](#) 0.11 (1075 words, 24054 total)
([Chains](#), [Profile](#))

Case: 10
[QPT2.txt](#) 0.11 (553 words, 17424 total)
([Chains](#), [Profile](#))

Case: 9
[QPCNut.txt](#) 0.09 (1406 words, 41744 total)
([Chains](#), [Profile](#))

Case: 8
[QPCPos.txt](#) 0.08 (911 words, 17791 total)
([Chains](#), [Profile](#))

Case: 7
[QDCFev.txt](#) 0.08 (838 words, 18657 total)
([Chains](#), [Profile](#))

Case: 6
[QPCFev.txt](#) 0.08 (512 words, 16160 total)
([Chains](#), [Profile](#))

Case: 5
[QDCDel.txt](#) 0.07 (759 words, 13805 total)
([Chains](#), [Profile](#))

Case: 4
[QPCTra.txt](#) 0.07 (1530 words, 38419 total)
([Chains](#), [Profile](#))

Case: 3
[QPCCon.txt](#) 0.07 (679 words, 18391 total)
([Chains](#), [Profile](#))

Case: 2
[QPCDel.txt](#) 0.05 (339 words, 9490 total)
([Chains](#), [Profile](#))

Case: 1
[QPCPru.txt](#) 0.05 (476 words, 13136 total)
([Chains](#), [Profile](#))

Case: 0
[QPCSle.txt](#) 0.03 (323 words, 9852 total)
([Chains](#), [Profile](#))

Apéndice VII b: Perfil del texto QDT2

Document Profile: No file name (default), Score 0.00000, 196 Features, Total 46217 (0)

therapy_658_4268_n Percent: 21.07,	Value: 9737
sick-person_651_4202_n Percent: 7.02,	Value: 3246
show_522_3381_v Percent: 5.69,	Value: 2628
minuteness_196_1271_n Percent: 3.46,	Value: 1600
surgery_658_4267_n Percent: 3.02,	Value: 1396
medical_658_4274_a Percent: 2.72,	Value: 1255
experiment_461_2959_n Percent: 2.64,	Value: 1220
cancer_651_4195_n Percent: 2.56,	Value: 1183
respiratory-disease_651_4192_n Percent: 1.90,	Value: 880
piece_53_370_n Percent: 1.84,	Value: 850
existence_1_1_n Percent: 1.73,	Value: 799
fundamental_156_978_a Percent: 1.67,	Value: 770
remedial_658_4273_a Percent: 1.66,	Value: 769
durability_113_697_n Percent: 1.62,	Value: 750
swelling_253_1658_n Percent: 1.56,	Value: 721
transference_272_1809_n Percent: 1.50,	Value: 692
quantity_26_160_n Percent: 1.23,	Value: 567
time_108_672_n Percent: 1.20,	Value: 553
small_33_224_a Percent: 1.15,	Value: 533
medical-art_658_4266_n Percent: 1.11,	Value: 514
part_53_366_n Percent: 1.09,	Value: 504
limit_236_1565_n Percent: 1.08,	Value: 500
diagnostic_658_4258_n Percent: 0.97,	Value: 449
medicine_658_4257_n Percent: 0.94,	Value: 433
scission_46_315_n Percent: 0.88,	Value: 405
sick_651_4205_a Percent: 0.87,	Value: 402
effect_157_982_n Percent: 0.84,	Value: 388
distant_199_1297_a Percent: 0.84,	Value: 386
lateral_239_1582_a Percent: 0.76,	Value: 351
disease_651_4187_n Percent: 0.69,	Value: 317
blood_335_2239_n Percent: 0.68,	Value: 314
union_45_297_n Percent: 0.65,	Value: 301
attempting_671_4357_a Percent: 0.65,	Value: 300
regularity_81_541_n Percent: 0.65,	Value: 300
surveillance_457_2921_n Percent: 0.65,	Value: 300
party_708_4561_n Percent: 0.65,	Value: 300
be-near_200_1307_v Percent: 0.65,	Value: 300
careful_457_2922_a Percent: 0.57,	Value: 265
dramaturgy_594_3804_n Percent: 0.53,	Value: 245
remedy_658_4275_v Percent: 0.52,	Value: 242
ulcer_651_4198_n Percent: 0.51,	Value: 237
superior_34_235_a Percent: 0.49,	Value: 225
composition_56_391_n Percent: 0.48,	Value: 223
literal_558_3619_a Percent: 0.47,	Value: 216
date_108_674_n Percent: 0.43,	Value: 200
evidence_466_2993_n Percent: 0.43,	Value: 200
dissimilar_19_118_a Percent: 0.39,	Value: 182
production_164_1050_n Percent: 0.38,	Value: 175
speech_579_3706_n Percent: 0.38,	Value: 175
illness_651_4186_n Percent: 0.37,	Value: 172
pathology_651_4203_n Percent: 0.36,	Value: 168

important_638_4094_a	Percent: 0.36,	Value: 166
different_15_94_a	Percent: 0.35,	Value: 164
hanging-object_217_1409_n	Percent: 0.33,	Value: 154
separate_46_317_a	Percent: 0.32,	Value: 150
long-duration_113_696_n	Percent: 0.32,	Value: 150
be_1_6_v	Percent: 0.32,	Value: 150
useful_640_4108_a	Percent: 0.32,	Value: 150
towards_281_1902_r	Percent: 0.32,	Value: 150
age_131_818_n	Percent: 0.32,	Value: 150
set-apart_46_321_v	Percent: 0.32,	Value: 150
exhibit_522_3375_n	Percent: 0.29,	Value: 134
study_536_3481_n	Percent: 0.28,	Value: 131
circumstance_8_48_n	Percent: 0.26,	Value: 122
head_213_1386_n	Percent: 0.26,	Value: 120
require_627_4030_v	Percent: 0.26,	Value: 120
great_32_202_a	Percent: 0.26,	Value: 119
arrangement_62_427_n	Percent: 0.26,	Value: 119
cause_156_979_v	Percent: 0.23,	Value: 108
produce_164_1056_v	Percent: 0.22,	Value: 103
near_200_1309_r	Percent: 0.22,	Value: 102
mark_547_3559_v	Percent: 0.22,	Value: 100
electricity_160_1008_n	Percent: 0.22,	Value: 100
partition_231_1540_n	Percent: 0.22,	Value: 100
sanguineous_335_2241_a	Percent: 0.22,	Value: 100
be-situated_186_1172_v	Percent: 0.22,	Value: 100
undisguised_522_3378_a	Percent: 0.21,	Value: 99
causal_156_977_a	Percent: 0.21,	Value: 96
small-quantity_33_220_n	Percent: 0.21,	Value: 95
chosen_605_3885_a	Percent: 0.18,	Value: 84
classification_77_510_n	Percent: 0.18,	Value: 83
remainder_41_273_n	Percent: 0.18,	Value: 83
degree_27_165_n	Percent: 0.18,	Value: 83
unit_88_591_n	Percent: 0.17,	Value: 80
universal_79_525_a	Percent: 0.16,	Value: 76
comparison_462_2968_n	Percent: 0.16,	Value: 75
wound_655_4234_n	Percent: 0.16,	Value: 75
fluid_335_2238_n	Percent: 0.16,	Value: 75
instrument_628_4033_n	Percent: 0.16,	Value: 75
businesslike_622_3990_a	Percent: 0.16,	Value: 75
manifestation_522_3374_n	Percent: 0.16,	Value: 74
source_156_973_n	Percent: 0.15,	Value: 68
hard_326_2204_a	Percent: 0.13,	Value: 60
retinue_67_453_n	Percent: 0.13,	Value: 60
receipt_807_5142_n	Percent: 0.13,	Value: 60
cause_156_972_n	Percent: 0.13,	Value: 60
utility_640_4107_n	Percent: 0.12,	Value: 55
situation_186_1169_n	Percent: 0.11,	Value: 53
retreat_192_1225_n	Percent: 0.11,	Value: 53
literature_557_3611_n	Percent: 0.11,	Value: 50
excited_820_5223_a	Percent: 0.11,	Value: 50
modality_7_44_n	Percent: 0.11,	Value: 50
lasting_113_699_a	Percent: 0.11,	Value: 50
cardiovascular-disease_651_4193_n	Percent: 0.11,	Value: 50
unprosperous_731_4742_a	Percent: 0.11,	Value: 50

Total Percent: 97.47

Apéndice VII c: Versión de QDT2 con hipervínculos hacia las cadenas léxicas

----- NON- [small³ cell lung-cancer¹](#) CANCER -----

***** GENERAL INFORMATION

Non- [small cell lung-cancer](#) cancer (NSCLC) is a heterogeneous aggregate of at least three distinct histologies of [lung-cancer](#) cancer including epidermoid or squamous [carcinoma](#), adenocarcinoma, and [large cell carcinoma](#). These histologies are often classified together because, when localized, all have the potential for [cure⁰](#) with [surgical](#) resection. Systemic [chemotherapy](#) can [produce²](#) objective [partial¹³](#) [responses](#) and palliation of [symptoms](#) for [short](#) durations. Local control can be achieved with radiation in a large number of [patients](#) with unresectable [disease](#), but [cure](#) is seen only in a [small](#) [minority](#) of [patients](#).

At [diagnosis](#), [patients](#) with NSCLC can be divided into three [groups](#) that [reflect](#) the extent of [disease](#) and [treatment](#) approach. The first [group](#) of [patients](#) has tumors that are [surgically](#) resectable (generally [stages](#) I and II). This is the [group](#) with the best [prognosis](#), depending on a [variety](#) of tumor and host factors. [patients](#) with resectable [disease](#) who have [medical](#) contraindications to surgery can be [considered⁹](#) for [curative radiotherapy](#). The second [group](#) [includes patients](#) with either locally (T -T) or regionally (N -N) advanced [lung-cancer](#) cancer who have a diverse natural history. This [group](#) is [treated](#) with [radiotherapy](#) or with [radiotherapy](#) in [combination](#) with other [therapy](#) modalities. Selected [patients](#) with T or N [disease](#) can be [treated effectively](#) with [surgical](#) resection alone. The [final](#) [group](#) of [patients](#) have [distant¹⁰](#) metastases (M) [found](#) at the time of [diagnosis](#). This [group](#) can be [treated](#) with [radiotherapy](#) or [chemotherapy](#) for palliation of [symptoms](#) from the [primary](#) tumor. [patients](#) with good [performance](#) status, women, and [patients](#) with [distant](#) metastases [confined](#) to a [single](#) site appear to live longer than others []. Cisplatin-based [chemotherapy](#) has been associated with short-term palliation of [symptoms](#) and a [small](#) survival advantage. Currently no [single chemotherapy regimen](#) can be recommended for routine use.

For [operable patients](#), [prognosis](#) is adversely influenced by the presence of pulmonary [symptoms](#), [large](#) tumor size (centimeters), and presence of the erbB- oncoprotein [-]. Other factors that have been identified as adverse prognostic factors in some series of [patients](#) with resectable non- [small cell lung-cancer](#) cancer [include](#) mutation of the K-ras gene, vascular invasion, and increased numbers of [blood⁸](#) vessels in the tumor [specimen](#) [, ,].

Since [treatment](#) is not satisfactory for almost all [patients](#) with NSCLC, with the possible exception of a [subset](#) of pathologic [stage](#) I (T , N , M) [patients treated surgically](#), eligible [patients](#) should be [considered](#) for [clinical](#) trials.

***** CELLULAR [classification](#)

Prior to [initiating treatment](#) of any [patient](#) with [lung-cancer](#) cancer, a [review](#) of pathologic material by an experienced [lung-cancer](#) cancer [pathologist](#) is [critical](#) since some [cases](#) of [small cell lung-cancer](#) cancer (which responds well to [chemotherapy](#)) can be [confused](#) on [microscopic](#) examination with non- [small cell carcinoma](#) []. Nonsquamous [cell cancers](#) may be more likely to recur after [surgical](#) resection of early stage I tumors than other types of non- [small cell lung cancers](#) []. Bronchoalveolar [carcinoma](#) [represents](#) %- % of adenocarcinomas and sometimes has a distinct [presentation](#) and biologic behavior [-]. Bronchoalveolar [cancer](#) may [present](#) as a more diffuse [lesion](#) than other types of [cancer](#); %- % of [patients](#) undergoing an attempt at [surgical](#) resection [present](#) with an infiltrate on their chest [radiograph](#). Bronchoalveolar [cancer](#) is more common in women and in [patients](#) who do not smoke cigarettes than other histologic types of [lung-cancer](#) cancer.

Histologic [classification](#) of non- [small cell lung-cancer](#) cancer:

squamous [cell](#) (epidermoid) [carcinoma](#) / spindle [cell](#) variant / adenocarcinoma / acinar / papillary / bronchoalveolar [,] / solid tumor with mucin / [large cell carcinoma](#) / [giant cell](#) / clear [cell](#) / adenosquamous [carcinoma](#) / undifferentiated [carcinoma](#)

***** [stage](#) INFORMATION

Since determination of [stage](#) has [important therapeutic](#) and prognostic implications, [careful initial diagnostic](#) evaluation to define location and extent of [primary](#) and metastatic tumor involvement is [critical](#) for the appropriate care of [patients](#).

stage has a **critical role** in the selection of **therapy**. The **stage** of **disease** is based on a **combination** of **clinical** (physical examination, radiologic, and laboratory studies) and pathologic (**biopsy** of **lymph** nodes, **bronchoscopy**, mediastinoscopy, or anterior mediastinotomy []). The distinction between **clinical stage** and pathologic **stage** should be **considered** when evaluating reports of survival **outcome**. **surgical staging** of the mediastinum is **considered** standard if **accurate** evaluation of the nodal status is needed to **determine therapy**. The **radiology diagnostic oncology group** reported that the sensitivity and specificity of **computerized tomography** (CT) **scanning** is only % and %, respectively []. Magnetic resonance **imaging** (MRI) does not appear to improve the accuracy of **staging** []. Early evaluation of the **role** of positron emission **tomography** (PET) suggests that the **combination** of CT and PET may have greater sensitivity and specificity than CT alone []. A report evaluating the **staging** of , **patients** undergoing tumor resection **found** that **clinical staging** by radiologic studies accurately assessed the T **stage** in % of **patients** and the N **stage** in only % of **patients**. **errors in clinical staging** were equally divided between overstaging and understaging [].

----- The Revised International **staging** System for **lung-cancer** Cancer

The Revised International System for **staging lung-cancer** Cancer was adopted in by the American Joint Committee on **cancer** and the Union Internationale Contre le **cancer** []. These revisions were **made** to provide greater specificity for **patient groups**. **stage** I is divided into two **categories** by the size of the tumor; IA, T N M and IB, T N M . **stage** II is divided into two **categories** by the size of the tumor and by the nodal status; IIA, T N M and IIB, T N M . T N has been moved from **stage** IIIA in the version of the **staging** system to **stage** IIB. The other change has been to clarify the **classification** of multiple tumor **nodules**⁴. Satellite tumor **nodules** in the same lobe as the **primary lesion** that are not **lymph nodes** should be classified as T **lesions**. Intrapulmonary ipsilateral metastasis in a lobe other than the lobe containing the **primary lesions** should be classified as an M **lesion** (**stage** IV). The American **joint**¹¹ **committee** on **cancer** (AJCC) has designated **staging** by TNM **classification** [].

----- TNM definitions **primary** tumor (T)

TX: **primary** tumor cannot be assessed, or tumor proven by the presence of malignant **cells** in sputum or **bronchial** washings but not visualized by **imaging** or **bronchoscopy**. T : No **evidence** of **primary** tumor. **tis carcinoma** in situ. T : A tumor that is cm or less in greatest dimension, surrounded by lung or visceral pleura, and without bronchoscopic **evidence** of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus)*. T : A tumor with any of the following features of size or extent: More than cm in greatest dimension **involves** the main bronchus, cm or more **distal** to the carina Invades the visceral pleura Associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not **involve** the entire lung. T : A tumor of any size that directly invades any of the following: chest wall (including superior sulcus tumors), diaphragm, mediastinal **pleura**¹², **parietal** pericardium; or tumor in the main bronchus less than cm **distal** to the carina but without involvement of the carina; or associated atelectasis or obstructive pneumonitis of the entire lung. T : A tumor of any size that invades any of the following: mediastinum, heart, great vessels, trachea, esophagus, **vertebral** body, carina; or separate tumor **nodules** in the same lobe; or tumor with a malignant **pleural** effusion **.

*Note: The uncommon **superficial** tumor of any size with its invasive **component limited** to the **bronchial wall**, which may extend proximal to the main bronchus, is also classified as T . **Note: Most **pleural** effusions associated with **lung-cancer** cancer are **due-to** to tumor. However, there are a few **patients** in whom multiple cytopathologic examinations of **pleural fluid** are **negative** for tumor. In these **cases**, **fluid** is non- **bloody** and is not an exudate. When these elements and **clinical** judgement dictate that the effusion is not related to the tumor, the effusion should be excluded as a **staging element** and the **patient** should be staged as T , T , or T .

Regional **lymph nodes** (N)

NX: Regional **lymph nodes** cannot be assessed. N : No regional **lymph node metastasis**⁵. N : **metastasis** to ipsilateral peribronchial and/or ipsilateral hilar **lymph nodes**, and intrapulmonary **nodes** including **involvement** by direct extension of the **primary** tumor. N : **metastasis** to ipsilateral mediastinal and/or subcarinal **lymph node(s)**. N : **metastasis** to contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular **lymph node(s)**.

distant metastasis (M)

MX: **distant metastasis** cannot be assessed. M : No **distant metastasis**. M : **distant metastasis present**.

Note: M **includes** separate tumor **nodule(s)** in a different lobe (ipsilateral or contralateral).

Specify sites according to the following notations:

BRA = brain / eye = **eye** / HEP = hepatic / LYM = **lymph nodes** / MAR = bone marrow / OSS = *osseous* OTH = other / OVR = ovary / PER = peritoneal PLE = **pleura** / PUL = pulmonary / SKI = skin
----- AJCC **stage** groupings

Occult **carcinoma** TX, N, M / **stage** : **tis**, N, M / **stage** IA: T, N, M / **stage** IB: T, N, M / **stage** IIA: T, N, M / **stage** IIB: T, N, M / T, N, M **stage** IIIA: T, N, M / T, N, M / T, N, M / T, N, M **stage** IIIB: Any T, N, M / T, Any N, M **stage** IV: Any T, Any N, M.

***** **treatment** OPTION OVERVIEW

In non- **small cell lung-cancer** cancer (NSCLC), results of standard **treatment** are **poor** in all but the most localized **cancers**. All newly diagnosed **patients** with NSCLC are potential candidates for studies evaluating new forms of **treatment**. **surgery** is the major potentially **curative therapeutic** option for this **disease**; **radiotherapy** can **produce cure** in a **small minority** and palliation in the **majority** of **patients**. In advanced- **stage disease**, **chemotherapy** offers **modest** improvements in **median** survival although overall survival is **poor** [,]. Where studied, **chemotherapy** has been reported to **produce** short-term improvement in **disease-related symptoms**. In one **study**, symptomatic **relief** with **combination chemotherapy** was significant but **independent** of objective **response** [,]. The **impact** of **chemotherapy** on **quality of life** **requires** more **study**.

Current areas under evaluation **include** combining **local** (**surgery**), regional (**radiotherapy**), and systemic (**chemotherapy** and **immunotherapy**) **treatments** and developing more **effective** systemic **therapy**. Several new **agents**, including paclitaxel (Taxol), docetaxel (Taxotere), topotecan, irinotecan, vinorelbine, and gemcitabine have been **shown** to be active in the **treatment** of advanced NSCLC. Chemoprevention of second **primary cancers** of the upper aerodigestive tract is also under active investigation in early-stage **lung-cancer** cancer [].

The **designations** in PDQ that **treatments** are "standard" or "under **clinical** evaluation" are not to be used as a **basis** for reimbursement determinations.

***** OCCULT NON- **small cell lung-cancer** CANCER

----- TX, N, M In occult **lung-cancer** cancer, a **diagnostic** evaluation often **includes** **chest x-ray** and selective **bronchoscopy** with **close** follow-up (e.g, computed **tomography** scan), when needed, to define the site and **nature** of the **primary** tumor; tumors discovered in this **fashion** are generally early stage and curable by **surgery**. After **discovery** of the **primary** tumor, **treatment** is determined by **establishing** the **stage** of the **patient's** tumor. **therapy** is identical to that recommended for other non- **small cell lung-cancer** cancer **patients** with similar **stage disease**-----

stage NON- **small cell lung-cancer** CANCER

----- **tis**, N, M **stage** non- **small cell lung-cancer** cancer (NSCLC) is the same as **carcinoma** in situ of the lung. Because these tumors are by definition noninvasive and incapable of metastasizing, they should be curable with **surgical** resection; however, there is a high incidence of second **primary cancers**, many of which are unresectable. Endoscopic **phototherapy** with a hematoporphyrin **derivative** has been described as an alternative to **surgical** resection in **carefully** selected **patients** [-]. This investigational **treatment** seems to be most **effective** for very early **central** tumors that extend less than centimeter within the bronchus []. Efficacy of this **treatment** modality in the management of early NSCLC remains to be proven.

treatment options:

surgical resection using the least extensive technique possible (segmentectomy or **wedge** resection) to preserve maximum normal pulmonary tissue since these **patients** are at high risk for second lung **cancers**.

Endoscopic photodynamic **therapy** [,].

***** **stage** I NON- **small cell lung-cancer** CANCER

----- T, N, M or T, N, M **surgery** is the **treatment** of **choice** for **patients** with **stage** I non- **small cell lung-cancer** cancer (NSCLC). **careful** preoperative assessment of the **patient's** overall **medical condition**, especially the **patient's** pulmonary reserve, is **critical** in considering the benefits of **surgery**. The immediate postoperative mortality rate is

age-related, but %-% with lobectomy can be expected []. **patients** with impaired pulmonary function may be **considered** for *segmental* or **wedge** resection of the **primary** tumor; the **lung-cancer** Cancer Study Group has conducted a randomized **study** (LCSG-) to *compare* lobectomy with **limited** resection for **patients** with **stage I cancer** of the lung. The results of this **study** **show** a reduction in **local** recurrence for **patients** **treated** with lobectomy compared with those **treated** with **limited** excision but no significant difference in overall survival []. Similar results have been reported from a nonrandomized **comparison** of anatomic segmentectomy and lobectomy []. A survival advantage was noted with lobectomy for **patients** with tumors greater than centimeters, but not for those with tumors **smaller** than centimeters. However, the rate of local/regional recurrence was significantly less after lobectomy, regardless of **primary** tumor size. Another study of **stage I patients** **showed** that those **treated** with wedge or segment resections had a local recurrence rate of % (of) despite having undergone complete resections []. Exercise testing may **aid** in the selection of **patients** with impaired pulmonary function who can tolerate lung resection []. The availability of video-assisted thoracoscopic wedge resection permits **limited** resections in **patients** with poor pulmonary function who are not usually **considered** candidates for lobectomy [].

inoperable patients with **stage I disease** and with sufficient pulmonary reserve may be **considered** for **radiotherapy** with **curative** intent. In one report of **patients** older than years of age who had resectable lesions **smaller** than centimeters but who were **medically inoperable** or who refused **surgery**, survival at years following **radiotherapy** with **curative** intent was comparable to a historical control **group** of **patients** of similar age resected with **curative** intent []. In the two largest **retrospective radiotherapy** series, **inoperable patients** **treated** with definitive **radiotherapy** achieved -year survival rates of % and %. Both series **found** that **patients** with T , N tumors had better **outcomes**, with -year survival rates of % and % in this subgroup [,].

primary radiotherapy should **consist-of** of *approximately* , cGy delivered with megavoltage equipment to the midplane of the known tumor volume using conventional fractionation. A boost to the cone-down **field** of the **primary** tumor is frequently used to further **enhance** local control. **Careful treatment planning** with precise definition of target volume and avoidance of **critical** normal **structures** to the extent possible is needed for optimal results and requires the use of a simulator.

Many **patients** **treated surgically** subsequently **develop** regional or distant metastases []. Therefore, **patients** should be considered for entry into **clinical** trials evaluating adjuvant **treatment** with **chemotherapy** or **radiotherapy** following **surgery**. Trials of adjuvant **chemotherapy regimens** have failed to **demonstrate** a consistent benefit. **smokers** who undergo complete resection of **stage I NSCLC** are also at risk for second malignant tumors. In the **lung-cancer** Cancer Study Group trial of **stage T , N** resected **patients**, the rate of nonpulmonary second **cancers** was . % per year and . % per year for new lung **cancers** []. Others have reported even higher risks of second tumors in long-term survivors, including rates of % for second lung **cancers** and % for all second **cancers** []. A randomized trial of vitamin A versus observation in resected **stage I patients** **showed** a trend toward decreased second **primary cancers** in the vitamin A arm with no difference in overall survival rates []. An ongoing intergroup **clinical** trial will evaluate the role of isotretinoin in the chemoprevention of second **cancers** in **patients** resected for **stage I NSCLC** [].

treatment options:

Lobectomy or segmental, wedge, or sleeve resection as appropriate.

radiotherapy with **curative** intent (for potentially resectable **patients** who have **medical** contraindications to **surgery**).

clinical trials of adjuvant **chemotherapy** following resection [,].

Adjuvant chemoprevention trials [,].

Endoscopic photodynamic **therapy** (under **clinical** evaluation in highly selected T , N , M **patients**) [].

***** **stage II NON- small cell lung-cancer** CANCER

----- T , N , M or T , N , M or T , N , M **surgery** is the **treatment** of choice for **patients** with **stage II non- small cell lung-cancer** cancer (NSCLC). Careful preoperative assessment of the **patient's** overall **medical condition**, especially the **patient's** pulmonary reserve, is **critical** in considering the benefits of **surgery**. The immediate postoperative mortality rate is age-related, but up to %-% with pneumonectomy or %-% with lobectomy can be expected.

inoperable patients with **stage II disease** and with sufficient pulmonary reserve may be considered for **radiotherapy** with **curative** intent []. Among **patients** with excellent **performance** status, up to a % -year survival

rate may be expected if a course of **radiotherapy** with **curative** intent can be completed. In the largest **retrospective** series reported to date, **patients** with **medically inoperable** NSCLC **treated** with definitive **radiotherapy** achieved a -year overall survival rate of %; however, the **patients** with T tumors achieved an actuarial **disease-free** survival rate of %. This **retrospective** study also suggested that improved **disease-free** survival was obtained with **radiotherapy doses** greater than , cGy []. **primary radiotherapy** should **consist-of** of **approximately** , cGy delivered with megavoltage equipment to the midplane of the volume of known tumor using conventional fractionation. A boost to the cone-down field of the **primary** tumor is frequently used to further **enhance** local control. Careful **treatment planning** with precise definition of target volume and avoidance of **critical** normal **structures** to the extent possible is needed for optimal results and requires the use of a simulator.

Many **patients treated surgically** subsequently **develop** regional or distant metastases []. Therefore, **patients** should be considered for entry into **clinical** trials evaluating the use of adjuvant **treatment** with **chemotherapy** or **radiotherapy** following **surgery** []. One controlled trial has failed to **demonstrate** an overall survival benefit for **patients** with carefully staged squamous **cell carcinoma** receiving postoperative irradiation, although local recurrences were significantly **reduced** []. In two controlled trials in carefully staged, **surgically** resected **patients**, adjuvant **combination chemotherapy** with cisplatin, doxorubicin, and cyclophosphamide **produced modestly** increased **disease-free** survival and a trend toward improved overall survival, especially in the first year after **surgery** [,]. Based on these data, participation in **clinical** trials evaluating adjuvant **chemotherapy** after **surgical** resection should be **encouraged**.

treatment options:

Lobectomy, pneumonectomy, or segmental, wedge, or sleeve resection as appropriate.

radiotherapy with **curative** intent (for potentially **operable patients** who have **medical** contraindications to **surgery**).

radiotherapy combined with **curative surgery** [].

Adjuvant **chemotherapy** combined with other modalities [, -].

***** **stage IIIA NON- small cell lung-cancer** CANCER

Some **citations** in the text of this section are followed by a level of **evidence**. The PDQ editorial boards use a formal ranking system to help the reader **judge** the strength of **evidence** linked to the reported results of a **therapeutic** strategy. Refer to the PDQ levels of **evidence** summary for more information.

---- T , N , M or T , N , M or T , N , M or T , N , M Depending on **clinical** circumstances, the principal forms of **treatment** that are considered for **patients** with **stage III non- small cell lung-cancer** cancer (NSCLC) are **radiotherapy**, **chemotherapy**, **surgery**, and combinations of these modalities. Although the majority of these **patients** do not **achieve** a complete response to **radiotherapy**, there is a reproducible long-term survival benefit in %- % of **patients treated** with standard **fractionation** to , cGy, and significant palliation often results. **patients** with excellent **performance** status and those who require a thoracotomy to prove that **surgically** unresectable tumor is **present** are most likely to benefit from **radiotherapy** []. Because of the poor long-term results, these **patients** should be considered for **clinical trials**⁶. **trials** examining **fractionation** schedules, endobronchial laser **therapy**, brachytherapy, and combined modality approaches may **lead-to** improvement in the control of this regional **disease** []. One prospective randomized **clinical** study **showed** that **radiotherapy** given as three daily **fractions** improved overall survival compared to **radiotherapy** given as one daily **fraction** [].[Level of **evidence** iiA].

The addition of **chemotherapy** to **radiotherapy** has been reported to improve survival in prospective **clinical** studies that have used modern cisplatin- **based chemotherapy regimens** [-]. A meta-analysis of **patient** data from randomized **clinical** trials **showed** that cisplatin- **based** combinations plus **radiotherapy** **resulted** in % reduction in the risk of death compared with **radiotherapy** alone []. The optimal sequencing of modalities and schedule of **drug** administration remains to be determined and is under study in ongoing **clinical** trials [].

patients with N **disease apparent** on chest **radiograph** and documented by **biopsy** or discovered by prethoracotomy exploration have a -year survival rate of only about %. The use of preoperative (neoadjuvant) **chemotherapy** has been **shown** to be **effective** in these **clinical** situations in two **small** randomized studies of a total of **patients** with **stage IIIa NSCLC** [,]. The **patients** randomized to three cycles of cisplatin- **based chemotherapy** followed by **surgery** had a **median** survival more than three times as long as **patients treated** with **surgery** but no **chemotherapy** in both these studies. Two additional single-arm studies have evaluated either two to four cycles of **combination**

chemotherapy or **combination chemotherapy** plus chest irradiation for **patients** with histologically confirmed N stage IIIa NSCLC []. Sixty-five percent to % of **patients** were able to have a resection of their **cancer**, and %- % were alive at years. These results are encouraging, and combined-modality therapy with neoadjuvant **chemotherapy** with **surgery** and/or chest **radiotherapy** should be considered for **patients** with good **performance** status who have stage IIIa NSCLC.

Although most **retrospective** studies suggest that postoperative **radiotherapy** can improve local control for node-positive **patients** whose tumors were resected, it remains controversial whether it can improve survival [,]. One controlled trial in **patients** with completely resected **stage** II or III squamous **cell lung-cancer** failed to **demonstrate** a survival benefit for **patients** who received postoperative irradiation, although local recurrences were significantly **reduced** []. In two controlled trials with carefully staged **surgically** resected **patients**, adjuvant **combination chemotherapy** with cisplatin, doxorubicin, and cyclophosphamide **produced modestly** increased **disease-free** survival and a trend toward improved survival, especially in the first year after **surgery** [-]. Based on these data, participation in **clinical** trials evaluating adjuvant **chemotherapy** after **surgical** resection should be **encouraged**.

No consistent benefit from any **form** of **immunotherapy** has been demonstrated thus far in the **treatment** of NSCLC.

treatment options: **surgery** alone in highly selected **cases** [-].

chemotherapy combined with other modalities [- , -].

surgery and postoperative **radiotherapy** [, ,].

radiotherapy alone [,].

----- Superior sulcus tumor (T , N or N , M) Another category that merits a special approach is that of superior sulcus tumors, a locally invasive problem usually with a **reduced** tendency for distant metastases. Consequently, local **therapy** has **curative** potential, especially for T , N **disease**. **radiotherapy** alone, **radiotherapy** preceded or followed by **surgery**, or **surgery** alone (in highly selected **cases**) may be **curative** in some **patients**, with a -year survival rate of % or more in some studies []. **patients** with more invasive tumors of this area, or true Pancoast tumors, have a worse **prognosis** and generally do not benefit from **primary surgical** management. Follow-up **surgery** may be used to verify complete response in the **radiotherapy** field and to resect necrotic tissue.

treatment options:

radiotherapy and **surgery**.

radiotherapy alone.

surgery alone (selected **cases**).

chemotherapy combined with other modalities.

Brachytherapy [].

clinical trials of combined modality **therapy** [].

----- Chest wall tumor (T , N or N , M) Selected **patients** with bulky **primary** tumors that directly invade the chest wall can obtain long-term survival with **surgical** management provided that their tumor is completely resected.

treatment options:

surgery [,].

surgery and **radiotherapy**.

radiotherapy alone.

chemotherapy combined with other modalities.

***** stage IIIB NON- small cell lung-cancer CANCER

patients with stage IIIB non- small cell lung-cancer cancer (NSCLC) do not benefit from surgery alone and are best managed by initial chemotherapy, chemotherapy plus radiotherapy, or radiotherapy alone, depending on sites of tumor involvement and performance status. Most patients with excellent performance status should be considered for combined modality therapy. However, patients with malignant pleural effusion are rarely candidates for radiotherapy, and should generally be treated similarly to stage IV patients (see separate section of this summary on treatment of stage IV disease). Many randomized studies of unresectable patients with stage III NSCLC show that treatment with neoadjuvant or concurrent cisplatin- based chemotherapy and chest irradiation is associated with improved survival compared to treatment with radiotherapy alone [-]. A meta-analysis of patient data from randomized clinical trials showed that cisplatin- based combinations plus radiotherapy resulted in % reduction in the risk of death compared with radiotherapy alone [].

patients with stage IIIB disease with poor performance status are candidates for chest irradiation to palliate pulmonary symptoms (e.g, cough, shortness of breath, or local chest-pain pain). No consistent benefit from any form of immunotherapy has been demonstrated thus far.

---- T or N , M An occasional patient with supraclavicular node involvement who is otherwise a good candidate for irradiation with curative intent will survive years. Although the majority of these patients do not achieve a complete response to radiotherapy, significant palliation often results. patients with excellent performance status and those who are found to have advanced- stage disease at the time of resection are most likely to benefit from radiotherapy []. Adjuvant systemic chemotherapy with radiotherapy has been tested in randomized trials for patients with inoperable or unresectable locoregional NSCLC [- ,]. Some patients have shown a modest survival advantage with adjuvant chemotherapy. The addition of chemotherapy to radiotherapy has been reported to improve long-term survival in some,[, ,], but not all,[]. prospective clinical studies. A meta-analysis of patient data from randomized clinical trials showed an absolute survival benefit of % at years with the addition of cisplatin- based chemotherapy to radiotherapy []. One study showed improvement in local control and survival when cisplatin was given daily with concurrent radiotherapy but not when it was given weekly []. It is not yet clear whether the schedule of drug administration is responsible for the variation in results; the optimal sequencing of modalities remains to be determined and is under study in ongoing clinical trials [].

Because of the poor overall results, these patients should be considered for clinical trials⁷; trials examining fractionation schedules, radiosensitizers, radiolabeled antibodies, and combined modality approaches may lead-to to improvement in the control of regional disease.

patients with NSCLC can present with superior vena cava syndrome. Refer to the PDQ supportive care summary on superior vena cava syndrome for more information. Regardless of stage, this problem should generally be managed with radiotherapy with or without chemotherapy.

treatment options:

radiotherapy alone.

chemotherapy combined with radiotherapy [- ,].

chemotherapy and concurrent radiotherapy followed by resection [,].

chemotherapy alone.

***** stage IV NON- small cell lung-cancer CANCER

---- Any T, any N, M Cisplatin-containing and carboplatin-containing combination chemotherapy regimens produce objective response rates (including a few complete responses) that are higher than those achieved with single- agent chemotherapy. Although toxic effects may vary, outcome is similar with most cisplatin-containing regimens; a randomized trial comparing five cisplatin-containing regimens showed no significant difference in response, duration of response, or survival []. patients with good performance status and a limited number of sites of distant metastases have superior response and survival when given chemotherapy when compared to other patients []. A prospective randomized comparison of vinorelbine plus cisplatin versus vindesine plus cisplatin versus single agent vinorelbine has reported improved response rate (%) and median survival (weeks) with the vinorelbine plus cisplatin regimen []. Two small phase II studies reported that paclitaxel (Taxol) has single- agent activity in stage IV patients, with response rates in the range of %- % [,]. Reports of paclitaxel combinations have shown relatively

high response rates, significant year survival, and palliation of **lung-cancer** cancer **symptoms** []. With the paclitaxel plus carboplatin **regimen**, response rates have been in the range of %- % with -year survival rates of %- % [,]. The combination of cisplatin and paclitaxel was **shown** to have a higher response rate and higher year survival rate than the combination of cisplatin and etoposide []. Additional **clinical** studies should better define the role of these newer combination **chemotherapy regimens** in the **treatment** of advanced non- **small cell lung-cancer** cancer []. Meta-analyses have **shown** that **chemotherapy produces modest** benefits in short-term survival compared to supportive care alone in **patients** with **inoperable stages** IIIb and IV **disease** [-].

Although these results support further evaluation of chemotherapeutic approaches for both metastatic and locally advanced non- **small cell lung-cancer** cancer (NSCLC), efficacy of current programs is such that no **specific regimen** can be regarded as standard **therapy**. Appropriate **patients** should be **encouraged** to participate in **clinical** trials. Outside of a **clinical** trial setting, **chemotherapy** should be given only to **patients** with good **performance** status and evaluable tumor **lesions** who desire such **treatment** after being fully informed of its anticipated risks and **limited** benefits.

radiotherapy may be **effective** in palliating symptomatic local involvement with NSCLC such as tracheal, esophageal, or **bronchial compression**, bone or brain metastases, **pain**, vocal cord paralysis, hemoptysis, or superior vena cava **syndrome**. In some **cases**, endobronchial laser **therapy** and/or brachytherapy has been used to alleviate proximal obstructing **lesions** []. Such **therapeutic** intervention may be **critical** in the prolongation of an acceptable lifestyle in an otherwise functional **patient**. In the **rare patient** with synchronous **presentation** of a resectable **primary** tumor in the lung and a single brain metastasis, **surgical** resection of the solitary brain **lesion** is indicated with resection of the **primary** tumor and appropriate postoperative **chemotherapy** and/or irradiation of the **primary** tumor site and with postoperative whole-brain irradiation delivered in daily **fractions** of - cGy to avoid long-term **toxic** effects to normal brain tissue [,].

In asymptomatic **patients** kept under close observation, **treatment** may often be appropriately deferred until **symptoms** or **signs** of progressive tumor **develop**.

treatment options:

External-beam **radiotherapy**, primarily for **palliative** relief of local symptomatic tumor **growth**.

chemotherapy. The following **regimens produce** similar survival outcomes: cisplatin plus vinblastine []. cisplatin plus vinblastine plus mitomycin []. cisplatin plus vinorelbine []. cisplatin plus vindesine []. cisplatin plus paclitaxel []. carboplatin plus paclitaxel [,].

clinical trials evaluating the role of new **chemotherapy regimens**. Refer to the **clinical** trials section of PDQ for a list of **clinical** trials. The **clinical** trials in PDQ are also available on CancerNet (<http://cancernet.nci.nih.gov>).

Endobronchial laser **therapy** and/or brachytherapy for obstructing **lesions** [].

***** RECURRENT NON- **small cell lung-cancer** CANCER

Many **patients** with recurrent non- **small cell lung-cancer** cancer (NSCLC) are eligible for **clinical** trials. **radiotherapy** may provide excellent palliation of **symptoms** from a localized tumor mass. **patients** who **present** with a solitary cerebral metastasis after resection of a **primary** NSCLC **lesion** and who have no **evidence** of extracranial tumor can **achieve** prolonged **disease-free** survival with **surgical** excision of the brain metastasis and postoperative whole-brain irradiation [,]. Unresectable brain metastases in this setting may be **treated** radiosurgically []. Because of the **small** potential for long-term survival, **radiotherapy** should be delivered by conventional methods in daily **doses** of - cGy, while higher daily **doses** over a **shorter** period of time (hypofractionated schemes) should be avoided because of the high risk of **toxic** effects observed with such **treatments** []. Most **patients** not suitable for **surgical** resection should receive conventional whole-brain **radiotherapy**. Selected **patients** with good **performance** status and **small** metastases can be considered for stereotactic radiosurgery [].

approximately one half of **patients treated** with resection and postoperative **radiotherapy** will **develop** recurrence in the brain; some of these **patients** will be suitable for additional **treatment** []. In those selected **patients** with good **performance** status and without progressive metastases outside of the brain, **treatment** options include reoperation or stereotactic radiosurgery [,]. For most **patients**, conventional **radiotherapy** can be considered; however, the **palliative** benefit of this **treatment** is **limited** [].

A solitary pulmonary metastasis from an initially resected bronchogenic **carcinoma** is unusual. The lung is frequently the site of second **primary** malignancies in **patients** with **primary** lung **cancers**. **determining** whether the new **lesion**

is a new **primary cancer** or a metastasis may be difficult. Studies have indicated that in the majority of **patients** the new **lesion** is a second **primary** tumor, and following resection some **patients** may **achieve** long-term survival. Thus, if the first **primary** tumor has been controlled, the second **primary** tumor should be resected if possible [,].

The use of **chemotherapy** has **produced** objective responses and **small** improvement in survival for **patients** with metastatic **disease** []. In studies that have examined symptomatic response, improvement in subjective **symptoms** has been reported to occur more frequently than objective response [,]. Informed **patients** with good **performance** status and symptomatic recurrence can be offered **treatment** with a cisplatin- *based chemotherapy regimen* for palliation of **symptoms**.

treatment options:

palliative radiotherapy.

chemotherapy alone.

surgical resection of isolated cerebral metastasis (highly selected **patients**) [].

Laser **therapy** or interstitial **radiotherapy** for endobronchial **lesions** [].

Stereotactic radiosurgery (highly selected **patients**) [, ,].

Apéndice VII d: Representación de la cadena 1 del texto QDT2

lung-cancer, lung-cancer, lung-cancer, carcinoma, carcinoma, symptoms, patients, disease, patients, diagnosis, patients, disease, patients, prognosis, patients, disease, patients, lung-cancer, patients, disease, patients, diagnosis, symptoms, patients, patients, confined, symptoms, patients, prognosis, symptoms, patients, lung-cancer, patients, patients, patient, lung-cancer, lung-cancer, critical, cases, lung-cancer, carcinoma, cancers, cancers, carcinoma, cancer, lesion, cancer, patients, cancer, patients, lung-cancer, lung-cancer, carcinoma, carcinoma, carcinoma, carcinoma, critical, patients, critical, disease, patients, patients, patients, lung-cancer, lung-cancer, cancer, cancer, patient, lesion, lesions, lesions, lesion, cancer, bronchial, carcinoma, bronchial, lung-cancer, patients, cases, patient, carcinoma, lung-cancer, cancers, patients, disease, patients, disease, disease, cancers, lung-cancer, lung-cancer, lung-cancer, patient, lung-cancer, patients, disease, lung-cancer, lung-cancer, carcinoma, cancers, patients, patients, cancers, lung-cancer, patients, lung-cancer, patient, condition, patient, critical, patients, lung-cancer, patients, cancer, patients, patients, patients, patients, inoperable, patients, disease, patients, inoperable, patients, inoperable, patients, patients, critical, patients, patients, smokers, lung-cancer, patients, cancers, cancers, cancers, cancers, patients, cancers, cancers, patients, patients, patients, lung-cancer, patients, lung-cancer, patient, condition, patient, critical, inoperable, patients, disease, patients, patients, inoperable, patients, disease, disease, critical, patients, patients, patients, carcinoma, patients, disease, patients, lung-cancer, patients, lung-cancer, patients, patients, patients, patients, disease, patient, patients, disease, patients, patients, patients, cancer, patients, patients, patients, lung-cancer, patients, patients, disease, cases, disease, cases, patients, patients, cases, patients, lung-cancer, patients, lung-cancer, patients, patients, patients, disease, patients, patient, patients, disease, cough, chest-pain, patient, patients, patients, disease, patients, inoperable, patients, patient, patients, disease, patients, lung-cancer, toxic, patients, patients, patients, lung-cancer, lung-cancer, patients, inoperable, disease, lung-cancer, patients, patients, lesions, bronchial, pain, cases, lesions, critical, patient, patient, lesion, toxic, patients, lesions, lung-cancer, patients, lung-cancer, patients, lesion, disease, toxic, patients, patients, patients, patients, patients, patients, carcinoma, patients, cancers, lesion, cancer, patients, lesion, patients, patients, disease, patients, patients, lesions, patients,

```
:lung-cancer      :Wd  4, Sent  1, Para  1, ???->  0, Value 0
  respiratory-disease 651 4192 n
:lung-cancer      :Wd 11, Sent  1, Para  3, ID->  4, Value 0
  respiratory-disease 651 4192 n
:lung-cancer      :Wd 25, Sent  1, Para  3, ID-> 11, Value 0
  respiratory-disease 651 4192 n
:carcinoma        :Wd 31, Sent  1, Para  3, GRP-> 25, Value 0
  cancer 651 4195 n
:carcinoma        :Wd 36, Sent  1, Para  3, ID-> 31, Value 0
  cancer 651 4195 n
:symptoms         :Wd 65, Sent  1, Para  3, GRP-> 36, Value 0
  illness 651 4186 n
:patients         :Wd 81, Sent  1, Para  3, GRP-> 65, Value 0
  sick-person 651 4202 n
:disease          :Wd 84, Sent  1, Para  3, GRP-> 81, Value 0
  disease 651 4187 n
:patients         :Wd 95, Sent  1, Para  4, ID-> 81, Value 0
  sick-person 651 4202 n
:diagnosis        :Wd 97, Sent  1, Para  4, GRP-> 95, Value 0
  pathology 651 4203 n
:patients         :Wd 98, Sent  1, Para  4, ID-> 95, Value 0
  sick-person 651 4202 n
:disease          :Wd 112, Sent  1, Para  4, ID-> 84, Value 0
  disease 651 4187 n
:patients         :Wd 120, Sent  1, Para  4, ID-> 98, Value 0
  sick-person 651 4202 n
:prognosis        :Wd 139, Sent  1, Para  4, CAT-> 97, Value 0
  pathology 651 4203 n
:patients         :Wd 149, Sent  1, Para  4, ID->120, Value 0
  sick-person 651 4202 n
:disease          :Wd 152, Sent  1, Para  4, ID->112, Value 0
  disease 651 4187 n
:patients         :Wd 169, Sent  1, Para  4, ID->149, Value 0
  sick-person 651 4202 n
```

:lung-cancer :Wd 180, Sent 1, Para 4, ID-> 25, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 204, Sent 1, Para 4, ID->169, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 209, Sent 1, Para 4, ID->152, Value 0
[disease 651 4187 n](#)

:patients :Wd 222, Sent 1, Para 4, ID->204, Value 0
[sick-person 651 4202 n](#)

:diagnosis :Wd 232, Sent 1, Para 4, ID-> 97, Value 0
[pathology 651 4203 n](#)

:symptoms :Wd 245, Sent 1, Para 4, ID-> 65, Value 0
[illness 651 4186 n](#)

:patients :Wd 250, Sent 1, Para 4, ID->222, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 258, Sent 1, Para 4, ID->250, Value 0
[sick-person 651 4202 n](#)

:confined :Wd 262, Sent 1, Para 4, GRP->258, Value 0
[sick 651 4205 a](#)

:symptoms :Wd 286, Sent 1, Para 4, ID->245, Value 0
[illness 651 4186 n](#)

:patients :Wd 305, Sent 1, Para 5, ID->258, Value 0
[sick-person 651 4202 n](#)

:prognosis :Wd 306, Sent 1, Para 5, ID->139, Value 0
[pathology 651 4203 n](#)

:symptoms :Wd 315, Sent 1, Para 5, ID->286, Value 0
[illness 651 4186 n](#)

:patients :Wd 343, Sent 1, Para 5, ID->305, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 349, Sent 1, Para 5, ID->180, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 380, Sent 1, Para 6, ID->343, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 397, Sent 1, Para 6, ID->380, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 401, Sent 1, Para 6, ID->397, Value 0
[sick-person 651 4202 n](#)

:patient :Wd 417, Sent 1, Para 9, CAT->401, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 419, Sent 1, Para 9, ID->349, Value 0
[respiratory-disease 651 4192 n](#)

:lung-cancer :Wd 429, Sent 1, Para 9, ID->419, Value 0
[respiratory-disease 651 4192 n](#)

:critical :Wd 433, Sent 1, Para 9, CAT->262, Value 0
[sick 651 4205 a](#)

:cases :Wd 436, Sent 1, Para 9, CAT->417, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 440, Sent 1, Para 9, ID->429, Value 0
[respiratory-disease 651 4192 n](#)

:carcinoma :Wd 457, Sent 1, Para 9, ID-> 36, Value 0
[cancer 651 4195 n](#)

:cancers :Wd 462, Sent 1, Para 9, CAT->457, Value 0
[cancer 651 4195 n](#)

:cancers :Wd 485, Sent 1, Para 9, ID->462, Value 0
[cancer 651 4195 n](#)

:carcinoma :Wd 489, Sent 1, Para 9, ID->457, Value 0
[cancer 651 4195 n](#)

:cancer :Wd 505, Sent 1, Para 9, CAT->489, Value 0
[cancer 651 4195 n](#)

:lesion :Wd 512, Sent 1, Para 9, GRP->505, Value 0
[ulcer 651 4198 n](#)

:cancer :Wd 517, Sent 1, Para 9, ID->505, Value 0
[cancer 651 4195 n](#)

:patients :Wd 519, Sent 1, Para 9, ID->401, Value 0
[sick-person 651 4202 n](#)
:cancer :Wd 535, Sent 1, Para 9, ID->517, Value 0
[cancer 651 4195 n](#)
:patients :Wd 543, Sent 1, Para 9, ID->519, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 554, Sent 1, Para 9, ID->440, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 562, Sent 1, Para 10, ID->554, Value 0
[respiratory-disease 651 4192 n](#)
:carcinoma :Wd 567, Sent 2, Para 11, ID->489, Value 0
[cancer 651 4195 n](#)
:carcinoma :Wd 590, Sent 2, Para 11, ID->567, Value 0
[cancer 651 4195 n](#)
:carcinoma :Wd 599, Sent 2, Para 11, ID->590, Value 0
[cancer 651 4195 n](#)
:carcinoma :Wd 602, Sent 2, Para 12, ID->599, Value 0
[cancer 651 4195 n](#)
:critical :Wd 632, Sent 2, Para 14, ID->433, Value 0
[sick 651 4205 a](#)
:patients :Wd 638, Sent 2, Para 15, ID->543, Value 0
[sick-person 651 4202 n](#)
:critical :Wd 642, Sent 2, Para 15, ID->632, Value 0
[sick 651 4205 a](#)
:disease :Wd 652, Sent 2, Para 15, ID->209, Value 0
[disease 651 4187 n](#)
:patients :Wd 788, Sent 2, Para 15, ID->638, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 806, Sent 2, Para 15, ID->788, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 814, Sent 2, Para 15, ID->806, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 834, Sent 2, Para 16, ID->562, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 842, Sent 2, Para 17, ID->834, Value 0
[respiratory-disease 651 4192 n](#)
:cancer :Wd 853, Sent 2, Para 17, ID->535, Value 0
[cancer 651 4195 n](#)
:cancer :Wd 860, Sent 2, Para 17, ID->853, Value 0
[cancer 651 4195 n](#)
:patient :Wd 872, Sent 2, Para 17, ID->417, Value 0
[sick-person 651 4202 n](#)
:lesion :Wd 964, Sent 2, Para 17, ID->512, Value 0
[ulcer 651 4198 n](#)
:lesions :Wd 975, Sent 2, Para 17, CAT->964, Value 0
[ulcer 651 4198 n](#)
:lesions :Wd 989, Sent 2, Para 17, ID->975, Value 0
[ulcer 651 4198 n](#)
:lesion :Wd 996, Sent 2, Para 17, ID->964, Value 0
[ulcer 651 4198 n](#)
:cancer :Wd 1004, Sent 2, Para 18, ID->860, Value 0
[cancer 651 4195 n](#)
:bronchial :Wd 1037, Sent 3, Para 20, GRP->1004, Value 0
[diseased 651 4206 a](#)
:carcinoma :Wd 1054, Sent 5, Para 20, ID->602, Value 0
[cancer 651 4195 n](#)
:bronchial :Wd 1254, Sent 13, Para 21, ID->1037, Value 0
[diseased 651 4206 a](#)
:lung-cancer :Wd 1275, Sent 14, Para 21, ID->842, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 1286, Sent 14, Para 21, ID->814, Value 0
[sick-person 651 4202 n](#)

:cases :Wd 1301, Sent 14, Para 21, CAT->1286, Value 0
[sick-person 651 4202 n](#)

:patient :Wd 1338, Sent 14, Para 21, ID->872, Value 0
[sick-person 651 4202 n](#)

:carcinoma :Wd 1507, Sent 25, Para 30, ID->1054, Value 0
[cancer 651 4195 n](#)

:lung-cancer :Wd 1587, Sent 33, Para 33, ID->1275, Value 0
[respiratory-disease 651 4192 n](#)

:cancers :Wd 1602, Sent 33, Para 33, CAT->1507, Value 0
[cancer 651 4195 n](#)

:patients :Wd 1606, Sent 33, Para 33, ID->1286, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 1629, Sent 33, Para 33, GRP->1606, Value 0
[disease 651 4187 n](#)

:patients :Wd 1644, Sent 33, Para 33, ID->1606, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 1648, Sent 33, Para 33, ID->1629, Value 0
[disease 651 4187 n](#)

:disease :Wd 1675, Sent 33, Para 33, ID->1648, Value 0
[disease 651 4187 n](#)

:cancers :Wd 1757, Sent 33, Para 34, ID->1602, Value 0
[cancer 651 4195 n](#)

:lung-cancer :Wd 1771, Sent 33, Para 34, ID->1587, Value 0
[respiratory-disease 651 4192 n](#)

:lung-cancer :Wd 1803, Sent 33, Para 37, ID->1771, Value 0
[respiratory-disease 651 4192 n](#)

:lung-cancer :Wd 1810, Sent 33, Para 38, ID->1803, Value 0
[respiratory-disease 651 4192 n](#)

:patient :Wd 1872, Sent 33, Para 38, ID->1338, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 1886, Sent 33, Para 38, ID->1810, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 1888, Sent 33, Para 38, ID->1644, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 1892, Sent 33, Para 39, ID->1675, Value 0
[disease 651 4187 n](#)

:lung-cancer :Wd 1897, Sent 33, Para 39, ID->1886, Value 0
[respiratory-disease 651 4192 n](#)

:lung-cancer :Wd 1906, Sent 33, Para 40, ID->1897, Value 0
[respiratory-disease 651 4192 n](#)

:carcinoma :Wd 1913, Sent 33, Para 40, ID->1507, Value 0
[cancer 651 4195 n](#)

:cancers :Wd 1946, Sent 33, Para 40, ID->1757, Value 0
[cancer 651 4195 n](#)

:patients :Wd 1970, Sent 33, Para 40, ID->1888, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 2033, Sent 34, Para 42, ID->1970, Value 0
[sick-person 651 4202 n](#)

:cancers :Wd 2041, Sent 34, Para 43, ID->1946, Value 0
[cancer 651 4195 n](#)

:lung-cancer :Wd 2053, Sent 34, Para 45, ID->1906, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 2069, Sent 34, Para 46, ID->2033, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 2076, Sent 34, Para 46, ID->2053, Value 0
[respiratory-disease 651 4192 n](#)

:patient :Wd 2084, Sent 34, Para 46, ID->1872, Value 0
[sick-person 651 4202 n](#)

:condition :Wd 2088, Sent 34, Para 46, GRP->2084, Value 0
[illness 651 4186 n](#)

:patient :Wd 2091, Sent 34, Para 46, ID->2084, Value 0
[sick-person 651 4202 n](#)

:critical :Wd 2096, Sent 34, Para 46, GRP->2091, Value 0
[sick 651 4205 a](#)
:patients :Wd 2119, Sent 34, Para 46, ID->2069, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 2137, Sent 34, Para 46, ID->2076, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 2154, Sent 34, Para 46, ID->2119, Value 0
[sick-person 651 4202 n](#)
:cancer :Wd 2158, Sent 34, Para 46, CAT->2041, Value 0
[cancer 651 4195 n](#)
:patients :Wd 2174, Sent 34, Para 46, ID->2154, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2219, Sent 34, Para 46, ID->2174, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2255, Sent 34, Para 46, ID->2219, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2287, Sent 34, Para 46, ID->2255, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2311, Sent 34, Para 46, ID->2287, Value 0
[sick-person 651 4202 n](#)
:inoperable :Wd 2326, Sent 34, Para 47, CAT->2096, Value 0
[sick 651 4205 a](#)
:patients :Wd 2327, Sent 34, Para 47, ID->2311, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 2331, Sent 34, Para 47, ID->1892, Value 0
[disease 651 4187 n](#)
:patients :Wd 2349, Sent 34, Para 47, ID->2327, Value 0
[sick-person 651 4202 n](#)
:inoperable :Wd 2366, Sent 34, Para 47, ID->2326, Value 0
[sick 651 4205 a](#)
:patients :Wd 2387, Sent 34, Para 47, ID->2349, Value 0
[sick-person 651 4202 n](#)
:inoperable :Wd 2404, Sent 34, Para 47, ID->2366, Value 0
[sick 651 4205 a](#)
:patients :Wd 2405, Sent 34, Para 47, ID->2387, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2420, Sent 34, Para 47, ID->2405, Value 0
[sick-person 651 4202 n](#)
:critical :Wd 2492, Sent 34, Para 48, ID->2096, Value 0
[sick 651 4205 a](#)
:patients :Wd 2512, Sent 34, Para 49, ID->2420, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 2524, Sent 34, Para 49, ID->2512, Value 0
[sick-person 651 4202 n](#)
:smokers :Wd 2554, Sent 34, Para 49, CAT->2137, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 2573, Sent 34, Para 49, ID->2137, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 2583, Sent 34, Para 49, ID->2524, Value 0
[sick-person 651 4202 n](#)
:cancers :Wd 2589, Sent 34, Para 49, ID->2041, Value 0
[cancer 651 4195 n](#)
:cancers :Wd 2599, Sent 34, Para 49, ID->2589, Value 0
[cancer 651 4195 n](#)
:cancers :Wd 2621, Sent 34, Para 49, ID->2599, Value 0
[cancer 651 4195 n](#)
:cancers :Wd 2626, Sent 34, Para 49, ID->2621, Value 0
[cancer 651 4195 n](#)
:patients :Wd 2641, Sent 34, Para 49, ID->2583, Value 0
[sick-person 651 4202 n](#)
:cancers :Wd 2649, Sent 34, Para 49, ID->2626, Value 0
[cancer 651 4195 n](#)

:cancers :Wd 2681, Sent 34, Para 49, ID->2649, Value 0
[cancer 651 4195 n](#)

:patients :Wd 2683, Sent 34, Para 49, ID->2641, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 2709, Sent 35, Para 52, ID->2683, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 2742, Sent 35, Para 55, ID->2709, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 2751, Sent 35, Para 57, ID->2573, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 2771, Sent 35, Para 58, ID->2742, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 2778, Sent 35, Para 58, ID->2751, Value 0
[respiratory-disease 651 4192 n](#)

:patient :Wd 2786, Sent 35, Para 58, ID->2091, Value 0
[sick-person 651 4202 n](#)

:condition :Wd 2790, Sent 35, Para 58, ID->2088, Value 0
[illness 651 4186 n](#)

:patient :Wd 2793, Sent 35, Para 58, ID->2786, Value 0
[sick-person 651 4202 n](#)

:critical :Wd 2798, Sent 35, Para 58, ID->2492, Value 0
[sick 651 4205 a](#)

:inoperable :Wd 2824, Sent 35, Para 59, ID->2404, Value 0
[sick 651 4205 a](#)

:patients :Wd 2825, Sent 35, Para 59, ID->2771, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 2829, Sent 35, Para 59, ID->2331, Value 0
[disease 651 4187 n](#)

:patients :Wd 2846, Sent 35, Para 59, ID->2825, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 2879, Sent 35, Para 59, ID->2846, Value 0
[sick-person 651 4202 n](#)

:inoperable :Wd 2882, Sent 35, Para 59, ID->2824, Value 0
[sick 651 4205 a](#)

:patients :Wd 2897, Sent 35, Para 59, ID->2879, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 2904, Sent 35, Para 59, ID->2829, Value 0
[disease 651 4187 n](#)

:disease :Wd 2916, Sent 35, Para 59, ID->2904, Value 0
[disease 651 4187 n](#)

:critical :Wd 2984, Sent 35, Para 59, ID->2798, Value 0
[sick 651 4205 a](#)

:patients :Wd 3004, Sent 35, Para 60, ID->2897, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 3016, Sent 35, Para 60, ID->3004, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 3051, Sent 35, Para 60, ID->3016, Value 0
[sick-person 651 4202 n](#)

:carcinoma :Wd 3057, Sent 35, Para 60, CAT->2681, Value 0
[cancer 651 4195 n](#)

:patients :Wd 3078, Sent 35, Para 60, ID->3051, Value 0
[sick-person 651 4202 n](#)

:disease :Wd 3090, Sent 35, Para 60, ID->2916, Value 0
[disease 651 4187 n](#)

:patients :Wd 3145, Sent 36, Para 63, ID->3078, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 3173, Sent 36, Para 67, ID->2778, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 3256, Sent 36, Para 69, ID->3145, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 3263, Sent 36, Para 69, ID->3173, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 3280, Sent 36, Para 69, ID->3256, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3299, Sent 36, Para 69, ID->3280, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3311, Sent 36, Para 69, ID->3299, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3347, Sent 36, Para 69, ID->3311, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 3376, Sent 36, Para 69, ID->3090, Value 0
[disease 651 4187 n](#)
:patient :Wd 3439, Sent 37, Para 70, ID->2793, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3490, Sent 37, Para 71, ID->3347, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 3493, Sent 37, Para 71, ID->3376, Value 0
[disease 651 4187 n](#)
:patients :Wd 3540, Sent 37, Para 71, ID->3490, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3548, Sent 37, Para 71, ID->3540, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3571, Sent 37, Para 71, ID->3548, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3604, Sent 37, Para 71, ID->3571, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3619, Sent 37, Para 71, ID->3604, Value 0
[sick-person 651 4202 n](#)
:cancer :Wd 3628, Sent 37, Para 71, CAT->3493, Value 0
[disease 651 4187 n](#)
:patients :Wd 3655, Sent 37, Para 71, ID->3619, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3680, Sent 37, Para 72, ID->3655, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3699, Sent 37, Para 72, ID->3680, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 3709, Sent 37, Para 72, ID->3263, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 3718, Sent 37, Para 72, ID->3699, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3740, Sent 37, Para 72, ID->3718, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 3752, Sent 37, Para 72, ID->3493, Value 0
[disease 651 4187 n](#)
:cases :Wd 3812, Sent 38, Para 74, CAT->3740, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 3875, Sent 38, Para 78, ID->3752, Value 0
[disease 651 4187 n](#)
:cases :Wd 3890, Sent 38, Para 78, ID->3812, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3896, Sent 38, Para 78, ID->3740, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3910, Sent 38, Para 78, ID->3896, Value 0
[sick-person 651 4202 n](#)
:cases :Wd 3964, Sent 39, Para 83, ID->3890, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 3990, Sent 39, Para 86, ID->3910, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 4037, Sent 40, Para 93, ID->3709, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 4039, Sent 40, Para 94, ID->3990, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 4046, Sent 40, Para 94, ID->4037, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 4078, Sent 40, Para 94, ID->4039, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4091, Sent 40, Para 94, ID->4078, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4110, Sent 40, Para 94, ID->4091, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 4122, Sent 40, Para 94, ID->3875, Value 0
[disease 651 4187 n](#)
:patients :Wd 4128, Sent 40, Para 94, ID->4110, Value 0
[sick-person 651 4202 n](#)
:patient :Wd 4163, Sent 40, Para 94, ID->3439, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4190, Sent 40, Para 95, ID->4128, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 4194, Sent 40, Para 95, ID->4122, Value 0
[disease 651 4187 n](#)
:cough :Wd 4210, Sent 40, Para 95, CAT->4046, Value 0
[respiratory-disease 651 4192 n](#)
:chest-pain :Wd 4216, Sent 40, Para 95, GRP->4210, Value 0
[cardiovascular-disease 651 4193 n](#)
:patient :Wd 4237, Sent 40, Para 96, ID->4163, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4261, Sent 40, Para 96, ID->4190, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4274, Sent 40, Para 96, ID->4261, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 4288, Sent 40, Para 96, ID->4194, Value 0
[disease 651 4187 n](#)
:patients :Wd 4315, Sent 40, Para 96, ID->4274, Value 0
[sick-person 651 4202 n](#)
:inoperable :Wd 4317, Sent 40, Para 96, GRP->4315, Value 0
[sick 651 4205 a](#)
:patients :Wd 4325, Sent 40, Para 96, ID->4315, Value 0
[sick-person 651 4202 n](#)
:patient :Wd 4365, Sent 40, Para 96, ID->4237, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4461, Sent 40, Para 97, ID->4325, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 4488, Sent 40, Para 98, ID->4288, Value 0
[disease 651 4187 n](#)
:patients :Wd 4489, Sent 40, Para 98, ID->4461, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 4556, Sent 41, Para 105, ID->4046, Value 0
[respiratory-disease 651 4192 n](#)
:toxic :Wd 4591, Sent 41, Para 106, GRP->4556, Value 0
[diseased 651 4206 a](#)
:patients :Wd 4624, Sent 41, Para 106, ID->4489, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4650, Sent 41, Para 106, ID->4624, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 4702, Sent 41, Para 106, ID->4650, Value 0
[sick-person 651 4202 n](#)
:lung-cancer :Wd 4728, Sent 41, Para 106, ID->4556, Value 0
[respiratory-disease 651 4192 n](#)
:lung-cancer :Wd 4804, Sent 41, Para 106, ID->4728, Value 0
[respiratory-disease 651 4192 n](#)
:patients :Wd 4827, Sent 41, Para 106, ID->4702, Value 0
[sick-person 651 4202 n](#)
:inoperable :Wd 4829, Sent 41, Para 106, ID->4317, Value 0
[sick 651 4205 a](#)
:disease :Wd 4834, Sent 41, Para 106, ID->4488, Value 0
[disease 651 4187 n](#)

:lung-cancer :Wd 4855, Sent 41, Para 107, ID->4804, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 4875, Sent 41, Para 107, ID->4827, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 4896, Sent 41, Para 107, ID->4875, Value 0
[sick-person 651 4202 n](#)

:lesions :Wd 4904, Sent 41, Para 107, GRP->4896, Value 0
[ulcer 651 4198 n](#)

:bronchial :Wd 4936, Sent 41, Para 108, CAT->4591, Value 0
[diseased 651 4206 a](#)

:pain :Wd 4942, Sent 41, Para 108, GRP->4936, Value 0
[illness 651 4186 n](#)

:cases :Wd 4954, Sent 41, Para 108, CAT->4896, Value 0
[sick-person 651 4202 n](#)

:lesions :Wd 4967, Sent 41, Para 108, ID->4904, Value 0
[ulcer 651 4198 n](#)

:critical :Wd 4975, Sent 41, Para 108, CAT->4829, Value 0
[sick 651 4205 a](#)

:patient :Wd 4987, Sent 41, Para 108, ID->4365, Value 0
[sick-person 651 4202 n](#)

:patient :Wd 4991, Sent 41, Para 108, ID->4987, Value 0
[sick-person 651 4202 n](#)

:lesion :Wd 5014, Sent 41, Para 108, CAT->4967, Value 0
[ulcer 651 4198 n](#)

:toxic :Wd 5050, Sent 41, Para 108, ID->4591, Value 0
[diseased 651 4206 a](#)

:patients :Wd 5060, Sent 41, Para 109, ID->4896, Value 0
[sick-person 651 4202 n](#)

:lesions :Wd 5177, Sent 43, Para 114, ID->4967, Value 0
[ulcer 651 4198 n](#)

:lung-cancer :Wd 5185, Sent 43, Para 116, ID->4855, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 5188, Sent 43, Para 117, ID->5060, Value 0
[sick-person 651 4202 n](#)

:lung-cancer :Wd 5194, Sent 43, Para 117, ID->5185, Value 0
[respiratory-disease 651 4192 n](#)

:patients :Wd 5214, Sent 43, Para 118, ID->5188, Value 0
[sick-person 651 4202 n](#)

:lesion :Wd 5228, Sent 43, Para 118, ID->5014, Value 0
[ulcer 651 4198 n](#)

:disease :Wd 5240, Sent 43, Para 118, ID->4834, Value 0
[disease 651 4187 n](#)

:toxic :Wd 5311, Sent 43, Para 118, ID->5050, Value 0
[diseased 651 4206 a](#)

:patients :Wd 5320, Sent 43, Para 118, ID->5214, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 5333, Sent 43, Para 118, ID->5320, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 5353, Sent 43, Para 119, ID->5333, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 5369, Sent 43, Para 119, ID->5353, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 5381, Sent 43, Para 119, ID->5369, Value 0
[sick-person 651 4202 n](#)

:patients :Wd 5405, Sent 43, Para 119, ID->5381, Value 0
[sick-person 651 4202 n](#)

:carcinoma :Wd 5431, Sent 43, Para 120, GRP->5405, Value 0
[cancer 651 4195 n](#)

:patients :Wd 5445, Sent 43, Para 120, ID->5405, Value 0
[sick-person 651 4202 n](#)

:cancers :Wd 5449, Sent 43, Para 120, CAT->5431, Value 0
[cancer 651 4195 n](#)

:lesion :Wd 5454, Sent 43, Para 120, ID->5228, Value 0
[ulcer 651 4198 n](#)
:cancer :Wd 5459, Sent 43, Para 120, CAT->5449, Value 0
[cancer 651 4195 n](#)
:patients :Wd 5474, Sent 43, Para 120, ID->5445, Value 0
[sick-person 651 4202 n](#)
:lesion :Wd 5477, Sent 43, Para 120, ID->5454, Value 0
[ulcer 651 4198 n](#)
:patients :Wd 5487, Sent 43, Para 120, ID->5474, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 5527, Sent 43, Para 121, ID->5487, Value 0
[sick-person 651 4202 n](#)
:disease :Wd 5530, Sent 43, Para 121, ID->5240, Value 0
[disease 651 4187 n](#)
:patients :Wd 5557, Sent 43, Para 121, ID->5527, Value 0
[sick-person 651 4202 n](#)
:patients :Wd 5593, Sent 44, Para 125, ID->5557, Value 0
[sick-person 651 4202 n](#)
:lesions :Wd 5603, Sent 44, Para 126, ID->5177, Value 0
[ulcer 651 4198 n](#)
:patients :Wd 5610, Sent 44, Para 127, ID->5593, Value 0
[sick-person 651 4202 n](#)

Return to [Source file](#)

APÉNDICE VIIe: Representación tabular de las cadenas reconocidas por *Hesperus*

**TREATMENT SUMMARIES FOR HEALTH PROFESSIONALS
MALIGNANT MESOTHELIOMA (QDT1): 15 CHAINS**

H 0: Experiment (5, 1)

TRIALS	5
--------	---

H 1: Treatment, diagnosis, pathology of malignat mesothelioma (96, 29)

DISEASE	13
PATIENTS	12
CLINICAL	9
SURGICAL	8
THERAPY	8
CHEMOTHERAPY	5
PALLIATIVE	5
TREATED	5
PROGNOSIS	3
SURGERY	3
CARCINOMAS	2
CURE	2
DIAGNOSIS	2
PAIN	2
REGIMENS	2

SYMPTOMS	2
ANAPLASTIC	1
BIOPSY	1
CAPSULE	1
CARCINOMA	1
CASES	1
CONFINED	1
CURATIVE	1
DIAGNOSTIC	1
HISTORY-OF	1
LESIONS	1
SYMPTOM	1
TOXICITY	1
TREATMENTS	1

H 2: ? (23, 7)

2 a: Staging (18, 3)

STAGE (stand n)	12
STAGES (stand n)	3
STAGING (frame n)	3
TOTAL	18

2 b: Improvement (5, 4)

IMPROVED (restored a)	2
FOUND (restored a)	1
PROVIDE (support v)	1

RELIEF (recuperation n)	1
-------------------------	---

H 3: ? (17, 5)

3 a: Treatment (12, 1)

TREATMENT	12
-----------	----

3 b: Change (5, 4)

STAINS	2
INTERPRETATION	1
MODIFICATION	1
VARIABILITY	1

H 4: Survival (15, 3)

SURVIVAL	11
LONG-TERM	3
PROLONGED	1

H 5: Pleura (10, 3)

PLEURAL	6
PLEURA	3
PARIETAL	1

H 6: ? (20, 13)

DIFFICULT (puzzling a, moot a)	3
PROPOSED (moot a)	3
ACID (sullen a)	2
HIGHER (exam n)	2
INFORMATION (enquiry n)	2
CANDIDATES (exam n)	1
DISCUSSION (enquiry n)	1

EXAMINATION (enquiry n)	1
INVOLVEMENT (difficulty n)	1
OPERATIVE (secret-service n)	1
PROBLEMS (difficulty n)	1
UNCLEAR (puzzling a)	1
UNKNOWN (unintelligible a)	1

H 7: Risk factor (2,2)

ASBESTOS (incombustibility n)	1
EXPOSURE (refrigeration n)	1

H 8: Tests (2, 2)

ELECTRON	1
MICROSCOPY	1

H 9: Time (5, 4)

TIME	2
PERIOD	1
PERIODIC	1
TEMPORARY	1

H 10: ? (11, 8)

MARGINS (edition n)	2
NEGATIVE (numerical a)	2
SERIES (edition n)	2
DESCRIPTION (reading-matter n)	1
DESCRIPTIONS	1

(reading-matter n)	
FACTOR (numerical element n)	1
POSITIVE (numerical a)	1
SELECTION (textbook n)	1

H 11: ? (15, 10)

AGENT (intermediary n)	3
DIAPHRAGM (middle n)	3
SPECIMENS (example n)	2
ADEQUATE (middling n)	1
AGENTS (intermediary n)	1
HEART (middle n)	1

INSTANCES (example n)	1
INTERVENTION (interjacency n)	1
MEDIAN (middling n)	1
WALL (partition n)	1

H 12: Size, increase, importance (18, 17)

IMPORTANT (great a)	2
ADDITION (increase n)	1
EXTENSION (increase n)	1
EXTENSIVE	1
GENERALLY (greatly r)	1
GREAT (large a)	1
GROSS (consummate a)	1
GROWS (grow v)	1

HIGH (great a)	1
IMPROVE (grow v)	1
INTEREST (increment n)	1
LARGE (great a)	1
MAJOR (great a)	1
PRODUCE (increment)	1
PROGRESSION (increase n)	1
VALUABLE (great a)	1

WIDE (great a)	1
----------------	---

H 13: Part-whole, composition (16, 10)

COMBINATION (union n)	3
INCLUDE (add / join v)	3
INCLUDING (in-addition r)	2
NODES (joint n)	2
COMBINED (joined a)	1
COMPOSED-OF	1

(composing a)	
IMPACT (affix v)	1
MAKING (composing a)	1
NATURE (composition n)	1
NODE (joint n)	1

H 14: Order, categorization (10, 8)

SYSTEM (arrangement n)	3
ADMINISTRATION (arrangement n)	1
CLASSIFICATION (arrangement n)	1
GROUP (arrangement n)	1

LIMITED (comparative a)	1
MAJORITY (degree n)	1
SIZE (degree n)	1
STRUCTURES (structure n)	1
SYSTEMS (arrangement n)	1

NON-SMALL CELL LUNG CANCER (QDT2): 13 CHAINS

H 0: Treatment (369, 39)

RADIOTHERAPY	65
CHEMOTHERAPY	54
TREATMENT	37
CLINICAL	35
SURGERY	29
SURGICAL	16
TREATED	16
THERAPY	15
CURATIVE	13
REGIMENS	8
SURGICALLY	7
SYMPTOMS	7
MEDICAL	5
REGIMEN	5
THERAPEUTIC	4
BRONCHOSCOPY	3
CURE	3
DIAGNOSTIC	3
DOSES	3
IMMUNOTHERAPY	3
PALLIATIVE	3
PROGNOSIS	3

SYNDROME	3
TOMOGRAPHY	3
TREATMENTS	3
BIOPSY	2
DIAGNOSIS	2
DRUG	2
MEDICALLY	2
PROGNOSIS	2
OPERABLE	2
RADIOGRAPH	2
AID	1
ANTIBODIES	1
ONCOLOGY	1
PALLIATE	1
PATHOLOGIST	1
PHOTOTHERAPY	1
RADIOLOGY	1
SCANNING	1
SPECIFIC	1
TOTAL	369

H 1: Pathology of lung cancer (276, 23)

Pathology of lung cancer

LUNG-CANCER	37
DISEASE	25
CANCERS	13
CARCINOMA	13
CANCER	9
CRITICAL	8
INOPERABLE	7
LESION	7
LESIONS	6
SYMPTOMS	4

BRONCHIAL	3
TOXIC	3
CONDITION	2
DIAGNOSIS	2
PROGNOSIS	2
CHEST-PAIN	1
CONFINED	1
PAIN	1
TOTAL	144

Patient

PATIENTS	110
PATIENT	14
CASES	6

Smoking

COUGH	1
SMOKERS	1

H 2: ? (261, 76)

2 a: Staging (71, 4)

STAGE (show v)	56
STAGING (dramaturgy n)	10
TIS (produce v)	3
STAGES (show v)	2
TOTAL	71

2 b: Cause and effect, evidence, importance (193, 73)

Cause and effect, evidence

PERFORMANCE (production n)	12
SHOWED (show v)	8
BASED (source n)	7
COMBINATION (composition n)	7
EVIDENCE (exhibit n)	7
PRESENT (show v)	6
SHOWN (shown a)	6
FREE (undisguised a)	5

PRODUCE (cause v)	5
ACHIEVE (produce v)	4
DEVELOP (manifest v)	4
EFFECTIVE (causal a)	4
FOUND (cause v)	4
RETROSPECTIVE (exhibit n)	4
AGENT (producer n)	3
DEMONSTRATE (show v)	3
ENCOURAGED (conduce v)	3

INCLUDES (contain v)	3
PRODUCED (shown a)	3
COMPARISON (manifestation n)	2
CONSIST-OF (contain v)	2
ENHANCE (manifest v)	2
FORM (produce v)	2
IMAGING (show v)	2
INCLUDE (contain v)	2
LEAD-TO (conduce v)	2
PLANNING (production n)	2
PRESENTATION (manifestation n)	2
RESULTED (product n)	2
ROLE (acting n)	2
SHOW (show v)	2
STRUCTURES (production n)	2
AGENTS (cause n)	1
APPARENT (manifest a)	1
BASIS (reason-why n)	1
CITATIONS (exhibit n)	1
CLEAR (manifest a)	1
COMPUTERIZED (productive a)	1
DERIVATIVE (effect n)	1
DETERMINE (cause v)	1
DETERMINING (cause v)	1

DISCOVERY (manifestation n)	1
DUE-TO (caused a)	1
EFFECTIVELY (causal a)	1
ESTABLISHING (produce v)	1
FASHION (produce v)	1
FINAL (causal a)	1
GROWTH (production n)	1
INITIATING (cause v)	1
INVOLVE (contain v)	1
INVOLVES (contain v)	1
JUDGE (cause v)	1
MADE (produced a)	1
OUTCOME (effect n)	1
OUTCOMES (effect n)	1
PRODUCES (produce v)	1
REFLECT (show v)	1
REPRESENTS (show v)	1
RESPONSE (effect n)	1
RESPONSES (effect n)	1
RESPONSIBLE (causal a)	1
REVIEW (stage-show n)	1
SIGNS (manifestation n)	1
SPECIMEN (exhibit n)	1
VARIETY (stage-show n)	1
TOTAL	261

Importance

PRIMARY (fundamental a)	31
CENTRAL (fundamental a)	1
IMPORTANT (fundamental a)	1

H 3: Minuteness / quantity / size (108, 24)

CELL	34
SMALL (21 in non-small cell lung cancer)	33
LIMITED	6
APPROXIMATELY	3
FRACTIONATION	3
LARGE	3
MODEST	3
REDUCED	3
FRACTIONS	2
MEDIAN	2

MODESTLY	2
SMALLER	2
BREATH	1
CELLS	1
COMPRESSION	1
FRACTION	1
INITIAL	1
MICROSCOPIC	1
RARE	1
RAY	1
SHORT	1
SHORTER	1

SHORTNESS	1
SUPERFICIAL	1

TOTAL	108
-------	-----

H 4: Hard body parts (18, 7)

NODES	7
NODE	4
NODULES	3
BONE	1
CHEST	1

NODULE	1
OSSEOUS	1
TOTAL	18

H 5: Transference (26, 3)

METASTASIS	15
METASTASES	10
REQUIRES	1

H 6: Research, experimentation (34, 2)

TRIALS	27
TRIAL	7
TOTAL	34

H 7: Body fluids (26, 13)

LYMPH	9
CLASSIFICATION	4
CATEGORIES	2
FLUID	2
BLOOD	1
BLOODY	1
DESIGNATIONS	1
FIELD	1

GROUP	1
LIFE	1
NATURE	1
QUALITY	1
SUBSET	1
TOTAL	26

H 8: ? (14, 5)

CONSIDERED (careful a)	7
CAREFUL (careful a)	3
EYE (surveillance n)	2
ACCURATE (careful a)	1

CAREFULLY (carefully r)	1
TOTAL	14

H 9: Distance (12, 5)

DISTANT	6
DISTAL	2

SINGLE	2
CONFUSED	1

GIANT	1
-------	---

H 10: ? (20, 15)

TOTAL	12
-------	----

STUDY (retreat n)	4
LOCAL (tavern n)	2
WEDGE (affix v)	2
CLOSE (housing n)	1
COMBINATION	1
COMMITTEE	1
COMPARE (join v)	1
IMPACT	1
INDEPENDENT	1

(political.-party n)	
INVOLVEMENT (union n)	1
JOINT (tavern n)	1
MARROW (vegetable n)	1
RELIEF (relief n)	1
VERTEBRAL (supporting a)	1
WALL (prop n)	1
TOTAL	20

H 11: Pleura (9, 3)

PLEURA	4
PLEURAL	4
PARIETAL	1
TOTAL	9

H 12: ? (21, 12)

12 a: Part / whole

GROUP	7
GROUPS	2
MINORITY	2
CHOICE	1
COMPONENT	1
ELEMENT	1

MAJORITY	1
PARTIAL	1
SEGMENTAL	1
TOTAL	17

12 b: Imperfection

POOR	2
NEGATIVE	1
ERRORS	1
TOTAL	4

H 13: Survival (75, 11)

SURVIVAL	45
LONG-TERM	8
YEAR	6
YEARS	5
REMAINS	4

DEATH	2
LIFE	1
OCCUR	1
PROLONGED	1
SURVIVE	1

SURVIVORS	1
-----------	---

TOTAL	75
-------	----

SMALL CELL LUNG CANCER (QDT3): 8 CHAINS

H 0: Treatment and diagnosis (225, 27)

CHEMOTHERAPY	51
TREATMENT	35
REGIMENS	23
RADIOTHERAPY	21
TREATED	18
THERAPY	14
CLINICAL	13
DRUG	11
SURGICAL	8
DOSES	5
PALLIATIVE	4
PROGNOSIS	3
SYMPTOMS	3
DOSE	2
PROPHYLACTIC	2

BRONCHOSCOPE	1
CURE	1
MEDICAL	1
MEDICALLY	1
PALLIATED	1
PATHOLOGIST	1
PHYSICIANS	1
REGIMEN	1
THERAPEUTIC	1
TRANSPLANT	1
TRANSPLANTS	1
TREATMENTS	1
TOTAL	225

H 1: Minuteness / Size (192, 39)

SMALL	59
CELL	56
LIMITED	19
MEDIAN	6
DECLINE	4
POOR	4
DECLINING	2
IMPAIRED	2
IMPAIRMENT	2
LOW	2
MICROSCOPY	2
MINIMAL	2
RARE	2
REDUCED	2
TYPE	2
TYPES	2
YIELD	2
APPROXIMATELY	1
COMPRESSION	1
DETAIL	1
ELECTRON	1

FAILING	1
FEMALE	1
FOLD	1
FRACTIONATION	1
GRANULES	1
LIGHT	1
MIXED	1
MODERATE	1
MODERATELY	1
MODEST	1
OBSTRUCTING	1
POORLY	1
REQUIRES	1
SIMPLE	1
SUGGESTION	1
TOLERABLE	1
UNSATISFACTORY	1
YIELDS	1
TOTAL	192

H 2: Pathology of lung cancer (201, 23)

Pathology of lung cancer

LUNG-CANCER	43
DISEASE	21
CARCINOMA	13
DIAGNOSIS	9
PROGNOSIS	6
CANCER	5
CONFINED	4
CARCINOMAS	3
AFFECTED	2
BRONCHIAL	2
LESIONS	2

MORBIDITY	2
SYNDROME	2
TOXICITY	2
CANCERS	1
CASES	1
DYING	1
PATHOLOGICALLY	1
POORLY	1
SYMPTOMS	1
TOXIC	1
TOTAL	123

Patient

PATIENTS	75
PATIENT	3
TOTAL	78

H 3: ? (133, 40)

3 a: Staging (41, 2)

STAGE (dramatize v)	36
STAGING (dramaturgy n)	5
TOTAL	41

3b: ? (92, 38)

Increase

ADDITIONAL (increasing a)	6
INCREASED (increasing a)	3
ADDITION (increase n)	2
CONTRIBUTE-TO (augment v)	2
EXCEEDS (grow v)	2
ADDING (increase n)	1
INCREASING (increasing a)	1
INTENSIVE (increasing a)	1
MATURE (grow v)	1

PROGRESSED (increase n)	1
PROGRESSIVE (increasing a)	1
PROLONGATION (increase n)	1
PROLONGS (augment v)	1
SPREAD (increase n)	1
TOTAL	24

Improvement

IMPROVEMENT (increase n)	7
IMPROVE (grow v)	3
IMPROVEMENTS (increase n)	2

PERFORMANCE (dramaturgy n)	2
LIVE (dramatic a)	1
TOTAL	15

Evidence

SHOWN (dramatic a)	8
PRESENT (dramatize v)	4
SHOWED (dramatize v)	2
PRESENTING (dramatize v)	1

SHOWING (dramatic a)	1
TOTAL	16

Part-whole

COMBINATION (composition n)	12
INCLUDE (contain v)	6
COMPOSED-OF (composing a)	1
INCLUSION (composition n)	1

INCORPORATION (composition n)	1
PERIOD (composition n)	1
TOTAL	22

Cause

PRODUCE (dramatize v)	6
PRODUCES (dramatize v)	2

Research, categorization

REVIEW (stage-show n)	2
ART (composition n)	1
MODELS (composition n)	1

ROLE (acting n)	1
TOTAL	5

Change

VARIETY (stage-show n)	1
------------------------	---

Miscelanea

CYCLE (stage-play n)	1
----------------------	---

H 4: Existence, survival, time (55, 17)

SURVIVAL	32
ORIGIN	3
YEARS	3
CONTINUE	2
LONG-TERM	2
OBTAIN	2
ENTITY	1
OCCUR	1
OCCURRED	1

OCCURS	1
POSSIBLY	1
PRESENCE	1
REMAIN	1
REPRESENT	1
SERIES	1
SURVIVALS	1
SURVIVORS	1
TOTAL	55

H 5: Importance, body parts (53, 19)

5 a: Body parts

CHEST	17
ABDOMEN	1
CAV	1
INTRAVENOUS	1

LIVER	1
MARROW	1
TOTAL	22

5b: Importance

CENTRAL	10
SIGNIFICANT	4
IMPORTANT	3
CORNERSTONE	2
HIGH	2
PRIMARILY	2
PRIMARY	2
CONSIDERABLE	1

HIGHLY	1
INTRINSIC	1
MAJOR	1
MATERIAL	1
SIGNIFICANTLY	1
TOTAL	31

H 6: Largeness, large quantity (56, 26)

EXTENSIVE	11
COMPLETE	8
MAJORITY	5
SUPERIOR	4
RARELY	3
AREAS	2
EXTREME	2
GREATER	2
PROLONGED	2
CLEAR	1
EXCEEDINGLY	1

EXCELLENT	1
EXTENSIVELY	1
EXTREMELY	1
GREATLY	1
HIGHLY	1
LONG	1
MASSIVE	1
POSITIVE	1
STRONGLY	1
SUBSTANTIAL	1
TOTAL	1

UNEQUIVOCALLY	1
UNIVERSALLY	1
UNUSUAL	1

WIDESPREAD	1
TOTAL	56

H 7: Smoking (3, 3)

CIGARETTE	1
JOINT	1
SMOKERS	1

TREATMENT SUMMARIES FOR PATIENTS

MESOTHELIOMA (QPT1): 7 CHAINS

H 0: Disease, death, harm, badness (81, 5)

CANCER	52
MALIGNANT	18
KILL	5
DISEASE	3

PAIN	3
TOTAL	81

H 1: Treatment and diagnosis, hospital (74, 23)

CLINICAL	16
SURGERY	13
CHEMOTHERAPY	11
DRUGS	5
SYMPTOMS	5
TEST	4
DRUG	3
HOSPITAL	2
ANSWER	1
ANSWERS	1
BIOPSIES	1
BIOPSY	1
DOSE	1

HOSPITALS	1
INJECTED	1
PILL	1
PROGNOSIS	1
SCREENING	1
SHRINK	1
SPECIALIST	1
TESTING	1
TESTS	1
TREAT	1
TOTAL	74

H 2: Body parts (54, 11)

CHEST	24
ABDOMEN	14
PLEURA	6
CAVITY	2

CELLS	2
ABDOMINAL	1
CASES	1
DRAIN	1

OPENING	1
SAC	1
SWELLING	1

TOTAL	54
-------	----

H 3: Treatment, agency (47, 10)

Treatment

TREATMENT	27
TREATMENTS	7
TREATED	3
TREAT	1
TOTAL	38

Agency

SERVICE	4
EFFECTIVENESS	1
JOBS	1
OPERATION	1
WORK	1
WORKED	1
TOTAL	10

H 4: Therapy, dispersion (28, 6)

RADIATION	11
THERAPY	11
RAYS	3
SPREAD	1
TYPE	1
X-RAY	1
TOTAL	28

H 5: Knowledge, instruction (32, 9)

INFORMATION	14
LEARN	6
CURRENT	4
PROFESSIONALS	3
DETECTION	1
DISCOVERED	1

EXPERTS	1
STUDY	1
TRAINED	1
TOTAL	32

H 6: ? (51, 18)

CELLS (party n)	11
BODY (corporation n)	7
STAGE (rostrum n)	8
FOUND (produce v)	5
STAGES (rostrum n)	4
ENERGY (causal-means n)	2
INSTITUTE (Corporation n)	2
MADE (produced a)	2
ENTERS (join-a-party v)	1
EXPLANATION (reason-why n)	1

FORM (produce v)	1
LABORATORY (classroom n)	1
MACHINE (causal-means n)	1
MEANS (causal-means n)	1
PLAN (produce v)	1
PRODUCE (produce v)	1
START (start n)	1
TOOL (causal-means n)	1
TOTAL	51

NON-SMALL CELL LUNG CANCER (QPT2): 4 CHAINS

H 0: Learn (1, 1)

LEARN	1
-------	---

H 1: Treatment and diagnosis (130, 27)

THERAPY	34
CHEMOTHERAPY	18
TREATMENT	17
SURGERY	16
CLINICAL	6
OPERATED	5
TREATED	5
OPERATION	3
TREAT	3
CURE	2
DOCTOR	2
DRUGS	2
TESTS	2
TREATMENTS	2
BRAIN-SURGERY	1

CURED	1
DISCUSSED	1
DOSE	1
DRUG	1
FOCUSED	1
HEALTH	1
MEDICAL	1
PUTTING	1
RADIOTHERAPY	1
SHRINK	1
SYMPTOMS	1
TOP	1
TOTAL	130

H 2: Pathology of lung cancer (85, 10)

LUNG-CANCER	35
CANCER	20
PATIENTS	13
CARCINOMA	6
CANCERS	3
PATIENT	3

DISEASE	2
DIAGNOSIS	1
PAIN	1
PROGNOSIS	1
TOTAL	85

H 3: Size, minuteness (70, 11)

CELL	25
SMALL	24
CELLS	12
LIGHT	2
COMMON	1
LARGE	1
NARROW	1

NECK	1
SIDE	1
THIN	1
VEIN	1
TOTAL	70

SMALL CELL LUNG CANCER (QPT3): 9 CHAINS

H 0: Learn (2, 2)

LEARN	2
--------------	---

H 1: Drugs (6, 2)

DRUGS	4
DRUG	2
TOTAL	6

H 2: Experiment (15, 4)

TRIALS	6
TESTING	4
TEST	3
TRIAL	2
TOTAL	15

H 3: Treatment, diagnosis and medicine (105, 23)

RADIATION	15
THERAPY	15
TREATMENT	13
CHEMOTHERAPY	10
CLINICAL	8
SURGERY	8
IRRADIATION	5
PROPHYLACTIC	5
TREATMENTS	4
SYMPTOMS	3
TEST	3
BIOPSY	2

BRONCHOSCOPE	2
RAYS	2
TREATED	2
BRONCHOSCOPY	1
DOSE	1
HOSPITAL	1
PILL	1
SURGICAL	1
TESTS	1
THERAPIES	1
TREAT	1

TOTAL	105
-------	-----

H 4: Minuteness, size (73, 13)

SMALL	24
CELL	21
CELLS	14
MICROSCOPE	3
CUT	2
LIMITED	2
BEAN	1
BREATH	1

NECK	1
PIECES	1
PRESSURE	1
SHORTNESS	1
VEIN	1
TOTAL	73

H 5: Pathology of lung cancer (78, 13)

LUNG-CANCER	27
CANCER	20
PATIENT	9
KILL	5
DISEASE	4
PATIENTS	4
PAIN	2
SYMPTOMS	2

CHEST-PAIN	1
CONDITION	1
COUGH	1
PROGNOSIS	1
SWELLING	1
TOTAL	78

H 6: Health care (17, 6)

DOCTOR (preserve v)	9
HIGH (pungent a)	3
SERVICE (preserve v)	1

STORE (preserve v)	1
TOTAL	14

H 7: Staging, location (21, 6)

STAGE (serial-place n)	10
STAGES (serial-place n)	4
REMOVES (serial-place n)	3
PLACES (serial-place n)	2

PLACE (serial-place n)	1
REMOVE (serial-place n)	1
TOTAL	21

H 8: Respiratory tract (26, 10)

LUNGS	8
CALLED	7
CHEST	3
INSIDE	2
BREATHING	1

CALL	1
COUGHING	1
SKIN	1
SOUND	1
WHEEZING	1

TOTAL	26
-------	----

H 9: Smoking

SMOKE (preserve v)	2
CIGARETTES (tobacco n)	1

APÉNDICE VIII: Lista de frecuencia del Corpus *Brown* con los 75 lexemas más frecuentes

N	Word	Freq.	%
1	THE	69.459	5.80
2	OF	36.214	3.02
3	AND	28.793	2.40
4	TO	26.069	2.18
5	A	23.385	1.95
6	IN	21.306	1.78
7	THAT	10.534	0.88
8	IS	10.004	0.84
9	WAS	9.762	0.81
10	HE	9.494	0.79
11	FOR	9.450	0.79
12	IT	8.703	0.73
13	WITH	7.250	0.61
14	AS	7.215	0.60
15	HIS	6.960	0.58
16	ON	6.719	0.56
17	BE	6.339	0.53
18	AT	5.358	0.45
19	BY	5.317	0.44
20	I	5.155	0.43
21	HAD	5.120	0.43
22	THIS	5.120	0.43
23	NOT	4.570	0.38
24	ARE	4.360	0.36
25	BUT	4.358	0.36

26	FROM	4.356	0.36
27	OR	4.197	0.35
28	HAVE	3.910	0.33
29	AN	3.723	0.31
30	THEY	3.594	0.30
31	WHICH	3.543	0.30
32	ONE	3.411	0.28
33	WERE	3.274	0.27
34	YOU	3.234	0.27
35	ALL	3.053	0.25
36	HER	3.009	0.25
37	SHE	2.848	0.24
38	WOULD	2.711	0.23
39	THERE	2.710	0.23
40	THEIR	2.653	0.22
41	WE	2.635	0.22
42	HIM	2.591	0.22
43	BEEN	2.467	0.21
44	HAS	2.425	0.20
45	WHEN	2.320	0.19
46	WHO	2.235	0.19
47	WILL	2.230	0.19
48	MORE	2.206	0.18
49	NO	2.203	0.18
50	IF	2.170	0.18
51	OUT	2.151	0.18

52	SO	2.014	0.17
53	UP	1.958	0.16
54	SAID	1.955	0.16
55	WHAT	1.881	0.16
56	ITS	1.845	0.15
57	ABOUT	1.810	0.15
58	THAN	1.789	0.15
59	INTO	1.785	0.15
60	THEM	1.770	0.15
61	CAN	1.751	0.15
62	ONLY	1.732	0.14
63	OTHER	1.698	0.14
64	TIME	1.681	0.14
65	NEW	1.620	0.14
66	SOME	1.611	0.13
67	COULD	1.588	0.13
68	THESE	1.568	0.13
69	TWO	1.508	0.13
70	MAY	1.392	0.12
71	FIRST	1.376	0.11
72	THEN	1.363	0.11
73	DO	1.340	0.11
74	ANY	1.336	0.11
75	LIKE	1.331	0.11

**APÉNDICE IX: Lista de frecuencia del *Corpus de Referencia del Español Actual*
(CREA) con los 75 lexemas más frecuentes**

N	Word	Freq.	%
1	De	5928345	6.30%
2	La	3777244	4.02%
3	Que	2898840	3.08%
4	El	2761603	2.94%
5	Y	2546060	2.71%
6	En	2516211	2.68%
7	A	2020396	2.15%
8	Los	1570946	1.67%
9	Se	1238281	1.32%
10	Del	1104825	1.17%
11	Un	1017524	1.08%
12	Las	1004783	1.07%
13	No	958539	1.02%
14	Por	953044	1.01%
15	Con	897848	0.95%
16	Una	818491	0.87%
17	Su	687024	0.73%
18	Para	625239	0.66%
19	Es	616943	0.66%
20	Al	585384	0.62%
21	Lo	558611	0.59%
22	Como	470893	0.50%
23	Más	405955	0.43%
24	O	312528	0.33%
25	Pero	295166	0.31%

26	Sus	282465	0.30%
27	Le	279818	0.30%
28	Me	269732	0.29%
29	Ha	229444	0.24%
30	Si	209953	0.22%
31	Sin	188452	0.20%
32	Ya	178428	0.19%
33	Sobre	173505	0.18%
34	Este	163004	0.17%
35	Cuando	162628	0.17%
36	Todo	159026	0.17%
37	Entre	157071	0.17%
38	Había	153421	0.16%
39	Era	151182	0.16%
40	Ser	140578	0.15%
41	Fue	138728	0.15%
42	Esta	138687	0.15%
43	También	137531	0.15%
44	Dos	137506	0.15%
45	Mi	134097	0.14%
46	Son	133919	0.14%
47	Qué	132312	0.14%
48	Muy	131185	0.14%
49	Hasta	127222	0.14%
50	Yo	124131	0.13%
51	Porque	121062	0.13%

52	Está	119724	0.13%
53	Desde	119708	0.13%
54	Años	118669	0.13%
55	Sólo	108952	0.12%
56	Vez	104113	0.11%
57	Hay	103509	0.11%
58	Ni	103241	0.11%
59	Han	99838	0.11%
60	Todos	97723	0.10%
61	Así	95947	0.10%
62	Nos	95667	0.10%
63	Puede	94966	0.10%
64	Él	92429	0.10%
65	Tiene	89504	0.10%
66	Te	89196	0.09%
67	Parte	86854	0.09%
68	Ahora	84038	0.09%
69	Bien	84022	0.09%
70	Uno	83832	0.09%
71	Ese	83582	0.09%
72	Mismo	81683	0.09%
73	Tiempo	81177	0.09%
74	Donde	80151	0.09%
75	Eso	77931	0.08%

Apéndice X: Palabras clave identificadas con la aplicación *Keywords* de *Wordsmith Tools* en los textos sobre tratamiento para profesionales de la salud del PDQ (QDT).

N	WORD	FREQ. DTREAT.LST %	FREQ. PCARE.LST %	KEYNESS	P		
1	STAGE	103	1.03	17	0.03	312.4	0.000000
2	LUNG	101	1.01	17	0.03	305.1	0.000000
3	RADIOTHERAPY	86	0.86	5		301.0	0.000000
4	SURVIVAL	89	0.89	12	0.02	280.0	0.000000
5	CELL	91	0.91	26	0.04	241.7	0.000000
6	CHEMOTHERAPY	110	1.10	91	0.15	183.2	0.000000
7	SMALL	69	0.69	31	0.05	156.7	0.000000
8	TUMOR	77	0.77	50	0.08	147.5	0.000000
9	PATIENTS	197	1.97	398	0.65	139.7	0.000000
10	RESECTION	38	0.38	5		119.9	0.000000
11	TRIALS	41	0.41	12	0.02	108.0	0.000000
12	NON-SMALL	26	0.26	0		102.1	0.000000
13	NSCLC	26	0.26	0		102.1	0.000000
14	CLINICAL	57	0.57	45	0.07	97.6	0.000000
15	CHEST	30	0.30	7	0.01	84.0	0.000000
16	REGIMENS	33	0.33	14	0.02	76.6	0.000000
17	TREATED	39	0.39	26	0.04	73.6	0.000000
18	CARCINOMA	27	0.27	8	0.01	70.8	0.000000
19	SURGICAL	32	0.32	18	0.03	65.8	0.000000
20	STAGING	18	0.18	1		63.1	0.000000
21	METASTASES	23	0.23	7	0.01	59.8	0.000000
22	MESOTHELIOMA	14	0.14	0		55.0	0.000000
23	DISTANT	13	0.13	0		51.0	0.000000
24	PULMONARY	18	0.18	4		51.0	0.000000
25	SURGERY	34	0.34	35	0.06	48.5	0.000000
26	METASTASIS	15	0.15	2		47.2	0.000000
27	PCI	12	0.12	0		47.1	0.000000
28	LIMITED	27	0.27	21	0.03	46.6	0.000000
29	BENEFIT	18	0.18	6		45.5	0.000000
30	PRIMARY	33	0.33	36	0.06	45.0	0.000000
31	RANDOMIZED	17	0.17	5		44.7	0.000000
32	CISPLATIN	21	0.21	11	0.02	44.6	0.000000
33	STANDARD	19	0.19	8	0.01	44.2	0.000000
34	PALLIATION	14	0.14	2		43.5	0.000000
35	ADJUVANT	14	0.14	2		43.5	0.000000
36	PLEURAL	11	0.11	0		43.2	0.000000
37	RESECTED	11	0.11	0		43.2	0.000000
38	IRRADIATION	24	0.24	19	0.03	41.0	0.000000
39	SECOND	15	0.15	4		40.5	0.000000
40	II	13	0.13	2		39.9	0.000000
41	LOBECTOMY	10	0.10	0		39.3	0.000000
42	PROGNOSIS	15	0.15	5		37.9	0.000000
43	SELECTED	15	0.15	5		37.9	0.000000
44	OVERALL	15	0.15	5		37.9	0.000000

45	CURATIVE	14	0.14	4		37.1	0.000000
46	ETOPOSIDE	12	0.12	2		36.2	0.000000
47	PERFORMANCE	18	0.18	11	0.02	35.5	0.000000
48	REGIONAL	9	0.09	0		35.3	0.000000
49	PROSPECTIVE	9	0.09	0		35.3	0.000000
50	ENDOBRONCHIAL	9	0.09	0		35.3	0.000000
51	STUDIES	30	0.30	40	0.07	34.3	0.000000
52	COMBINED	17	0.17	10	0.02	34.2	0.000000
53	NODES	13	0.13	4		33.7	0.000000
54	EXTENSIVE	13	0.13	4		33.7	0.000000
55	COMBINATION	27	0.27	35	0.06	31.7	0.000000
56	IPSILATERAL	8	0.08	0		31.4	0.000000
57	RETROSPECTIVE	8	0.08	0		31.4	0.000000
58	CISPLATIN-BASED	8	0.08	0		31.4	0.000000
59	LONG-TERM	16	0.16	10	0.02	31.2	0.000000
60	RATES	17	0.17	12	0.02	31.0	0.000000
61	LYMPH	14	0.14	7	0.01	30.4	0.000000
62	PALLIATIVE	12	0.12	4		30.3	0.000000
63	CGY	11	0.11	3		29.5	0.000000
64	LOCAL	20	0.20	20	0.03	29.1	0.000000
65	SHOWED	10	0.10	2		29.1	0.000000
66	CONSIDERED	22	0.22	25	0.04	29.0	0.000000
67	CARE	3	0.03	167	0.27	32.2	0.000000