

Diseño y técnicas de explotación de un corpus oral para el análisis de parámetros de calidad en interpretación

JOSÉ MANUEL PAZOS BRETaña

OLALLA GARCÍA BECERRA

RAFAEL BARRANCO-DROEGE

Universidad de Granada

Resumen

La investigación en el campo de la evaluación de la calidad en interpretación simultánea puede enriquecerse significativamente con las ventajas que aporta el uso de un corpus estructurado. En esta contribución presentamos las razones que subyacen a la creación del corpus oral ECIS, sus características y técnicas de aplicación que bien pudieran resultar útiles para otros campos.

Palabras clave: corpus oral, interpretación simultánea, evaluación de la calidad

Abstract

Research in simultaneous interpreting quality assessment could draw significant benefits from a structured corpus-assisted approach. In this paper we discuss the reasons behind the creation of ECIS - an oral corpus-, its characteristics and some implementation techniques that might well prove useful in other fields.

Keywords: oral corpus, simultaneous interpreting, quality assessment

1. ANTECEDENTES Y GÉNESIS DEL CORPUS¹

El grupo *Investigación pedagógica y social sobre interpretación de conferencias y comunitaria* (HUM-560, Junta de Andalucía) desarrolla su trabajo en la Universidad de Granada desde el año 1995. Su directora, Ángela Collados Aís, inició, en 1998, una nueva línea de investigación en el ámbito de la calidad de la interpretación de conferencias con un estudio experimental en el que se analizaba la incidencia de la entonación monótona en la evaluación de la calidad que realizaban los usuarios (Collados Aís, 1998).

El primer estudio empírico de la calidad en interpretación simultánea (IS) se remonta a Bühler (1986), quien propuso dieciséis parámetros para sistematizar las expectativas de los usuarios respecto a la calidad. Entre los parámetros sugeridos por esta investigadora (op. cit.) destacan los siguientes: acento, agradabilidad de la voz, fluidez, cohesión lógica, transmisión

¹ Este trabajo de investigación se presenta en el marco de los proyectos de investigación HUM2007-62434/FILO (Ministerio de Ciencia y Tecnología, España) y P07-HUM-02730 (Junta de Andalucía, España).

correcta del sentido, transmisión completa del discurso, gramaticalidad, terminología y estilo. Este catálogo de parámetros ha sido utilizado, con ligeras modificaciones, en la mayoría de los estudios posteriores, tanto de expectativas (por ejemplo, Kurz, 1989, 1993), como de evaluación de la IS por los usuarios (entre otros, Gile, 1990).

Desde 1998, el grupo de investigación ha ido completando una serie de estudios, que se han centrado en once parámetros de calidad, añadiendo dicción y entonación a los mencionados (Collados Aís, 2001; Pradas Macías, 2003, 2004, 2006; Collados Aís & Pradas Macías & Stévaux & García Becerra, 2007; Barranco-Droege, 2009; Stévaux, en preparación).

Los resultados de estos estudios pusieron de manifiesto la necesidad de realizar un trabajo de carácter observacional en el que se analizaran los distintos parámetros en una situación real de interpretación. Para ello, en el marco del Proyecto de Investigación del Ministerio de Ciencia y Tecnología *Evaluación de la calidad en interpretación simultánea. Parámetros de incidencia* (BFF2002-00579), se llevó a cabo la recopilación de material para la elaboración de un corpus oral de interpretaciones reales de una sesión plenaria del Parlamento Europeo. Este material constituye el núcleo sobre el que se construye ECIS. Durante la realización de dicho proyecto, se estudiaron por primera vez en su contexto real los once parámetros de calidad, incorporando, además, dos aspectos nuevos: la influencia de los llamados *problem triggers*² de la interpretación y las estrategias de resolución de los intérpretes a partir de tres combinaciones lingüísticas (alemán-español, francés-español e inglés-español). Se analizaron, también, otros posibles factores que pudiesen afectar a la calidad y su evaluación, como pueden ser las primeras impresiones (García Becerra, 2006; 2008).

En la etapa actual³ se están efectuando un análisis y una cuantificación pormenorizados de los parámetros, factores y funciones relevantes del proceso de la IS, maximizadores y minimizadores de la calidad, a través de estudios empíricos con intérpretes y usuarios que, desde metodologías y enfoques complementarios, contribuyan a una definición de la calidad de la interpretación de conferencias a partir de las combinaciones lingüísticas ya mencionadas. Esta línea de investigación hizo necesario emprender la creación de un corpus

² Elementos que pueden desencadenar problemas en la interpretación (Gile, 1995).

³ Una vez finalizada la primera etapa, el grupo recibió, en 2007, la concesión de dos proyectos: un proyecto I+D financiado por el Ministerio de Ciencia y Tecnología (HUM2007-62434/FILO) y un Proyecto de Excelencia de la Junta de Andalucía (P07-HUM-02730).

oral que permitiera la extracción de todos los datos necesarios y que, además, presentase una organización sistemática de esta información.

2. PREPARACIÓN TÉCNICA DEL MATERIAL

El corpus ECIS se ha elaborado partiendo de la grabación en vídeo de una sesión plenaria del Parlamento Europeo celebrada en Bruselas los días 10, 11 y 12 de marzo de 2003. De esta sesión se han obtenido un total de 44 discursos originales, presentados en inglés, francés o alemán. La duración de los discursos varía entre 1 y 8 minutos aproximadamente. A cada discurso original (DO) se asocian las IS realizadas a las distintas lenguas de trabajo. Dado que el interés principal del proyecto es, en primera instancia, el estudio de la incidencia de los parámetros que influyen en la calidad de la interpretación al español, ésta se ha incluido en todos los casos. Además se han incorporado al corpus –atendiendo a las futuras necesidades de los miembros del equipo– las interpretaciones realizadas al alemán y al francés de aquellos discursos originales realizados en lengua inglesa.

En resumen, el material en bruto del corpus consta de un total de 44 discursos originales (15 en inglés, 14 en alemán y 15 en francés) y de 44 interpretaciones al español, 15 al alemán y 15 al francés.

Tras la grabación de los discursos se ha procedido a aislar los segmentos relevantes y extraer el canal de audio correspondiente a los DO y a las IS, generándose ficheros individuales para cada uno. Debido a que las características técnicas de los DO y las IS no coincidían, ha sido preciso realizar un resamplado de los DO. Tanto el demultiplexado como el resamplado se han realizado mediante el programa Audacity⁴, obteniéndose así el conjunto de ficheros normalizados⁵.

El último paso de la preparación de los ficheros ha consistido en la sincronización de las IS y los DO. En este proceso, se han editado cada DO y sus correspondientes IS para sincronizar los inicios con el fin de obtener el contexto de “simultaneidad” de las interpretaciones.

⁴ Audacity. Versión 1.2.6. [Software para OS X y otros sistemas] s.l. s.n. <http://audacity.sourceforge.net>.

⁵ Contenedor *wav*, códec *araw* y pista *mono* a 22.050 Hz. 16 bits.

3. ANÁLISIS DEL MATERIAL SONORO Y EXTRACCIÓN DE INFORMACIÓN

3.1. *Evaluación de las interpretaciones*

Una tarea fundamental, puesto que sus resultados se utilizarán después como *tertium comparationis*, es la evaluación de las interpretaciones. Para llevar a cabo esta tarea, los miembros del grupo de investigación han utilizado los cuestionarios empleados en trabajos anteriores (por ejemplo, Collados Aís, 1998). Tras la escucha de las grabaciones de las IS, se han valorado, en una escala de 1 a 5, los 11 parámetros de calidad. Además, se han introducido las variables profesionalidad y fiabilidad, así como una valoración global de la actuación del intérprete.

3.2. *Transcripción, segmentación y anotación*

Cada fichero DO y sus IS asociadas se han transcrito ortográficamente, utilizando el programa Praat⁶. Esta transcripción se ha llevado a cabo segmentando los ficheros sonoros, de tal manera que es posible, mediante Praat (Fig. 1), acceder a cada fragmento y obtener, dado el caso, toda la información fonética para cada segmento en cuestión, además de para la totalidad del discurso. Durante la segmentación se han marcado, además, las pausas discursivas y silencios (vid. 3.5.).

⁶ Boersma, Paul and Weenink, David (1992–2010). Praat: A system for doing phonetics by computer <http://www.praat.org>.

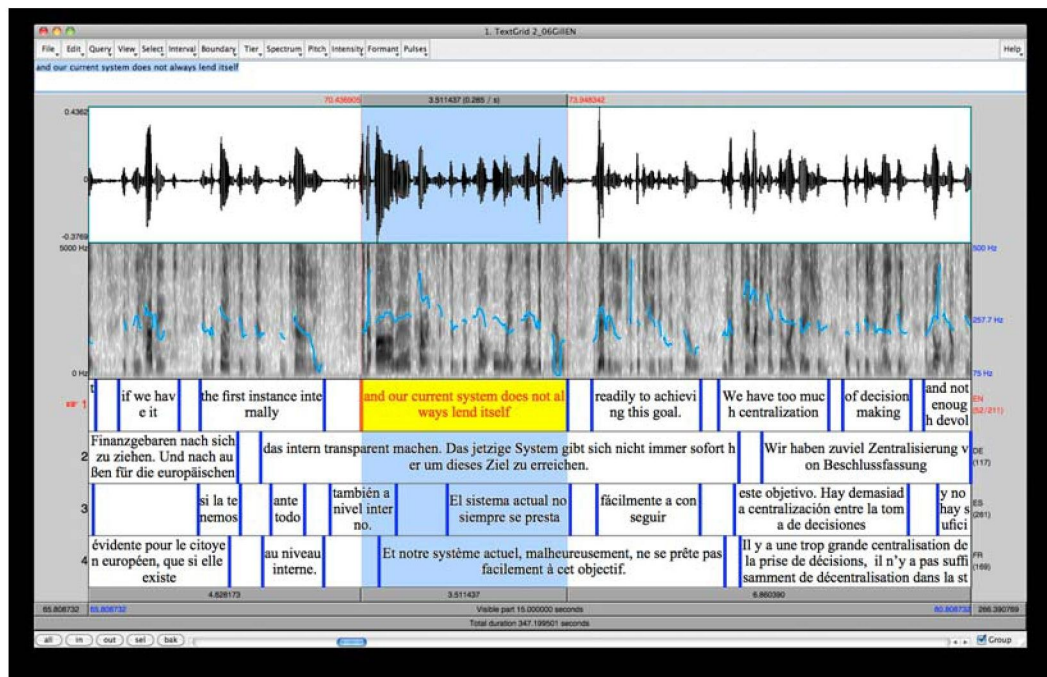


Figura 1: Instantánea de una ventana de Praat en la que se muestran el oscilograma, el espectrograma y la curva melódica correspondientes a un discurso (en este caso un DO en inglés) y la transcripción segmentada del mismo y de sus interpretaciones al alemán, español y francés.

3.3. Análisis cuantitativo

Hasta ahora se han descrito procesos cualitativos como la evaluación y la anotación. En la siguiente etapa del análisis del corpus para su explotación, se procederá a extraer del mismo aquella información cuantitativa relevante para la naturaleza de las investigaciones que se van a realizar. Esta información procederá, fundamentalmente, del análisis fonético acústico, del recuento de las palabras y sílabas articuladas y del etiquetado *POS*.

3.4. Análisis fonético acústico

Un parámetro acústico cuya cuantificación puede mostrar la relación entre la calidad de la interpretación y la entonación es la frecuencia fundamental. Una de las hipótesis de trabajo (Collados Aís, 2008) es que la entonación monótona se percibe, a partir de cierto umbral, como un elemento que influye de manera negativa en la recepción de una interpretación.

La *frecuencia fundamental*⁷, en concreto su desviación típica, nos permite obtener un índice para medir este fenómeno. Praat nos proporciona un listado de los valores de la frecuencia fundamental de cada DO e IS, así como el valor de su desviación típica.

⁷ La frecuencia fundamental (F_0) es la frecuencia más baja del espectro armónico tal que las frecuencias del resto de armónicos pueden expresarse como múltiplos de esta frecuencia fundamental.

3.5. Recuento de sílabas y palabras

La velocidad articulatoria o cadencia es un parámetro al que se le atribuye en algunos estudios (por ejemplo, Pio, 2003) una influencia determinante en la interpretación. Ello, además, desde una doble vertiente: por un lado, la velocidad de articulación del ponente en el DO y, por otro, la del intérprete (que puede estar, o no, en relación con la del DO).

Este parámetro se calcula de dos maneras distintas: medido en palabras por minuto y en sílabas por minuto. El hecho de calcularlo en sílabas por minuto permite obviar cierto desequilibrio que podría producirse según la combinación de lenguas DO/IS, y que se debe a la propia idiosincrasia tipológico-morfológica de cada lengua (Luque Durán, 2001: 51) –por ejemplo, el número de palabras de un discurso en alemán puede diferir del de su interpretación al español, simplemente por el hecho de ser una lengua más sintética (alemán) que analítica (español)–. El cálculo se realiza de manera sencilla haciendo un recuento del número de palabras y sílabas⁸ de cada DO e IS dividiendo estas cantidades por la duración.

El hecho de haber tenido en cuenta las pausas y silencios durante la segmentación nos permite obtener lo que denominamos cadencias efectivas. Entendemos por tales el número de palabras (cadencia de palabras efectiva: CW_{ef}) y de sílabas (cadencia silábica efectiva: CS_{ef}) articuladas por segundo durante el discurso, sin tener en cuenta la duración de las pausas relevantes y silencios. El cálculo de las cadencias efectivas se realiza de manera similar al de la cadencia normal, pero restando de la duración los intervalos silentes que obtenemos utilizando Praat.

3.6. Etiquetado POS

Otro nivel que puede tener una influencia sobre la calidad de la interpretación es el morfológico-sintáctico. La existencia de una concentración de elementos léxicos pertenecientes a ciertas categorías gramaticales (por ejemplo, textos predominantemente nominales, verbales, etc.) bien pudiera reflejarse sobre los parámetros evaluados y cuantificados.

Para poder medir y evaluar esta hipotética influencia se ha realizado una anotación POS de la transcripción de los DO e IS. Se han utilizado para ello, según las lenguas, las aplicaciones *Freeling* (Carreras et al., 2004) y *Morphy* (Lezius, 2000). Los *tagsets* de ambos programas difieren un poco entre sí; el utilizado por Freeling está basado más consistentemente en EAGLES que el de Morphy. Sin embargo, para nuestro propósito,

ambos son más que adecuados y no se produce ningún conflicto, ya que sólo necesitamos para nuestro análisis la categoría gramatical general a la que pertenece cada elemento léxico. En el estudio, se considera principalmente la influencia de cuatro categorías: nombres, verbos, adjetivos y adverbios, aunque la disponibilidad inmediata de la información correspondiente al resto de categorías nos ha hecho considerar la pertinencia de incluir éstas también. Asimismo, se está contemplando la posibilidad de realizar análisis complementarios similares a los descritos en Pazos Bretaña & Pamies Bertrán (2008).

4. EXPLOTACIÓN DEL CORPUS Y ESTUDIOS DE LOS PARÁMETROS

Toda la información recopilada en las diferentes fases de elaboración y análisis del corpus se han recogido en una base de datos relacional convencional (MS Access) que permite exportar la información a hojas de cálculo cuando las circunstancias y necesidades lo requieran, bien para la realización de análisis estadísticos suplementarios, bien para la generación de tablas y gráficos⁹.

Uno de los mayores problemas con los que nos encontramos a la hora de realizar un análisis estadístico de la influencia de las distintas variables, es que al existir un gran número de éstas y no poder conocer de antemano las relaciones que pueden existir entre algunas –si es que efectivamente existen– el volumen ingente de cálculos necesarios y datos resultantes enseguida alcanza una complejidad tal que no sólo puede ser difícil de mostrar, sino que incluso ciertas relaciones bien pudieran pasar desapercibidas.

Idealmente, sería necesario poder observar la variación e interacción de una manera gráfica e intuitiva en la que se muestren las variaciones e interacciones de los distintos parámetros, tanto aisladamente como en combinación con otros. Ello nos permitiría obtener y observar patrones relacionales y de comportamiento que de otra manera no serían detectados o requerirían de un esfuerzo mucho mayor para serlo.

Afortunadamente, existe una herramienta que cumple con estos requisitos. Se trata de *Trendalyzer*, desarrollado por la fundación *Gapminder* en un proyecto dirigido por el matemático y médico sueco Hans Rosling. Este software fue creado para permitir la visualización animada de la evolución de diversos conjuntos de datos a lo largo del tiempo.

⁸ Existen varias herramientas para realizar el recuento silábico. Por ejemplo, *Morphy* y *Contador de sílabas* de Lexiquetos. A ambas puede accederse en <http://wolfganglezius.de/doku.php?id=cl:morphy> y <http://lexiquetos.org/silio/> respectivamente.

⁹ Véase apéndice 2.

En 2007, Google Inc. adquirió los derechos sobre este software, renombrándolo *Motion Chart* y lo puso a disposición del público, integrándolo en su *API*¹⁰ de visualización.

Si bien la aplicación es una herramienta para analizar la evolución de variables en un marco temporal, es posible, realizando ciertos ajustes en la estructuración de los datos, utilizarla para el análisis de variables “estáticas”, i.e. sin dependencia temporal.

Hemos volcado el conjunto de datos obtenidos en las fases de evaluación y análisis en una hoja de cálculo para poder estudiarlos con Trendalyzer (Fig. 2), y los resultados han sido muy satisfactorios. Al no existir una variación temporal de las variables, el efecto que se consigue es el equivalente a una animación en la que cada fotograma representa el gráfico de la combinación de los parámetros elegidos para una IS (o DO). La animación muestra la evolución de la combinación de variables elegidas en todo el conjunto, pudiéndose reconocer a simple vista la existencia o no de patrones de regularidad.

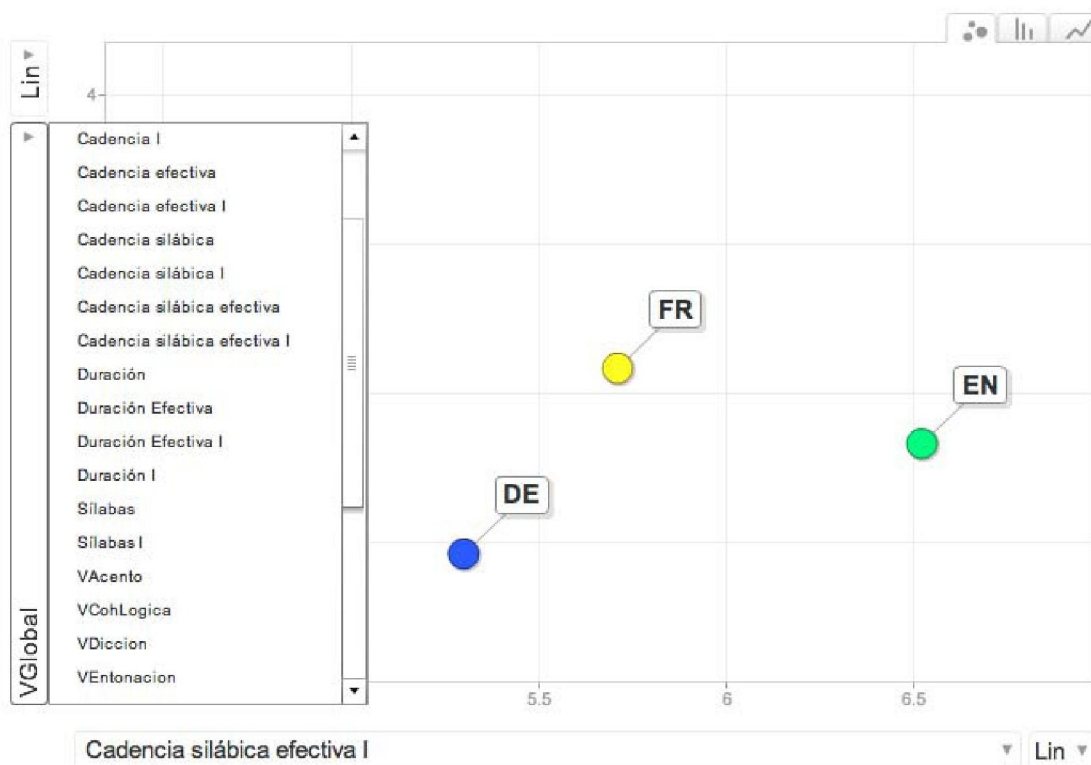


Figura 2: Vista de un momento de la animación del conjunto de datos correspondiente a la Valoración Global de las interpretaciones función de su CS_{ef} (la etiqueta muestra la lengua del DO). Se muestra asimismo parte del conjunto de variables disponibles en el eje de ordenadas. El mismo conjunto de variables está también disponible en el eje de abscisas.

¹⁰ Application programming interface.

La aplicación permite mostrar la interacción de cualquier par de variables, es decir, en la agrupación que hemos realizado en este primer experimento (47 variables¹¹), resultan un total de 1.081 combinaciones posibles. Si a ello se añade la posibilidad de comparar tres parámetros simultáneamente, representándose la tercera mediante una variación en el diámetro de los puntos, resultan en total 16.215 combinaciones.

5. CONCLUSIÓN

Hemos presentado un ejemplo de elaboración, tratamiento y aplicación de un corpus oral para el estudio de parámetros de calidad en IS sobre la base del desarrollo de ECIS.

Este corpus se distingue de otros *corpora* actualmente en uso en el campo de la interpretación¹² no sólo en el enfoque metodológico, sino en que ha sido diseñado atendiendo a los requisitos específicos de la investigación de la calidad en IS. Por primera vez en los estudios de interpretación asistidos por corpus, se aplica un enfoque comprensivo que introduce la combinación de técnicas de análisis fonético-acústico, morfológico y estadístico. Otra importante innovación metodológica es la inclusión de una herramienta de análisis intuitivo, *Trendalyzer*, que es de una excepcional utilidad para un análisis de grandes volúmenes de datos que permite poner de manifiesto relaciones entre las distintas variables.

Si bien ECIS ha sido creado a la medida de las necesidades de nuestro grupo de trabajo, centradas en el estudio de los denominados parámetros “verbales¹³” y “no verbales¹⁴”, creemos que, con las variaciones y adaptaciones específicas necesarias, la metodología puede ser aplicada a las investigaciones en otros campos del lenguaje.

¹¹ Véase apéndice 1.

¹² Monti et al. (2005), Meyer (2008) y Bendazzoli & Sandrelli (2009).

¹³ Cohesión lógica, transmisión correcta, transmisión completa, gramaticalidad, terminología y estilo.

¹⁴ Acento, agradabilidad de la voz, fluidez, entonación y dicción.

APÉNDICE I

Tabla A1: Listado de los parámetros medidos y contabilizados en los discursos originales (DO) y las interpretaciones simultáneas (IS)

Lengua	Lengua del DO
Sexo	Sexo del orador
Tipo de Exposición	Espontánea, leída o semi-espontánea
Duración	Duración del DO en segundos
Duración Efectiva	Duración del DO en s. sin contabilizar intervalos silentes
Palabras	Número de palabras del DO
Cadencia	Nº de palabras articuladas por minuto en el DO
Cadencia efectiva	Nº de palabras articuladas por minuto en el DO sin contabilizar la duración de los intervalos silentes
Sílabas	Nº de sílabas del DO
Cadencia silábica	Nº de sílabas articuladas por minuto en el DO
Cadencia silábica efectiva	Nº de sílabas articuladas por minuto en el DO sin contabilizar la duración de los intervalos silentes
Frecuencia Fundamental	Media de la frecuencia fundamental del DO
Desviación Típica F0	Desviación estándar típica de la frecuencia fundamental del DO
Sustantivos	Nº de sustantivos del DO
Verbos	Nº de verbos del DO
Adjetivos	Nº de adjetivos del DO
Adverbios	Nº de Adverbios del DO
Lengua I	Lengua de la IS
Sexo I	Sexo del intérprete
Duración I	Duración de la IS en segundos
Duración Efectiva I	Duración del IS en s. sin contabilizar intervalos silentes
Palabras	Número de palabras de la IS
Cadencia I	Nº de palabras articuladas por minuto en la IS
Cadencia efectiva I	Nº de palabras articuladas por minuto en el IS sin contabilizar la duración de los intervalos silentes
Sílabas I	Nº de sílabas de la IS
Cadencia silábica I	Nº de sílabas articuladas por minuto en la IS
Cadencia silábica efectiva I	Nº de sílabas articuladas por minuto en la IS sin contabilizar la duración de los intervalos silentes
Frecuencia Fundamental I	Media de la frecuencia fundamental de la IS
Desviación Típica F0 I	Desviación estándar típica de la frecuencia fundamental de la IS
Sustantivos I	Nº de sustantivos de la IS
Verbos I	Nº de verbos de la IS
Adjetivos I	Nº de adjetivos de la IS
Adverbios I	Nº de Adverbios de la IS
VGlobal	Valoración global de la IS
VAcento	Valoración del acento del intérprete en la IS
VVoz	Valoración de la voz del intérprete en la IS
VFluidez	Valoración de la fluidez del intérprete en la IS
VCohLogica	Valoración de la coherencia lógica de la IS
VTransCorrecta	Valoración de la transmisión correcta del sentido en la IS
VTransCompleta	Valoración de la transmisión completa del sentido en la IS
VTerminologia	Valoración de terminología usada en la IS
VEstilo	Valoración del estilo de la IS

VEntonacion	Valoración de la entonación en la IS
VDiccion	Valoración de la dicción en la IS
VGramaticalidad	Valoración de la gramaticalidad de la IS
VProfesionalidad	Valoración de la profesionalidad del intérprete
VFiabilidad	Valoración de la fiabilidad de la IS

APÉNDICE 2.

Tabla B1: Ejemplo de la información extraída de un discurso original (DO)

ID	1
Código	208
Fecha	11/03/03
Sesión	Mañana
Lengua	DE
Orador	Ceyhun
Sexo	H
Temática	Inmigración
Tipo de Exposición	Espontánea
Duración	132,83
Duración Efectiva	110,08
Palabras	288
Cadencia	130,1
Cadencia efectiva	157
Sílabas	602
Cadencia silábica	271,9
Cadencia silábica efectiva	328,1
Frecuencia Fundamental	181,6
Desviación Típica F0	29,52
Sustantivos	82
Verbos	33
Adjetivos	24
Adverbios	37
Comentarios	

Tabla B2: Ejemplo de la información extraída de la interpretación simultánea (IS) del DO de la tabla A1

ID	1
Lengua I	ES
Sexo I	H
Duración I	134,47
Duración Efectiva I	102,75
Palabras	265
Cadencia I	118,2
Cadencia efectiva I	154,7
Sílabas I	545
Cadencia silábica I	243,2
Cadencia silábica efectiva I	318,3
Frecuencia Fundamental I	103,47
Desviación Típica F0 I	33,99
Sustantivos I	69
Verbos I	36
Adjetivos I	28
Adverbios I	15
VGlobal	2,46
VAceto	4,86
VVoz	3,21
VFluidez	2
VCohLogica	2,71
VTransCorrecta	2,57
VTransCompleta	2,57
VTerminologia	3
VEstilo	2,29
VEntonacion	2,21
VDiccion	2,5
VGramaticalidad	2,96
VProfesionalidad	2,57
VFiabilidad	2,5
Comentarios I	

REFERENCIAS BIBLIOGRÁFICAS

- Barranco-Droege, R. (2009). *La transmisión correcta del sentido como parámetro de calidad en interpretación simultánea*. [Trabajo de investigación tutelada] Granada: Universidad de Granada.
- Bendazzoli, C. & Sandrelli, A. (2009). Corpus-based interpreting studies: Early work and future prospects. *Tradumàtica* 7.

- Consultado en <http://webs2002.uab.es/tradumatica/revista/num7/articles/08/08art.htm>
- Bühler, H. (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua* 5:4, 231-235.
- Carreras, X. & Chao, I. & Padró, L. & Padró, M. (2004). FreeLing: An Open-Source Suite of Language Analyzers. En *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'04)*. Lisboa, Portugal
- Collados Aís, A. (1998). La evaluación de la calidad en interpretación simultánea. La importancia de la comunicación no verbal. Granada: Comares Interlingua.
- Collados Aís, A. (2001). The Evaluation of Quality in Simultaneous Interpreting. The importance of Nonverbal Communication. En F. Pöchhacker & M. Shlesinger (Eds), *The Interpreting Studies Reader*, (pp. 236-336). Londres: Routledge.
- Collados Aís, A.; Pradas Macías, E. M.; Stévaux, E. & García Becerra, O. (Eds.) (2007). *Evaluación de la calidad en interpretación simultánea: parámetros de incidencia*. Granada: Comares Interlingua.
- Collados Aís, A. (2008). Evaluación de la calidad en interpretación simultánea: contrastes de exposición e inferencias emocionales. Evaluación de la evaluación. En G. Hansen; A. Chesterman & H. Gezymisch-Arbogast (Eds.), *Efforts and Models in Interpreting and Translation Research* (pp. 193-214) Amsterdam/Philadelphia: John Benjamins.
- García Becerra, O. (2006). La incidencia de las primeras impresiones en la evaluación de la calidad en interpretación simultánea: estudio piloto. [Trabajo de investigación para la obtención del DEA] Granada: Universidad de Granada.
- García Becerra, O. (2008). La incidencia de las primeras impresiones en la evaluación de la calidad en interpretación simultánea: estudio piloto. En M. M. Fernández Sánchez & R. Muñoz Marín (Eds.), *Aproximaciones cognitivas al estudio de la traducción y la interpretación*. (pp. 301-326) Granada: Comares Interlingua.
- Gile, D. (1990). L'évaluation de la qualité de l'interprétation par les délégués: une étude de cas, *The Interpreters' Newsletter* 3, 66-71.
- Gile, D. (1995). Basic concepts and models of for interpreter and translator training. Ámsterdam: John Benjamins.
- Kurz, I. (1989). Conference interpreting - user expectations. En D. Hammond (Ed.), *Coming of Age. Proceedings of the 30th Conference of the ATA* (pp. 143-148). Medford, N.J.: Learned Information Inc.
- Kurz, I. (1993). Conference interpretation: expectations of different user groups. *The Interpreters' Newsletter* 5, 13-21.

- Lezius, W. (2000). Morphy - German Morphology, Part-of-Speech Tagging and Applications. En U. Heid; S. Evert; E. Lehmann & C. Rohrer (Eds.), *Proceedings of the 9th EURALEX International Congress* (pp. 619-623). Stuttgart, Alemania.
- Luque Durán, J. de D. (2001). Aspectos universales y particulares de las lenguas del mundo. Granada: Granada Lingvistica.
- Meyer, B. (2008). Interpreting proper names: Different interventions in simultaneous and consecutive interpreting? *Trans-kom 1:1*.
- Consultado en http://www.trans-kom.eu/ihv_01_01_2008.html
- Monti, C.; Bendazzoli, C.; Sandrelli, A. & Russo, M. (2005). Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus). *Meta* 50:4.
- Consultado en <http://www.erudit.org/revue/meta/2005/v50/n4/>
- Pazos Breña, J. M. & Pamies Bertrán, A. (2008). Combined statistical and grammatical criteria for the retrieval of phraseological units in an electronic corpus. En S. Granger & F. Meunier (Eds.), *Phraseology. An interdisciplinary perspective* (pp. 391-406). Ámsterdam/Philadelphia: John Benjamins
- Pio, S. (2003). "The relation between st delivery rate and quality in simultaneous interpretation" *The Interpreters' Newsletter* 12, 69-100.
- Pradas Macías, E. M. (2003). Repercusión del intraparámetro pausas silenciosas en la fluidez: Influencia en las expectativas y en la evaluación de la calidad en interpretación simultánea. [Tesis doctoral sin publicar] Universidad de Granada. Granada.
- Pradas Macías, E. M. (2004). La fluidez y sus pausas: enfoque desde la interpretación de conferencias. Granada: Comares.
- Pradas Macías, E. M. (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting* 8:1, 25-43.
- Stévaux, É. (en prep.). *El acento como parámetro de calidad en interpretación simultánea*. [Tesis doctoral] Granada: Universidad de Granada.

Language Windowing through Corpora.

Visualización del lenguaje a través de corpus

Part II

L-Z

Editors

Isabel Moskowich-Spiegel Fandiño

Begoña Crespo García

Inés Lareo Martín

Paula Lojo Sandino

Universidade da Coruña

A Coruña 2010

ISBN: 978-84-9749-401-4

Cover designed by Inés Lareo and Alejandro González

Contents/Contenidos Part II

Contents/Contenidos	i
In search of cross-linguistic anchor phenomena for translation quality assessment	481
<i>Belén Labrador</i>	481
<i>Tel</i> , comparativo atípico del francés : una gramática de usos.....	491
<i>Sarah Leroy</i>	491
<i>Sylvain Loiseau</i>	491
The science of astronomy: passive constructions in eighteenth-century texts	505
<i>Paula Lojo Sandino</i>	505
El paradigma derivativo de los adverbios en el inglés antiguo.....	518
<i>Gema Maiz Villalta</i>	518
Obtaining computational resources for languages with scarce resources from closely related computationally-developed languages. The Galician and Portuguese case.....	529
<i>Paulo Malvar Fernández, José Ramon Pichel Campos, Oscar Senra Gómez,</i>	529
<i>Pablo Gamallo Otero,</i>	529
<i>Alberto García,</i>	529
La importancia de la confección y el uso de un corpus para la investigación llevada a cabo en la tesis “sintaxis y semántica de la pasiva preposicional”	537
<i>Ana Isabel Martín Doña</i>	537
Reaction Object constructions in English. A corpus-based study	551
<i>Montserrat Martínez Vázquez</i>	551
Corpus of Interpreting Discourse = Speech Corpus + Parallel Corpus?.....	563
<i>Mikhail Mikhailov</i>	563
Paradigmas derivativos del inglés antiguo organizados en torno a adjetivos básicos	573
<i>Carmen Novo Urraca</i>	573
Lexical evidential verbs in English computing scientific articles	585
<i>Ivalla Ortega Barrera</i>	585
<i>Margarita Esther Sánchez Cuervo</i>	585
Pride – Stolz – orgullo: A corpus-based approach to the expression of emotion concepts in a foreign language.....	593
<i>Ulrike Oster</i>	593
Variations in the use of “I” in casually spoken English.....	611

<i>Michael Pace-Sigge</i>	611
Diseño y técnicas de explotación de un corpus oral para el análisis de parámetros de calidad en interpretación.....	627
<i>José Manuel Pazos Breña</i>	627
<i>Olalla García Becerra</i>	627
<i>Rafael Barranco-Droege</i>	627
El uso de aunque y pero por hablantes nativos y aprendices suecos	641
<i>Aymé Pino Rodríguez</i>	641
Fragmentation of parallel sentences.....	653
<i>Sergey B. Potemkin</i>	653
How to work with smaller corpora of indigenous languages	661
<i>Regina Pustet</i>	661
The appraisal of lexical content in ESP coursebooks against corpus-driven and frequency vocabulary lists	671
<i>Camino Rea Rizzo</i>	671
Does valency theory offer a holistic approach to teaching language?.....	689
<i>Renate Reichardt</i>	689
Sintaxis de un tipo de cláusula interrogativa a través de datos de corpus	703
<i>Iria del Río Gayo</i>	703
The problem of <i>false friends</i> in learner language: Evidence from two learner corpora	717
<i>María Luisa Roca Varela</i>	717
The discourse of Americans in Brazilian cookbooks: a proposal for an analysis based on Corpus Linguistics	731
<i>Rozane Rodrigues Rebechi</i>	731
<i>For example</i> and <i>for instance</i> as markers of exemplification in Present-day English: A corpus-based study.....	747
<i>Paula Rodríguez Abruñeiras</i>	747
Corpus léxico de onomatopeyas españolas.....	759
<i>Jorge Rodríguez Guzmán</i>	759
Lancashire English in diachronic perspective:	771
evidence from the Salamanca Corpus.....	771
<i>Javier Ruano-García</i>	771
Achieving representativeness through the parameters of spoken language and discursive features: the case of the Spoken Turkish Corpus.....	789

<i>Şükriye Ruhi</i>	789
<i>Hale Işık-Güler</i>	789
<i>Çiler Hatipoğlu</i>	789
<i>Betil Eröz-Tuğa</i>	789
<i>Derya Çokal Karadaş</i>	789
Los elementos de prolongación copulativos en textos científicos ingleses del siglo XVIII..	801
<i>Estefanía Sánchez Barreiro</i>	801
<i>I just come in Hong Kong by myself: Tense in spoken Hong Kong English</i>	817
<i>Elena Seoane</i>	817
<i>Cristina Suárez-Gómez</i>	817
The expression of politeness in research articles: Authorial presence vs. authorial invisibility in the discussion.....	829
<i>Carmen Soler Monreal</i>	829
<i>Luz Gil Salom</i>	829
Variación en el uso de conectores causales en alemán según tipos textuales de la lengua hablada	843
<i>Oliver Strunk</i>	843
<i>Claudia Bucher</i>	843
Linguistic features in a nanotechnology corpus.....	861
<i>Keith Stuart</i>	861
<i>Ana Botella</i>	861
The concept of ‘circumcollocate’ and its significance for lexicography: A discussion with particular reference to the Japanese language.....	873
<i>Tadaharu Tanomura</i>	873
Diátesis léxica de <i>gehen</i> y <i>kommen</i> en un corpus de lengua oral en alemán	881
<i>Eduard Tapia Yepes</i>	881
THAI-NEST: A framework for Thai named entity tagging specification and tools	895
<i>Thanaruk Theeramunkong, Monthika Boriboon, Choochart Haruechaiyasak,</i>	895
<i>Nichnan Kittiphattanabawon, Krit Kosawat, Chutamanee Onsuwan,</i>	895
<i>Issariyapol Siriwat, Thawatchai Suwanapong, and Nattapong Tongtep</i>	895
The semantic function of affixation in a corpus of Old English derived nouns	909
<i>Roberto Torre Alonso</i>	909
Método general de lematización con una gramática mínima y un diccionario óptimo.	
Aplicación a un corpus dialectal escrito	919

<i>Hiroto Ueda</i>	919
<i>Maria-Pilar Perea</i>	919
Doublets and nominalization in Early Modern scientific English.....	933
<i>Vera Vázquez López</i>	933
El verbo débil como base de la derivación léxica en inglés antiguo	945
<i>Raquel Vea Escarza</i>	945
Using multilingual parallel corpora for contrastive studies and translation studies: A case study of the verbs of sitting, standing, and lying	961
<i>Åke Viberg</i>	961
Subordinación sustantiva en redacciones de estudiantes de licenciatura en educación secundaria	979
<i>Irma Guadalupe Villasana Mercado</i>	979