



CAPÍTULO 9

CLASIFICACIÓN VERBAL, ETIQUETACIÓN SEMÁNTICA E INFORMACIÓN LEXICOGRÁFICA EN EL PROYECTO ADESSE

Gael dos Santos Vaamonde
y Fita González Domínguez
Universidade de Vigo

RESUMEN

Esta contribución pretende ofrecer una presentación muy general del proyecto ADESSE¹ (*Alternancias de Diátesis y esquemas Sintáctico-Semánticos del Español*), que se está desarrollando actualmente en la Universidad de Vigo, y cuyo objetivo fundamental es el de ofrecer una base de datos para el estudio empírico de la interacción entre verbos y construcciones en español. Tras esbozar brevemente las características esenciales del proyecto y la clasificación verbal propuesta, se prestará una atención especial a los aspectos más relevantes en el proceso de anotación de papeles semánticos, por ser éste uno de los objetivos principales del proyecto pero también por ser de interés en el trabajo que se presenta. Además, se dedicará un apartado al tratamiento de las entradas verbales y a la separación de acepciones y subacepciones dentro de un mismo sentido verbal, haciendo especial hincapié en los criterios sobre los que se apoya el proyecto para tal fin y en los problemas más frecuentes que surgen a la hora de delimitar diferentes sentidos de un verbo. Dicho apartado muestra cómo se realiza en el proyecto la incorporación de información lexicográfica y el análisis más detallado de las diferentes entradas verbales.

¹ El proyecto ADESSE (<http://webs.uvigo.es/adesse/index.html>) está financiado por el Ministerio de Educación y Ciencia (HUM2005-01573). Contó con una ayuda inicial de la Universidad de Vigo y recibió igualmente financiación por parte del Ministerio de Ciencia y Tecnología (BFF2002-01197) y por la Xunta de Galicia (PGIDIT03PXIC30201PN). El investigador principal del proyecto es José M. García-Miguel Gallego. Otros miembros del equipo investigador son, además de los firmantes de este texto, Francisco Albertuz, Inmaculada Anaya, Lourdes Costas, Vanessa Dacosta, Susana Martínez, María del Carmen Méndez, Antonio Rifón y Ángel Rodríguez Gallardo

1. CARACTERÍSTICAS FUNDAMENTALES DEL PROYECTO ADESSE

1.1. Objetivos del proyecto y trasfondo teórico

El punto de partida sobre el que se construye ADESSE es la BDS (Base de datos Sintácticos del Español Actual), que contiene el análisis sintáctico de casi 160.000 cláusulas del español, sobre un corpus bastante heterogéneo que alcanza aproximadamente el millón y medio de palabras. La información más relevante que se etiquetó en la BDS para cada registro (cláusula) incluye:

- Características generales de la cláusula: tipo de cláusula, modalidad, forma verbal, voz, polaridad...
- Funciones nucleares de la cláusula: Sujeto, Objeto Directo, Objeto Indirecto, Complemento Preposicional, Complemento Agente y Predicativo.
- Rasgos de cada función nuclear: categoría sintáctica, rasgos semánticos inherentes, determinación, número...

La idea principal del proyecto ADESSE consiste en recuperar toda la información sintáctica almacenada en la BDS para añadir información semántica relevante, concretada básicamente en tres apartados: diferenciación de sentidos verbales, clasificación semántica de cada sentido verbal y etiquetación de papeles semánticos. El objetivo final es el de conseguir una base de datos con la que poder obtener múltiples correlaciones de rasgos sintácticos y semánticos, tales como qué verbos y/o qué construcciones son más o menos frecuentes en determinada clase semántica, qué (clases de) verbos son más o menos frecuentes en determinada construcción sintáctica, qué alternativas funcionales presenta un papel semántico concreto para un determinado verbo o una determinada clase...

Sin entrar a desarrollar los diferentes enfoques teóricos que están detrás de la estructura y la metodología de nuestra base de datos, baste apuntar simplemente que, en la línea de algunos trabajos sobre Gramática de Construcciones, partimos de la independencia y compatibilidad semántica entre significado verbal y significado construccional (García-Miguel, 1995; Goldberg, 1995). Si el significado aportado por el verbo evoca cierta estructura o marco conceptual², que incluye el conocimiento cultural sobre las entidades en interacción, por su parte, el esquema

² Cf. Fillmore (1985) y FrameNet: <http://framenet.icsi.berkeley.edu/>

construccional presenta un significado más abstracto que indica puntos de referencia o moldes de presentación de la información:

En conjunto, qué argumentos o participantes resultan expresados, y cómo se expresan deriva de la integración de significado verbal y significado construccional. Las alternancias de diátesis proporcionan conceptualizaciones alternativas que muestran especialmente diferencias de perfil, atención y focalización (García-Miguel et al.,2005b).

1.2. Clases verbales en ADESSE

La clasificación semántica de ADESSE se caracteriza por presentar una base conceptual y un claro diseño de jerarquización. Conviene aclarar, además, que se trata de una clasificación flexible, en el sentido de que ha sido modificada y revisada a medida que se iban clasificando un mayor número de verbos y en función de las necesidades de agrupación y reagrupación que éstos fueron imponiendo. Por tanto, no se trata de una clasificación definitiva y el número y tipo de clases, así como los verbos integrados en ella, están sujetos a posibles alteraciones para dar respuesta a diferentes problemas que puedan surgir y que, por otro lado, son inherentes a cualquier intento de clasificación verbal.

Partiendo de la idea ya apuntada sobre la naturaleza del significado verbal y asumiendo que en la comprensión del lenguaje operan ciertos procesos de categorización y esquematización, nuestra clasificación pretende representar diferentes niveles de generalización a partir de los marcos conceptuales activados por verbos individuales y/o de diferentes dominios cognitivos en los que se integran grupos de verbos de significado relativamente homogéneo. Por poner un ejemplo, entendemos que el significado de verbos como ver o escuchar representa en ambos casos un evento relacionado con la percepción, sea visual o auditiva, que implica en todo caso la asunción de dos participantes en la escena (llamémosles Perceptor y Percibido). Además, en un nivel más alto de abstracción, podemos agrupar dichos verbos con otros como entender o desear, que no remiten a un evento perceptivo (sino cognitivo y volitivo, respectivamente) pero sí a un tipo de proceso de carácter mental, al igual que ver y escuchar y cualquier verbo de percepción³.

³ Para un estudio de los verbos de percepción en ADESSE, véase García-Miguel (2005a).

Por motivos de espacio, recogemos en la Tabla 1 únicamente los dos niveles más altos de la clasificación tal y como está configurada en la fase actual del proyecto. Para una descripción más detallada, aunque ligeramente obsoleta, del diseño, la estructura y los fundamentos semánticos de la clasificación verbal de ADESSE pueden consultarse en el trabajo de Albertuz.

Tabla 1. Clasificación verbal de ADESSE

1 MENTAL	11 Sensación
	12 Percepción
	13 Cognición
	14 Elección
2 RELACIONAL	21 Atribución
	22 Posesión
3 MATERIAL	31 Espacio
	32 Cambio
	33 Otros hechos
	34 Comportamiento
4 COMUNICACIÓN	41 Valoración
	42 Mandato
5 EXISTENCIA	51 Fase
	52 Tiempo
	53 Vida
6 MODULACIÓN	61 Causativo
	62 Disposición
	63 Aceptación

El nivel de generalización más alto se corresponde grosso modo con los tipos de proceso descritos por Halliday (1994). Algunas otras clasificaciones que se han tenido en cuenta, en mayor o menor grado, son Faber & Mairal (1999), [FrameNet](#) (Johnson et al. 2001), Dixon (1991), [WordNet](#) (Fellbaum, 1998), Levin (1993) y Vázquez et al. (2000)⁴

Por último, el fichaje numérico responde a la idea jerárquica de la clasificación. El primer dígito se refiere al nivel más general, el segundo dígito al segundo nivel

⁴ Véase Albertuz para una explicación de los puntos de encuentro y las divergencias entre ADESSE y otras clasificaciones verbales.

en orden de categorización, y así sucesivamente. Por ejemplo, el verbo poner pertenece a la clase Localización, que recibe la etiqueta numérica 312. De izquierda a derecha, el 3 remite a procesos de tipo Material, el 1 al grupo de verbos que se integran en el Espacio y el 2 al conjunto de los verbos que, dentro de Espacio, conforman la clase de Localización.

Por supuesto, nos enfrentamos a diversos problemas que se derivan irremediablemente de las limitaciones propias de una clasificación jerárquica de tipo semántico, como son ciertos límites borrosos entre grupos verbales o verbos que comparten rasgos de más de una clase. Para este último caso se ha contemplado la posibilidad de una **clasificación múltiple** (p. ej., parir está tratado como un verbo de Fisiología y de Creación), aunque somos conscientes de que algunas de las cuestiones a las que nos enfrentamos merecerán revisiones futuras que irán perfilando y ajustando la clasificación actual del proyecto.

2. LA ANOTACIÓN DE PAPELES SEMÁNTICOS

El objetivo primordial del proyecto ADESSE es el de la **anotación de papeles semánticos**. Sin embargo, el proceso de etiquetación de roles no se ha llevado a cabo para cada una de las cláusulas que componen el corpus. Partimos de la idea básica de que cada nivel de análisis hereda la información establecida en el nivel superior. Teniendo esto en cuenta, cada esquema sintáctico de un verbo hereda los papeles que han sido previamente definidos para dicho sentido verbal, que a su vez hereda por defecto los papeles considerados básicos para la clase en la que se integra. En el nivel de análisis más bajo cada expresión particular presenta la estructura semántica definida para el esquema, de tal forma que podemos acometer la anotación de los 3995 sentidos verbales que presenta actualmente la base, así como de los 12555 esquemas sintácticos diferentes, salvando el esfuerzo de tener que etiquetar cada una de las casi 160.000 cláusulas que conforman el corpus de la BDS, siempre y cuando seamos conscientes de que los niveles inferiores pueden incrementar o contradecir la información heredada.

Volviendo a la clasificación semántica, a cada tipo de proceso diferenciado (=clase conceptual) se le asignan una serie de papeles propios del dominio cognitivo que se describe. La Tabla 2 muestra la relación de roles para algunos de los grupos verbales que han sido creados en ADESSE, a saber:

Tabla 2. Etiquetas de papeles en ADESSE.

Dominio cognitivo	1	2	3	4
<i>Sensación</i>	Experimentador	Estímulo		
<i>Percepción</i>	Perceptor	Percibido		
<i>Posesión: transferencia</i>	Poseedor-inicial	Posesión	Poseedor-final	
<i>Espacio: desplazamiento</i>	Móvil	Origen	Dirección	Trayecto
<i>Hecho: emisión</i>	Fuente	Emisión		
<i>Comportamiento</i>	Actor	Manera		
<i>Comunicación</i>	Emisor	Mensaje	Receptor	

Por otro lado, cada entrada verbal incorpora un conjunto de argumentos que permite dar cuenta de todos los actantes posibles en la escena designada (su potencial valencial). Por defecto, los argumentos verbales heredan la etiqueta de rol correspondiente en función de la clase a la que el verbo pertenece. La Figura 1 ilustra lo dicho tomando como ejemplo un verbo de Comunicación:

<i>Comunicación</i>	Emisor	Mensaje	Receptor
↑	↑	↑	↑
Avisar	A1	A2	A3

Figura 1. Herencia de roles para el verbo avisar.

Por último, se etiquetan los esquemas sintáctico-semánticos de cada verbo (sus realizaciones valenciales) con un índice numérico para cada función de cada esquema. Con ello, se pueden establecer fácilmente las correspondencias entre argumentos de diferentes esquemas sintácticos, como se puede ver en la Figura 2:

VOZ	Esquema	S	D	I	O	O2	A	P	Ejs	OBSERV
Activa	S D	1	3						41	[ES Dupl: S=A1-D=A2/ S=A1-D=A3]
Activa	S D I	1	2	3					10	
Activa	S D R(de)	1	3		2				7	
Activa	S R(de)	1			2				2	
Activa	S	1							1	
Activa	S:Q	1							1	
Activa	S D	1	2						1	[ES Dupl: S=A1-D=A2/ S=A1-D=A3]
Activa	S D:Q	1	3						1	
Activa	S R(para)	1			7				1	
CPmpasiva	S	2							1	
CPimpersona	S D	1	3						2	

Figura 2. Etiquetación de roles para el verbo avisar.

En resumen, el proceso de anotación de papeles semánticos en ADESSE para un verbo dado ha de seguir los pasos siguientes:

- Observación de los ejemplos registrados en el corpus y del análisis lexicográfico en los diccionarios del español
- Definición de los sentidos verbales y de los roles verbales
- Clasificación del verbo y herencia de roles
- Anotación de roles en cada esquema sintáctico
- Anotación de usos figurados y otro tipo de observaciones

3. LA INFORMACIÓN LEXICOGRÁFICA EN ADESSE

En ADESSE se diferencian dos fases relacionadas con la labor **lexicográfica**; la primera de ellas se ocupa de la delimitación de macroacepciones y la segunda de la división de **acepciones** y subacepciones. Esta última es en la que nos centraremos en este último apartado.

A diferencia de otras bases de datos como WordNet, en la que se defiende una alta **polisemia** verbal, introduciendo una nueva acepción por cada sentido, en ADESSE, debido a que el interés principal se centra en la estructura argumental y no en las entradas lexicográficas, las macroacepciones verbales se han separado lo mínimo posible.

En la segunda fase, el objetivo principal es la creación de acepciones y subacepciones, tomando siempre como referencia obras lexicográficas como el DRAE, el DEA, el CLAVE y el DUE (referenciados en la sección de bibliografía de este capítulo) y en ocasiones el diccionario de José Cuervo o WordNet, para la organización y reelaboración de nuestras propias acepciones, teniendo en cuenta sólo los ejemplos recogidos en la BDS.

Por lo general, los diccionarios presentan las acepciones de forma individual sin establecer ningún tipo de relación entre ellas, lo que provoca que, normalmente, el número de acepciones sea mayor que el recogido en ADESSE. Por lo tanto, una de las labores que se está realizando es la de reorganizar la información lexicográfica verbal de los diccionarios, estableciendo y jerarquizando las relaciones existentes entre los diferentes sentidos. A diferencia de lo que ocurría con las

macroacepciones, en las acepciones y subacepciones no se pretende tanto simplificar y agrupar los significados, como establecer relaciones entre ellos.

Por ejemplo, en el DEA, el verbo *escapar* diferencia 18 acepciones, mientras que nosotros hemos reorganizado los significados en 3 acepciones, en las que hemos incluido subacepciones que recogen el sentido inicial especializándose en algún matiz o aspecto particular:

Escapar

1. Irse de prisa o a escondidas.
 - 1.1. Huir, irse de prisa o a escondidas.
 - 1.1.1. Abandonar un lugar durante un (breve) periodo de tiempo con la intención de volver.
 - 1.1.2. Derramar(se), soltar(se) o salir por una abertura alguna cosa.
 - 1.2. Realizar una acción de forma involuntaria (extensión metafórica).
 - 1.2.1. Emitir un sonido.
 - 1.2.2. Producir una información.
 - 1.3. Ignorar o no tener en cuenta una información o hecho (extensión metafórica).
 - 1.3.1. Pasar inadvertida una información para alguien.
 - 1.3.2. Resultar algo desconocido e incomprensible.
 - 1.3.3. Dejar pasar un suceso sin aprovecharlo.

Uno de los puntos más problemáticos es establecer los criterios empleados para la separación de macroacepciones, acepciones y subacepciones. A continuación se describe cómo hacerlo, teniendo en cuenta que son diferentes los empleados en uno y otro nivel de jerarquización.

Las *alternancias de diátesis* con diferencias semánticas que atribuimos a la construcción y no específicamente al verbo (como los de la elección de voz activa/media/pasiva, la alternancia locativa, alternancia causativa/incoativa, etc.) son siempre interpretados como variaciones diatéticas de un mismo verbo. Este es uno de los objetivos principales de ADESSE: registrar alternancias construccionales de un mismo verbo y comprobar qué añade la construcción al significado básico del verbo. Este parámetro no da lugar a entradas verbales diferentes en ADESSE en ninguno de los tres niveles de jerarquización.

Otro criterio no diferenciador de macroacepciones son las *alternativas paradigmáticas*. Por ejemplo, el diccionario CLAVE enumera tres acepciones diferentes del verbo *vestir* en función del complemento (quién hace el vestido, quién paga el vestido...) mientras que en ADESSE se separa únicamente una en la que se incluyen todos los casos con el significado de *poner o llevar puesta ropa*. En cambio, las *alternancias paradigmáticas* pueden crear nuevas acepciones y subacepciones cuando en una determinada estructura uno de los participantes posee una mayor relevancia que implica una diferencia de significado. Por ejemplo, no tiene el mismo significado el verbo *vestir* en *Ana viste a su muñeca* que si decimos *Las hojas nuevas visten ya los árboles*. También puede implicar un cambio de sentido la variación de los rasgos semántico-inherentes de los participantes (animado/inanimado, concreto/abstracto...).

Las *extensiones metafóricas y metonímicas* que puedan derivarse del sentido básico del verbo no distinguen macroacepciones. Por ejemplo, el verbo *asegurar* puede utilizarse con el sentido figurado de *hacer que algo quede garantizado* (*Halard, cuyo resultado le aseguró 2.782.000 pesetas* [2VOZ: 47, 2, 1, 009]) en relación con el sentido básico *poner firme. Hacer algo seguro*. Sin embargo, es uno de los criterios fundamentales para la separación de acepciones y subacepciones.

Las acepciones los usos figurados no poseen ningún sentido verbal que nos permita relacionarlos con uno de los significados más generales ya existentes.

Beber

- 1.1. Ingerir alimentos líquidos.
- 1.2. Obtener información de algo o alguien. (extensión metafórica)

En las subacepciones presentan el sentido inicial de la acepción, pero mediante un uso metafórico, por lo tanto existe una relación de inclusión entre los significados.

Marchar

- 1.1. Ir(se) abandonar un lugar.
 - 1.1.1. Morir(se).

Para la diferenciación de macroacepciones se han tenido en cuenta los casos de *homonimia* y cuando los diferentes sentidos de un verbo se corresponden con diferentes clases semánticas reconocidas en ADESSE; por ejemplo *saber 1 (conocer)* pertenece a la clase de *conocimiento* y *saber 2(sabor)* a la de *apariencia*.

La *conciencia lingüística* de los etiquetadores y las *obras lexicográficas* como el DEA, DRAE, CLAVE, DUE, DCR y WordNet son dos factores externos que están presentes en todos los niveles de jerarquización.

Una vez que se completa la información lexicográfica de cada verbo, todos los ejemplos son catalogados dentro de la acepción o subacepción correspondiente y, después, se selecciona un ejemplo para que esté visible al lado de su definición, ayudando así al lector a tener una comprensión más fácil y rápida. Por último, se hace la correspondencia de las acepciones y subacepciones de ADESSE con las de los demás diccionarios citados.

Así, ADESSE pretende, tomando siempre como referencia otras obras lexicográficas, elaborar sus propias acepciones y subacepciones de forma clara y estableciendo, en la mayor medida de lo posible, relaciones semánticas que permitan agrupar los significados por su proximidad, de manera que en ellas se puedan incluir cada uno de los ejemplos del corpus ARTHUS (*adesse.uvigo.es/data/corpus.php*).

4. REFERENCIAS BIBLIOGRÁFICAS

ADESSE. Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español (Universidad de Vigo): <<http://webs.uvigo.es/adesse/>>

ALBERTUZ, FRANCISCO J. (en prensa): «Sintaxis, semántica y clases de verbos: Clasificación verbal en el proyecto ADESSE». *Actas del VI Congreso de Lingüística General. Santiago de Compostela*, 3-7 mayo 2004: <<http://webs.uvigo.es/adesse/textos/Albertuz-CLG6.pdf/>>

BDS. Base de Datos Sintácticos del Español Actual (Universidad de Santiago): <<http://www.bds.usc.es/>>

CLAVE: *Diccionario de Uso del Español Actual*. Madrid: S. M. 2002 (5.^a edición).

DCR: *Diccionario de Construcción y Régimen de la Lengua Castellana*. 2002 (edición electrónica).

DEA: *Diccionario del Español Actual*. Aguilar: Madrid 1999 (1.^a edición).

DRAE: *Diccionario de la Real Academia Española*. Madrid: Espasa Calpe 2001 (22.^a edición)

- DIXON, ROBERT M. W. (1991): *A New Approach to English Grammar, on Semantic Principles*. Oxford: Oxford University Press.
- DUE: *Diccionario de Uso del español*. Madrid: Gredos 2001 (edición electrónica)
- FABER, PAMELA y RICARDO MAIRAL (1999): *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.
- FELLBAUM, CHRISTIANE (1998): «A Semantic Network of English Verbs». Christiane Fellbaum, ed. *WordNet. An Electronic Lexical Database*. Cambridge (Massachusetts): The MIT Press, pp. 69-104.
- FILLMORE, CHARLES (1985): «Frames and the Semantics of Understanding». *Quaderni di Semantica*, 6/2, pp. 222-254.
- FrameNet (University of California at Berkeley): <<http://www.icsi.berkeley.edu/framenet/>>
- GARCÍA-MIGUEL, JOSÉ M. (1995): *Las relaciones gramaticales entre predicado y participantes*. Universidad de Santiago de Compostela.
- GARCÍA-MIGUEL, JOSÉ M. (2005a): Aproximación empírica a la interacción de verbos y esquemas construccionales, ejemplificada con los verbos de percepción». *Estudios de Lingüística Universidad de Alicante (ELUA)*, 19, 169-191: http://webs.uvigo.es/adesse/textos/Verbos-construcciones_ELUA_final.pdf
- GARCÍA-MIGUEL, JOSÉ M.; LOURDES COSTAS y SUSANA MARTÍNEZ (2005b): «Diátesis verbales y esquemas construccionales. Verbos, clases semánticas y esquemas sintáctico-semánticos en el proyecto ADESSE», Wotjak, Gerd & Juan Cuartero Otaol, eds. *Entre semántica léxica, teoría del léxico y sintaxis*. Frankfurt am Main: Peter Lang, pp. 373-384: <http://webs.uvigo.es/adesse/textos/VICILH03.pdf>
- HALLIDAY, M. A. K. (1994): *An Introduction to Functional Grammar*. London: Edward Arnold.
- JOHNSON, CHRISTOPHER R., CHARLES J. FILLMORE, ESTHER J. WOOD, JOSEF RUPPENHOFER, MARGARET URBAN, MIRIAN R.L. PETRUCK & COLLIN F. BAKER (2001): *The FrameNet Project: Tools for Lexicon Building* (Vers. 0.7):
- LEVIN, BETH (1993): *English Verb Classes and Alternations: a preliminary Investigation*. Chicago and London.
- VÁZQUEZ, GLORIA; ANA FERNÁNDEZ y M.^a ANTONIA MARTÍ (2000): *Clasificación verbal. Alternancias de diátesis*. Lleida: Universitat de Lleida.

WordNet. A Lexical Database for the English Language (Princeton University): <<http://www.cogsci.princeton.edu/~wn/>>