

DATOS PARA UN ANÁLISIS CONTRASTIVO DE VALENCIAS VERBALES: DICCIONARIOS Y CORPUS

Barbara Lübke (barbara.lubke@usc.es)

Pia Sloth Poulsen (piasloth.poulsen@usc.es)

Victoria Vázquez Rozas (victoria.vazquez@usc.es)

Universidad de Santiago

Resumen

Se presenta el *Diccionario contrastivo de valencias verbales español-alemán* (DCVVEA) y se describe brevemente su estructura y el proceso de elaboración, ilustrándolos con una unidad de descripción lexicográfica, en este caso una acepción con un esquema actancial específico, que junto con otras constituye una entrada verbal. La comunicación se centra en la discusión de los instrumentos empíricos y analíticos utilizados en el proceso descriptivo, y en particular sobre las ventajas y las limitaciones del uso de corpus como base material para la elaboración de un diccionario de este tipo. Se abordan asimismo algunas de las implicaciones metodológicas de la vertiente específicamente contrastiva del proyecto.

Palabras clave: valencia verbal, lingüística de corpus, español, alemán, lexicografía contrastiva.

I. INTRODUCCIÓN

En esta comunicación presentaremos las características del proyecto del *Diccionario contrastivo de valencias verbales español-alemán* (DCVVEA), que se está elaborando en la Universidad de Santiago.¹ Se trata de un diccionario sintagmático dirigido en primer lugar a usuarios hispanohablantes estudiantes de alemán como lengua extranjera, aunque por sus características puede ser también útil como fuente de datos para otras investigaciones contrastivas sobre aspectos léxicos y sintácticos de las construcciones verbales.

Los diccionarios sintácticos o ‘de construcción y régimen’, como se denomina la gran obra sobre la lengua española de Rufino José Cuervo, nacen con el objetivo de proporcionar datos sintácticos y semánticos más amplios y detallados de los que habitualmente aportan los diccionarios generales, y de responder con ello a la necesidad de información específica sobre las posibilidades combinatorias de las unidades léxicas. Los primeros diccionarios de valencias se crearon en el curso de la evolución de la teoría de valencias a partir de los años sesenta en Alemania. Las obras pioneras de la lexicografía valencial alemana, el *Wörterbuch zur Valenz und Distribution deutscher Verben* de Gerhard Helbig y Wolfgang Schenkel (1969) y el *Kleines Valenzlexikon deutscher Verben* de Ulrich Engel y Helmut Schumacher (1976), se elaboraron en primer lugar para demostrar y comprobar los supuestos teóricos de la lingüística valencial de Tesnière e introdujeron sendos modelos de análisis y descripción de la valencia verbal, que a continuación sirvieron como base tanto para procedimientos lexicográficos más complejos en los diccionarios sucesores como para el desarrollo de la teoría en general.

Desde el principio los diccionarios de valencias fueron concebidos también para responder a las necesidades del aprendizaje de una lengua extranjera, si bien es verdad que las primeras obras de este tipo prestaron poca atención a su presentación didáctica, y que el acceso a las informaciones recopiladas resulta en algunos casos difícil, sobre todo si pensamos en un usuario que aún no posee conocimientos avanzados de la lengua de llegada.

Para remediar este problema, pronto se crearon diccionarios de valencias bilingües que incluyen equivalentes, ofrecen la traducción de los ejemplos o presentan la introducción y las instrucciones en la lengua materna del destinatario. Los primeros pasos en esta dirección los dieron el *Französisches Verblexikon* de Winfried Busse y Jean-Pierre Dubost (1977), dirigido a aprendices germanohablantes del francés, y el *Diccionario de valencias verbales. Alemán-Español*, elaborado en México por Dieter y Marlene Rall y Oscar Zorilla (1980) para usuarios hispanohablantes que aprenden alemán. El paso siguiente lo representa la elaboración de diccionarios contrastivos a partir de los años ochenta, i. e. diccionarios que ofrecen una descripción valencial para los lexemas correspondientes de dos o más lenguasⁱⁱ; en esta línea se sitúa también nuestro proyecto de *Diccionario contrastivo de valencias verbales español-alemán*.

Si comparamos los diccionarios de valencias más recientes, como por ejemplo el *Valency Dictionary of English* de Thomas Herbst y otros (2004) o el *VALBU – Valenzwörterbuch deutscher Verben* de Helmut Schumacher y otros (2004), con las obras de la primera generación, no sólo observamos, como es natural, un avance en la metodología, en las bases teóricas y en la complejidad y precisión de los modelos de descripción, sino también en el acercamiento a los datos empíricos y a los usos reales de la lengua, posible en gran medida gracias a la creación de los corpus informatizados y los progresos de la lingüística de corpus, que ponen al alcance de los lexicógrafos las herramientas y métodos para el registro de datos representativos en grandes dimensiones. Si los primeros diccionarios de valencias se sustentan casi exclusivamente en la introspección y en el análisis de ejemplos inventados, ya a partir de finales de los años ochenta se observa la tendencia de operar cada vez más con datos empíricos e integrar material auténtico en la descripción lexicográfica. Así, diccionarios como *Verben in Feldern* (1986) y *VALBU* (2004) elaborados en el Instituto de la Lengua Alemana toman los ejemplos ilustrativos de los corpus de textos del IDS, el PONS *Wörterbuch der italienischen Verben* de Peter Blumenthal y Giovanni Rovere (1998) está basado en la explotación de un corpus de textos italianos mayoritariamente periodísticos y el *Valency Dictionary of English* (2003) surge como resultado del análisis del *Cobuild Bank of English*.

En nuestro proyecto utilizamos como fuente de datos empíricos la *Base de datos sintácticos del español actual* (BDS), elaborada en los años noventa por investigadores de la USC bajo la dirección de Guillermo Rojo (<http://www.bds.usc.es>). La BDS contiene los resultados del análisis sintáctico de casi 160.000 cláusulas de un corpus textual de español contemporáneo, el corpus ARTHUS (*Archivo de textos hispánicos de la Universidad de Santiago de Compostela*) de casi un millón y medio de palabras, que incluye muestras de discurso narrativo (37,19%), ensayístico (17,78%), teatral (14,66%), periodístico (11,51%) y oral (11,51%), mayoritariamente representativos del español peninsular (78,77%) frente al español de América (21,23%).

En el apartado siguiente presentamos la estructura y contenido de una unidad de descripción del diccionario, una acepción con un esquema actancial específico, que junto con otras constituye una entrada lexicográfica. A continuación (apartado 3) hablaremos de los instrumentos empíricos y analíticos utilizados en el proceso descriptivo, y en particular sobre las ventajas y las limitaciones del uso de corpus como base material para la elaboración de un diccionario de este tipo. Finalmente (apartado 4) comentaremos algunas de las implicaciones metodológicas de la vertiente específicamente contrastiva del proyecto.

II. CONFIGURACIÓN DE LAS ENTRADAS DEL DCVVEA

El ejemplo incluido a continuación corresponde a una de las cinco acepciones registradas en el DCVVEA para el verbo *suponer*. Los datos que aquí aparecen en un formato de tabla se introducen en una base de datos informatizada cuyos campos corresponden a las diferentes indicaciones descriptivas que hemos considerado pertinentes. La versión electrónica que estará accesible en internet al término del proyecto facilitará la recuperación de la información atendiendo a diferentes criterios de búsqueda.

<p>suponer 2</p> <p><i>Implicar, llevar consigo o significar</i></p>	<p>su_j acu (dat) / (prp:para)</p> <p>ORACIONAL: su_j: <i>que+s, inf</i> acu: <i>que+s, inf</i></p> <p>su_j: – acu: <i>inmat</i> dat: <i>anim</i> prp: <i>anim</i></p>	<p>su_j acu (prp:für+A) dat ⇒ prp</p> <p>ORACIONAL: su_j: <i>dass, zu+inf</i> acu: <i>dass, zu+inf</i></p> <p>su_j: – acu: <i>inmat</i> prp: <i>anim</i></p>	<p>bedeuten 1</p> <p>bedeutet, bedeutete, hat bedeutet</p> <p>P: imposible</p>
<p>su_j acu No me atrevo a presentir el grado de entrega, inmolación, arrasamiento personal que puede suponer el amor al hijo. JOVENES: 168, 19</p>		<p>su_j acu Ich wage nicht, mir den Grad an Hingabe, Aufopferung, Selbstaufgabe vorzustellen, den die Liebe zu einem Kind bedeutet.</p>	
<p>su_j:OR:que acu Tampoco supone ninguna sorpresa el que la primera incursión importante de los islamistas en este terreno fuese tan impresionante que, de haberse llevado a cabo conforme a lo que parece ser su planificación original, habría figurado, junto a Lockerbie, en los anales de las atrocidades aéreas. CREA</p>		<p>su_j:OR:dass acu Es bedeutet auch keine Überraschung, dass das erste schwerwiegende Eingreifen der Islamisten auf diesem Gebiet so eindrucksvoll war, dass es, wenn es so wie offenbar ursprünglich geplant durchgeführt worden wäre, neben Lockerbie in die Analen der Gräueltaten im Flugverkehr eingegangen wäre.</p>	
<p>su_j:OR:inf acu «Tú no sabes, Julián, lo que supone haber vivido aquí desde el principio.» JOVENES: 109, 37</p>		<p>su_j:OR:zu+inf acu “Du weißt nicht, Julián, was es bedeutet, hier von Anfang an gelebt zu haben.“</p>	
<p>su_j acu:OR:que La tan magnificada creación histórico-político-cultural romana supuso, entre otras cosas, que, [...], César llegase a matar hasta un millón de galos. RATON: 217, 3</p>		<p>su_j acu:OR:dass Die so gelobte historisch-politisch-kulturelle Leistung Roms bedeutete unter anderem, dass Cäsar bis zu einer Million Gallier töten würde.</p>	
<p>su_j acu dat [...] el año pasado hice gimnasia, tuve que ir veinte días, me suponía una pérdida de tiempo tremenda [...]. MADRID: 383, 30</p>		<p>su_j acu prp Letztes Jahr habe ich Gymnastik gemacht, ich musste 20 Tage lang hingehen, es hat für mich einen riesigen Zeitverlust bedeutet.</p>	
<p>su_j acu:OR:inf dat Total que me supuso repetir un curso. SEVILLA: 68, 43</p>		<p>su_j acu:OR:zu+:inf prp Kurz, es bedeutete für mich, ein Schuljahr zu wiederholen.</p>	
<p>su_j acu prp</p>		<p>su_j acu prp</p>	

A Jano le impresionó aquella extraña coincidencia, el hecho de que Adriana pudiera suponer algún día para Marescu lo que Francesca había supuesto para él [...]. CARTA: 68, 2	Jano beeindruckte diese seltsame Übereinstimmung, die Tatsache, dass Adriana für Marescu eines Tages das bedeuten könnte, was Francesca für ihn bedeutet hatte.
	<p>subj acu (prp:für+A) dat ⇒ prp</p> <p>ORACIONAL: subj: zu+inf, dass</p> <p>subj:- acu: <i>inmat</i> prp: <i>anim</i></p> <p>darstellen 1 <i>Significar</i> stellt dar, stellte dar, hat dargestellt P: imposible</p>
subj acu [...] los niños, el jardinero, el ama, dormidos todos apenas suponían compañía o apoyo [...]. JOVENES: 38, 29	subj acu Die Kinder, der Gärtner, die Haushälterin, die alle schliefen, stellten kaum Gesellschaft oder Unterstützung dar.
subj:OR:que acu Tampoco supone ninguna sorpresa el que la primera incursión importante de los islamistas en este terreno fuese tan impresionante que, de haberse llevado a cabo conforme a lo que parece ser su planificación original, habría figurado, junto a Lockerbie, en los anales de las atrocidades aéreas. CREA	subj:OR:dass acu Es stellt auch keine Überraschung dar, dass das erste schwerwiegende Eingreifen der Islamisten auf diesem Gebiet so eindrucksvoll war, dass es, wenn es so wie offenbar ursprünglich geplant durchgeführt worden wäre, neben Lockerbie in die Analen der Gräueltaten im Flugverkehr eingegangen wäre.
subj:OR:inf acu [...] meter un entrenamiento a diario y un ejercicio y demás, pues suponía un gran sacrificio. MADRID: 91, 32	subj:OR:zu +inf acu Ein Training täglich und eine Übung und alles übrige einzubauen, stellte ein großes Opfer dar.
subj acu dat [...] el año pasado hice gimnasia, tuve que ir veinte días, me suponía una pérdida de tiempo tremenda [...]. MADRID: 383, 30	subj acu prp Letztes Jahr habe ich Gymnastik gemacht, ich musste 20 Tage lang hingehen, es hat für mich einen riesigen Zeitverlust dargestellt.

Tabla 1. Ejemplo de descripción valencial del DCVVEA correspondiente a una acepción del verbo *suponer*

Según se ve en la tabla, la descripción de cada acepción verbal española incluye una definición de significado y el esquema valencial. Las indicaciones funcionales que integran los esquemas son comunes a las partes española y alemana, aunque por supuesto los mecanismos de codificación son diferentes en ambas lenguas. Las abreviaturas que aparecen en la tabla 1 tienen los siguientes valores: *subj*= sujeto; *acu*= acusativo, complemento directo; *dat*= dativo, complemento indirecto; *prp*= complemento preposicional, en este caso introducido por *para*. Los paréntesis marcan facultatividad (el caso de *dat* y *prp* en la tabla), frente a los actantes *subj* y *acu* que son obligatorios. Finalmente, en el esquema también se indica mediante la barra inclinada “/” que hay una alternancia entre el complemento indirecto y el preposicional, es decir, que puede realizarse uno u otro pero no ambos.

La descripción valencial señala la posibilidad de realización oracional de los actantes. Como se ve en el ejemplo, tanto el sujeto como el complemento directo se pueden realizar en forma de cláusula introducida por *que* (+s marca el modo subjuntivo) o de cláusula de

infinitivo. Se dan también los rasgos semánticos de cada uno de los actantes. En este caso, el sujeto no tiene restricciones semánticas, el complemento directo se limita a lo *inmaterial* y los complementos indirecto y preposicional a actantes animados.

En la descripción del equivalente alemán se opera de la misma manera, aunque naturalmente con algunas particularidades determinadas por las características gramaticales de esta lengua. Se indican “metataxis”, es decir, cambios en las correspondencias entre actantes semánticos y sintácticos (en este ejemplo *dat* \Leftrightarrow *prp*), así como los casos regidos por las preposiciones alemanas (*prp:für*+A – complemento preposicional introducido por *für* + acusativo). Se dan también las formas básicas del paradigma y las posibilidades de construcción pasiva del equivalente alemán.

La descripción valencial del español se sustenta en el análisis previo de los ejemplos del corpus, y la identificación de equivalentes y su consiguiente análisis valencial toma como base la traducción al alemán de dichos ejemplos. El diccionario incluye una selección de los usos registrados como ilustración de las posibilidades combinatorias de los verbos en ambas lenguas.

III. LOS DATOS PARA EL ANÁLISIS VALENCIAL EN EL DCVVEA

El objetivo de un diccionario contrastivo de valencias es ofrecer a los usuarios información relevante sobre las diferencias de combinatoria verbal entre dos lenguas, en este caso el español y el alemán. Los esquemas actanciales, que expresan de una forma sintética las indicaciones funcionales y categoriales pertinentes, constituyen un sistema adecuado para representar la valencia verbal. No obstante, la estructura actancial de los verbos no es accesible a la observación directa del analista, sino que requiere el empleo de algún método heurístico. Como señalamos en la introducción al trazar la trayectoria de los diccionarios de valencias a lo largo de las últimas décadas, los procedimientos para determinar el número y carácter de los elementos previstos en el plan constructivo del verbo oscilan entre el método introspectivo y el análisis de corpus textuales.

El recurso a la intuición lingüística como recurso analítico no tiene por qué entrar en contradicción con el manejo de corpus en la descripción valencial. La capacidad que poseen los usuarios de las lenguas para dirigir su atención consciente a ciertos aspectos de su propio lenguaje resulta ser un instrumento de inestimable valor en el análisis lingüístico, independientemente de la relevancia teórica y descriptiva que se atribuya a los datos de uso real. De hecho, ciertos aspectos del lenguaje resultan simplemente inaccesibles por otra vía que no sea la introspección. Es lo que ocurre con el significado, especialmente con el contenido de las palabras (léxicas) y oraciones, o con el valor pragmático de los enunciados. Asimismo, la intuición de los hablantes se manifiesta en sus juicios de aceptabilidad, “gramaticalidad” o “normatividad” (los conceptos varían dependiendo la perspectiva adoptadaⁱⁱⁱ) sobre las oraciones, es decir, sus opiniones sobre la adecuación o no de una expresión a los patrones gramaticales de la lengua, lo que no excluye un margen de imprecisión y diferencias de criterio en algunos casos.

Se entenderá así que la información esquemática que proporciona un diccionario de valencias pudiera idealmente extraerse del saber lingüístico del propio lingüista, o mejor dicho, de la observación introspectiva a la que el analista-hablante se somete. La capacidad de abstracción y generalización que constatamos en las intuiciones de los usuarios (por ejemplo, en su aptitud para asignar significados a las palabras u oraciones aisladas, o en la esquematización que supone valorar la “normatividad” de las expresiones), se transforma en el caso del usuario-lingüista en competencia para identificar los patrones constructivos de las unidades léxicas y para producir secuencias que ejemplifican las diferentes posibilidades

valenciales. El aprovechamiento de esta capacidad se plasma en los primeros diccionarios de valencias, que como contrapartida al mérito de su novedad, ponen de manifiesto las limitaciones del método introspectivo.

Porque ocurre que no todos los aspectos del lenguaje resultan igualmente accesibles a la conciencia del usuario. Talmy (2007: xiv), hace referencia a la imposibilidad de activar intencionalmente la gama completa de significados asociados con una forma dada, independientemente de que el hablante pueda reconocerlos sin problemas si se le ofrecen en los contextos adecuados: “Apparently, our cognition is organized in a way that allows particular senses of a word to come into consciousness in the relevant contexts, but not as a full connected set under introspection – though introspection does yield a few. The full set can again typically be achieved only with the aid of analytic procedures, like those giving rise to dictionaries through a kind of corpus research.”

Algo similar puede decirse de nuestro conocimiento sobre las diferentes alternativas de valencia (y de significado) asociadas a cada lexema verbal. Apelar a la competencia del usuario lleva a restringir y simplificar las opciones realmente existentes de esquemas y realizaciones actanciales, de ahí que los diccionarios de valencias elaborados más recientemente hayan incorporado datos de uso registrado en corpus para garantizar el requisito de la representatividad.

Asumiendo que en general se establecen distinciones de significado (acepciones lexicográficas) correlativas a los diferentes esquemas valenciales, una forma de ampliar el sustento empírico de las descripciones valenciales es recurrir a los diccionarios generales para completar la lista de acepciones. Para elaborar las entradas de nuestro diccionario hemos elegido el *Diccionario del español actual* (DEA) como base lexicográfica, pues por la estructuración de sus artículos, el detalle de las definiciones de significado, la inclusión de datos del contorno sintagmático y la incorporación de ejemplos textuales, reúne la información más adecuada a nuestros propósitos. No obstante, los objetivos de la descripción valencial conducen a modificar la propuesta de acepciones del DEA en bastantes casos. A continuación se exponen las líneas generales del proceso de adaptación que seguimos.

En general hemos prescindido de usos claramente especializados y de los calificados como “raros” en el DEA, de acepciones ligadas a registros muy marcados por su estilo (poético, vulgar, etc.), o restringidas a contextos muy particulares.

La noción de valencia se asienta sobre un principio de composicionalidad sintáctica: los esquemas actanciales constituyen marcos organizativos abstractos constituidos por espacios funcionales disponibles para las elecciones léxicas y morfológicas de los usuarios. En nuestro diccionario, los esquemas actanciales restringen dichas elecciones al incorporar las características semánticas generales de los actantes, así como indicaciones de tipo categorial respecto a sus posibles realizaciones oracionales. Pero la información que ofrece el DEA incluye restricciones de tipo gramatical y pragmático que van más allá de lo que puede recoger una descripción valencial y que son el resultado de procesos de fijación o rutinización sintáctica asociados a ciertas funciones comunicativas.

Por ejemplo, las acepciones 10 y 11 del verbo *andar* en el DEA se refieren exclusivamente a unas formas determinadas del paradigma verbal que expresan una intención ilocutiva de tipo directivo:

10 (col) *En imperat, se usa para exhortar al oyente a que haga algo que en el contexto se especifica o a veces se sobrentiende.* | Matute *Memoria* 109: Anda, renegado, Judas. [...]. SFerlosio *Jarama* 41: Se reían los otros desde el grupo y Fernando azuzaba: “¡Anda con él!”.

11 andando. (col) *Se usa para exhortar a otro a ponerse en marcha o a actuar.* | Torrente *DJuan* 301: -Tienes que llevar estas cartas.. [...]. Andando. Gala *Hotelito* 59: Hermosas como soles os quiero ver. Andando.

De la misma manera, la acepción 11 del verbo *tirar* solo aparece, según la información del DEA, en construcción pasiva –en realidad sólo en la forma de participio:

11 Arrastrar [algo] tirando [20] de ello. *Solo en constr pasiva*. | F. J. FTascón SYa 3.12.7, 29: [...] marinero como Lohrengrin, de frágil esquife, tirado por cisnes, surcando en el cielo azul.

Este tipo de restricciones constructivas rompe con la función generalizadora de los patrones actanciales y revela las limitaciones de la metodología valencial, que difícilmente puede darle un tratamiento adecuado. El concepto de “fórmula oracional” que proponen los autores del DEA (cf. vol I, págs. xix-xx) puede aplicarse a estos usos, o al menos a los casos más fijados, que cuando se incluyen en nuestro diccionario se relegan a una sección complementaria de expresiones fraseológicas añadida a algunas entradas.

Similar tratamiento reciben ciertas construcciones verbales que restringen léxicamente alguno de sus actantes y que en el DEA se reconocen como acepciones independientes. Es lo que ocurre, por ejemplo, con la acepción 13 del verbo *coger*, cuyo sujeto sólo puede ser “el sol”:

13 Ejercer sus efectos [sobre alguien (*cd*) un agente físico, esp. el sol]. | * ¡Cómo te ha cogido el sol!;

o con la acepción 6b de *caer*:

b) Terminarse [el día o la tarde]. | Cela *Pirineo* 272: Un coro de grajos .. se ríe a carcajadas del día que va cayendo pausado y triste como un anciano;

o con las acepciones 3a-c del verbo *levantar*, que sólo se refieren a “la voz”:

3 Subir la intensidad o volumen [de la voz (*cd*)]. *Tb fig*. | Diego *Abc* 3.5.75, sn: [...] se ve uno forzado insensiblemente a levantar la voz [...]. **b)** ~ **el gallo**, ~ **la voz** → GALLO, VOZ.

c) *pr* Subir de volumen [la voz (*suj*)]. | MMolina *Jinete* 271: [...] Lo hacía siempre que oía levantarse las voces.

Dado que la descripción actancial constituye el eje organizativo de un diccionario de valencias, con frecuencia se opta por unir acepciones del DEA, siempre que el significado lo permita y los esquemas actanciales puedan compatibilizarse. La relativa flexibilidad del sistema de representación sintáctica facilita estas fusiones pues su resultado puede quedar reflejado en el plano valencial mediante indicaciones de opcionalidad de los actantes o de alternancias entre ellos, como queda ilustrado en el caso de *suponer* 2 (*supra* Tabla 2).

A pesar de que la información reunida en el DEA, fundamentada en una amplia base documental, parece ofrecer sustento suficiente para un diccionario de valencias representativo de los usos verbales del español, el hecho es que este no es propiamente un diccionario basado en corpus^{iv}, ni tiene como objetivo central la caracterización detallada de la valencia de los verbos, por lo cual en el proceso de elaboración de las entradas del DCVVEA tomamos desde un principio la decisión de sustentar las descripciones en los ejemplos de un corpus textual.

Tal opción no implica que nuestro proyecto lexicográfico pueda considerarse realmente “basado en corpus”. De hecho su carácter aplicado y su finalidad didáctica, que conlleva una cierta simplificación e idealización de las descripciones para facilitar al usuario el acceso a la información de una forma cómoda, nos obligan a optar por una metodología ecléctica en la que la adecuación descriptiva a los datos de uso debe combinarse con las ventajas prácticas de un marco teórico que exige una clasificación formal de las posibilidades constructivas de los verbos. Por otra parte, a diferencia de otros repertorios valenciales, el DCVVEA estructura sus entradas en acepciones que asocian una definición de significado a un esquema valencial, con el objetivo de simplificar la identificación de las equivalencias interlingüísticas.

Con todo, la información que ofrecen los corpus resulta imprescindible para garantizar la adecuación descriptiva del diccionario. Por una parte, los corpus nos enfrentan a la continuidad dominante en el uso, una continuidad que queda oculta en las categorizaciones discretas (y necesariamente artificiales) de las descripciones gramaticales y léxicas. De ahí que tantas veces resulte problemático clasificar en una u otra acepción de significado un uso registrado en el corpus.

Las acepciones 1 y 2 del DEA del verbo *mantener* se definen de la siguiente manera:

1 Alimentar o proporcionar alimento [a alguien (*cd*)] durante más o menos tiempo. *Tb fig.* | * Sus padres la mantienen, pero ella se paga lo demás.

2 Proporcionar o costear [a alguien (*cd*)] todo cuanto precisa para vivir. | J. L. Montoya *Sev* 1.11.75, 11: El que no pueda mantener hijos, que no los traiga.

En nuestro diccionario, sin embargo, hemos optado por unir las en una sola, a falta de criterios lexicográficos que permitan asignar a una u otra acepción los usos registrados. De hecho, ni siquiera las inferencias contextuales ayudan en este caso (nótese también que en DEA se ejemplifica *mantener* 1 con una secuencia inventada).

Por otra parte, aunque ningún corpus nos dará todas las posibilidades de significado y realización valencial de cada verbo, su empleo proporciona un panorama realista, también en términos de frecuencias, de la explotación efectiva que hacen los hablantes de los recursos lingüísticos de que disponen. Como consecuencia, se refuerza la base empírica de las descripciones y se incrementa la representatividad y utilidad del diccionario.

Pero un análisis valencial estrictamente basado en corpus plantea dificultades teóricas y prácticas que pueden entrar en conflicto con la consecución de unos objetivos descriptivos de carácter aplicado. En primer término surge el problema de la extracción de los datos sintácticos, semánticos y léxicos precisos a partir de muestras textuales no analizadas sintácticamente, y en muchos casos tampoco lematizadas. La identificación de los esquemas actanciales de los verbos en un corpus no etiquetado, aparte de ser muy laboriosa, está sometida a una constante toma de decisiones de categorización sintáctica, especialmente funcional, que a su vez requiere la manipulación de las expresiones textuales por parte del analista. Dicho en otras palabras, la estructura valencial no se revela de modo transparente en el corpus sino que es el lingüista el que propone y aplica los criterios –sus criterios– de identificación.

En cuanto a la necesidad de tener en cuenta todos los ejemplares (*token*) registrados en el corpus para, entre otras cosas, valorar mediante la frecuencia relativa el nivel de uso de cada tipo de esquema valencial (*type*), no parece estar entre las prioridades de la teoría valencial, que aspira a ofrecer un repertorio de posibilidades constructivas más que un registro fiel de las tendencias de uso de dichas posibilidades en una muestra, por otra parte necesariamente limitada y sólo parcialmente representativa de una lengua.

En todo caso, el dilema no se ha planteado en esos términos para el DCVVEA, puesto que la incorporación de los datos de corpus es posible gracias a la *Base de Datos Sintácticos* (BDS) ya mencionada, que nos permite contar con los datos sintácticos y la frecuencia de los esquemas funcionales de cada verbo registrado en el corpus ARTHUS. Sin entrar en los detalles, baste decir que la BDS proporciona los rasgos más relevantes para la descripción valencial de las entradas verbales del diccionario pues contiene indicaciones sobre (i) la voz de la cláusula –activa, media (con subcategorías) o pasiva–; (ii) las funciones sintácticas de los actantes centrales; (iii) el tipo de unidad que representa a cada actante; (iv) la clase semántica general de cada elemento funcional (animado, inanimado...). A modo de ejemplo, véase la información sobre esquemas correspondiente al verbo *suponer* en la tabla 2:

Suponer [297 casos; 10 esquemas; 32 subesquemas]

VOZ	ESQUEMA	Nº	%
Activa	S	19	6.40%
Activa	SD	222	74.75%
Activa	SD PD	5	1.68%
Activa	SDI	9	3.03%
Media	S	10	3.37%
Media	S SP	1	0.34%
Media	SD	7	2.36%
Pasiva	S	16	5.39%
Pasiva	S PS	4	1.35%
Pasiva	S I	4	1.35%

Tabla 2. Los esquemas de *suponer* en la BDS (S= sujeto; D= complemento directo; PD= predicativo del complemento directo; I= complemento indirecto; SP= predicativo del sujeto).

Otra de las opciones de búsqueda de la aplicación en línea da información más específica o de “subesquemas”. La tabla 3 contiene los correspondientes a uno de los esquemas de *suponer 2*:

Activa SDI (F=9; 3.03% s/verbo)

San DnanInan (F=1; 11.11% s/esq.)

SnanDcf Ian (F=1; 11.11% s/esq.)

SnanDnanIan (F=7; 77.78% s/esq.)

Tabla 3. Los subesquemas que realizan la forma activa del esquema SDI de *suponer 2*. Abreviaturas no incluidas en la tabla 1: an= animado; nan= no animado; cf: cláusula de infinitivo.

Además de registrar la configuración sintáctica de los esquemas valenciales realizados en el corpus, la BDS ofrece la ventaja añadida de permitir el acceso inmediato a los ejemplos textuales de cada esquema. Esta opción es especialmente útil en la elaboración del DCVVEA, porque la conexión entre el español y el alemán se apoya en la traducción de los ejemplos del corpus.

En cuanto a los datos sobre frecuencia de uso, aunque no se refleja directamente en las entradas del diccionario, sí se emplea en su elaboración, ya que no sólo permite valorar la rentabilidad de uso de los diferentes esquemas, sino que también orienta la selección de los ejemplos que se incluyen para cada acepción, garantizando que no faltarán muestras de los que resulten más comunes en el corpus^v. La utilidad empírica de la BDS se completa con la posibilidad de acceder a los textos completos del corpus (no incluida en la interfaz de internet) para así considerar la información contextual tanto en el análisis de las cláusulas como en su traducción al alemán.

Con todo, el aprovechamiento de la BDS como fuente de datos para la redacción del diccionario de valencias no sería posible si el modelo teórico-descriptivo en que está basada no fuese compatible con la teoría de valencias. Aunque hay ciertas diferencias en las categorías analíticas funcionales, formales y semánticas, hay divergencias terminológicas y se aplican distintos criterios de análisis en algunos casos, la convergencia de ambos

planteamientos sintácticos en los aspectos esenciales del análisis facilita el transvase de información de la BDS al diccionario.

La BDS, no obstante, como fuente de datos para el DCVVEA presenta ciertas limitaciones derivadas, por una parte, de su alcance básicamente sintáctico, y por otra, de la extensión del corpus analizado.

Al no establecer acepciones de significado para los verbos, la única vía de clasificación y organización de los datos en la BDS viene dada por los esquemas y subesquemas. Sin embargo sabemos que las definiciones semánticas facilitan al usuario la tarea de consulta del diccionario (recuérdese lo dicho anteriormente sobre la accesibilidad cognitiva del significado léxico), y en el caso particular del DCVVEA son relevantes metodológicamente en la clasificación de los ejemplos del corpus, en su traducción y en la identificación de los equivalentes alemanes. Las consecuencias de manejar una clasificación estrictamente sintáctica pueden ilustrarse con unos ejemplos del verbo *suponer*. Atendiendo a los esquemas que presentan, los ejemplos (1-3) de la BDS son del mismo tipo, SD, y presentan igualmente un complemento directo de realización oracional:

- (1) La tan magnificada creación histórico-político-cultural romana supuso, entre otras cosas, que, [...], César llegase a matar hasta un millón de galos. RATON: 217, 3
- (2) [...] supongamos que fuera un genio... BAIRES: 457, 8
- (3) ¡Supón que llegan dos coches, con amigos del de la UVI! ¡Con cadenas, con navajas, con botellas rotas...! OCHENTA: 27, 18

Pero ejemplifican acepciones distintas en el DCVVEA: (1) es un ejemplo de *suponer 2* ('Implicar, llevar consigo o significar'), mientras que (2) y (3) lo son de *suponer 5* ('Plantear [algo] hipotéticamente como base de un argumento posterior'), una diferencia que se plasma asimismo en los distintos equivalentes alemanes de uno y otros ejemplos, *bedeuten* para (1) y *annehmen* para (2) y (3):

- (1') Die so gelobte historisch-politisch-kulturelle Leistung Roms bedeutete unter anderem, dass Cäsar bis zu einer Million Gallier töten würde.
- (2') Nehmen wir an, er ist ein Genie ...
- (3') Nimm mal an, dass zwei Wagen kommen, mit Freunden von dem von der Intensivstation! Mit Ketten, mit Messern, mit abgeschlagenen Flaschen ...!

En la actualidad un equipo de la Universidad de Vigo dirigido por José M^a García-Miguel está incorporando, en el marco del proyecto ADESSE (<http://webs.uvigo.es/adesse/>), la información semántica a los datos de la BDS. Se incluyen acepciones, clases semánticas verbales y roles semánticos actanciales, lo cual sin duda facilitará en el futuro la elaboración de descripciones contrastivas como la que estamos llevando a cabo. Aunque con una base textual diferente, restringida a textos de *El Periódico*, y una concepción diferente, el proyecto SENSEM (<http://grial.uab.es/search/>), desarrollado por el grupo interuniversitario catalán GRIAL, ofrece datos semánticos sobre los 250 verbos más frecuentes del corpus que son de utilidad en nuestro trabajo.

La segunda limitación que mencionamos más arriba, el tamaño del corpus analizado en la BDS (1.449.005 palabras), reduce la representatividad de los datos registrados, sobre todo para verbos menos frecuentes en el uso. Incluso en la relación de los 200 verbos más frecuentes nos encontramos con bastantes casos en los que nuestro corpus no ofrece ejemplo alguno de un cierto (sub)esquema actancial o una acepción de cuya existencia no nos permiten dudar ni los diccionarios ni nuestra propia intuición como hablantes de español. En tales casos recurrimos a la búsqueda de ejemplos textuales en el CREA, que en contraste con ARTHUS-BDS, ofrece una amplia y variada muestra del español actual, pero tiene para nuestros propósitos el inconveniente de ser un corpus no etiquetado.

IV. LA ASIGNACIÓN DE EQUIVALENTES EN EL DCVVEA

Un objetivo primordial de un diccionario bilingüe o contrastivo es el de registrar las equivalencias de significado en dos lenguas emparejadas. Dado que las lexicalizaciones y las categorizaciones conceptuales son parcialmente arbitrarias, cada lengua ostenta su estructuración particular del léxico y raras veces podemos asumir una correspondencia semántica biunívoca entre dos unidades léxicas de sistemas diferentes. De ahí que las relaciones de equivalencia, que casi siempre se manifiestan en una red de divergencia y convergencia interlingüísticas entre múltiples lexemas, adquieran a menudo una gran complejidad. Un diccionario unidireccional no puede registrar estas relaciones en todas sus dimensiones y de igual forma desde la perspectiva de ambas lenguas, sino que inevitablemente representa las unidades léxicas de la lengua de llegada en función de la estructuración de la lengua de partida. En nuestro diccionario, cuya lengua de partida es el español, se asocia una lista de equivalentes alemanes –plenos o parciales– a un lema español, pero no podemos indicar de manera directa las relaciones inversas de equivalencia que puedan existir entre cada uno de estos verbos alemanes y otros verbos españoles. De la misma manera, la descripción de un verbo alemán está ligada a las acepciones de los verbos españoles a las que corresponde como equivalente y por ello fraccionada dependiendo de la estructuración del léxico español. Estas desventajas se pueden remediar sólo en parte a través de un índice de verbos alemanes que remita a las acepciones correspondientes y facilite el acceso al diccionario partiendo del alemán.

Tomar el español como lengua de partida parece entrar en conflicto con los objetivos de un diccionario dirigido a usuarios que requieren ante todo información sobre el alemán como lengua extranjera. Sin embargo, esta concepción es la más coherente si consideramos que un diccionario de valencias debe atender en primer lugar a la función activa o productiva. La dirección español – alemán es la que se adapta a la perspectiva natural de un usuario hispanohablante que quiere informarse en una situación de producción de textos y que aún no dispone de conocimientos suficientes como para encontrar las palabras con las que puede expresar sus intenciones en la lengua extranjera. En esta situación, sólo un diccionario que parte de la lengua materna le da acceso a las informaciones precisas. El DCVVEA ofrecerá una serie de equivalentes alemanes, contrastará las particularidades de la combinatoria verbal en ambas lenguas y aportará ejemplos auténticos que documentan las construcciones posibles en cada acepción.

Para determinar los equivalentes alemanes de un verbo español y para detectar el grado de sinonimia interlingüística entre ellos, nos basamos en la traducción de los ejemplos de corpus. Cuando todos los ejemplos correspondientes a un verbo español se traducen con un único equivalente alemán, se puede asumir un alto grado de equivalencia semántica entre ambos, o incluso una equivalencia plena. Sin embargo, con más frecuencia se da el caso de que la traducción de una parte de los ejemplos conduce a un determinado verbo alemán y la de otros ejemplos a otro(s) equivalente(s), lo que indica que la equivalencia es sólo parcial. Por regla general, las diferencias interlingüísticas en el significado de los verbos se reflejan en distintas propiedades de las valencias, es decir en la combinatoria sintáctica o en las restricciones de selección que el verbo impone a sus actantes, pero se indican en nuestro diccionario también mediante una definición específica que se asigna a los verbos alemanes.

No hay, por supuesto, ningún mecanismo en el proceso de la traducción que relacione de manera fija o automática un lexema verbal de la lengua de partida con otro de la lengua de llegada, y en un sentido traductológico la equivalencia no existe ni entre unidades léxicas ni entre oraciones aisladas sino que se establece únicamente en el plano textual. Las traducciones que realizamos para el diccionario no representan más que una entre varias posibilidades de reproducir en alemán el contenido de los ejemplos españoles. Pero contamos

con la ventaja de poder recurrir al contexto inmediato e incluso más amplio para precisar la interpretación de los ejemplos y llegar a unas traducciones lo más fieles posible al texto original, respetando el contenido semántico, el estilo y, en la medida de lo posible, también la estructura sintáctica, pero siempre buscando una reproducción natural, acorde con las características gramaticales, idiomáticas y pragmáticas de la lengua alemana. Con todo, cuando asociamos una lista de equivalentes -que nunca será exhaustiva- a un lema español, en realidad únicamente afirmamos que los verbos alemanes pueden expresar *en algunos contextos* el significado del verbo español.

Proceder de esta manera implica que los datos en que se sustenta la descripción de la valencia de los verbos alemanes provienen en un principio de ejemplos traducidos del español y no de textos originales alemanes. Es evidente que las traducciones siempre están en cierta medida condicionadas por la lengua de partida y, además, dependen en buena parte de la competencia lingüística del traductor que en nuestro caso coincide con el analista, ya que sólo para una pequeña parte de nuestro corpus de referencia existen traducciones independientes (en concreto las de cuatro textos narrativos). Por estas razones resulta imprescindible contrastar las traducciones y completar los resultados del análisis con datos de corpus de textos alemanes. Recurrimos al *Deutsches Referenzkorpus* (DeReKo), el conjunto de los corpus del Instituto de la Lengua Alemana (<http://www.ids-mannheim.de/cosmas2>), y a los corpus del proyecto DWDS de la *Berlin-Brandenburgische Akademie der Wissenschaften* (<http://www.dwds.de>) para respaldar y precisar la descripción valencial de los verbos alemanes a través de la integración de datos originales.

REFERENCIAS BIBLIOGRÁFICAS

- Busse, W. & Dubost, J.-P. (1977). *Französisches Verblexikon. Die Konstruktion der Verben im Französischen*. Stuttgart: Klett.
- Blumenthal, P. & Rovere, G. (1998). *PONS. Wörterbuch der italienischen Verben. Konstruktionen, Bedeutungen, Übersetzungen*. Stuttgart: Klett.
- Engel, U. & Schumacher, H. (1976). *Kleines Valenzlexikon deutscher Verben*. Tübinga: Narr.
- Gries, S. T. (2006). Introduction. En S. T. Gries & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics. Corpus-based Approaches to Syntax and Lexis*. Berlín/New York: Mouton de Gruyter, pp. 1-17.
- Helbig, G. & Schenkel, W. (1969). *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig: VEB Bibliographisches Institut.
- Herbst, T., Heath, D., Roe, I. & Götz, D. (2004). *A Valency Dictionary of English: a corpus-based analysis of the complementation patterns of English verbs, nouns and adjectives*. Berlín: Mouton de Gruyter.
- Itkonen, Esa (2008). The central role of normativity in language and linguistics. En J. Zlatev, T. P. Racine, C. Shina y E. Itkonen (Eds.), *The Shared Mind. Perspectives on intersubjectivity*. Amsterdam: John Benjamins, pp. 279-305.
- Rall, D., Rall, M. & Zorilla, O. (1980). *Diccionario de valencias verbales: alemán-español*. Tübinga: Gunter Narr.
- Schumacher, H. (2003). Kontrastive zweisprachige Valenzwörterbücher. En V. Agel (Ed.), *Dependenz und Valenz: ein internationales Handbuch der zeitgenössischen Forschung*. Berlín: de Gruyter, pp. 1435 – 1446.
- Schumacher, H. (Ed.) (1986). *Verben in Feldern. Valenzwörterbuch zur Syntax und Semantik deutscher Verben*. Berlín: de Gruyter.
- Schumacher, H., Kubczak J., Schmidt, R. & de Ruiter, V. (2004). *VALBU - Valenzwörterbuch deutscher Verben*. Tübinga: Narr.

- Seco, M., de Andrés, Olimpia & Ramos, G. (1999). *Diccionario del español actual*. Madrid: Santillana-Aguilar.
- Talmy, L. (2007). Foreword. In Gonzalez-Marquez, M., Mittelberg, *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, pp. xi–xxi.

ⁱ Este proyecto ha sido subvencionado por la Dirección Xeral de I+D+I de la Xunta de Galicia (código PGIDIT06PXIB204123PR) y por el Ministerio de Ciencia e Innovación, parcialmente con fondos FEDER (HUM2006-05776/FILO). El equipo que elabora el proyecto está formado por M^a José Domínguez Vázquez, Ulrich Engel, Barbara Lübke, Meike Meliss, Mónica Mirazo Balsa, Gemma Paredes Suárez, Alejandro Pastor Lara, Pia Sloth Poulsen, Victoria Vázquez Rozas y Martina Silva López, y cuenta con las valiosas contribuciones de Guillermo Rojo y Dietrich Rall.

ⁱⁱ Para información detallada al respecto véase Schumacher (2003).

ⁱⁱⁱ La noción de normatividad y su carácter intersubjetivo es analizada en Itkonen (2008).

^{iv} La diferencia radica en que en los estudios basados en corpus (i.e., “corpus-based” y “corpus-driven”), “the analysis is, or at least attempts to be, systematic and exhaustive, meaning that the corpus does not simply serve as a database of examples from which some can be chosen *ad libitum* and others neglected, but that the whole (sample of the) corpus is taken into consideration so that even less frequent patterns must somehow be integrated or at least addressed.” (Gries 2006: 4).

^v Hay que decir también que las frecuencia totales de los verbos registrados en la BDS fueron determinantes para elaborar la lista de los que, por su alta frecuencia, formarían parte del diccionario.