

Lo fragmentario y lo global: una propuesta metodológica de corte holístico para la investigación en lingüística¹

JOSÉ ANTONIO CALAÑAS CONTINENTE

Universitat de València

Recibido: 3 de julio de 2009

Aceptado: 4 de octubre de 2009

RESUMEN

Los estudios sobre lenguas han experimentado una evolución similar a la de todas las ciencias: en los inicios, quizá por desconocimiento del objeto de estudio, se mantiene una visión de conjunto que va dejando paso a una especialización progresiva que permite una mayor profundidad y alcance de las conclusiones a las que se llega. Tras décadas de especialización progresiva en la investigación filológica pensamos que puede resultar interesante volver a una de la lengua en su conjunto, ya que la pérdida de esa visión holística lleva en ocasiones a formular teorías con un nivel de abstracción tan alto que ocasiona incluso errores en la interpretación del fenómeno lingüístico que se esté estudiando. A modo de ejemplo presentaremos el proyecto VerbLexGen en el que se propone una perspectiva integradora de corte holístico para estructurar un diccionario especializado de verbos con proverbios.

Palabras clave: Metodología, holístico.

*Holistic outlook or specialisation? Lexicogenic processes, semantics
and syntax in the VerbLexGen project*

ABSTRACT

Studies on languages have undergone a similar evolution to that of all the sciences: in the early days there is an overall view that gradually gives way to progressive specialisation, enabling more in-depth and far-reaching conclusions to be drawn. After decades of growing specialisation in philological research we think it may be interesting to return to an overall view of the language, as this loss of holistic perspective can sometimes lead to the formula-

¹ Este artículo es parte de la investigación realizada en el proyecto I+D+i GV/2007/171 titulado “Verbos alemanes con proverbio: Diseño e implementación de una base de datos para el desarrollo de herramientas léxicas alemán-catalán-castellano” financiado por la Generalitat Valenciana.

tion of theories with such a high degree of abstraction that they can even cause errors in the interpretation of the linguistic phenomenon being studied. As an example we will present the VerbLexGen, which puts forward an all-embracing holistic type of view for structuring a specialised verb dictionary with proverbs.

Key words: methodology, holistic.

1. Introducción

La búsqueda del conocimiento ha adoptado diversas metodologías desde que el ser humano adquiere consciencia y reflexiona sobre la realidad que le rodea. La necesidad adquirida con el paso del tiempo de analizar y comprender, de explicar de manera objetiva el funcionamiento del mundo ha motivado que el ser humano busque herramientas descriptivas cada vez más potentes contribuyendo a que, por un lado, se plantee reflexionar sobre realidades de complejidad creciente y, por otro, a que las herramientas descriptivas hayan ido adquiriendo una precisión cada vez mayor.

En un primer momento, el ser humano contempla la realidad de manera global y, ante la carencia de un instrumental explicativo adecuado, recurre a los mitos sublimando así su propia incapacidad. Conforme se evoluciona, se llega a la conclusión de que la observación de fenómenos complejos en su totalidad dificulta la explicación objetiva, abriendo así el camino a la segmentación de la realidad como recurso a la hora de comprender y desentrañar los misterios del entorno. Esa segmentación se hace cada vez mayor: ya no sólo sabemos que el agua se da en la naturaleza en estado líquido, sólido y gaseoso: sabemos también en qué condiciones cambia de un estado a otro, sabemos que está compuesta por dos elementos, hidrógeno y oxígeno, que estos dos elementos están en una proporción de un átomo de oxígeno por cada dos de hidrógeno, que esos átomos se unen en su orbital externo para compartir electrones y alcanzar el equilibrio de carga... Es decir, se ha pasado de contemplar la naturaleza a ver la relación entre tres entidades diferentes, se llega a conocer las leyes que regulan el aspecto del agua, se continúa segmentando para entender la naturaleza molecular, de ésta se llega al nivel de los átomos y de ahí al subatómico.

En el estudio de las lenguas se ha vivido una evolución similar: de las obras literarias y de la historia de las lenguas en sí se pasa a ir desentrañando el funcionamiento de estas en todas sus dimensiones. Con este fin se ha ido dotando a la lingüística de herramientas cada vez más potentes para poder lograr así la formulación de descripciones e incluso de leyes de alto potencial predictivo. Este proceso también ha llevado desde la contemplación del fenómeno completo hacia la segmentación del objeto de estudio. El cénit de ese proceso de segmentación se puede ver en el apogeo del estructuralismo, donde se describe la producción de cada uno de los sonidos o donde se describen, por ejemplo, los morfemas sin tener en cuenta el valor de la enunciación completa.

En este artículo queremos hacer una propuesta de descripción –y de codificación o representación– de un segmento de la realidad de la lengua alemana sin perder de vista la lengua en su conjunto: se plantea la construcción de un diccionario de verbos clasificados ateniéndonos a un criterio lexicogénico, pero sin dejar de lado ni su relación semántica ni la interficie morfológica-sintaxis-semántica. Reivindicamos con ello la reconsideración de la lengua en su conjunto para que las explicaciones tengan una validez mayor.

2. La investigación en Lingüística

La Lingüística no es la única disciplina científica que se ocupa de la lengua. En su manual de Lingüística, Linke / Nussbaumer / Portmann (⁵2004: 2 ss.)² nos hacen ver que existen tres tipos de reflexión sobre el lenguaje a la hora de considerarlo objeto científico: la reflexión sobre el lenguaje como fenómeno individual, la reflexión sobre el lenguaje escrito como forma fijada de ese fenómeno y la reflexión hecha sobre el lenguaje desde otras disciplinas.

La reflexión primaria sobre el lenguaje se centra en este en su calidad de fenómeno inherente a la naturaleza humana. El lenguaje como fenómeno y las lenguas como manifestaciones individuales de este son consideradas como la expresión más genuina de la identidad humana y atrae de este modo la atención de los estudiosos.

En el nivel siguiente, nos encontramos con la lengua escrita. La escritura es al mismo tiempo resultado y motivo de la reflexión descrita en el párrafo anterior: el fenómeno del lenguaje lleva a la búsqueda de una manera de fijar lo que se transmite y la reflexión se produce generalmente a partir del lenguaje ya fijado, es decir, puesto por escrito. Recuérdese que la investigación sobre las manifestaciones orales de las distintas lenguas es relativamente reciente y su auge va en paralelo con el perfeccionamiento de los distintos sistemas de grabación y proceso del sonido.

La omnipresencia del lenguaje –de las lenguas– en el mundo se refleja también en su presencia como objeto o como herramienta en otras disciplinas científicas. La reflexión sobre el lenguaje no se limita a la gramaticografía, si bien las gramáticas constituyen el producto más inmediato de esa reflexión. Esa aproximación científica al lenguaje se puede hacer desde la Filosofía (investigación sobre el significado), la Teología (exégesis e interpretación de la tradición bíblica), la Filología (diacronía y estética de las manifestaciones literarias), Derecho (interpretación de jurisprudencia), Psicología, Antropología y Biología (cambios evolutivos en el ser humano y comprensión de los procesos de aprendizaje de lenguas) y en Matemáticas e Informática (lenguajes simbólicos, generación de lenguaje natural, etc.).

Además del problema (común a todas las disciplinas científicas) de acotar el objeto de estudio, el lingüista ha de enfrentarse a un segundo nivel de decisión tras optar por las lenguas. Tras este paso debe decidir entre el estudio de la lengua en su manifestación oral o en la escrita. Si bien en un primer momento de los estudios lin-

² LINKE, A. / NUSSBAUMER, M. / PORTMANN, P.R.: *Studienbuch Linguistik*. Tübingen: Niemeyer ⁵2004.

güísticos la opción por la lengua escrita estaba clara (era la única que se podía fijar para diseccionarla), en la actualidad cobra cada vez más importancia el estudio de la lengua oral. Esta perspectiva resulta aún más necesaria si se tiene en cuenta que las lenguas son, en primer lugar, un instrumento del ser humano para la comunicación interpersonal, que es en primer término hablada. El código escrito responde sólo a una convención posterior cuya necesidad surge, precisamente, del deseo de dar a lo que se hablaba una forma constante y reconocible en el tiempo, fuera de la inmediatez de la situación comunicativa inicial.

3. Enfoques metodológicos

La consolidación de cualquier disciplina como ciencia (en el sentido tradicional de ciencia positiva) requiere no sólo acotar el objeto de estudio, sino también adoptar un método de investigación cuya aplicación pueda hacerse de manera sistemática. En la historia de la ciencia encontramos una polarización entre los métodos hipotético-deductivos y los inductivos; se considera que la ciencia moderna empieza cuando se pasa del método inductivo, que deriva reglas generales a partir de la observación de fenómenos aislados (procede desde lo particular a lo general) al hipotético-deductivo, que pasa de lo general, la hipótesis, a los casos particulares (y procede, por tanto, en el sentido inverso).

Las diferentes escuelas lingüísticas han seguido también este patrón de comportamiento. El método inductivo se reconoce fácilmente en los primeros estudios de historia de las lenguas que se conocen, ya que parten de los textos que han llegado hasta nuestros días para así formular reglas que explican la génesis de las lenguas en la actualidad. Con el avance de los estudios y con el intento de lograr conclusiones de una amplitud cada vez mayor, se pasa a formular hipótesis que permitan un alto valor de aplicación. Hoy en día se intenta partir de la hipótesis y desde ahí se inicia la búsqueda de ejemplos, de muestras de lengua real que confirmen o revoquen esa hipótesis inicial. Junto a estos cambios de método se observa también una tendencia a la hipersegmentación del objeto de estudio: de las lenguas completas objeto de la genética lingüística se pasa a unidades mínimas en el estructuralismo (fonemas, morfemas, etc.), volviendo poco a poco a unidades mayores (sintagma, oración, texto).

Si bien la especialización contribuye a que sea posible llegar a una profundidad apreciable en el saber, no debe perderse de vista que esa profundidad de campo reduce sensiblemente la amplitud de la perspectiva y, de ese modo, se deja de describir, de ver la vertebración de todas las unidades que componen el sistema. El significado no surge en la palabra, sino en el uso que se hace de ella, un uso que no es posible de manera aislada, sino que ha de producirse en combinación con otros elementos, con una sintaxis concreta, una entonación adecuada, un contexto productivo, etc.

Proponemos, pues, desde este artículo volver a una perspectiva más amplia intentando mantener una profundidad suficiente. Los numerosos fenómenos que intervienen en la comunicación humana se producen de manera simultánea y ejer-

cen una influencia mutua sin perder por ello su efectividad. Tratemos, pues, de iniciar una descripción lingüística sin segmentar excesivamente, sin compartimentar en subdisciplinas de manera limitadora. Aun así, optar por una perspectiva holística no es óbice para decidir el punto inicial de la descripción: lo importante aquí es el resultado final, y el significado verbal que surge de los procesos lexicogénicos que afectan al verbo depende no sólo de la semántica verbal o de los preverbios que se utilicen, sino también de la sintaxis oracional, los patrones de complementación verbales, la entonación, etc.

4. El proyecto *VerbLexGen*

En estas coordenadas holísticas se desarrolla el trabajo del equipo investigador del proyecto *VerbLexGen*, compuesto por investigadores de las Universidades de Valencia, Valladolid, Complutense de Madrid y Politécnica de Valencia. El acrónimo intenta dar ya una idea de cuál es el objetivo final del trabajo: la compilación en un sistema informático de un catálogo de *verbos* alemanes ordenados en función de criterios *lexicogénicos* con el que se puedan desarrollar herramientas lexicográficas multilingües orientadas a la investigación contrastiva en formación de palabras y a la explotación de los datos en el campo de la traducción o la didáctica así como en sistemas informatizados de generación y recuperación de lenguaje natural. A continuación detallaremos algunos de los ejes en los que se desarrollará la labor del equipo.

a. *Filosofía del proyecto*

Aun tratándose de un proyecto eminentemente lexicográfico, el punto de partida radica en la concepción de la lengua como unidad: la naturaleza del significado consiste en la combinación de múltiples factores que determinan el punto concreto del espacio semántico, que se activa de modo similar a un sistema de coordenadas que sirve para ubicar un punto en el espacio. Surge así una imagen en la que el significado es la combinación de parámetros a lo largo de un eje paradigmático, uno sintagmático y un tercero de tipo cognitivo que incorpora la información sobre posibles usos metafóricos. Las coordenadas del significado se complementan con la información de tipo pragmático para configurar una definición lo más completa posible que requiere un aparato descriptivo de gran simplicidad³.

Si bien no es este artículo el lugar más adecuado para discutir la relevancia de diccionarios y otros productos lexicográficos, la función de estos se encuentra muy presente en la investigación previa a *VerbLexGen*⁴. Sí debe resaltarse el hecho de

³ Cfr. CALAÑAS CONTINENTE, J.A.: *El dominio léxico 'Existencia' en alemán*. Frankfurt a. M. et al.: Peter Lang 2002 (en especial el capítulo 2).

⁴ *Ibíd.*, capítulo 1.

que el diccionario/lexicón se concibe como el lugar “físico” en el que se sistematiza todo el conocimiento necesario para poder hacer un uso adecuado de la lengua. Un diccionario enriquecido de este modo pone a disposición del usuario información enciclopédica, léxica, sintáctica, pragmática, morfológica, de manera que cada entrada constituye en sí misma una pequeña gramática léxica.

Siguiendo nuestro planteamiento de no tratar cada aspecto de la lengua en un compartimento estanco diferente, parece pertinente comenzar la compilación de este tipo de compendios léxicos por la clase lexemática del verbo. El verbo es la unidad en torno a la cual se vertebra la totalidad de la enunciación: será este elemento el que regule la cantidad de participantes en la acción, el orden en que deben aparecer, etc. Parece, pues, pertinente acometer la compilación de un diccionario de verbos que detalle toda esta información.

En el caso de la lengua alemana, además, no resulta fácil pasar por alto la enorme relevancia que tienen los mecanismos lexicogénicos en el proceso de etiquetado de la realidad. A diferencia del castellano, que dispone de un número mayor de lexemas verbales, el alemán cubre sus necesidades de expresión de realidades verbales con un número relativamente pequeño de lexemas independientes que son sometidos a distintos procesos de afijación, derivación, etc. hasta lograr una precisión semántica altísima: es conocido el hecho de que un verbo, que llamaremos *Basisverb*, se puede combinar con una cantidad variable de preverbios, separables o no, hasta lograr esa precisión de significado⁵. Es más, en cualquier caso, el hablante tiene la posibilidad de utilizar esos mecanismos para expresar con exactitud y sin excesivos circunloquios lo que quiere decir en la situación comunicativa en que se encuentre.

Todas estas reflexiones se unen al interés indiscutible del estudio contrastivo de lenguas para configurar el marco en el que se desarrolla VerbLexGen. A continuación comentaremos los objetivos principales del proyecto.

b. *Objetivos*

El proyecto VerbLexGen tiene tres objetivos fundamentales, todos ellos, como es lógico, relacionados entre sí y encaminados a obtener un corpus de verbos organizado de forma que constituya un punto de partida para otros trabajos de investigación. Los objetivos en esta primera fase son estos:

- Implementación de una base de datos léxica con verbos de las lenguas alemana, castellana y catalana. Esta herramienta se construirá a partir de software

⁵ Vid. CALAÑAS CONTINENTE, J. A.: “Wortbildung an der Schnittstelle von Syntax und Semantik: Rahmenbedingungen für eine lexikologisch-lexikografische Aufgabenstellung”, en AICHINGER, L. M. / MELISS, M. / DOMÍNGUEZ VÁZQUEZ, M.J. (eds.): *Wortbildung heute: Tendenzen und Kontraste der deutschen Gegenwartssprache*. Tübingen: Gunter Narr (=Studien zur Deutschen Sprache. Forschungen des IDS Bd. 44) 2008, 61-73.

comercial⁶ de manera que se puedan realizar búsquedas atendiendo a cualquier criterio que necesite el investigador. La base de datos se construye a partir de dos conjuntos limitados: el de preverbios separables y no separables y el de verbos monolexemáticos. Cada registro de preverbios incluirá todas las acepciones disponibles numeradas para poder definir con exactitud la aportación de este prefijo al significado global del nuevo verbo. Los verbos monolexemáticos con los que se trabajará en esta primera fase proceden del léxico básico del alemán, ya que partiremos de esta lengua para realizar la totalidad del trabajo. Como resultado se dispondrá de una ventana en la que se nos llevará a todas las combinaciones posibles de este verbo base con los preverbios lexicalizados y se dotarán las entradas de la plantilla léxica correspondiente, una herramienta con la que se representará la interficie sintaxis-semántica que puede emplearse para la creación de aplicaciones informáticas de análisis y generación de lenguaje natural⁷.

- Compilación de un diccionario/lexicón trilingüe de verbos. Partiendo de un primer listado de verbos alemanes, se buscarán las equivalencias exactas en castellano y catalán como paso previo al estudio contrastivo de los mecanismos que se utilizan en las lenguas de nuestro estudio para responder a las necesidades de etiquetado del significado verbal.
- Este paso es consecuencia del anterior: catalogación de los medios disponibles en castellano y catalán para realizar las funciones lexicogénicas propias de los preverbios con que se combinan los verbos alemanes objeto de estudio.

Hay que resaltar que este proyecto comienza con un número bastante reducido de lenguas, todas ellas indoeuropeas, por cuestiones de índole práctica. Sin embargo, la ampliación a otras lenguas de familias diferentes queda abierta y constituirá el objetivo de fases posteriores de puesta en valor y de mejora del sistema de consulta y búsqueda de equivalencias de VerbLexGen.

c. Algunos ejemplos

En esta sección ejemplificaremos los distintos campos que contendrá la base de datos sobre la que se construirán las aplicaciones de VerbLexGen. En la figura 1 tenemos la tabla en la que se recogen las posibles aportaciones semánticas de cada uno de los preverbios al significado total del verbo compuesto. En la figura se puede observar la entrada correspondiente a «*ab-*». Junto con las acepciones también se refleja si el preverbo forma un verbo separable o no, algo que creemos que es relevante para la sistematicidad en la creación de significado.

⁶ Aunque se están valorando otras posibilidades, por el momento se está trabajando en una aplicación hecha sobre la plataforma FileMakerPro® que permita su ejecución tanto en las plataformas más habituales sin necesidad de instalar programas adicionales.

⁷ Vid. MAIRAL USÓN, R.: "Why the notion of lexical template?", en FERRER MORA, H. / CALAÑAS CONTINENTE, J.A. / PAVÓN VÁZQUEZ, V. (eds.), *Anglogermanica Online* 2002-1 (2002), 51-68. [Http://anglogermanica.uv.es](http://anglogermanica.uv.es) [última consulta: 27.08.2007].

The screenshot shows the FileMaker Pro interface with a table named 'VerbLexGen1.fp5'. The table has three columns: 'N° Pres...', 'Verbzusatz', 'Trennbar', and 'Bedeutung'. The data is as follows:

N° Pres...	Verbzusatz	Trennbar	Bedeutung
	ab 1	ja	jmd./etw. entfernt sich von einem Ort
	ab 2	ja	etw. (meist mit einem Werkzeug) wird von etw. getrennt
	ab 3	ja	etw. wird von einem Gegenstand entfernt
	ab 4	ja	Mit einem Gegenstand wird etwas gemacht, um etwas zu entfernen
	ab 5	ja	die Funktion von etw. wird vorübergehend beendet
	ab 6	ja	ein Vorbild oder Original wird imitiert
	ab 7	ja	Es geschieht das Gegenteil von der Bedeutung des Basisverbs

Figura 1. Acepciones del preverbio *ab-*

La productividad de los mecanismos lexicogénicos en la categoría verbal se ejemplifican a continuación con los derivados que se generan a partir de *schlagen*. Nótese que los verbos marcados con un asterisco son en realidad dos, ya que el preverbio puede ser separable o no:

abschlagen	herunterschlagen
anschlagen	hinschlagen
aufschlagen	hineinschlagen
ausschlagen	kaputtschlagen
beschlagen	losschlagen
breitschlagen	nachschlagen
dazuschlagen	niederschlagen
dazwischenschlagen	totschlagen
durchschlagen	überschlagen*
einschlagen	umschlagen*
entgegenschlagen	unterschlagen*
entschlagen	verschlagen
entzweischlagen	vorschlagen
erschlagen	wegschlagen
fehlschlagen	zerschlagen
herabschlagen	zurückschlagen
herumschlagen	zusammenschlagen

En la base de datos que se quiere implementar se recogerá el significado del verbo monolexemático que se constituirá en *definiendum* de los verbos compuestos. Esta propuesta de definición es coherente con los mecanismos descriptivos del modelo lexemático-funcional en el que los predicados de significado más específi-

co se definen en función de los archilexemas⁸. En este caso, se explorará en qué medida permanece el significado del verbo base en los derivados y se expresará en tablas que también incluirán la equivalencia en las distintas lenguas que se incluyan en la base de datos.

Verbzusatz	tr./untr.	Bedeutung	Spanisch
ab	ja	resultativ	separar a golpes, talar
an	ja	direktiv	clavar
be	nein	übertragen	herrar, guarnecer

Con el fin de permitir un uso informatizado del corpus que genera la base de datos VerbLexGen, se incluirá en un momento posterior la codificación en plantillas léxicas de los significados de los diferentes verbos. La plantilla léxica de *hämmern*, a título de ejemplo, sería la siguiente:

hämmern: [[**do'** (w, [**use.tool.**(Hammer).**in.**(b).**manner.for.**(d)' (w, x)] CAUSE [**do'** (x, [**move.toward'** (x, y) & INGR **be.in.contact.with'** (y, x)], a = x

Figura 2. **Plantilla léxica de *Hämmern***

5. Conclusión

En este artículo hemos iniciado una reflexión sobre las ventajas e inconvenientes de la segmentación de los diferentes objetos de estudio: un objeto de estudio más reducido, un subsistema, es susceptible de una descripción más detallada, el problema es que la realidad no la constituyen los subsistemas, sino el conjunto en su totalidad. Así, la mayor profundidad explicativa, el mayor potencial predictivo que se puede alcanzar concentrando la mayor intensidad en un espectro más reducido, queda empañado por la pérdida de amplitud, por el fracaso a la hora de integrar los resultados obtenidos con los subsistemas en el conjunto complejo de la realidad. Esto también es válido para la investigación en lingüística: la segmentación de la lengua en esos subsistemas nos hace perder la visión de conjunto, logrando aparatos descriptivos o predictivos de efectividad limitada.

Aun constituyendo un subsistema, el léxico merece una atención especial, ya que es el conjunto de herramientas tangibles con el que nos comunicamos. Son las piezas que nos permiten materializar la comunicación y están indisociablemente imbricadas en la concepción del mundo que da sentido a la comunicación interpersonal. En este artículo proponemos la creación de nuevas herramientas que permitan una sistematización de los mecanismos que actúan cuando se crea léxico o cuando se

⁸ Cfr. CALAÑAS CONTINENTE, 2002. Op. cit. en nota 3.

combinan unidades simples para responder a una necesidad comunicativa específica. Nuestra propuesta, el proyecto VerbLexGen, pretende ser una herramienta para caracterizar los mecanismos lexicogénicos que operan en la clase verbal con el valor añadido de ofrecer un instrumental para el análisis contrastivo interlingüístico y de contener definiciones formalizadas que describen (y hacen, en cierta medida, predecible) el comportamiento de la productividad léxica en la interfaz léxico-sintaxis.