

INFORME TÉCNICO DE PROYECTO INTERNACIONAL DE I + D:
ESTUDIO LÉXICO-TIPOLOGICO Y DESCRIPCIÓN
COMPUTACIONAL DE LOS MECANISMOS UNIVERSALES DE
LEXICOGÉNESIS EN LAS LENGUAS DEL MUNDO (2001-2003)

Lucía Luque Nadal
Universidad de Córdoba

FICHA TÉCNICA DEL PROYECTO DE I + D (2001-2003)

MODALIDAD: PROYECTO INTERNACIONAL DE I + D

Título del Proyecto: *Estudio léxico-tipológico y descripción computacional de los mecanismos universales de lexicogénesis en las lenguas del mundo*

INVESTIGADOR PRINCIPAL: Dr. Juan de Dios Luque Durán
(Universidad de Granada)

MIEMBROS DEL EQUIPO DE INVESTIGACIÓN:

Dr. Antonio Pamies Bertrán (Universidad de Granada)

Dr. Francisco José Manjón Pozas (Universidad de Granada)

Dr. José Manuel Martín Morillas (Universidad de Granada)

Dr. Antonio Hernández Fernández (Universidad de Jaén)

1. OBJETIVO PRINCIPAL

This research project aims to the analysis and study of those mechanisms that make possible word-creation and the match of new lexical elements with new concepts in the languages of the world. The core of the project stands on the isolation of a set of cognitive universal basic principles which rule the generation and creation of lexical elements. The research on this inventory of lexicogenetical keys will be useful for psychologists, educational staff, speech therapists and, of course, general linguists, comparative linguists, lexicologers, lexicographers and semanticists because it will probably help them to understand more thoroughly different aspects of the lexical level of the languages of the world.

Once the inventory of lexical keys is arranged, the second stage of the project will aim to the creation of a formal organization which enables it to be

used as a point of reference in the analysis of a certain number of lexical-genetical process. For that reason, the project includes the development of a basic computational environment -limited in this stage to a few lexemes- consisting of a database of interconnected lexical-genetical keys. This database would be used as a way of describing linguistically and cognitively any lexical-genetical process in an individual language.

2. OBJETIVOS PLANTEADOS

1) Elaboración y descripción de un inventario de claves de funcionamiento de los procesos de lexicogénesis y creación léxica en las lenguas del mundo. Este primer objetivo, que deberá desarrollarse a lo largo de los dos primeros años del proyecto, persigue crear un entramado teórico universal que permita clasificar cualquier lengua del mundo tomando como parámetro tipológico de clasificación el modo en que dicha lengua utiliza sus recursos de neogénesis léxica. Estos principios no serán de carácter estrictamente gramatical (morfosintáctico). Estos principios adoptados de la lingüística cognitiva son los que fundamentalmente se utilizarán para la descripción de las claves cognitivas profundas que organizan y mueven la creación léxica. Esto permitirá una comprensión de los procesos de lexicogénesis como procesos internos en los que mentalmente se conexionan una serie de conceptos y nociones semánticas cada vez más complejas que afloran -como una necesidad básica en la lengua- produciendo la creación de unidades léxicas nuevas.

2) Clasificación de las lenguas implicadas en el estudio en base a las claves lexicogénicas descritas. Este es uno de los puntos en los que el proyecto incide doblemente sobre un mismo aspecto. El inventario inicial de 52 lenguas que servirá como base para el estudio ha sido seleccionado por su relevancia y diversidad tipológica, lo que en principio confiere mayor validez, por su carácter más general, a las conclusiones a las que se puedan llegar.

3) Creación en el segundo y tercer año del proyecto de un entorno computacional que permita:

3.1) la presentación informática de los datos y de las claves de lexicogénesis. Esto permitirá una búsqueda múltiple, temática, semántica y formal de los diferentes procesos lexicogénicos y, por otro lado -lo que sería el objetivo más importante de esta elaboración computacional- permitir que mediante el programa finalmente se pueda de clasificar cualquier proceso lexicogénico individual o, en su globalidad, cualquier lengua según su comportamiento lexicogénico.

3.2) Eventualmente, el programa realizado, conforme mayor sea el número de lenguas y procesos lexicogénicos introducidos, debe estar preparado para una caracterización léxico-tipológica de lenguas -mediante índices cuantitativos- que nos permita una primera clasificación de las lenguas según una faceta concreta de su funcionamiento léxico: la creación de palabras.

3.3) Un objetivo adicional será que el entorno computacional tenga capacidad de elaboración lexicogénica propia, es decir, que sea capaz de generar léxico real de una lengua como muestra del potencial mayor o menor de dicha lengua en un determinado nivel de neogénesis léxica. En este proyecto, este objetivo sólo tendrá un aspecto inicial, con un número muy limitado de palabras (unas cuantas decenas). Evidentemente, para ello el programa contendrá un formulario básico y mediante la introducción de secuencias reales -palabras, morfemas- de una lengua en dicho formulario y especificando un proceso lexicogénico, el programa generará secuencias -léxicamente pertinentes o no, pero formalmente posibles- de una forma léxica determinada en una lengua determinada.

3. RESULTADOS DE INVESTIGACIÓN

1. La **creación de un corpus multilingüe** (español, francés, inglés y alemán) que cuenta en la actualidad con aprox. 250 millones de palabras. Al objeto de optimizar la eficacia de las herramientas de recuperación de información integradas, se han aplicado técnicas de indización a dos niveles al objeto de reducir el tamaño del índice sin mermar la calidad de los resultados.

El esfuerzo invertido en la ampliación del corpus está siendo muy importante, no sólo se ha incrementado el ya considerable tamaño, llegando prácticamente a duplicar el volumen inicial, sino que esta ampliación se está realizando de manera que el nuevo corpus atienda a las múltiples necesidades de consulta de la manera más eficaz y flexible posible. Esta ampliación conlleva aspectos como la reestructuración y reclasificación del ingente volumen de textos en categorías (teniendo en cuenta que esta clasificación ha de ser necesariamente flexible y fácilmente modificable), la homogeneización, i.e. la conversión a formatos universales XML, HTML y TXT fácilmente escalables y que facilitan la utilización de las herramientas de búsqueda (tanto de la herramienta de búsqueda e indización que se modificó y adaptó anteriormente, como las nuevas que necesariamente habrán de desarrollarse en futuros estadios del proyecto).

El desarrollo del corpus electrónico en ningún momento ha obviado la intención originaria del proyecto y, tal como ha sido programado, permitirá en

un futuro próximo la incorporación -que ya se está realizando paulatinamente- de lenguas de diversas familias lingüísticas y tipos lingüísticos de todo el mundo.

2. CONFERENCIAS Y PONENCIAS IMPARTIDAS POR LOS MIEMBROS DEL GRUPO DE INVESTIGACIÓN:

-Luque Durán, J. D: Profesor invitado para impartir dos conferencias en los Cursos Internacionales de la Universidad de Salamanca sobre el tema 'La variación léxico-conceptual en las lenguas del mundo'. Salamanca, 4-6 de Julio de 2001.

-Luque Durán, J. D.: 'Evolución del lenguaje desde la perspectiva de los procesos universales de lexicogénesis'. En *Congreso internacional sobre nuevas tendencias de la lingüística*. Granada, 20-22 noviembre 2001.

-Luque Durán, J. D.: "Lingüística y traducción. El estudio de las colocaciones desde un planteamiento traductológico y tipológico-lingüístico". Ponencia presentada al *II Simposio Internacional Traducción, Texto e Interferencias. La investigación y la práctica profesional de la Traducción e Interpretación en los albores del siglo XXI*. Málaga, 22, 23 y 24 de octubre de 2003

-Manjón Pozas, Francisco José (2001): La expresión léxica de la conceptualización del espacio: los sistemas de localización y deixis y la simbiosis entorno-lengua. *V Jornadas Internacionales de Lenguas y Culturas Amerindias* (Valencia, 8-12 de noviembre de 1999)

-Luque Durán, Juan de Dios (2001): "Deconstrucción y reconstrucción de la realidad: la captación y plasmación lingüística del mundo en algunas lenguas de América". *V Jornadas Internacionales de Lenguas y Culturas Amerindias* (Valencia, 8-12 de noviembre de 1999)

-Pamies Bertrán, Antonio: "Fraseología contrastiva y universales metafóricos". VII Encuentro de Hispanistas de Eslovaquia. 18-20 de noviembre 2002 . Universidad Comenius, Bratislava.

-Pamies Bertrán, Antonio: "Sémantique grammaticale de la possession dans les langues d'Europe". Journées d'études Eurom4. 13 de diciembre de 2001. Universidad de Niza.

-Pamies Bertrán, Antonio: "A relação forma-sentido na construções possessivas nas línguas do mundo", *III Congresso Internacional da Associação Brasileira de Lingüística*, Universidade Federal de Rio de Janeiro, 11-15 marzo 2003.

-Pamies Bertrán, Antonio: "Acceso automatizado a fraseologismos y colocaciones en corpus no etiquetado" *VI Congreso internacional de lingüística hispánica*. Universidad de Leipzig, 12 de octubre de 2003.

-Pamies Bertrán, Antonio: "La conceptualización de la posesión en las lenguas americanas". *Primer Simposio Internacional de Lingüística Amerindia*, Universidad Autónoma de Yucatán, 28 de febrero de 2003.

3. CONGRESOS Y SEMINARIOS ORGANIZADOS. DIFUSIÓN DE RESULTADOS DE INVESTIGACIÓN EN CONGRESOS, JORNADAS Y OTRAS REUNIONES CIENTÍFICAS:

1. *Congreso internacional sobre nuevas tendencias de la lingüística*. Universidad de Granada, 20-22 noviembre 2001.

2. *I Simposio sobre Lenguas Asiáticas*, Granada, Universidad de Granada, 10-14 de marzo de 2003.

3. *Curso de Mongol Literario*, Universidad de Granada, 10-14 de marzo de 2003.

4. RESULTADOS DE INVESTIGACIÓN (2): PUBLICACIONES

4.1 MONOGRAFÍAS

-Luque Durán, Juan de Dios (2002): *Aspectos universales y particulares del léxico de las lenguas del mundo*. Granada: Método.

La monografía del Dr. Juan de Dios Luque Durán "Aspectos universales y particulares del léxico de las lenguas del mundo" ha sido considerada por muchos autores y en muchos centros universitarios y de investigación como una obra fundamental en un campo científico emergente, como es el estudio sistemática de los lexicones de las lenguas. Para hacer más fácil su acceso a investigadores, ha sido publicada, parcialmente, en formato electrónico (acceso por Web) en la Universidad Complutense de Madrid, en el Hipertexto de Enseñanza en Lingüística de la UCM. Otras universidades y centros se interesaron por hacer algo similar. Por esta razón, el profesor Carlos Subirats Rüggeberg de la Universidad Autónoma de Barcelona propuso pedir al MCYT una subvención que permitiera poner el libro en la red, a disposición de todos los usuarios. Una vez recibida esta dotación del Ministerio se colocó el libro en la red Elies. ELiEs, como es sabido, es una revista electrónica y una base de datos que se edita en el Laboratorio de Lingüística Informática de la Universidad Autónoma de Barcelona.

Además este libro ha contado con numerosas reseñas entre las que cabe destacar:

- Reseña realizada por Julio Calvo Pérez en *LynX* n11 (2002), págs. 131-139.
- Reseña realizada por Francisco Osuna García, *Revista Española de Lingüística*, Vol.32, 1. (Enero- Junio de 2002).
- Reseña realizada por Ventura Salazar García en *ELUA* (Estudios de Lingüística, Universidad de Alicante) número 16 (2002).

4.2 ARTÍCULOS EN LIBROS DE ACTAS

- Juan de Dios Luque Durán: “Sobre el papel del lexicón en la emergencia y evolución de los lenguajes naturales”. *Language Design* 4 (2002): 113-150. ISSN: 1139-4218.
- Luque Durán, J.D. y Manjón Pozas, F. J. (2002): “Claves culturales del diseño de las lenguas: fundamentos de tipología fraseológica”. En *Novedad bibliográfica en la red: Investigaciones de vanguardia en la lingüística hispánica*. Estudios de Lingüística Española 16 (2002): <http://elies.rediris.es/elies16>. ISSN: 1139-8736.
- Antonio Pamies Bertrán: “Modelos icónicos y arquetipos: algunos problemas metalingüísticos en el ámbito de la fraseología”. *Language Design* 4 (2002): 9-20. ISSN: 1139-4218.

4.3 ARTÍCULOS EN REVISTAS

- Luque Durán, Juan de Dios (2003): “Las lenguas de África”. *El fingidor*, 19-20. *Monografía “Cultura negro-africana”*, pp. 2-5.

4.4 CAPÍTULOS DE LIBROS

- Luque Durán, Juan de Dios (2001): “Problemas de la caracterización genética y tipológica de las lenguas de África”. En García Marcos et al. (2001): *Miscelánea de estudios africanos*. Granada: Método Ediciones, pp. 95-132.
- Luque Durán, Juan de Dios (2001): “Deconstrucción y reconstrucción de la realidad: la captación y plasmación lingüística del mundo en algunas lenguas de América”. En Calvo Pérez, J. (ed.): *Contacto interlingüístico e intercultural en el mundo hispano*. Valencia: Instituto Valenciano de Lenguas y Culturas Amerindias. Universidad de Valencia, 259-288.
- Pamies Bertrán, Antonio e Iñesta Mena, E. V. (2002): *Fraseología y metáfora*. Granada: Método.

- Pamies Bertrán, Antonio (2002): "El concepto de equivalencia cognitiva en fraseología". En Sabio, J.A., Guatelli, J. & Guillén, G.: *Traductología y Lingüística Aplicada*, pp. 23-31. Granada: Método.
- Pamies Bertrán, Antonio (2001): "Sémantique cognitive, lexicologie multilingue et traduction", en Cherednychenko, O. (ed.): *Movni i kontseptual'nyj kartyny svitu*, pp. 344-352. Kiev: Universidad Taras Shevchenko.