

## La comparabilidad de los trabajos de disponibilidad léxica

### *Comparability of researchs on lexical availability*

**Florentino Paredes García**

*Universidad de Alcalá.*

florentino.paredes@uah.es

Paredes García, F. (2017). La comparabilidad de los trabajos de disponibilidad léxica. *Revista Nebrija de Lingüística Aplicada* (2017) 22/.

### **RESUMEN**

Este comentario del artículo de Jiménez Catalán contiene consideraciones acerca de la comparabilidad de los trabajos de disponibilidad léxica.

Palabras clave: disponibilidad léxica, comparabilidad, método comparativo

### **ABSTRACT**

*This commentary of Jiménez Catalán's article contains some considerations about the comparability of lexical availability works.*

*Keywords: lexical availability, comparability, comparative method*

Fecha de recepción: 27/02/2017

Fecha de aceptación: 25/03/2017

## **1. INTRODUCCIÓN**

El trabajo de Jiménez Catalán plantea algunas cuestiones relacionadas con los estudios de Disponibilidad Léxica (DL) que inciden en el valor e interés general de esta metodología de investigación cuando se consideran en conjunto las investigaciones, más allá de la utilidad de cada trabajo individualmente. Como se ve en el artículo, la autora elige un corpus de trabajos sobre léxico disponible y lo analiza comparando algunos aspectos metodológicos. Independientemente de que el conjunto de trabajos seleccionados por la autora sea heterogéneo y de que no quede claro en el artículo principal si se consideran al mismo nivel las investigaciones originales y los trabajos derivados de estas, la propuesta de la autora pone sobre el tapete uno de los asuntos que más importan a los trabajos científicos: hasta qué punto son comparables los resultados de los trabajos insertos en una misma línea de investigación. Hay que tener en cuenta que en ciencia la comparación de investigaciones, con ser útil y muy deseable, no necesariamente tiene que ser buscada por el investigador; a veces incluso las características de la investigación —por ejemplo, por iniciar una línea inexplorada o por otras razones diversas—, hacen que la comparación simplemente no sea posible.

Así pues, a reflexionar sobre la comparación y las condiciones que deben reunir las comparaciones entre trabajos de DL se van a dedicar las páginas que siguen, que han de entenderse como meras notas y apuntes al hilo del artículo citado.

---

## 2. EL MÉTODO COMPARATIVO

---

La comparación es el principal procedimiento de análisis científico, hasta el punto de que se ha afirmado que el método científico es “inevitablemente comparativo” (Grosser 1973: 19). Esta consideración ha de hacerse extensiva, por supuesto, a las humanidades y las ciencias sociales, uno de cuyos máximos exponentes llegó a identificar el método comparativo con la propia disciplina —la sociología en su caso— (Durkheim 1988). Ahora bien, siguiendo al *Diccionario Crítico de Ciencias Sociales* (Román Reyes 2009), cuando se habla de comparación en ciencia se deben diferenciar dos maneras de considerarla (Colino 2009): una de ellas consiste en entenderla en un sentido amplio, como una “actividad mental lógica, presente en multitud de situaciones de la vida humana, que consiste en observar semejanzas y diferencias en dos o más objetos”; la segunda, entendida en sentido estricto, se refiere a un método de análisis e investigación, “un procedimiento sistemático y ordenado para examinar relaciones, semejanzas y diferencias entre dos o más objetos o fenómenos, con la intención de extraer determinadas conclusiones”.

El método comparativo en las ciencias sociales es, por tanto, un modo particular de proceder que, partiendo de una determinada concepción de la realidad, trata de determinar la ocurrencia de los fenómenos, las relaciones que se establecen entre ellos, las condiciones que afectan a su realización y las causas y factores que inciden en su ocurrencia – tratando de establecer en qué medida actúan los factores individualmente o en posibles agrupaciones–; todo ello con el objetivo de explicar el fenómeno, extraer conclusiones y, a partir de ellas, establecer posibles generalizaciones basadas en las observaciones parciales. Los objetivos del análisis comparativo, por tanto, se pueden sintetizar en los siguientes puntos (Colino 2009):

- 1) la investigación de semejanzas y diferencias (variación) entre casos, 2) la observación de regularidades, que deben ser explicadas mediante 3) el descubrimiento de covariación o la interpretación de la diversidad, que suele hacerse a través de 4) la comprobación de hipótesis explicativas, que llevan a 5) la explicación de la complejidad de relaciones causales y el establecimiento de generalizaciones o pautas particulares, que permitirán finalmente 6) la generación de teorías o su refutación.

Para que entre dos o más entidades de la realidad social pueda llevarse a cabo el método comparativo, deben darse unas *condiciones de comparabilidad*, esto es unos requisitos que han de cumplir las unidades sociales y el tratamiento del análisis. Estos requisitos son los siguientes: a) las unidades sociales han de seleccionarse de forma sistemática, b) las realidades han de ser comparables en subconjuntos o totalmente (contextos homogéneos o heterogéneos), y c) ser consideradas como el contexto del análisis de la variación (semejanzas o diferencias) entre variables o relaciones.

Por último, la aplicación del método comparativo requiere un procedimiento específico, que se resume en tres fases o etapas (Gómez Díaz de León y Ayde de León 2014). La primera, consiste en determinar la estructura teórica sobre la que fundamentar las hipótesis de la investigación, definir las propiedades y características de los elementos susceptibles de comparación e identificar las variaciones y semejanzas del objeto de estudio. La segunda etapa consiste en la selección de las muestras que se van a comparar, prestando atención a que los casos sean efectivamente comparables y relevantes al menos en relación con cuatro aspectos: ámbito del objeto, contexto, tiempo y espacio. La tercera etapa consiste en el análisis de los casos a partir de la comparación de las variables, para determinar las diferencias o las semejanzas (concordancias) y extraer las conclusiones y posibles generalizaciones teóricas pertinentes.

---

## 3. LA COMPARABILIDAD DE LOS TRABAJOS DE DISPONIBILIDAD LÉXICA

---

A partir de lo expuesto en el apartado anterior, surge la pregunta de qué tipo de comparaciones permiten los trabajos de DL. La cantidad y variedad de trabajos adheridos a la DL hace que no se pueda hablar de método comparativo en sentido estricto, si bien muchas de las decisiones tomadas en el ámbito de esta orientación disciplinar han tratado de aproximarse a ese objetivo. Pero parece claro que las investigaciones actuales solo permiten la comparación en sentido lato, si bien con un alto grado de fiabilidad en los resultados del contraste.

Los estudios de DL se han aplicado a poblaciones muy variadas, entre los que se cuentan niños preescolares, estudiantes de educación primaria, estudiantes preuniversitarios, estudiantes de diversas carreras universitarias, jubilados, estudiantes de ELE, inmigrantes, etc. Los contextos también han sido muy diversos: áreas geográficas de distinta extensión, comunidades rurales y comunidades urbanas, contextos monolingües o multilingües, etc. Y también han sido distintos los momentos históricos en que se han realizado las investigaciones. Añádase, además, la diversidad de concepciones teóricas sobre el hecho lingüístico. Todo este conjunto de divergencias hace que no tenga cabida la aplicación del método en sentido restringido, que solo resulta posible si un investigador o un equipo de investigación orienta su trabajo toma una serie de decisiones y sigue las fases ya apuntadas. No obstante, algunas investigaciones de DL han logrado acercarse al método comparativo estricto al mantener una similitud notable en los planteamientos teóricos, en la muestra estudiada y en los procedimientos analíticos. Un ejemplo de este tipo es la serie de trabajos sobre la República Dominicana (Alba 2013), en los que la estabilidad metodológica seguida por el autor le permite obtener datos perfectamente equiparables sobre dos momentos históricos y trazar la evolución diacrónica de las unidades léxicas.

Pero conviene destacar también el alto grado de convergencia entre los estudios de DL, pues se comparten objetivos, metodología de recogida y análisis de los datos e incluso interpretación de los resultados. De ahí que las comparaciones entre investigaciones realizadas por autores distintos hayan sido muy frecuentes y productivas (*cf.*, a título de ejemplo simplemente, las contenidas en Arnal Purroy 2008); de hecho, la mayor parte de las investigaciones han dado lugar en alguna medida al contraste de los datos propios con los de procedencia ajena (como sucede también en el trabajo comentado de Jiménez Catalán). Las comparaciones suelen aparecer más a menudo en los estudios derivados que en las investigaciones primarias, lo cual es entendible, pues normalmente los diccionarios de disponibilidad, objeto básico de la investigación, ocupan una parte sustancial y central de los resultados publicados.

La comparabilidad de los trabajos se ve afectada por factores como los siguientes:

1. La fundamentación teórica de las investigaciones que se comparan.
2. El diseño de la investigación y, especialmente, las características de la muestra (tamaño, distribución).
3. El tratamiento de los datos.
4. Los instrumentos de análisis.

La fundamentación teórica y epistemológica de los estudios de disponibilidad léxica depende en gran medida de la concepción de cada investigador acerca del lenguaje y la lengua. En este sentido, no sería razonable esperar unanimidad en la manera de concebir el hecho lingüístico y, por ello, los intereses de cada investigador pueden derivar en aspectos particulares que dificulten la comparabilidad de los trabajos. Pero la posibilidad de comparar estudios se facilita en la medida en que todos ellos atienden al mismo nivel de análisis lingüístico, el léxico-semántico. Desde el punto de vista lingüístico, *grosso modo*, se puede decir que las investigaciones han ido evolucionando desde el planteamiento lexicológico que movió las investigaciones iniciales hasta derivar en planteamientos más generales, entre los que ha ido adquiriendo cada vez más peso la orientación cognitiva (Hernández Muñoz 2006, Šifrar Kalan 2016). La unidad básica que se maneja en los estudios es la palabra, aunque por lo general en los trabajos no suelen plantearse explícitamente reflexiones teóricas acerca de qué se incluye en este concepto; más bien se suele asumir un apriorismo respecto al concepto *palabra*. A mi juicio, los aspectos más descuidados en este sentido –importantes por las consecuencias metodológicas que tendrán–, atañen a la consideración de las unidades complejas (¿son unidades léxicas *azul claro*, *cortador de zanahorias* o *salir de paseo*?) y de otras unidades marginales, como las creaciones léxicas (Moliné Juste 2008, García Gondar 2011, Paredes en prensa). Algo mayor ha sido la atención recibida por los aspectos semánticos (Sánchez-Saus, 2011), si bien sigue siendo objeto de discusión si en los diccionarios de DL se deben incluir solo las unidades léxicas que contraen relaciones semánticas fuertes con el estímulo o si por el contrario los listados deben estar sometidos a menores restricciones en este sentido.

Todavía dentro de los aspectos teóricos, desde su origen los estudios de disponibilidad léxica han tomado como instrumento de análisis la estadística, herramienta metodológica que ha ido adquiriendo tanto peso que hoy se puede decir que la lingüística estadística aplicada al estudio del léxico se ha constituido como disciplina (Ávila Muñoz 2010: 15). Esta importancia teórica ha de tener su correlato en la exigencia de seguir manejando herramientas estadísticas cada vez más precisas para fundamentar los resultados y las conclusiones. Por fortuna, la incorporación casi sistemática de herramientas estadísticas cada vez más refinadas en los trabajos recientes de DL ha contribuido a aumentar la fiabilidad de las comparaciones realizadas.

La similitud en la metodología de la investigación es, sin duda, el factor más decisivo de cara a la comparabilidad de los estudios. En este punto hay que señalar que, aunque los estudios de DL han seguido de manera bastante fiel los criterios de los pioneros franceses, habría que diferenciar entre los estudios de DL en lengua materna y los de segundas lenguas. Para los primeros, la puesta en marcha del Proyecto Panhispánico de Disponibilidad Léxica (PPHDL) constituye sin duda el hito más importante de cara a garantizar los estudios contrastivos. Dirigido por H. López Morales, el macroproyecto tuvo entre sus objetivos desde el principio garantizar la comparabilidad con los trabajos originales por lo que gran parte de las decisiones

Sánchez-Saus, 2011), pues de ello depende el tipo de palabras que los encuestados asociarán con el estímulo.

El PPHDL también ha marcado directrices en relación con el tamaño de la muestra, que abogan por la representatividad de la muestra seleccionada en relación con el universo al que representan. Ciertamente es deseable que en las publicaciones se señalase siempre la representatividad de la muestra; no obstante, el número de informantes que se suele usar en los estudios de DL supera con mucho los criterios de representatividad sociolingüística, fijados en 0,025 % del universo estudiado (Labov 1966). Por debajo de ese nivel de representatividad deberían situarse solo los estudios exploratorios.

Un momento crucial de la investigación en DL, que afecta a los resultados obtenidos y, en consecuencia, a la comparabilidad con otros trabajos, es la fase que se conoce como “edición de los materiales”, donde el autor ha de tomar decisiones sobre las palabras anotadas por los encuestados. La principal decisión en este punto es sobre qué se debe mantener o qué se debe eliminar de los listados producidos. Las reuniones de coordinación mantenidas por los equipos que trabajan en el PPHDL han logrado una importante base de acuerdos sobre los criterios de edición de los materiales (Samper 1998), en asuntos como la repetición de las palabras o la unificación formal de las unidades que se incorporarán a los diccionarios de disponibilidad. Estos criterios, necesariamente amplios y generalizadores, han de combinarse con los que se han de tomar en relación con cada uno de los centros de interés. Respecto a la preservación de las unidades contenidas en los listados, los investigadores se han movido entre el criterio lato de mantener lo más fielmente posible las unidades léxicas anotadas y el criterio restrictivo de eliminar las menos relacionadas semánticamente con el estímulo o centro de interés. La diferencia entre mantener las todas las palabras asociadas al estímulo o preservar solo las asociaciones directas, como han demostrado Ávila y Sánchez (2010), tiene una influencia notable en los resultados, pues afecta a valores como las medias por sujeto y hasta al propio índice de disponibilidad de los vocablos, piedra angular en que se basa la DL. En este sentido, sería deseable una mayor unidad de criterios si se pretende que los resultados de los distintos territorios y de los diferentes grupos sean realmente equiparables.

En relación con las variables, muchas diferencias entre los estudios derivan de las propias características de la realidad estudiada. El investigador es responsable de incluir en su estudio aquellos factores o variables que considere relevantes para explicar el conocimiento léxico que tiene la población que desea estudiar. El PPHDL ha conservado también las variables explicativas manejadas: sexo, tipo de enseñanza recibida –pública o privada–, residencia urbana o rural, nivel sociocultural y edad (Moreno Fernández, Moreno Fernández y García de las Heras 2005), pero está claro que estos factores no pueden ser idénticos para todos los contextos. Pongamos algunos ejemplos. Si se trabaja con niños de preescolar, el momento del año en que nació el niño podría ser un factor explicativo de la capacidad léxica; en los estudios de español como lengua extranjera, es plausible suponer que factores como el conocimiento de otras lenguas o la distancia lingüística respecto a la lengua meta puedan influir en la capacidad léxica individual. Si se trabaja con la población general, pueden ser factores explicativos la profesión, el grado de la inclusión en el mercado lingüístico o la red social, entre muchos otros. En definitiva, resulta difícil –si no imposible– establecer de antemano la lista de posibles factores explicativos. No obstante lo anterior, los estudios de DL han venido señalando diferencias explicables por grupos sociales, que podrían ser constantes y que convendría mantener en futuros estudios. En las investigaciones sobre lengua materna, la clase social y el nivel educacional aparecen como factores que suelen originar diferencias (el sexo no siempre ha resultado significativo, en cambio); en lengua extranjera, parecen influir también la lengua materna, el tiempo de estudio de la lengua meta y el conocimiento de otras lenguas. En este punto, no obstante, conviene que al realizar estudios comparativos se tenga en cuenta el grado de similitud de los subgrupos manejados en cada investigación.

Los resultados cuantitativos son los aspectos más consistentes para establecer comparaciones fiables y los trabajos de DL ofrecen una serie de datos constantes, como el índice de disponibilidad, el índice de frecuencia, el número total de palabras, el número total de vocablos, etc. A este respecto, se debe tener presente que hay valores que se ven afectados por el tamaño de la muestra, como sucede con el número total de palabras o, en menor medida, el de vocablos. Otros valores estadísticos, en cambio, son independientes del número de encuestas, como sucede con las medias de palabras, pero influye el modo en que se han tratado los datos en la fase de “edición de los materiales”. Hay que destacar a este respecto el trabajo realizado por Callealta y Gallego (2016), quienes han incorporado modificaciones en la fórmula matemática usada tradicionalmente para el cálculo del índice de disponibilidad y en otros índices manejados en las investigaciones con el objetivo de lograr la comparabilidad de estudios con muestras de diferente tamaño y distintas características.

En conclusión, es mucho lo que se ha andado para favorecer la comparabilidad de los estudios de DL. No obstante, se ha de seguir trabajando para dotar de mayor coherencia teórica y metodológica las investigaciones, lo cual permita a) realizar comparaciones en nuevos aspectos y b) hacer más fiables los estudios comparativos. Estas consideraciones han de hacerse extensivas especialmente a las investigaciones sobre disponibilidad léxica en segundas lenguas.

## REFERENCIAS BIBLIOGRÁFICAS

- Alba, O. (2013): Variación diacrónica del léxico disponible dominicano. *Lingüística Española Actual*, 35(1): 127-148.
- Arnal Purroy, M. L. (ed.) (2008): *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución Fernando el Católico: 17-50.
- Ávila Muñoz, A. (2010): Aproximación a los estudios estadísticos del léxico. En Antonio M. Ávila y Juan A. Villena (eds.): *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Sarriá, 2010: 15-34.
- Ávila Muñoz, A. y Sánchez Sáez, J. M. (2010): La investigación sobre la disponibilidad léxica. Antecedentes y fundamentos. En Antonio M. Ávila y Juan A. Villena (eds.): *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Sarriá, 2010: 37-82.
- Callealta Barroso, J. y Gallego Gallego, D. J. (2016): Medidas de disponibilidad léxica: comparabilidad y normalización. *Boletín de Filología*, 51(1): 39-92.
- Colino, C. (2009): "Método comparativo ". En Román Reyes (Dir): *Diccionario Crítico de Ciencias Sociales. Terminología Científico-Social*, Tomo 1/2/3/4, Madrid-México: Plaza y Valdés.
- [http://pendientedemigracion.ucm.es/info/eurotheo/diccionario/M/metodocomparativo\\_a.htm](http://pendientedemigracion.ucm.es/info/eurotheo/diccionario/M/metodocomparativo_a.htm)
- Durkheim, E. (1895 [2001]): *Las reglas del método sociológico y otros escritos sobre Filosofía de las Ciencias Sociales*. Madrid: Alianza Editorial (ed. De Santiago González Noriega).
- García Gondar, F. (2011): La creatividad léxica a través de recursos morfológicos en el léxico disponible del español de Galicia. En B. López Meirama (ed.): *Estudios sobre disponibilidad léxica en el español de Galicia*, Santiago de Compostela: Universidade de Santiago de Compostela, 2011: 217-266.
- Gómez Díaz de León, C., Ayde de León de la Garza, E. (2014) Método Comparativo Estudio comparado. En *Métodos y técnicas cualitativas y cuantitativas aplicables a las investigaciones en ciencias sociales.*, México: Tirant Humanidades, 2014: 223-252.
- Grosser, A. (1973): *Politik erklären (Explicar la política)*. Múnich: Hanser.
- Hernández Muñoz, N. (2006): *Hacia una teoría cognitiva integrada de la disponibilidad léxica. El léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad de Salamanca.
- Labov, W. (1966): *The Social Stratification of English in New York City*. Oxford: Center for Applied Linguistic.
- Moliné Juste, A. B. (2008): Creación léxica en el vocabulario de los jóvenes aragoneses: aspectos morfológicos. En M. L. Arnal Purroy (ed.): *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*, Zaragoza: Institución Fernando el Católico / Excma. Diputación de Zaragoza, 2008: 163-194.
- Moreno Fernández, F., Moreno Fernández, J. E. y García de las Heras, A. (1995): Cálculo de disponibilidad léxica. El programa LexiDisp. *Lingüística*, 7: 243-249.
- Paredes García, F. (2012): Desarrollos teóricos y metodológicos recientes de los estudios de disponibilidad léxica. *Revista Nebrija de Lingüística Aplicada*, 6(4).
- Paredes García, F. (en prensa): Neologismos y creaciones léxicas en las pruebas de disponibilidad léxica de la Comunidad de Madrid, *Homenaje a O. A. Bern: Peter Lang*.

Samper Padilla, J. A. (1998): Criterios de edición del léxico disponible: sugerencias. *Lingüística*, 10: 311-332

Sánchez-Saus Laserna, M. (2011): *Bases semánticas para el estudio de los centros de interés del léxico disponible. Disponibilidad léxica de informantes extranjeros en las universidades andaluzas*. Tesis doctoral. Universidad de Cádiz.

Šifrar Kalan, M. (2016): La universalidad de los prototipos semánticos en el léxico disponible de español. *Verba hispanica*, 24: 147-165.