ESTUDIO DE DISPONIBILIDAD LÉXICA EN APRENDICES ITALIANOS DE ESPAÑOL: ANÁLISIS CUALITATIVO

ROBERTO RUBIO SÁNCHEZ
Universidad de Salamanca

RESUMEN

Este artículo presenta una investigación que forma parte del conjunto de estudios dedicados al léxico disponible. Partiendo de los criterios marcados por el Proyecto Panhispánico*, se ha elaborado un trabajo sobre disponibilidad léxica de estudiantes de español como lengua extranjera. La particularidad de este estudio es que se centra en aprendices de procedencia italiana, colectivo hasta ahora nunca analizado. En estas líneas se muestran los criterios metodológicos empleados, destacando las diferencias respecto al* Proyecto Panhispánico*. Por último, se presentan algunos resultados de carácter cualitativo con los que se muestran la utilidad de un estudio de estas características.*

Palabras clave: Disponibilidad léxica, ELE, español, italiano, metodología, resultados cualitativos

ABSTRACT

This paper presents a research about lexical knowledge base. Taking into account the criteria used by the Panhispanic Project*, we have created and applied a lexical availability test for students of Spanish as a Foreign Language. The uniqueness of this investigation stems from the fact that we work with Italian students, a group that has not been studied so far. In this study, we show the methodological criteria used; we also underline the differences between our research and the* Panhispanic Project*. Finally, we present some qualitative results, which illustrate the usefulness of studies such as this.*

Keywords: Lexical availability, Spanish, Italian, methodology, qualitative results

1. INTRODUCCIÓN

Este artículo presenta una investigación que forma parte del conjunto de estudios dedicados al léxico disponible. Tomando como punto de referencia los criterios marcados por el *Proyecto Panhispánico* (Bartol Hernández, 2006: 380-381) con informantes nativos, hemos decidido incorporar algunas novedades en lo que se refiere a los criterios metodológicos, dado que trabajamos la disponibilidad léxica de estudiantes de español como lengua extranjera (ELE).

El estudio del léxico disponible de estudiantes de ELE no es algo novedoso, sin embargo, nuestro trabajo tiene la peculiaridad de centrarse en aprendices de procedencia italiana –colectivo hasta ahora nunca estudiado– para poder desarrollar, en un futuro, un estudio específico de tipo comparado y contrastivo. La elección de estudiantes de origen italiano está motivada por la afinidad entre los dos idiomas, hecho que ocasiona numerosos problemas de interferencias léxico-semánticas (Schmidely, 2001: 55-57).

En este estudio se ha empleado una muestra reducida de informantes para observar la utilidad de un estudio de estas características. Por ello, planteamos este trabajo como un mero ejemplo de cómo se llevaría a cabo la investigación más amplia que pretendemos desarrollar en los próximos años.

Dadas las restricciones impuestas para la redacción de este artículo, nos centraremos en presentar los criterios metodológicos adoptados y en un breve comentario de los resultados cualitativos obtenidos de nuestra labor de procesamiento y análisis de datos.

2. METODOLOGÍA

2.1 La encuesta

2.1.1 El cuestionario y los informantes

Las encuestas de disponibilidad léxica se componen de dos partes bien distintas. La primera de ellas recoge los datos de los informantes

que más tarde se utilizarán como variables. La segunda, en cambio, consta de la prueba asociativa propiamente dicha, en la que se pide al encuestado que escriba todas las palabras que le vengan a la mente a partir de una serie de centros de interés dados. El encuestado dispondrá de dos minutos para completar cada centro (Samper Hernández, 2002: 20-21).

En nuestro estudio, las encuestas se han llevado a cabo en Salamanca a dos grupos de estudiantes italianos: veintidós del instituto Giordano Bruno de Perugia y catorce del instituto A. Tambosi de Trento. Un total de treinta y seis informantes de nivel preuniversitario que se encontraban en España para asistir a un curso de inmersión de ELE de una semana.

Las pruebas se han realizado el primer día de su estancia, siguiendo el procedimiento estándar aconsejado por el *Proyecto Panhispánico*. En primer lugar, se insistió en que se trataba de una prueba anónima y que de ninguna manera se enfrentaban a un examen. En segundo lugar, se les describió brevemente las dos partes de las que constaba la encuesta. Al tratarse de informantes extranjeros, decidimos ilustrar la segunda parte mediante un ejemplo – eligiendo un centro que no aparecía en la encuesta real – para asegurarnos de que entendieran bien el mecanismo de la prueba.

El trabajo de procesamiento de datos se ha llevado a cabo utilizando *DISPOLEX*[1], página web en que –además de alojar una potente base de datos documental e informativa de los trabajos que se vienen desarrollando– es un instrumento estadístico muy refinado que ha sido utilizado también para el procesamiento de los datos obtenidos mediante las encuestas.

A continuación se muestran las variables que hemos analizado en nuestro estudio.

2.1.2 Las variables

Las variables que se han tenido en cuenta en esta investigación difieren en cierta medida de las que se han empleado en la mayoría de los estudios de disponibilidad léxica sobre hablantes nativos de español (Fernández Juncal, 2008: 21), que son:

a) *Sexo*

b) *Nivel sociocultural*
c) *Residencia de los padres*
d) *Carácter del centro*
e) *Localización del centro*

Dadas las características de nuestro trabajo, hemos modificado algunas de las variables del modelo panhispánico por considerar más provechosas para este estudio las variables que exponemos a continuación.

En nuestro caso, hemos decidido mantener las variables a) *Sexo* y b) *Nivel sociocultural*, mientras que hemos prescindido de d) *Carácter del centro* y e) *Localización del centro*, al darse una respuesta unánime en las encuestas: se trataba en ambos casos de centros públicos de ubicación urbana. También prescindimos de la variable c) *Residencia de los padres* dado que no parecía demasiado relevante para este estudio.

Por otro lado, hemos introducido como novedosas las variables f) *Años de estudio del español* y g) *Estancia en país hispano*.

Por tanto, las variables consideradas en nuestra investigación son:
a) *Sexo*
b) *Nivel sociocultural*
c) *Años de estudio de español*
d) *Estancia en país hispano*

Otra variable que incluiremos en futuras investigaciones será e) *Nivel de español*, con la que examinaremos las distintas fases de aprendizaje de ELE. En este trabajo hemos prescindido de esta variable al tratarse de estudiantes que poseen el mismo nivel: B1[2].

2.1.3 Los centros de interés

La encuesta del *Proyecto Panhispánico* consta de dieciséis centros de interés (Bartol Hernández, 2004: 16):

1. *El cuerpo humano*
2. *La ropa*
3. *Partes de la casa sin muebles*
4. *Los muebles de la casa*
5. *Alimentos y bebidas*
6. *Objetos colocados en la mesa para la comida*

7. *La cocina y sus utensilios*
8. *La escuela: muebles y materiales*
9. *Calefacción e iluminación*
10. *La ciudad*
11. *El campo*
12. *Medios de transporte*
13. *Trabajos del campo y del jardín*
14. *Animales*
15. *Juegos y diversiones*
16. *Profesiones y oficios*

En nuestro caso, al trabajar con estudiantes cuya lengua materna difiere del español, hemos optado por reducir el número de centros de interés a catorce. De esta manera, hemos agrupado los centros 03. Partes de la casa sin los muebles, 04. Los muebles de la casa, 06. Objetos colocados en la mesa para la comida y 07. La cocina y sus utensilios, en uno más general denominado La casa (03. en la nueva lista). Además, hemos decidido eliminar los centros 09. Calefacción e iluminación, 11. El campo y 13. Trabajos del campo y del jardín por considerarlos poco pertinentes a los intereses de los jóvenes que hemos encuestado.

Por otro lado, hemos decidido incorporar cuatro centros de interés novedosos: 11. Partes del cuerpo (en italiano), 12. La ropa (en italiano), 13. España y 14. Italia. 11. y 12. nos servirán para realizar una comparación entre el número de vocablos que los encuestados saben en español y en su LM. En cambio, con la introducción de 13. y 14. pretendemos observar qué concepción tienen los encuestados de la cultura española e italiana.

Además, hemos modificado ligeramente el enunciado de otros tres centros de interés como hace Hernández Muñoz (2004: 28): 1. cambia de El cuerpo humano a Partes del cuerpo, 8. La escuela: muebles y materiales se reduce a La escuela; al igual que ocurre con 16. Profesiones y oficios que se queda solo con el término Profesiones.

Por tanto, nuestra encuesta queda reducida a los siguientes catorce centros de interés:

1. *Partes del cuerpo*

2. *La ropa*
3. *La casa*
4. *Alimentos y bebidas*
5. *La escuela*
6. *La ciudad*
7. *Medios de transporte*
8. *Animales*
9. *Juegos y diversión*
10. *Profesiones*
11. *Partes del cuerpo* (en italiano)
12. *La ropa* (en italiano)
13. *España*
14. *Italia*

3. DATOS CUALITATIVOS

Los datos que hemos podido recopilar hasta ahora nos han permitido realizar un análisis de los resultados, tanto cuantitativo como cualitativo, comprobando el interés que puede tener un estudio de este tipo. Dadas las restricciones impuestas a este artículo, aquí se presentan únicamente resultados de carácter cualitativo.

3.1. (1) *Partes del cuerpo* frente a (11) *Partes del cuerpo (en italiano)*

Si contrastamos estos dos centros de interés, se puede observar que *Partes del cuerpo (en italiano)* presenta más del doble de palabras totales respecto a su homónimo en español –1086 frente a 485– y prácticamente el doble de palabras diferentes –99 frente a 50. No obstante, llama la atención que los índices de cohesión de ambos centros son los más elevados de toda la muestra (0.30 en *Partes del cuerpo (en italiano)* y 0.27 en *Partes del cuerpo*).

A continuación se muestra una tabla en la que se exponen las diez palabras con mayor disponibilidad en ambos centros:

Rango	**Partes del cuerpo**	**Partes del cuerpo (en italiano)**
1	ojos/jo	braccia/o 'brazos/zo'[3]

2	cabeza	mano/i ‘mano/s’
3	pies/pie	occhi/o ‘ojos/jo’
4	manos/no	gambe/a ‘piernas/na’
5	orejas/ja	piedi/e ‘pies/pie’
6	nariz	bocca ‘boca’
7	boca	naso ‘nariz’
8	cara	orecchie ‘orejas’
9	pierna/s	dita ‘dedos’
10	pelo	capelli ‘pelo’

Figura 5. Las diez palabras con mayor promedio en *Partes del cuerpo* y *Partes del cuerpo (en italiano)*

Se puede observar que siete de las diez palabras coinciden, aunque no aparecen en el mismo orden –salvo el caso de *pelo* y *capelli* que las encontramos en el décimo puesto. Por otro lado, resulta interesante ver que las dos palabras con mayor promedio en italiano, no aparecen entre las diez más frecuentes en español.

3.2. (2) *La ropa* frente a (12) *La ropa (en italiano)*

Al igual que ocurre en el caso anterior, el centro de interés en italiano presenta más del doble de palabras totales y casi el doble de palabras diferentes que su homónimo en español: *La ropa (en italiano)* recopila 662 palabras totales y 116 palabras diferentes, frente a las 299 totales y las 67 diferentes que nos encontramos en *La ropa*. Sin embargo, en este caso llama la atención la gran distancia que encontramos en el rango o posición relativa de estos campos, ya que mientras que *La ropa (en italiano)* ocupa el segundo lugar, *La ropa* se presenta en el penúltimo puesto.

Seguidamente exponemos una tabla que presenta las diez palabras con mayor promedio en estos centros de interés:

Rango	**La ropa**	**La ropa (en italiano)**
1	camiseta/s	pantaloni ‘pantalones’
2	zapatos/to	scarpe ‘zapatos’
3	pantalones/lón	camicia ‘camisa’
4	bufanda	jeans ‘jeans’

5	chaqueta	mutande ‘bragas/calzoncillos’
6	vaqueros	maglietta/e ‘camiseta’
7	zapatillas	canottiera/e ‘camiseta/s de tirantes’
8	jeans	felpa/e ‘sudadera/s’
9	camisa	sciarpa ‘bufanda’
10	(ropa) interior	maglia ‘jersey’

Figura 6: Las diez palabras con mayor promedio en *La ropa* y *La ropa (en italiano)*

Del mismo modo que ocurría en el apartado anterior, en este caso también coinciden bastantes palabras –seis de diez– aunque tampoco aparecen en el mismo orden, salvo el caso de *zapatos/to* y *scarpe* que se muestran en segundo lugar.

3.3. (13) *España* frente a (14) *Italia*

Los centros de interés *España* e *Italia* ocupan el tercer y séptimo lugar respectivamente en cuanto a rango se refiere, teniendo en cuenta el número de palabras de cada centro. El primero presenta 622 palabras totales y 199 diferentes, mientras que el segundo muestra 462 totales y 168 diferentes. Es destacable que estos centros de interés poseen los índices de cohesión menores –0.9 en el caso de *España*, 0.8 en el de *Italia*– junto a *Juegos y diversiones* (0.7), lo que indica que se trata de dos centros cuyas áreas temáticas poseen límites poco nítidos y por ello son más abiertos.

A continuación mostramos una tabla con las diez palabras con mayor disponibilidad en cada campo:

Rango	**España**	**Italia**
1	Madrid	Roma
2	Salamanca	Pizza
3	Barcelona	Pasta
4	corrida	Milán
5	paella	mar
6	toros/ro	Perugia
7	fiesta/s	Coliseo
8	comunidades (autónomas)	montañas/ña
9	Castilla y León	playas/ya

10	Andalucía	Sicilia

Figura 7: Las diez palabras con mayor promedio en *España* e *Italia*

Las respuestas en *España* se podrían clasificar en dos apartados: palabras relacionadas con cultura geográfica (1. *Madrid*, 2. *Salamanca*, 3. *Barcelona*, 8. *comunidades (autónomas)*, 9. *Castilla y León* y 10. *Andalucía*) y estereotipos de comida y fiestas (4. *corrida*, 5. *paella*, 6. *toros/ro* y 7. *fiesta/s*); mientras que las respuestas en *Italia* las podríamos clasificar en 3 bloques: ciudades y regiones (1. *Roma*, 4. *Milán*, 6. *Perugia* y 10. *Sicilia*), elementos naturales (5. *mar*, 8. *montañas/ña* y 9. *playa/s*) y por último, estereotipos (2. *pizza*, 3. *pasta* y 7. *Coliseo*). Llama la atención que los informantes hagan referencias a estereotipos en su propia cultura.

No obstante, tenemos que considerar que este trabajo se basa en un número reducido de encuestas –36 informantes– y por ello los resultados obtenidos no son concluyentes, pero sí representativos, pues muestran unas tendencias que corroboraremos con futuras investigaciones.

4. CONCLUSIONES

Con este trabajo hemos intentado poner de manifiesto el enorme campo de posibilidades que la disponibilidad léxica abre a la investigación.

Los datos que hemos podido recopilar hasta ahora nos han permitido realizar un análisis de las variables consideradas, comprobando el interés que puede tener un estudio de este tipo.

No dudamos que, siguiendo con nuestra investigación, podremos alcanzar una visión más exhaustiva de la disponibilidad léxica en aprendices italianos de ELE.

Algunas aplicaciones de este estudio serán la creación de un corpus léxico actualizado y agrupado en campos nocionales, que pueda servir como base para la selección de vocabulario para la clase de ELE, adecuándolo a los requerimientos del MCER, y, por otro lado, comprobar qué diferencias pueden existir en las asociaciones que se lleva a cabo en el lexicón mental de los aprendices de español como L2 frente a lo que pasa con los nativos.

NOTAS

[1] www.dispolex.com
[2] Según la nivelación propuesta por el Consejo de Europa en el *Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza, evaluación* (2000).
[3] Para la traducción al español de los términos italianos hemos recurrido al diccionario *Lo spagnolo minore* de la editorial Zanichelli (2002).

REFERENCIAS BIBLIOGRÁFICAS

Bartol Hernández, J.A. 2004. *Léxico disponible de Soria. Estudio y diccionarios*. Burgos: Instituto Castellano y Leonés de la Lengua.

Bartol Hernández, J.A. 2006. "La disponibilidad léxica", *Revista Española de Lingüística*, 36: 379-391. Madrid: SEL.

Consejo de Europa. 2000. *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. En línea en: http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/ [Consulta: marzo 2014]

Equipo DispoLex. 2003. *Base de datos de disponibilidad léxica DispoLex*. En línea: www.dispolex.com [Consulta: marzo 2014]

Fernández Juncal, C. 2008. *Léxico disponible de Burgos*. Burgos: Instituto Castellano y Leonés de la Lengua.

Hernández Muñoz, N. 2004. *El léxico disponible de los estudiantes conquenses*. Salamanca: Ediciones Universidad de Salamanca.

Lo spagnolo minore. 2002. *Dizionario spagnolo italiano. Italiano spagnolo*. Milano: Zanichelli.

Samper Hernández, M. 2002. *Disponibilidad léxica en alumnos de español como lengua extranjera*. Málaga: ASELE

Schmidely, J. (coord.). 2001. *De una a cuatro lenguas. Del español al portugués, al italiano y al francés*. Madrid: Arco libros.