

## Palabras Vocabulario Léxico

La lexicología aplicada a la didáctica y a la diacronía  
editado por Florencio del Barrio de la Rosa

# Acercamiento al léxico disponible de 173 estudiantes italianos preuniversitarios de Español como Lengua Extranjera

Roberto Rubio  
(Universidad de Salamanca, España)

**Abstract** The research we present in this paper is 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. In particular, we work with Italian students, a group that has a related language as mother tongue. In this study, we show the methodological criteria we used; we also underline the differences between our research and the Panhispanic Project. Finally, we present some quantitative and qualitative results, which illustrate the usefulness of studies such as this.

**Sumario** 1 Introducción. – 2 Metodología. – 2.1 La encuesta y los informantes. – 2.2 Los centros de interés. – 2.3 Las variables. – 3 Datos cuantitativos. – 3.1 Índices generales. – 3.2 Índices según las variables. – 4 Datos cualitativos. Partes del cuerpo (en español y en italiano). – 5 Conclusiones.

**Keywords** Lexical availability. Vocabulary. Spanish as a Second Language. Romance languages. Italian students.

## 1 Introducción

Los estudios de disponibilidad léxica<sup>1</sup> se iniciaron en Francia a mediados del siglo XX y llegaron al mundo hispánico en la década de los setenta de la mano de Humberto López Morales. Desde entonces, la mayoría de estos trabajos con informantes nativos de español se realiza siguiendo los criterios marcados por el Proyecto Panhispanico. Sin embargo, desde los inicios de este nuevo siglo se están elaborando trabajos sobre la disponibilidad léxica de estudiantes de Español como Lengua Extranjera (ELE de ahora en adelante). Las primeras investigaciones en este contexto fueron llevadas a cabo por Carcedo González (2000) con estudiantes finlandeses

<sup>1</sup> El autor ha realizado el presente trabajo en el periodo de disfrute de una Ayuda para la Formación de Personal Investigador financiada por la Junta de Castilla y León y el Fondo Social Europeo.

---

### VenPalabras 1

DOI 10.14277/6969-169-0/VP-1-9 | Submission 2016-10-31 | Acceptance 2017-01-23  
ISBN [ebook] 978-88-6969-169-0 | ISBN [print] 978-88-6969-170-6 | © 2017

de escuela secundaria y Samper Hernández (2002) con extranjeros universitarios en Salamanca.

En el mundo hispanohablante, los trabajos de análisis léxico con informantes cuya Lengua Materna (a partir de ahora LM) es distinta a la española constituyen un ámbito de interés cada vez más relevante. Así lo evidencian los estudios realizados por López González (2009), Rodríguez Menduiña (2010), Lin (2012), Gamazo Carretero (2014) y del Barrio de la Rosa (2016), entre otros, quienes trabajan con aprendices polacos, estadounidenses, chinos, portugueses e italianos, respectivamente.

Esta investigación, en concreto, se centra en estudiantes de origen italiano con el fin de desarrollar un estudio específico de tipo comparado y contrastivo. La elección de estudiantes italianos se debe a la afinidad de los dos idiomas, hecho que ocasiona numerosos problemas de interferencias léxico-semánticas (Schmidely 2001, 55-7).

A continuación, trataremos la metodología empleada, realizaremos un análisis de los datos cuantitativos y, para finalizar, haremos un breve comentario de los resultados cualitativos obtenidos del procesamiento y análisis de datos de nuestra muestra.

## **2 Metodología**

Los criterios seguidos para la realización de este estudio son similares a los empleados en investigaciones recientes desarrolladas en este campo (Bartol Hernández 2006, 380-1). No obstante, como mostraremos en los apartados 2.2 y 2.3, hemos introducido algunos cambios para adaptar dicha metodología a nuestras necesidades específicas.

### **2.1 La encuesta y los informantes**

La encuesta consiste en un cuestionario anónimo que consta de dos partes: en la primera, el informante responde a una serie de preguntas que posteriormente codificamos y utilizamos como variables en el estudio; en la segunda, tiene lugar la prueba asociativa propiamente dicha, en la que se proporciona al encuestado una serie de centros de interés para que escriba todas las unidades léxicas que le vengan a la mente en un tiempo determinado. El encuestado dispone de dos minutos para completar cada centro (Samper Hernández 2002, 20-1).

En nuestro estudio, hemos realizado la encuesta a 173 informantes de nivel preuniversitario. Los centros que han participado en esta investigación son: Istituto Marcelline de Lecce (con 11 informantes), Liceo Leonardo da Vinci de Bérgamo (74), I.I.S.S. Alessandro Greppi de Monticello Brianza, Lecco (52) y Liceo Malpighi de Bolonia (36). Las pruebas se han llevado a cabo en las aulas entre abril de 2015 y marzo de 2016.

La herramienta que hemos utilizado para el procesamiento de datos ha sido Dispalex (<http://www.dispalex.com>), página web que integra un instrumento estadístico refinado, además de una potente base de datos documental de los trabajos que se están desarrollando en este ámbito. Sin embargo, como señala Gallego Gallego (2014, 274), hay investigadores que actualmente utilizan el programa SPSS para complementar los análisis convencionales brindados por Dispalex. En este trabajo, también presentamos una pequeña prueba realizada con el paquete estadístico SPSS.

## 2.2 Los centros de interés

Hemos modificado algunos centros de interés del Proyecto Panhispánico (Bartol Hernández 2004, 16)<sup>2</sup> para adecuarlos a los intereses de nuestro trabajo. De este modo, al tratar con aprendices cuya LM difiere del español, hemos decidido disminuir a catorce el número de centros de interés. Por ello, hemos suprimido los centros que hemos estimado menos adecuados a los intereses de nuestra investigación: 3. Partes de la casa (sin los muebles), 4. Los muebles de la casa, 6. Objetos colocados en la mesa para la comida, 7. La cocina, 9. Calefacción e iluminación, 10. La ciudad, 11. El campo, 12. Medios de transporte y 13. Trabajos del campo y del jardín.

Por otra parte, hemos introducido seis centros de interés novedosos: 4. Acciones y actividades habituales, 5. España y los españoles, 9. Aspecto físico y carácter, 10. La lengua española, 12. Italia y los italianos, 13. Partes del cuerpo (en italiano) y 14. Comidas y bebidas (en italiano). La incorporación de los centros 4. y 9. persigue la producción de palabras de categoría gramatical diferente al sustantivo (verbos y adjetivos, respectivamente). La introducción de 10. tiene como objetivo examinar las relaciones efectuadas por los encuestados y su opinión acerca de la materia de estudio. Por otro lado, con la inclusión de 5. y 12. examinaremos la concepción que tienen los informantes de la cultura española e italiana.<sup>3</sup> En cambio, con la incorporación de 13. y 14. pretendemos efectuar una confrontación entre el número de vocablos que los encuestados saben en español y en su LM.

<sup>2</sup> 1. Partes del cuerpo, 2. La ropa, 3. Partes de la casa (sin los muebles), 4. Los muebles de la casa, 5. Alimentos y bebidas, 6. Objetos colocados en la mesa para la comida, 7. La cocina, 8. La escuela, 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

<sup>3</sup> Estos dos campos ofrecen características muy especiales motivadas por el objetivo final de su inclusión en el cuestionario: el estudio de la adquisición de la cultura hispana en alumnos de ELE. En estos dos centros de interés se preguntan conceptos, no palabras.

Por tanto, nuestra encuesta presenta los siguientes centros de interés:

1. Partes del cuerpo
2. La ropa
3. La escuela
4. Acciones y actividades habituales
5. España y los españoles
6. Comidas y bebidas
7. Animales
8. Juegos y diversiones
9. Aspecto físico y carácter
10. La lengua española
11. Profesiones
12. Italia y los italianos
13. Partes del cuerpo (en italiano)
14. Comidas y bebidas (en italiano)

### 2.3 Las variables

Hemos modificado también las variables utilizadas en la mayoría de los estudios del Proyecto Panhispánico (Fernández Juncal 2008, 21),<sup>4</sup> adaptándolas a los intereses de nuestro trabajo. De esta manera, hemos desechado las variables Localización del centro y Residencia de los padres, puesto que no consideramos estos elementos de especial importancia para nuestra investigación.

Por otro lado, hemos incorporado el parámetro Nivel de español. Por tanto, las variables utilizadas en este trabajo son:

- a. Sexo
- b. Nivel sociocultural
- c. Titularidad de centro
- d. Nivel de español

---

<sup>4</sup> Sexo, nivel sociocultural, residencia de los padres, carácter del centro y localización del centro.

### 3 Datos cuantitativos

#### 3.1 Índices generales

El número total de palabras obtenidas entre los 173 informantes se eleva a 34 543, lo que implica una media de 199,67 respuestas por alumno y de 2467,36 por centro de interés (14,26 palabras por informante y centro de interés).

En la tabla 1 se pueden observar los datos globales desglosados.

Tabla 1. Índices generales

Nº	CI	PT	PD	PI	IC
1	Partes del cuerpo	2333	198	13,49	0,07
2	La ropa	1736	239	10,03	0,04
3	La escuela	2368	351	13,69	0,04
4	Acciones y actividades habituales	2379	385	13,75	0,04
5	España y los españoles	2368	543	13,69	0,03
6	Comidas y bebidas	2567	362	14,84	0,04
7	Animales	1840	243	10,64	0,04
8	Juegos y diversiones	1672	402	9,66	0,02
9	Aspecto físico y carácter	2218	411	12,82	0,03
10	La lengua española	1964	578	11,35	0,02
11	Profesiones	2037	347	11,77	0,03
12	Italia y los italianos	2445	595	14,13	0,02
13	Partes del cuerpo (en italiano)	4360	280	25,20	0,09
14	Comidas y bebidas (en italiano)	4256	559	24,60	0,04

CI: Centro de interés; PT: Palabras totales; PD: Palabras diferentes; PI: Palabras por informante; IC: Índice de cohesión.

Los centros de interés más productivos son, lógicamente por responderse en la LM de los informantes, Partes del cuerpo (en italiano) y Comidas y bebidas (en italiano); seguidos por: Comidas y bebidas, Italia y los italianos, Acciones y actividades habituales, La escuela, España y los españoles, Partes del cuerpo y Aspecto físico y carácter. En cambio, los centros menos productivos han sido: Juegos y diversiones, La ropa, Animales, La lengua española y Profesiones.

Por otro lado, los centros con un índice de cohesión más elevado, es decir, aquellos que representan un campo más cerrado son: Partes del cuerpo (en italiano) y Partes del cuerpo. Sin embargo, los centros con índice de cohesión más bajo, por tanto, con límites menos nítidos y que, por ello, favorecen la aparición de un mayor número de asociaciones son: Juegos y diversiones, La lengua española e Italia y los italianos.

En cuanto a los centros de interés que se responden primero en español y más tarde en italiano, tenemos que Partes del cuerpo (en italiano) presenta un 46,5% más de palabras totales y un 29,3% más de palabras diferentes frente a su correspondiente en español; mientras que Comidas y bebidas (en italiano) presenta un 39,7% más de palabras totales y un 35,2% más de palabras diferentes respecto a su homónimo en español.

En los apartados 4.1 y 4.2 de este trabajo mostramos un pequeño análisis cualitativo de estos centros de interés teniendo en cuenta las 20 palabras con mayor índice de disponibilidad en cada centro.

### 3.2 Índices según las variables

#### 3.2.1 Variable sexo

Samper Padilla y Samper Hernández afirman que «todas las muestras, sin excepción, presentan una proporción mayor de mujeres que de hombres» (2006, 46). En este sentido, nuestro trabajo no es diferente. Así, esta variable ha mostrado un porcentaje mucho mayor de informantes de sexo femenino: 128 mujeres han realizado la encuesta frente a 45 hombres.

Dada la desigualdad que presenta esta variable en el número de informantes, hemos efectuado el análisis teniendo en cuenta las palabras por informante, ya que «es el único índice que no se ve afectado por el número de informantes de la muestra» (Fernández Juncal 2008, 42). De este modo, obtenemos que las mujeres presentan mayor media de respuestas globales, 15,3 palabras por mujer, frente a 11,29 palabras por hombre, lo que significa un 26,21% más (cf. tab. 2). Esta superioridad en la producción léxica por parte del sexo femenino también se puede observar en otras investigaciones: Bartol Hernández (2004), Fernández Smith et al. (2008), Fernández Juncal (2008, 2013). En estos casos, las diferencias son mínimas y se suelen alternar centros de interés donde presentan mayor producción los hombres, con centros más productivos por parte de las mujeres. Sin embargo, en nuestro trabajo, las mujeres aventajan a los hombres en el promedio de respuestas en todos los centros de interés, como ocurre en el trabajo de Sánchez-Saus Laserna (2016).

Tabla 2. Promedio de respuestas según la variable sexo

Nº	CI	Respuestas por varón	Respuestas por mujer
1	CUE	10,2	14,64
2	ROP	7,44	10,95
3	ESC	10,2	14,91
4	ACC	10,71	14,82
5	ESP	11,84	14,34
6	COM	10,69	16,3
7	ANI	7,22	11,84
8	JUE	8,04	10,23
9	ASP	9,11	14,13
10	LEN	8,07	12,51
11	PRO	9	12,75
12	ITA	10,69	15,34
13	CUE ITA	23,56	25,78
14	COM ITA	21,27	25,77

CUE: Partes del cuerpo; ROP: La ropa; ESC: La escuela; ACC: Acciones y actividades habituales; ESP: España y los españoles; COM: Comidas y bebidas; ANI: Animales; JUE: Juegos y diversiones; ASP: Aspecto físico y carácter; LEN: La lengua española; PRO: Profesiones; ITA: Italia y los italianos; CUE ITA: Partes del cuerpo (en italiano); COM ITA: Comidas y bebidas (en italiano).

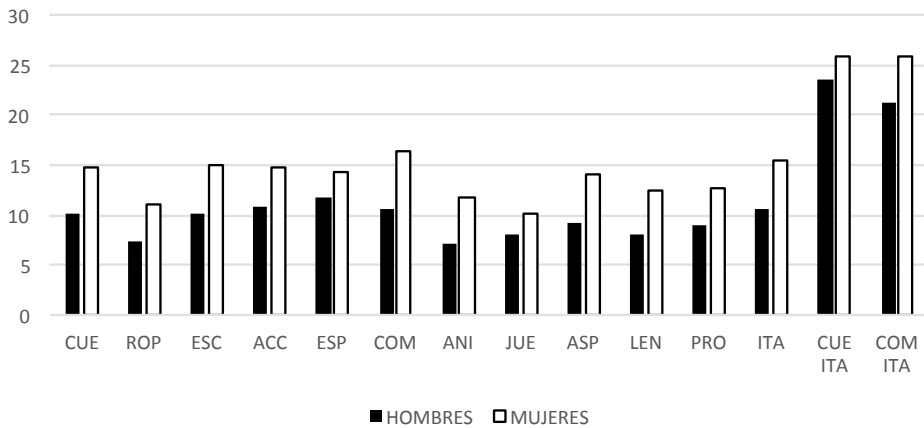


Figura 1. Palabras por informante según el sexo

En nuestra investigación, Comidas y bebidas, Aspecto físico y carácter, La escuela, Italia y los italianos y Animales son los centros que presentan las mayores diferencias de producción. En ellos, las mujeres aventajan en más de 4,5 palabras los resultados obtenidos por hombres. No obstante, podemos observar que, pese a que se mantiene mayor producción por parte del sociolecto femenino, las diferencias se reducen en los centros donde los informantes responden en su lengua materna. De este modo, en Partes del cuerpo (en italiano), la diferencia se reduce a 2,22 frente a las 4,44 que se dan en el homónimo en español. Mientras que en el centro Comidas y bebidas (en italiano), la diferencia es de 4,5 frente al 5,61 de su homónimo en español.

Por otro lado, al igual que en los estudios de Fernández Juncal (2013) o Gallego Gallego (2014), hemos recurrido al programa SPSS para comprobar la significación estadística de los resultados. Para ello, hemos hecho uso del análisis de varianza (ANOVA) de una vía. Esta prueba permite «refutar o ratificar la hipótesis nula que asigna un valor de  $p < 0,05$ » (Gallego Gallego 2014, 274). En otras palabras, esta prueba nos permite hallar el valor de  $p$  para determinar si los resultados son estadísticamente significativos, es decir, que no se deben al azar. De este modo, si el valor de  $p$  es inferior a 0,05, la hipótesis nula queda refutada y, por tanto, los resultados obtenidos son significativos. En cambio, si el valor de  $p$  es igual o superior a 0,05, debemos aceptar la hipótesis nula, o lo que es lo mismo, que los resultados se deben a causas accidentales (cf. tab. 3<sup>5</sup>).

Tabla 3. Valor de  $p$  en los centros de interés según variable sexo

Nº	CI	Valor de $p$
1	CUE	0,00050*
2	ROP	0,00673*
3	ESC	0,00262*
4	ACC	0,00010*
5	ESP	0,10850
6	COM	0,00344*
7	ANI	0,04641*
8	JUE	0,02051*
9	ASP	0,00001*
10	LEN	0,00231*
11	PRO	0,01809*
12	ITA	0,00165*
13	CUE ITA	0,05241
14	COM ITA	0,01900*

5 Se señalan con asterisco los centros que demuestran la relevancia de la variable.



En cuanto a la variable sexo, los trabajos de Galloso Camacho (2002), Gómez Devís (2004) y Hernández Muñoz (2006) muestran que las diferencias entre los promedios de las producciones léxicas de los sociolectos masculino y femenino tienen significación estadística. En nuestro trabajo, las diferencias sí son significativas en doce de los catorce centros. Tan solo en 5. España y los españoles y 13. Partes del cuerpo (en italiano) debemos considerar que las diferencias se deben a causas accidentales.

### 3.2.2 Variable nivel sociocultural

Las pautas seguidas para la codificación de esta variable aparecen detalladas en el trabajo de Fernández Juncal (2008, 21). Teniendo en cuenta los estudios y las profesiones de los padres se halla el nivel sociocultural. En nuestro estudio, la distribución de los encuestados queda de la siguiente manera: 14 informantes de nivel sociocultural bajo, 88 de nivel medio-bajo, 59 de nivel medio-alto y 13 de nivel alto.

La mayoría de los estudios de disponibilidad léxica afirman que existe una relación entre los resultados lingüísticos y el nivel sociocultural: los informantes ofrecen mayor número de respuestas cuanto mayor es su nivel sociocultural. Sin embargo, al igual que ocurre en el estudio realizado por Rubio Sánchez (2016, 117-8) con estudiantes portugueses, los resultados obtenidos con informantes italianos difieren de los obtenidos en las investigaciones del Proyecto Panhispánico. En nuestro caso, tenemos que el número medio de palabras por informante para el nivel sociocultural bajo es de 14,36, que aventaja al nivel medio-bajo y al medio-alto que tienen una media de 14,17 y 14,02 respectivamente. Todos ellos son superados por el nivel alto que posee una media de 15,88 palabras por informante (cf. tab. 4 y fig. 2).

Si analizamos cada uno de los centros de interés, vemos que el nivel alto presenta mayor producción léxica en doce de los catorce centros. Sin embargo, en Animales y Comidas y bebidas (en italiano), el nivel bajo es el que muestra mayor productividad léxica.

Tabla 4. Promedio de respuestas según el nivel sociocultural

N°	CI	Bajo	Medio-Bajo	Medio-Alto	Alto
1	CUE	13,29	13,74	12,62	15,85
2	ROP	9,07	9,69	10,31	12,15
3	ESC	13,07	13,52	13,71	15,38
4	ACC	14,57	13,59	13,6	14,62
5	ESP	13,43	13,55	13,6	15,31
6	COM	14,93	14,57	14,76	16,92
7	ANI	12,71	10,69	10,02	10,77
8	JUE	9,86	9,38	9,69	11,31
9	ASP	12,14	12,77	12,74	14,23
10	LEN	11,29	11,47	10,55	14,23
11	PRO	11,57	11,72	11,83	12,15
12	ITA	13,21	14,24	13,74	16,15
13	CUE ITA	25,29	24,91	25,24	26,92
14	COM ITA	26,64	24,49	23,9	26,31

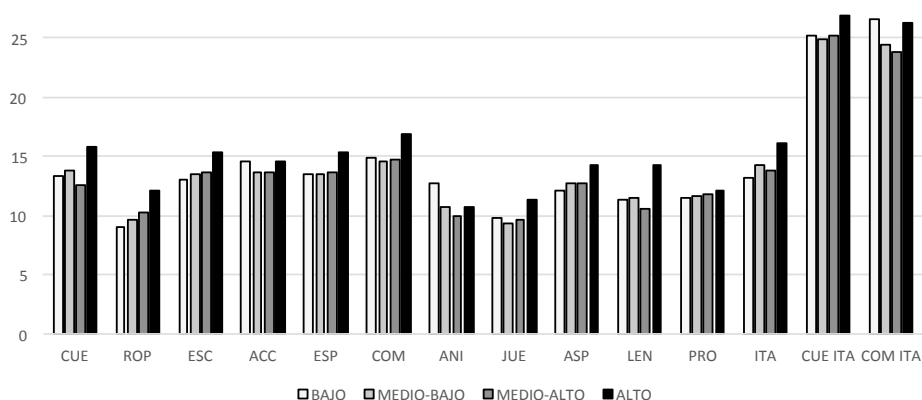


Figura 2. Palabras por informante según el nivel sociocultural

Para comprobar si estas diferencias tienen significación estadística, procedemos a comparar la diferencia entre las medias de los cuatro grupos. Con este propósito, utilizamos nuevamente el análisis de varianza (ANOVA) de una vía (cf. tab. 5). De esta forma, tenemos cinco centros donde las diferencias entre el promedio de respuestas de los diferentes sociolectos son significativas: Partes del cuerpo, La ropa, Comidas y bebidas, Juegos y diversiones y La lengua española.

Tabla 5. Valor de p en los centros de interés según variable nivel sociocultural

Nº	CI	Valor de p
1	CUE	0,01120*
2	ROP	0,01229*
3	ESC	0,31964
4	ACC	0,41960
5	ESP	0,28313
6	COM	0,03681*
7	ANI	0,50904
8	JUE	0,04924*
9	ASP	0,14450
10	LEN	0,01249*
11	PRO	0,87546
12	ITA	0,16480
13	CUE ITA	0,33117
14	COM ITA	0,32065

### 3.2.3 Variable titularidad del centro

Mateo García (1998, 53) afirma que «implícitamente se considera que entre las instituciones educativas públicas y privadas existen diferencias significativas». Además, como señala López González (2014, 116) tras analizar muchas de las investigaciones realizadas en España, «los resultados casi siempre coinciden en señalar un mayor bagaje léxico para los alumnos de instituciones privadas». Sin embargo, en nuestra muestra, que presenta 52 informantes procedentes de centro público frente a 121 de centro privado, tenemos que el promedio de palabras por informante de centro público supera en más de cinco respuestas el resultado obtenido por los informantes de centro privado: 17,22 frente a 12,99. Esto supone un 24,56% de diferencia en favor de los informantes de escuela pública (cf. tab. 6 y fig. 3). Por tanto, al contrario de lo que ocurre en los estudios desarrollados en España, nuestra investigación nos muestra que los encuestados pertenecientes a centros públicos italianos presentan mayor producción léxica que los relacionados con las instituciones privadas.

Tabla 6. Promedio de respuestas según la titularidad del centro

Nº	CI	Público	Privado
1	CUE	17,54	11,74
2	ROP	12,77	8,86
3	ESC	16,52	12,47
4	ACC	16,33	12,64
5	ESP	15,37	12,97
6	COM	19,9	12,66
7	ANI	15,98	8,34
8	JUE	11,1	9,05
9	ASP	16,1	11,41
10	LEN	14,04	10,2
11	PRO	14,92	10,42
12	ITA	16,83	12,98
13	CUE ITA	26,63	24,59
14	COM ITA	27,12	23,52

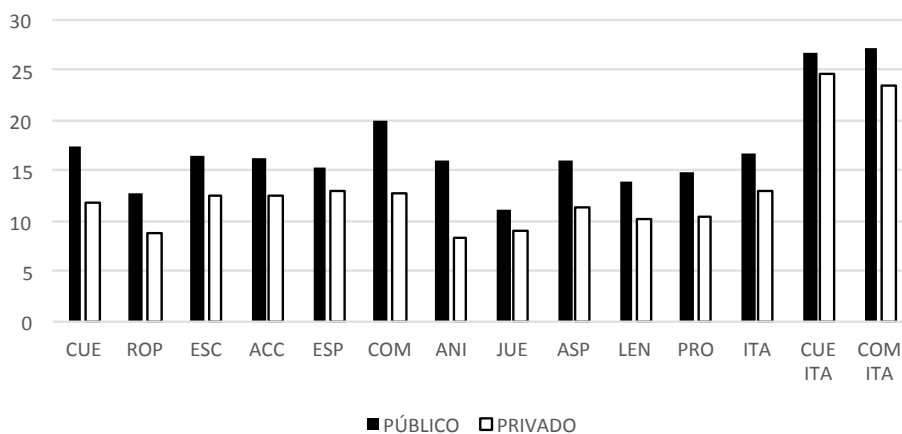


Figura 3. Palabras por informante según la titularidad del centro

Para comprobar si la superioridad léxica de los centros públicos es estadísticamente significativa, recurrimos nuevamente al análisis de varianza de una vía. Al realizar la comparación entre las medias con ANOVA, los resultados nos muestran una significación de  $p < 0,05$  en 13 de los 14 centros de interés (cf. tab. 7). Por tanto, estas diferencias sí son significativas. Solamente, el campo Juegos y diversiones presenta una  $p > 0,05$ , por lo que las diferencias presentadas en este centro de interés se deben a causas accidentales.

Tabla 7. Valor de p en los centros de interés según variable titularidad del centro

Nº	CI	Valor de p
1	CUE	0,00000*
2	ROP	0,00005*
3	ESC	0,00013*
4	ACC	0,00004*
5	ESP	0,00738*
6	COM	0,00000*
7	ANI	0,00000*
8	JUE	0,09160
9	ASP	0,00126*
10	LEN	0,01243*
11	PRO	0,00000*
12	ITA	0,01822*
13	CUE ITA	0,07944*
14	COM ITA	0,02956*

### 3.2.4 Variable nivel de español

Esta variable presenta cuatro variantes: estudiantes de nivel inicial (A), de nivel intermedio (B1), de nivel avanzado (B2) y estudiantes que no supieron decir su nivel de español (NSNC: no sabe/no contesta). La muestra que recogemos en estas páginas está constituida por 13 informantes de nivel A, 109 de nivel B1, 36 de nivel B2 y 15 NSNC. Como era lógico, los informantes de nivel A (11,99) poseen menos promedio de respuestas que los de B1 (13,99), que a su vez presentan menos contestaciones que los de nivel B2 (16,09). En cuanto a la categoría NSNC, el promedio de respuestas es de 13,93, lo que la sitúa entorno al nivel B1 (cf. tab. 8 y fig. 4).

Tabla 8. Promedio de palabras según la variable nivel de español

Nº	CI	A	B1	B2	NSNC
1	CUE	10,36	13,01	16,17	13,4
2	ROP	8,57	9,91	11,36	9,13
3	ESC	10,57	13,91	14,47	13,13
4	ACC	12,5	13,19	15,31	15,27
5	ESP	11,57	13,49	14,75	14,53
6	COM	12,93	14,21	17,69	14,27
7	ANI	9,14	10,42	12,5	9,13
8	JUE	7,71	9,3	11,17	10,53
9	ASP	9,93	12,29	15,97	11,8
10	LEN	7,57	11,24	13,56	10,4
11	PRO	9,43	11,48	13,56	11,8
12	ITA	9,93	13,96	16,5	13,6
13	CUE ITA	24,14	24,92	26,5	25,13
14	COM ITA	23,43	24,6	25,75	22,93

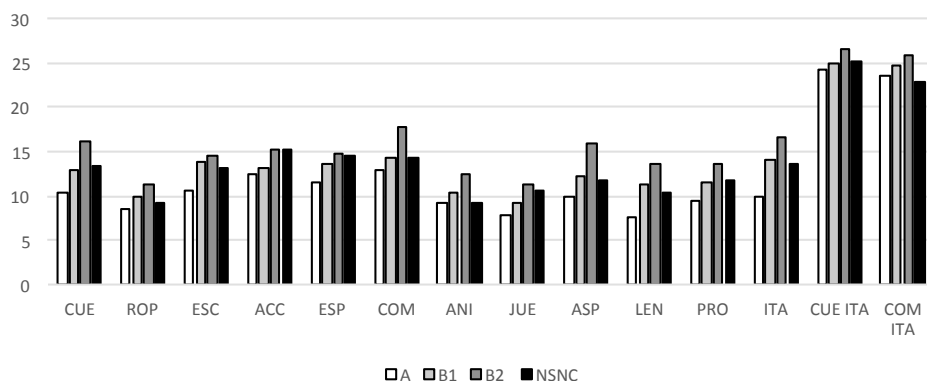


Figura 4. Palabras por informantes según el nivel de español

Para realizar el análisis estadístico de esta variable, hemos optado por la eliminación del grupo NSNC. Esto se debe a que los datos podrían verse alterados, puesto que los integrantes de este grupo deberían estar asociados a uno de los niveles, pero no han sabido decir a cuál. De este modo, al aplicar la prueba ANOVA de una vía - teniendo en cuenta el promedio de respuestas de los grupos A, B1 y B2 - vemos diferencias significativas en 6 de los 14 centros de interés: Partes del cuerpo, Comidas y bebidas, Aspecto físico y carácter, La lengua española, Profesiones y Italia y los italianos. Para los demás centros, las diferencias no resultan significativas (cf. tab. 9).

Tabla 9. Valor de p en los centros de interés según variable nivel de español

Nº	CI	Valor de p
1	CUE	0,00067*
2	ROP	0,23719
3	ESC	0,10185
4	ACC	0,05992
5	ESP	0,14435
6	COM	0,01469*
7	ANI	0,14875
8	JUE	0,06760
9	ASP	0,00028*
10	LEN	0,00878*
11	PRO	0,01545*
12	ITA	0,00218*
13	CUE ITA	0,15252
14	COM ITA	0,38843

#### 4 Datos cualitativos. Partes del cuerpo (en español y en italiano)

Debido a limitada extensión de este artículo, en este punto nos centraremos solamente en uno de los centros de interés respondidos tanto en español como en italiano: Partes del cuerpo.

Para analizar cualitativamente los listados obtenidos, tomamos como referencia los dos criterios señalados por López González (2014, 156-7): la descripción lexicológica de las unidades que componen los listados y el valor del índice de disponibilidad léxica de cada vocablo.

De este modo, al analizar las categorías gramaticales que aparecen en los listados, tenemos que el 70% de los vocablos del centro respondido en español son sustantivos, mientras que el 13% son adjetivos, el 12,5% son verbos y el 4,5% son construcciones sintagmáticas. En cambio, el centro de interés contestado en italiano presenta un 87,1% de sustantivos, un 6,1% de construcciones sintagmáticas, un 3,8% de verbos y un 3% de adjetivos.

Por otro lado, como sugieren Ávila Muñoz y Villena Ponsoda (2010, 183-184), podemos clasificar las unidades léxicas de los listados teniendo en cuenta el grado de compatibilidad con los centros de interés.

De este modo, los 20 vocablos que aparecen en la tabla 10 representan los 'elementos nucleares' del centro Partes del cuerpo (en español y en italiano). Así denominan los estudiosos malagueños a los ejemplares prototípicos, es decir, aquellos que ocupan las posiciones más altas en los listados de disponibilidad.

El segundo grupo lo forman las 'asociaciones denotacionales', elementos que forman parte del centro de interés - sin ser nucleares - y que aparecen

por asociaciones o evocaciones individuales pero de carácter colectivo. En el centro de interés Partes del cuerpo serían vocablos como: *tacto, gusto, vista*, etc.

Las ‘derivaciones’ constituyen el tercer bloque, que integra aquellos elementos que ya han sido mencionados pero presentan algún aumento cuantitativo o cualitativo. En Partes del cuerpo se pueden observar los siguientes: *dedo de la mano, dedo del pie, pelo largo, pelo corto*, etc.

El último grupo está formado por las «asociaciones individuales». Estos elementos no tienen relación directa con el centro de interés y están alejados de los prototípicos nucleares. En el caso de Partes del cuerpo tenemos: *medicina, doctor, enfermedad, pendiente, gafas*, etc.

Tabla 10. Los 20 vocablos con mayor disponibilidad en Partes del cuerpo (en español y en italiano)

Rango	Partes del cuerpo	Partes del cuerpo (en italiano)
1	ojo	<i>occhio</i> ‘ojo’
2	mano	<i>mano</i> ‘mano’
3	cabeza	<i>naso</i> ‘nariz’
4	boca	<i>gamba</i> ‘pierna’
5	pierna	<i>bocca</i> ‘boca’
6	pie	<i>dito</i> ‘dedo’
7	nariz	<i>braccio</i> ‘brazo’
8	pelo	<i>piede</i> ‘pie’
9	brazo	<i>orecchio</i> ‘oreja’
10	dedo	<i>testa</i> ‘cabeza’
11	espalda	<i>capello</i> ‘pelo’
12	oreja	<i>unghia</i> ‘uña’
13	diente	<i>ginocchio</i> ‘rodilla’
14	cara	<i>spalla</i> ‘hombro’
15	corazón	<i>collo</i> ‘cuello’
16	lengua	<i>schiena</i> ‘espalda’
17	uña	<i>pancia</i> ‘barriga’
18	rodilla	<i>gomito</i> ‘codo’
19	cuello	<i>dente</i> ‘diente’
20	barriga	<i>sopracciglio</i> ‘ceja’

Si analizamos los 20 vocablos más disponibles de ambos centros, vemos que el 85% de las unidades coinciden. Así, las palabras que se repiten son: *ojo, mano, cabeza, boca, pierna, pie, nariz, pelo, brazo, dedo, espalda, oreja, diente, uña, rodilla, cuello y barriga*. El resto, es decir, el 15% de las palabras, solamente aparecen en una columna: *cara, corazón, lengua, hombro, codo y ceja*.

Al buscar los vocablos del centro completado en español en el Plan Curricular (PCIC en adelante) del Instituto Cervantes (2006), observamos



que el 65% están clasificados como nivel A (*ojo, mano, cabeza, pierna, pie, nariz, pelo, brazo, dedo, espalda, cara, boca, diente*), 20% como B1 (*corazón, rodilla, cuello, barriga*) y 5% como B2 (*uña*). El otro 10% (*lengua y oreja*) no aparecen en el PCIC. Por tanto, los informantes que han declarado tener un nivel A dominan palabras de nivel superior.

Por otro lado, si tenemos en cuenta todo el listado de léxico disponible de Partes del cuerpo, los informantes han actualizado unidades léxicas (*índice, ombligo, palma de la mano, pulgar y vientre*) que el PCIC ha clasificado en el nivel C, nivel superior al de nuestros informantes. Este hecho nos confirma que los informantes italo-fonos actualizan, al menos en este centro de interés, palabras de nivel superior.

## 5 Conclusiones

En estas páginas hemos presentado un estudio de disponibilidad léxica en estudiantes italianos preuniversitarios de ELE. Tras mostrar brevemente la metodología seguida en la investigación, hemos realizado - con la ayuda de Dispolex y el paquete estadístico SPSS - un análisis cuantitativo de las variables consideradas en este trabajo. Los datos nos han revelado que las variables donde se dan más diferencias significativas son la variable sexo y la titularidad del centro.

Por otro lado, dadas las limitaciones de este artículo, nos hemos centrado en el análisis cualitativo de uno de los centros de interés respondidos en español y en italiano: Partes del cuerpo. En este punto hemos analizado las unidades léxicas que forman los listados de léxico disponible de ambos centros de interés y las hemos clasificado según la categoría gramatical. Hemos constatado que la gran mayoría de vocablos que actualizan los informantes en estos centros son sustantivos. Además, hemos clasificado los vocablos según su compatibilidad con los centros de interés siguiendo la clasificación de Ávila Muñoz y Villena Ponsoda (2010).

Por último, hemos buscado los 20 vocablos más disponibles en el listado de Partes del cuerpo en español en el PCIC. Este análisis nos ha permitido observar que los informantes italo-fonos dominan unidades léxicas de un nivel superior al que poseen.

## Bibliografía

- Ávila Muñoz, Antonio Manuel; Villena Ponsoda, Juan Andrés (2010). *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Sarriá.
- Bartol Hernández, José Antonio (2004). *Léxico disponible de Soria. Estudio y diccionarios*. Burgos: Instituto Castellano y Leonés de la Lengua.
- Bartol Hernández, José Antonio (2006). «La disponibilidad léxica». *Revista Española de Lingüística*, 36, 379-91.
- Carcedo González, Alberto (2000). *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Turku: Turun Yliopisto.
- del Barrio de la Rosa, Florencio (2016). «Algunas observaciones sobre la disponibilidad léxica en estudiantes italofonos de español» [online]. Sainz González, Eugenia et al. (eds.), *Geométrica explosión: Estudios de lengua y literatura en homenaje a René Lenarduzzi*. Venezia: Edizioni Ca' Foscari, 127-43. DOI 10.14277/6969-068-6/RiB-1-8.
- Fernández Juncal, Carmen (2008). *Léxico disponible de Burgos*. Burgos: Instituto Castellano y Leonés de la Lengua.
- Fernández Juncal, Carmen (2013). *Léxico disponible de Cantabria*. Salamanca: Ediciones Universidad de Salamanca.
- Fernández Smith, Gérard et al. (2008). *Léxico disponible de Melilla: estudio sociolingüístico y repertorios léxicos*. Madrid: Arco Libros.
- Gallego Gallego, Diego Javier (2014). *Léxico disponible de estudiantes de español como lengua extranjera en la Comunidad de Madrid* [tesis doctoral]. Alcalá de Henares: Universidad de Alcalá.
- Gamazo Carretero, Elena (2014). «Estereotipos en el léxico disponible de universitarios portugueses» [online]. *El español como lengua extranjera en Portugal: retos de la enseñanza de lenguas cercanas*, 28-41. URL <http://goo.gl/CVjIKy> (2016-10-31).
- Galoso Camacho, María Victoria (2002). *El léxico disponible en el nivel preuniversitario. Provincia de Zamora*. Huelva: Universidad de Huelva.
- Gómez Devís, María Begoña (2004). *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: metodología, análisis sociolingüístico y aplicaciones* [Tesis Doctoral] [online]. Valencia: Universidad de Valencia. URL <http://roderic.uv.es/bitstream/handle/10550/15248/gomez.pdf?sequence=1> (2017-05-18).
- Hernández Muñoz, Natividad (2006). *Hacia una teoría cognitiva integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad de Salamanca.
- Instituto Cervantes (2006). *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español* [online]. Madrid: Instituto Cervantes - Biblioteca Nueva. URL [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/plan\\_curricular/default.htm](http://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/default.htm) (2016-10-31).

- Mateo García, María Victoria (1998). *Disponibilidad léxica en el COU almeriense. Estudio de estratificación social*. Granada: Universidad de Granada.
- Lin, Jing (2012). «La disponibilidad léxica en los estudiantes chinos de Español como Lengua Extranjera». *MarcoELE*, 14. URL [http://marcoele.com/descargas/14/lin-disponibilidad\\_lexica.pdf](http://marcoele.com/descargas/14/lin-disponibilidad_lexica.pdf) (2016-10-31).
- López González, Antonio María (2009). «La disponibilidad léxica en las secciones bilingües de español de Polonia». Barrientos Clavero, Agustín et al. (eds.), *El profesor de español LE/L2 Actas del XIX Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera (ASELE)* (Cáceres, 24-27 de septiembre de 2008). Cáceres: Universidad de Extremadura, 567-81.
- López González, Antonio María (2014). *Disponibilidad léxica. Teoría, método y análisis*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Rodríguez Menduïña, Pompeyo (2010). *Estereotipos culturales sobre España en estudiantes estadounidenses desde la disponibilidad léxica* [Tesis doctoral]. Salamanca: Universidad de Salamanca.
- Rubio Sánchez, Roberto (2016). «Estudio de disponibilidad léxica en 60 aprendices portugueses de español». Domínguez, Noemí; Fernández, Carmen; García, Juan Luis (eds.), *Innovación y desarrollo en Español como Lengua Extranjera*. Salamanca: Ediciones Universidad de Salamanca, 107-28.
- Samper Hernández, Marta (2002). *Disponibilidad léxica en alumnos de español como lengua extranjera*. Málaga: ASELE.
- Samper Padilla, José Antonio; Samper Hernández, Marta (2006). «Aportaciones recientes de los estudios de disponibilidad léxica». *LynX: Panorámica de Estudios Lingüísticos*, 5, 5-95.
- Sánchez-Saus Laserna, Marta (2016). *Léxico disponible de los estudiantes de español como lengua extranjera en las universidades andaluzas*. Sevilla: Universidad de Sevilla.
- Schmidely, Jack (coord.) (2001). *De una a cuatro lenguas. Del español al portugués, al italiano y al francés*. Madrid: Arco libros.

