



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

## Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in  
Lingue Moderne per la Comunicazione e la Cooperazione Internazionale  
Classe LM-38

Tesi di Laurea

### *Estudio sobre disponibilidad léxica en alumnos de ELE en la Universidad de Padua*

Relatore  
Prof. Carmen Castillo Peña

Laureando  
Alessia Paolini  
n° matr.1082350 / LMLCC

Anno Accademico 2016 / 2017



# Agredecimientos

Mis sinceros agradecimientos van dirigidos a:

Mi tutora, la profesora Carmen Castillo Peña, por su apoyo constante, inmensa dedicación y por haberme dado una gran inyección de confianza.

José Antonio Bartol Hernández por su ayuda en relación con el programa informático *DispoLex*.

Los alumnos del curso de ELE por haber participado en la realización del estudio.

Mis padres, mi hermana y mi familia por haber tenido fe en mí siempre a lo largo de este camino lleno de avatares.

Davide, mi fuerte compañero de aventuras, por su paciencia inagotable y por haberme soportado, en todos los sentidos.

Carla, Jennifer y Martina R. por haber sido mi fuente de motivación durante todos estos meses, por las sonrisas, la alegría y la amistad fiel que nos une.

Martina D., mi amiga por correspondencia, conocida gracias a esta tesis.

Mis ex compañeras de trabajo, que me han animado a no darme por vencida.

# Índice

<b>Agradecimientos</b>	<b>i</b>
<b>Introducción</b>	<b>1</b>
<b>BLOQUE I: MARCO TEÓRICO</b>	<b>7</b>
<b>Capítulo I: cuestiones preliminares</b>	<b>7</b>
1.1 Palabras y vocablos	7
1.2 Lexicón mental y competencia léxica	9
1.3 Léxico y frecuencia	10
1.4 Léxico disponible	12
1.5 Léxico común	14
1.6 Léxico fundamental	15
1.7 Estadística y lingüística de corpus	15
1.7.1 Estudios estadísticos de léxico y diccionarios léxicos	15
1.7.2 Lingüística de corpus y corpus del español en la actualidad	17
<b>Capítulo II: estudios de disponibilidad léxica y cuestiones metodológicas</b>	<b>21</b>
2.1 Estudios cuantitativos sobre léxico	21
2.2 Estudios de disponibilidad léxica en hablantes nativos de español	21
2.3 El Proyecto Panhispánico	22
2.4 Los centros de interés	23
2.5 La prueba	24
2.6 El número de encuestados	24
2.7 Condicionantes extralingüísticos	25
2.7.1 Factor sexo	25
2.7.2 Edad	25
2.7.3 Nivel sociocultural	26
2.7.4 Zona geográfica y tipo de centro	26
2.8 Estudios de disponibilidad léxica en ELE	27
2.9 Condicionantes extralingüísticos en los estudios de disponibilidad léxica en estudiantes de ELE	28
2.10 Gradación de la disponibilidad	28
2.11 Aplicaciones de los trabajos de disponibilidad léxica	29
2.11.1 Dialectología	30
2.11.2 Sociolingüística	30
2.11.3 Psicolingüística	31
2.11.4 Etnolingüística	31
2.11.5 Didáctica de lenguas	32
<b>BLOQUE II: PRESENTACIÓN Y ANÁLISIS DE LOS DATOS</b>	<b>35</b>
<b>Capítulo III: Investigación de disponibilidad léxica en alumnos de ELE en la Universidad de Padua</b>	<b>35</b>
3.1 Objetivos del trabajo	35
3.2 La muestra	36

3.3	Condicionantes extralingüísticos considerados	37
3.3.1	Variabes ‘sexo’ y ‘edad’	38
3.3.2	Variable ‘carrera de estudios’ y ‘zona geográfica de procedencia’	39
3.3.3	Variable ‘nivel sociocultural’	41
3.3.4	Variabes ‘lengua materna’ y ‘conocimiento de otras lenguas’	41
3.3.5	Variable ‘años de estudio del español’	43
3.4	La encuesta	44
3.5	Edición de errores	46
3.6	Toma de decisiones	48
3.6.1	‘Las partes del cuerpo’	48
3.6.2	‘La ropa’	49
3.6.3	‘Las partes de la casa’	50
3.6.4	‘Comidas y bebidas’	51
3.6.5	‘La escuela’	52
3.6.6	‘La ciudad’	53
3.6.7	‘Animales’	53
3.6.8	‘Profesiones y oficios’	54
<b>Capítulo IV:</b>	<b>Análisis y comentario de resultados cuantitativos</b>	<b>57</b>
4.1	Resultados cuantitativos generales	57
4.2	Resultados de la encuesta según las variables analizadas	62
4.2.1	Promedio de palabras según las variables analizadas	62
4.2.2	Producción de vocablos según las variables analizadas	73
<b>Capítulo V:</b>	<b>Análisis y comentario de resultados cualitativos</b>	<b>79</b>
5.1	Palabras más disponibles por cada centro de interés	79
5.1.1	‘Partes del cuerpo’	79
5.1.2	‘La ropa’	81
5.1.3	‘Partes de la casa’	83
5.1.4	‘Comidas y bebidas’	84
5.1.5	‘La escuela’	86
5.1.6	‘La ciudad’	88
5.1.7	‘Animales’	90
5.1.8	‘Profesiones y oficios’	92
<b>Conclusiones</b>		<b>97</b>
<b>Anexo I</b>		<b>103</b>
<b>Anexo II</b>		<b>105</b>
<b>Bibliografía</b>		<b>141</b>
<b>Sitios web consultados</b>		<b>145</b>
<b>Riassunto</b>		<b>147</b>

# Introducción

A mediados del siglo pasado en Francia surgieron los primeros estudios sobre disponibilidad léxica, una corriente lingüística que tenía como objetivo primario el perfeccionamiento de la selección del vocabulario francés en una comunidad de hablantes no nativos para la sucesiva elaboración de programas de enseñanza de lengua adecuados. En el ámbito hispánico el impulsor fue Humberto López Morales, el cual con su investigación sobre el léxico disponible de una comunidad de hablantes en Puerto Rico realizó el primer trabajo de este tipo en la lengua española y puso en marcha el Proyecto Panhispánico sobre disponibilidad léxica, en el cual confluyen las investigaciones comunes a España e Hispanoamérica a fin de identificar el léxico disponible de estos hablantes y elaborar diccionarios de léxico que se puedan aplicar a diferentes ámbitos de estudio como el de la dialectología, la sociolingüística, la etnolingüística y la didáctica de lenguas.

A pesar de que los principales trabajos sobre la lengua española se hayan enfocado en su mayoría sobre el estudio del léxico disponible de comunidades de hablantes nativos de español, en las últimas décadas se han llevado a cabo también un buen número de investigaciones sobre estudiantes de español como lengua extranjera en los distintos países donde aprenden el idioma. Al considerar los trabajos realizados, no hemos detectado la presencia de investigaciones sobre este tema en Italia, lo cual nos ha impulsado a acercarnos a estos estudios.

En especial modo, el presente trabajo está basado sobre el material léxico recogido mediante una prueba de disponibilidad léxica realizada a finales del mes de octubre 2016, en la cual ha participado una muestra no preestratificada de alumnos de español como lengua extranjera de la Universidad de Padua. La mayoría de estos estudiantes proceden del II año de grado en “Mediación Lingüística y Cultural” y “Lenguas, Literaturas y Culturas Modernas” que han cursado las clases de la asignatura de español de la profesora tutora de esta tesis. Ellos han constituido una muestra de 112 informantes, limitada numericamente en comparación con las investigaciones estándar sobre hablantes nativos, pero que se pone en línea con nuestro propósito de aportar una contribución original, de carácter introductorio, dentro del panorama de investigación

sobre disponibilidad léxica con un enfoque sobre el vocabulario de alumnos italianos matriculados en la universidad.

Los objetivos específicos de esta investigación han tenido en cuenta los siguientes aspectos:

- 1 Aportar la cantidad absoluta y relativa de respuestas de los informantes en cada centro de interés.
- 2 Aportar la cantidad absoluta y relativa de vocablos en cada centro de interés.
- 3 Observar la influencia de las variables en cada centro de interés tanto en relación con el número de palabras como en la cantidad de vocablos.
- 4 Aportar los datos cuantitativos correspondientes a los índices de media de palabras por informante e índice de cohesión en cada centro de interés.
- 5 Enumerar los vocablos que ocupan los primeros lugares en cada centro de interés y realizar un análisis descriptivo de las unidades léxicas aportadas en cada campo léxico considerado.
- 6 Adjuntar en los anexos de este trabajo los vocabularios resultantes por cada centro de interés, según su índice de disponibilidad, aparición, frecuencia relativa y acumulada.
- 7 Debido al carácter introductorio de esta investigación, se optó por renunciar a establecer una segunda muestra para comparar los resultados ya que este objetivo se podrá perseguir en investigaciones sucesivas a la que se presenta en estas páginas.

Para alcanzar nuestros propósitos hemos seguido la metodología que se suele utilizar en las investigaciones sobre este campo de estudio, las cuales utilizan técnicas de la psicología asociativa para determinar el lexicón mental de un hablante y establecer su grado de competencia léxica. Los estudiantes, guiados por un encuestador, tienen que llevar a cabo una prueba en la cual se les pide actualizar en un tiempo cerrado de dos minutos todas las palabras que se les ocurren acerca de un determinado tema. Por lo tanto, las palabras más disponibles de un hablante serán aquellas que acuden de manera más rápida a su mente y aparecerán entre sus primeras respuestas, mientras que desde una perspectiva colectiva, serán las que se encontrarán en los primeros lugares de la

mayoría de los encuestados según un cálculo estadístico que tiene en cuenta la posición relativa de las palabras en el listado y su frecuencia, es decir el índice de disponibilidad.

Además, se han tomado en consideración también algunos datos personales de los participantes en la muestra, recogidos mediante una ficha personal en la cual hemos pedido que se proporcionase información sobre aquellos campos que sirven para determinar los factores extralingüísticos normalmente investigados en los estudios sobre disponibilidad léxica en nativos, con la adición de elementos específicos típicos de la metodología de estudio sobre no nativos.

Tras haber obtenido el material, hemos procedido con el trabajo de transcripción y edición de datos, respetando las normas comunes a los trabajos del Proyecto Panhispánico así como los criterios de las investigaciones sobre no nativos para futuros cotejos con nuevas investigaciones. Sucesivamente, hemos introducido todos los datos recogidos dentro de la herramienta informática Dispalex, un programa de cálculo de disponibilidad léxica disponible en internet y dirigido por la Universidad de Salamanca. Si bien este soporte ofrece numerosas posibilidades de cálculo y cotejo, en nuestra investigación nos hemos limitado a presentar algunas; en futuro será posible profundizar el estudio y las comparaciones con otras muestras.

A partir de estas bases, hemos organizado el presente trabajo en dos partes: la primera se ocupa de la tratación general sobre el marco teórico de los estudios de disponibilidad léxica, la segunda está vinculada con el aporte más innovador de nuestro estudio ya que está dedicada a la presentación y al análisis de los datos resultantes.

En el primer capítulo de esta investigación empezamos marcando algunas nociones de carácter conceptual en relación con el campo de la léxico-estadística y aportamos información sobre las distinciones terminológicas que se establecen, las bases psicolingüísticas sobre las cuales se funda el modelo teórico, los diferentes tipos de léxico que existen en la lengua española hasta llegar al planteamiento de qué es el léxico disponible de un hablante. Sigue en el mismo capítulo una descripción sobre los estudios relacionados con la ligüística estadística: los diccionarios de frecuencia y los corpus informáticos del español, que suscitan gran interés pedagógico para el estudio y la enseñanza del léxico.

A continuación, en el segundo capítulo exponemos el origen y el desarrollo de los estudios cuantitativos sobre léxico, haciendo hincapié en las investigaciones que



constituyen la base teórica principal de estos tipos de estudios. Además, abordamos los trabajos realizados en España en hablantes nativos y nos detenemos en las normas que a partir de 1999 han aportado relevantes indicaciones acerca de la metodología común de investigar sobre el léxico disponible en el ámbito hispánico. A través de estas pautas se estableció la caracterización de los dieciséis centros de interés que se pueden utilizar como estímulo para llevar a cabo la prueba, las indicaciones sobre la preparación y la realización de la encuesta, el número de participantes necesarios y cuales condicionantes extralingüísticos hay que considerar. La tratación continua con la exposición de los trabajos sobre disponibilidad léxica en ELE, algunos de los cuales se han servido de los datos de estudiantes nativos para después aplicar los resultados a la enseñanza de ELE, mientras que otros han empleado los métodos comunes a las investigaciones hispánicas para el estudio del léxico de alumnos de español como lengua extranjera en diferentes naciones. El capítulo se cierra con la descripción de la metodología de cálculo empleada y ofrece un breve enfoque sobre las diferentes áreas de aplicación de la disponibilidad léxica.

En el capítulo tres introducimos la metodología seguida en nuestra investigación para la realización de la prueba en clase con los 112 participantes y describimos las decisiones tomadas en relación con el establecimiento de las variables extralingüísticas consideradas y los correspondientes resultados, que presentamos con el soporte de algunos gráficos. Además, explicamos detalladamente qué tipo de encuesta hemos llevado a cabo y cómo se ha realizado, ponemos de manifiesto cuales centros de interés se han seleccionado y las razones. Por último, dedicamos un apartado a la descripción de las decisiones que se han tomado a la hora de editar y limpiar el material recogido mediante la prueba en los distintos campos léxicos encuestados.

El capítulo sucesivo aporta los datos cuantitativos resultantes de las pruebas de disponibilidad léxica de nuestros estudiantes desde una perspectiva general, mediante la descripción de los datos sobre el número de palabras totales, el promedio de palabras por informante en cada centro de interés, los vocablos totales y el índice de cohesión. Asimismo, se pone de manifiesto las veinte primeras palabras más disponibles que los informantes han actualizado y se procede con una presentación detallada de los resultados por promedio de palabras en cada campo léxico según cada factor extralingüístico considerado en la encuesta. Finalmente, mediante el soporte de tablas

númerica se exponen los datos referidos el número total de vocablos respondidos en cada centro de interés en relación con las variables analizadas.

En el capítulo final se analizan los centros de interés considerados en nuestra investigación mediante una descripción cualitativa de cada campo léxico con el objetivo de registrar las diferentes agrupaciones de palabras que constituyen los pilares de nuestro trabajo, es decir los centros de interés, y detectamos aquellos fenómenos lingüísticos que destacan las características cualitativas aportadas por nuestros informantes desde el punto de vista de la variación léxica, los cambios semánticos, los casos de importación léxica y los mecanismos de formación y derivación léxica.

Por último, pero no menos importante, completamos la investigación con la presentación del material a partir del cual ha sido posible empezar y llevar a cabo esta contribución experimental. En el anexo I de la investigación se añade el modelo de encuesta que ha sido empleado y entregado a los alumno para la aportación de sus datos personales útiles para el estudio según los factores extralingüísticos. En cambio, en el anexo II se ajuntan todos los listados de palabras actualizados por los informantes, editados y agrupados por centro de interés según el índice de disponibilidad, aparición, frecuencia relativa y acumulada. Estas unidades léxicas componen los vocabularios de disponibilidad, herramientas que en posteriores trabajos o actividades didácticas podrán servir como soporte al aprendizaje y enseñanza de léxico.

Esperamos que la investigación que hemos presentado, aun de carácter experimental e introductorio, cumpla los propósitos que hemos expuesto al principio y pueda representar un primer acercamiento a este tipo de estudio que lleve a una profundización del tema de disponibilidad léxica aplicada a ELE. Estamos seguros de que gracias a los datos obtenidos a partir de los recuentos estadísticos es posible hacer patente cómo influyen en el conocimiento léxico de los alumnos de español como lengua extranjera tanto los estímulos proporcionados como las variables extralingüísticas y poder potenciar así el estudio y desarrollo de las competencias comunicativas de los mismos.



# BLOQUE I: MARCO TEORÍCO

## Capítulo I

### Cuestiones preliminares

Para empezar nuestro trabajo basado sobre un estudio de lingüística aplicada a la enseñanza-aprendizaje del léxico en alumnos de LE son necesarias algunas delimitaciones conceptuales que establezcan a nivel terminológico sus nociones principales.

#### 1.1 Palabras y vocablos

En el lenguaje corriente los términos “léxico” y “vocabulario” suelen utilizarse como sinónimos para hablar del conjunto de palabras de un idioma o del conjunto de palabras que caracterizan el uso de una lengua por parte de un hablante o de una comunidad de hablantes (Casadei, 2003).

Aunque se trate de dos términos estrechamente relacionados entre sí, en nuestro ámbito de especialización responden a conceptos diferentes. Con referencia a la distinción realizada por Saussure entre *langue* (sistema lingüístico y norma) y *parole* (uso), se puede decir que el léxico se reserva para la lengua, mientras que el vocabulario para el discurso. Para Justicia (1995:16), “el léxico hace referencia al sistema de palabras que componen una lengua”, mientras que “cuando esas unidades virtuales de la lengua se actualizan en el discurso del hablante reciben el nombre de vocablos y el conjunto de los mismos lo llamamos vocabulario”.

El vocablo entonces es la denominación de cada una de los componentes del vocabulario, es decir, el conjunto de unidades efectivamente empleadas por un solo hablante o por un grupo en un acto de habla concreto. En un sentido más restrictivo y específico, define todas las unidades que aparecen en un corpus textual, escrito u oral.

El léxico, en cambio, es la suma de todos los vocablos que están a disposición del hablante en un momento determinado (Genouvrier y Peytard, 1970), un conjunto abierto e infinitamente más amplio de cualquier vocabulario. Es evidente, pues, que el

vocabulario es una parte o subconjunto del léxico y se denota una relación de inclusión entre estos dos conceptos. Con el término *léxico* se engloba, además del vocabulario activo y ya existente, todas las unidades virtuales – lexemas – que se pueden crear en cualquier momento a través de reglas de formación de palabras. La palabra, pues, dentro de este campo designa toda ocurrencia de un vocablo cualquiera. Dentro del conjunto léxico se distingue entre el léxico parcial, para referirse a léxicos específicos, repertorios sociales, diccionarios terminológicos, etc. y global, para designar de forma general una lengua.

En la léxico-estadística, es decir el sector de la lingüística que utiliza métodos estadísticos para el análisis cuantitativo de la lengua, la unidad básica de medición es la palabra, ya que, como señala Gómez (2004) desde una perspectiva formal se caracteriza por tener tres cualidades importantes:

- es fácil de identificar (está separada por espacios en blanco);
- no presenta rasgos que no se puedan aislar o contar;
- favorece la explicación de las propiedades distribucionales (formales) de sus constituyentes (flexión nominal y verbal).

Siguiendo esta línea, Morales (1986) individua tres planos de interpretación de la palabra: fonológico (unidades fónicas a nivel segmental), gramatical (forma y función en la estructura gramatical) y semántico (unión de forma y contenido).

Sin embargo, Dimitrijévic (1969) propone sustituir el término ‘palabra’ por el de ‘unidad léxica’, considerada como la unidad conceptual base en el aprendizaje del vocabulario, que sirve para proporcionar conceptos de cultura de la lengua y que puede constar de una o más palabras. El término implica la unión de una forma léxica de al menos una palabra con un significado unitario (Cerveró y Pichardo, 2000) o con un valor semántico identificable (Haensch, 1997).

A partir de Nattinger y DeCarrico (1992), que apoyaron la teoría según la cual el vocabulario que se actualiza no solo se almacena como lexemas individuales, sino que también abarca segmentos de frases o partes largas de enunciados, diferentes autores han señalado la necesidad de segmentar las cadenas habladas en unidades léxicas, entre ellos Lewis (1993, 1997, 2000), que con el término ‘lexical item’ define los segmentos prefabricados de multipalabras que se pueden combinar para producir textos, y

Boogards (2001), que habla de unidad léxica para referirse a la unión de una forma léxica y un sentido único.

En suma, como objeto de análisis de la estadística léxica podemos considerar la unidad léxica, el término más adecuado para el recuento de producciones de los hablantes. Las unidades léxicas, de hecho, permiten a los usuarios almacenar en su lexicón mental la información mediante la categorización en signos lingüísticos.

## **1.2 Lexicón mental y competencia léxica**

Como hemos afirmado anteriormente, desde un punto de vista estrictamente lingüístico, cuando hablamos de léxico entendemos el conjunto de palabras que componen el universo lingüístico de un idioma. Al lado de este concepto, desde una perspectiva psicolingüística, para poder entender mejor los procesos de producción y recepción de mensajes, se utiliza el término lexicón mental o léxico mental para definir el conjunto de unidades léxicas que un hablante tiene interiorizado y es capaz de reconocer y/o utilizar en la producción de mensajes orales o escritos.

Es importante diferenciar el léxico mental de lo que es el vocabulario de un hablante ya que por sus definiciones los dos conceptos pueden parecer similares. El primero hace referencia a un conocimiento individual, que comparte solo algunas unidades del vocabulario de una determinada comunidad de hablantes. Por otro lado, el vocabulario es común a un grupo de hablantes y el conocimiento de sus unidades es compartido y supraindividual.

En un sentido amplio y generalista, el lexicón mental puede entenderse como el almacén de palabras (representaciones mentales) que contiene la información semántica o conceptual, la información sintáctica y las unidades léxicas (Aitchison 1994, Altmann (ed.) 1995). Otros autores lo consideran como el centro de integración de todos los procesos del lenguaje ya que contiene unidades que median entre información acústica-fonética y la interpretación semántica y sintáctica (Marslen – Wilson, 1989).

En el modelo teórico generativista (Chomsky, 1995; Baralo, 1997) el concepto de lexicón mental se encuentra organizado en una estructura interna compleja que le permite regir un gran número de fenómenos diferentes de manera rápida. En palabras de Baralo (1997:60) “el lexicón mental está constituido por elementos sustantivos o formantes, que metafóricamente podríamos denominar el almacén o diccionario del

léxico”, es decir las unidades atómicas de la sintaxis. A estos componentes la autora añade otros que define como “elementos relacionales o reglas, que probablemente actúan de manera semejante al sistema computacional de la sintaxis; y, posiblemente, por principios abstractos que rigen la forma y funcionamiento de esas reglas”.

Según Baralo (1997) el lexicón se puede representar dividido en tres niveles: en primer lugar se encuentra una lista finita de morfemas, es decir los datos de entrada de las reglas de formación de palabras; en segunda posición, las reglas de formación de palabras; luego, el producto de estas reglas, que gracias a los procesos de formación de palabras, teóricamente puede estar formado por un número infinito de palabras, aunque en realidad el cerebro puede retener solo un conjunto de palabras “existentes”.

Es evidente que el lexicón es un constructo que da cuenta de la capacidad creativa del lenguaje y de las lenguas y que gracias a su “uso infinito de medios finitos” (Chomsky, 1986) permite la creación de nuevas palabras que pueden ser entendidas, procesadas y recreadas por cualquier hablante, nativo o no nativo, de una lengua.

Esta capacidad de construir conceptos léxicos a partir de dotaciones innatas, constituye un conocimiento intuitivo del hablante y, dentro del modelo teórico que nos sirve de marco para nuestra investigación, es sinónimo de competencia léxica, esto es, la capacidad de entender y utilizar unidades léxicas, pero también morfemas que le permiten interpretar o generar unidades no percibidas o producidas con anterioridad y de combinarlas con otras.

Baralo (1997) opina que los hablantes no nativos poseen el mismo tipo de competencia léxica y gramatical que poseen los hablantes nativos. Por lo tanto, desde el punto de vista de la didáctica es bueno considerar la importancia de enseñar el léxico no de forma aislada, sino dentro de contextos comunicativos significativos donde la información referencial favorecerá el aprendizaje y adquisición de nuevos elementos. El lexicón es dinámico y el conocimiento, comprensión y uso de las varias unidades se puede ampliar cada vez más a través de las distintas informaciones de input.

### **1.3 Léxico y frecuencia**

Tras la lectura de los apartados anteriores es evidente que si el léxico puede englobar el conjunto de todas las palabras de una lengua, nadie puede conocer y

dominar el léxico global perteneciente a un idioma. ¿Cuales serán entonces las palabras que dentro de este entero conjunto, hay que enseñar a un aprendiz de ELE?

Según Alvar Ezquerro (2004:21) para enseñar la lengua a estudiantes extranjeros, independientemente de su competencia léxica en ELE, es muy útil conocer cuál es su vocabulario y cuales son las voces más frecuentes que aparecen en una lengua ya que éstas son de alto rendimiento. La rentabilidad comunicativa es un elemento que se mide a través del concepto de cobertura: según Alvar Ezquerro (2003:100) las cinco mil palabras más frecuentes cubren el 90% de un texto y permiten al aprendiz desenvolverse en las situaciones cotidianas. Asimismo, Almela et al. (2005) señalan que es mucho más rentable aprender una palabra altamente usada y polisémica que otra que se utiliza menos. Como insiste Boogaards (2008:1234) existe una estrecha correlación entre frecuencia y polisemia: si una palabra tiene más sentidos, tiene también más ocurrencias. Además, Ávila Martín (2010:169) afirma que una palabra con alto grado de frecuencia a nivel comunicativo tendrá un alto grado de utilidad, mientras que a nivel pedagógico tendrá mayor interés que otra que se utiliza menos.

Los estudios de léxico-estadística, a partir de corpus de textos escritos u orales, se han ocupado de medir la frecuencia de aparición de sus unidades, que constituyen el llamado léxico frecuente. Gracias a estos recuentos, se ha puesto de manifiesto que si bien el índice de frecuencia<sup>1</sup> es un dato importante para establecer la importancia de una palabra, hay que considerar también otros factores como la dispersión y el uso.

Un primer planteamiento de lo que es la dispersión fue realizado por García Hoz (1953), quien tras dividir el universo léxico en diferentes campos, midió cuántas veces una palabra aparecía en la serie de fuentes y tipos de textos que formaban parte de la muestra total y observó si una palabra tenía la misma o diferente frecuencia en cada uno de estos campos. Surge, así, el concepto de ‘rango’<sup>2</sup>, es decir, el valor general de un término al establecer su presencia en los diversos tipos de texto que constituyen el corpus. En otras palabras, se trata de un índice que mide la distribución de la frecuencia. Este dato nos informa sobre el grado de estabilidad de la frecuencia y mide el grado de utilidad de una unidad léxica en una lengua (L.F. Lara 2006:170). En este sentido, las palabras que cubran mayor número de campos y que aparezcan en numerosos y diferentes tipos de obras y registros constituirán las más frecuentes. Estas últimas

---

<sup>1</sup> Número de veces que aparece una misma palabra en el recuento de un corpus.

<sup>2</sup> Lo que Juilliand (1964) llamó índice de dispersión.



difícilmente varían de corpus en corpus y se caracterizan por tener el mayor grado de estabilidad en la lengua, ya que constituyen su armazón y sirven para mantener la estructura de un mensaje. Por eso, las palabras frecuentes se incluyen entre las fundamentales que hay que enseñar y aprender.

Al lado de la dispersión, el índice de uso o de utilidad ha sido tomado como otro criterio estadístico muy útil para la medición de la frecuencia de palabras y la selección léxica. El índice de uso pondera la dispersión de cada palabra por su frecuencia ( $U = F \times D$ ). Según la fórmula, es lógico que el uso aumenta cuando la frecuencia es alta pero también cuando lo es la dispersión.

Sin embargo, la utilidad no siempre está determinada por la frecuencia. Como señala Boogards (1994) la utilidad de una palabra es una noción vaga, ni las palabras, ni ninguna otra cosa son útiles de una manera general e indeterminada, sino que tendrán mayor o menor utilidad dependiendo de su adecuación al objetivo que quieren alcanzar. Si un aprendiz de L2 necesita algunas palabras específicas de un determinado campo, tendrá especial interés en servirse de ellas aunque estas no figuren entre las palabras más frecuentes de una lengua. Si bien es cierto que el aprendiz para comprender los mensajes orales o escritos, depende de las palabras seleccionadas por los demás, es decir que para la competencia receptiva el factor frecuencia desempeña un papel importante, para la competencia productiva, este dato tiene menor importancia ya que los elementos a disposición y usados por el aprendiz no son necesariamente los más frecuentes. Cabe distinguir entonces entre vocabulario receptivo y productivo.

#### **1.4 Léxico disponible**

Conocer cuáles son las palabras más frecuentes en una lengua puede valer mucho en el proceso de decisión del léxico que se quiere enseñar e incluir en la enseñanza a alumnos de ELE. Sin embargo, como afirma Bou (1966:12) “una lista de frecuencias no es un ábrete sésamo para la enseñanza de vocabulario” pero queda patente que es una excelente ayuda. Cabe introducir entonces otra noción de selección léxica fundamental para nuestro trabajo: la disponibilidad léxica.

Los estudios de disponibilidad léxica se iniciaron en Francia a mediados de los años cincuenta, cuando se utilizó únicamente el criterio de frecuencia para realizar una recolección de palabras útiles a la creación de manuales de francés elemental para

extranjeros. Las investigaciones llevaron a la publicación en 1956 de la obra *Français Elementaire*<sup>3</sup>, considerada pionera de este tipo de estudios. El equipo de investigadores compuesto por G. Gougenheim, R. Michéa, P. Rivenc y A. Sauvageot descubrió que a partir de recuentos de palabras que aparecían en textos escritos, no estaban incluidas palabras conocidas y usadas comúnmente por los hablantes nativos (*fourchette, lettre, timbre...*).

Estos estudios demostraron que la frecuencia no podía ser el único dato sobre el cual basar la selección del léxico y llevaron a Michéa a formular una distinción entre palabras atemáticas –más estables y con frecuencia independiente del tema tratado– y palabras temáticas –con frecuencia vinculada a un tema concreto y por eso más inestables. A partir de esta diferenciación, el propio Michéa estableció la oposición entre léxico frecuente y léxico disponible. El primero se definió como aquél que recogía “las palabras de mayor estabilidad estadística” (H. López Morales, 1999:8), mientras que el léxico disponible se definió como “el que se usa en una situación comunicativa específica, actualizándose sólo ante la necesidad de comunicar algo referente a esa situación concreta” (H. López Morales, 1999:9). Las palabras disponibles se incluyen entre las que acuden primero a la memoria al hablar de un tema concreto y determinado y reflejan un vocabulario de tipo nocional. De hecho, este tipo de léxico está formado en su mayoría por un número abundante de sustantivos que aluden a realidades concretas. La parte activa de este vocabulario está compuesta por las palabras de alta disponibilidad, mientras que las palabras de baja disponibilidad forman parte del vocabulario activo de un número restringido de hablantes y del vocabulario pasivo del resto de los hablantes.

Como ha señalado H. López Morales, el léxico disponible es un léxico potencial, no actualizado, mientras que la frecuencia sólo trabaja con léxico actualizado. Es evidente que la frecuencia, por sí sola, no es el único índice válido de uso a la hora de trabajar sobre el léxico, ya que ha sido matizado con el de la disponibilidad. Las palabras empleadas para referirse a realidades cotidianas no son frecuentes pero disponibles si se activa un centro de interés. Se puede deducir por tanto que hay palabras muy útiles pero poco frecuentes y palabras con mayor frecuencia pero escaso contenido informativo, lo que determina una distinción clara entre frecuencia y disponibilidad.

---

<sup>3</sup> A partir de 1965 el título se convirtió en *Français Fondamental*.

Los estudios sobre disponibilidad sirvieron para poner de relieve esas palabras poco frecuentes pero de alto contenido informativo, necesarias para la comunicación y la selección léxica de material destinado al aprendizaje de un idioma.

### **1.5 Léxico común**

Frente a la consideración de que una persona no pueda pretender aprender las cerca de cien mil voces listadas en los diccionarios generales, conocer las unidades que componen el vocabulario básico de una comunidad, que como señala Haensch (1982:159) recoge las cinco mil palabras más usadas en la comunidad estudiada, y apuntar hacia el aprendizaje de esas cifras, puede llevar a un primer nivel de conocimiento de la lengua sin grandes pretensiones.

La cantidad de palabras que conoce un alumno tiene implicaciones en su competencia comunicativa y en su capacidad de comprensión de los textos. Se ha establecido la existencia de un umbral de conocimiento mínimo de vocabulario que permita la comprensión de un texto auténtico. Según estudios recientes una lista de palabras más frecuentes compuestas por 5000 unidades puede bastar para cubrir la comprensión de un texto al 99,7% y por lo tanto son suficientes para las necesidades básicas de cualquier hablante.

Estas palabras constituyen el vocabulario común de una lengua y, como afirma García Hoz (1953:385), son las que “necesariamente ha de utilizar el hablante español en cualquier manifestación de su vida”. Por tanto, resultan como las más generales, usuales e importantes, ya que se encuentran con mucha frecuencia en todo tipo de situación comunicativa y aparecen distribuidas en todos los tipos de texto, independientemente del tema de que se trate. Son las que sirven para mantener la estructura del mensaje, en su mayoría figuran palabras gramaticales (preposiciones, conjunciones, artículos, cuantificadores, numerales, indefinidos), las que no lo son pertenecen, en orden de frecuencia, a verbos, adjetivos y sustantivos de significado general. Los léxicos básicos entonces recogen las palabras más estables – más permanentes – en la lengua, seleccionadas a través del índice de uso, y ofrecen una proporción real de uso de las diferentes clases de palabras.

Además del criterio estadístico de uso, a lo largo de las investigaciones sobre léxicos básicos se ha utilizado el de frecuencia. De hecho, García Hoz, mediante unos

cálculos estadísticos relativos a medio millón de vocablos pertenecientes al lenguaje oficial, familiar, informativo y de cultura general, llega a establecer el vocabulario común de la lengua española, compuesto por las palabras que tienen sus frecuencias prácticamente distribuidas en partes iguales en todos los tipos de textos y situaciones comunicativas.

Asimismo, la selección de requisitos como una muestra homogénea y exhaustiva, además la toma en consideración de textos representativos de la sincronía de una lengua, son aspectos fundamentales para establecer el léxico básico de una lengua.

## **1.6 Léxico fundamental**

En la lingüística actual se considera léxico fundamental aquel que está formado por el léxico básico común y el léxico disponible (López Morales, 1999). Como hemos visto, mientras que los léxicos básicos recogen las palabras más estables en la lengua y ofrecen una ponderación de frecuencia y uso de las distintas categorías de palabras, los disponibles muestran las palabras más inestables que acuden más fácil y rápidamente a la mente de un hablante en una situación comunicativa determinada. Ambos métodos son complementarios y ofrecen el perfil léxico de una comunidad lingüística.

Basándose en García Hoz, Alvar Ezquerro (2003: 31) señala la necesidad de actualizar los datos del estudio para establecer que en la primera etapa del proceso de enseñanza de ELE se tendrá que introducir las palabras pertenecientes al vocabulario fundamental, seguidas por las restantes palabras que componen el vocabulario común, a las cuales se podrán añadir las otras que son más frecuentes de las que no aparecen en todos los tipos de vocabulario, así como las que expresan ideas afines y cuyas frecuencias se suman.

## **1.7 Estadística y lingüística de corpus**

### **1.7.1 Estudios estadísticos de léxico y diccionarios léxicos**

Los diccionarios de frecuencia son concebidos con una finalidad especialmente didáctica: “comprobar la frecuencia de palabras para poder usar en obras pedagógicas las palabras más frecuentes” (Haensch & Omeñaca 2004:154).

En la década de los años veinte se inician los primeros recuentos léxicos de la lengua española. La primera obra que se realiza es el *Common words in Spanish* (1920) de H. Keniston, una recolección de 1322 palabras usadas en el lenguaje cotidiano, procedentes de obras de teatro, periódicos, revistas, cuentos y novelas, destinada para uso especial de los enseñantes de español a extranjeros.

En 1927 se publica la obra de Milton A. Buchanan, con el título *Graded Spanish Word Book*, una recopilación basada en el lenguaje escrito que goza de gran prestigio en los años siguientes.

En los años 50 se empiezan a recopilar diferentes listas de palabras frecuentes destinadas al ámbito pedagógico. Uno de los primeros trabajos es el *Recuento de vocabulario español* de Rodríguez Bou, obra publicada en 1952 que incluye palabras del lenguaje oral y escrito de niños y adultos a partir de un corpus formado por composiciones y conversaciones transcritas de niños, periódicos, revistas, programas de radio, literatura religiosa y textos escolares. Poco después, en 1953 se publica el trabajo *Vocabulario usual, común y fundamental* de Victor García Hoz. La obra abarca alrededor de 15.000 palabras extraídas a partir de un corpus de 400.000 unidades pertenecientes a cuatro ámbitos: vida familiar, periódicos, documentos oficiales de estado y eclesiásticos y obras de literatura.

Al lado de los estudios estadísticos de léxico con interés pedagógico, en 1964 Charles Müller lleva a cabo un análisis de las unidades léxicas empleadas por autores literarios que se empleó en gran cantidad para estudios con fines estilísticos.

El mismo año se publica el diccionario más citado en el ámbito de la léxico estadística, es decir el *Frequency Dictionary of Spanish Word* de Juilland y Chang-Rodríguez basado en un corpus de 500.000 unidades procedentes de varios campos: drama, narrativa, literatura técnica y científica, periodismo. Los criterios que generaron el vocabulario básico de 5.000 palabras se basaron en la recolección entre 1920 y 1940 de 100.000 ocurrencias por cada género, por lo que además de indicar la frecuencia absoluta de cada palabra, se podían extraer datos sobre la frecuencia relativa en cada género y establecer relaciones de comparación de género a género por cada palabra, lo que influyó en establecer la importancia de cada palabra en el específico género.

Esta obra tuvo mucho seguimiento en ámbito hispánico y llevó a la realización de otras de gran importancia e interés didáctico. Se publicó así en 1986 *El léxico básico del*

*español de Puerto Rico*, obra de Amparo Morales, realizada a partir del léxico del español de Puerto Rico entre los años 1948 – 1970.

Más reciente y circunscrito a lenguajes específicos, como el de los adolescentes, es *El desarrollo del vocabulario. Diccionario de frecuencias* de Fernando Justicia (1995) que se basa en la recolección de datos de producciones escritas realizadas por niños desde 6 hasta 13 años. El corpus está compuesto por más de medio millón de ocurrencias y ha sido recogido entre 1985 y 1988 en las provincias de Almería, Granada, Jaén y Málaga.

Los últimos diccionarios de frecuencia del siglo XX son obras publicadas en la Universidad de Oviedo (Alameda *et al.* 1995) y en la Universidad de Barcelona (Sebastián *et al.* 2000).

### **1.7.2 Lingüística de corpus y corpus del español en la actualidad**

Además de averiguar la frecuencia de palabras consultando un diccionario de frecuencias, en la actualidad disponemos de otra vía, que corresponde a los datos proporcionados por los corpus de palabras. Si por un lado los diccionarios de frecuencia son instrumentos que, como hemos visto, disponen de una cierta antigüedad y no se han renovado en las últimas décadas, por otro lado los corpus del español tienen una estrecha relación con las nuevas tecnologías informáticas y sus aplicaciones son más ambiciosas debido a los progresos que la informática está experimentando.

En palabras de Alvar Ezquerro (2004:20) “un corpus está formado por una gran cantidad de datos lingüísticos [...] que se pretenden representativos de la lengua en alguna o varias de sus manifestaciones (lengua hablada, lengua escrita), niveles (culto, literario, coloquial, etc.), medios por los que se transmite (novela, teatro, prensa diaria, radio, etc.) especialidades o ámbitos (economía, política, deporte, historia, aficiones y coleccionismo, etc.)”. En general, se trata de almacenes masivos de datos, realizados por medio de herramientas informáticas, que pueden llegar a registrar millones de palabras representativas de la realidad lingüística de un momento determinado de la historia. Por tanto, la información proporcionada en un corpus acerca de una lengua tiene que ser extensa de manera que pueda representar todas las variedades relevantes de esta realidad en cuestión.

En la actualidad existen una gran cantidad de corpus del español con diferentes finalidades descriptivas. Frente a la dificultad de enumerar todos los trabajos realizados en este ámbito debido a su gran número, algunos llaman especial atención por los datos que revelan y por el interés pedagógico que suscitan.

Una descripción de la lengua oral es el *Léxico de frecuencia del español hablado en la ciudad de Málaga*, realizado por Antonio Manuel Ávila Muñoz en 1999. A partir de grabaciones de textos hablados por un número aproximado de 300 sujetos, la muestra total del corpus cuenta 500.000 unidades y el léxico básico obtenido es de alrededor de 5000 palabras. Asimismo, el trabajo de Marcial Terrádez Gurrea titulado *Frecuencias léxicas del español coloquial* (2001) está compuesto por un recuento de 100.000 palabras extraídas por el grupo Val.Es.Co. a partir de conversaciones coloquiales en Valencia y su área metropolitana.

Otros corpus que merecen mención son los que se han realizado para la elaboración de obras lexicográficas. Entre ellos destaca el *Corpus Cumbre* (2001), cuyo director fue Aquilino Sánchez Pérez, basado en veinte millones de palabras del español escrito y oral de España e Hispanoamérica. Los textos que forman el corpus proceden por un 65% de España y por un 35% de Hispanoamérica. Por lo que se refiere a España un 70% de las contribuciones son escritas y el 30% orales, mientras que en la parte de Hispanoamérica se encuentran un 60% de textos escritos y un 40% de textos orales. Este corpus se utilizó para la realización del *Gran diccionario de uso del español*, publicado en 2001, que presenta por primera vez la indicación de la frecuencia por cada palabra mediante la marcación numérica de la frecuencia.

A diferencia de una gran parte de corpus que pertenecen a editoriales u otros organismos, o que están publicados en papel y que no son fácilmente accesibles al gran público, otros están disponibles en la red y al alcance de todos. Entre ellos destaca el corpus de Mark Davis, disponible a la página electrónica <http://www.corpusdelespanol.org>. La versión original del corpus, realizada entre 2001 y 2002, permite hacer búsquedas entre más de 100 millones de palabras recogidas a partir de más de 20.000 textos del español pertenecientes a varios tipos de texto y registros. Dos tercios del corpus abarcan el lenguaje escrito, mientras que las demás unidades proceden del lenguaje oral (habla culta y coloquial). Se trata del primer corpus

que presenta una distribución equilibrada entre los textos de Hispanoamérica y España (43% de España y 57% de Hispanoamérica), datados entre los siglos XIII y XX.

Basándose en este corpus, Davies en 2006 publica *A frequency dictionary of Spanish*, un diccionario que recoge las 5.000 palabras más frecuentes del español del siglo XX (representadas por 20 millones de palabras en el corpus). En la obra se encuentran lemas, ordenados por frecuencia, con la indicación de la categoría gramatical, la traducción en inglés, un ejemplo extraído del corpus, el rango total de su frecuencia y el número de veces que aparece en el corpus.

Desde verano 2016 *El corpus del español* disponible en red se encuentra en una nueva versión, ajustada para el acceso con dispositivos móviles, que permite la realización de “corpus virtuales” a partir de un específico grupo de autores, fuentes o temas. El nuevo corpus se basa en un número aproximado de dos mil millones de palabras recogidas en páginas web de 21 países hispanohablantes (cantidad mayor con respecto al corpus original de 20 millones de palabras del siglo XX). Los textos, seleccionados entre 2013 – 2014, son representativos del español moderno y ofrecen la posibilidad de hacer comparaciones entre los diferentes dialectos del español.

Sin embargo, el corpus más extenso de la lengua española, disponible y abierto para todo el público que desee consultarlo, es el *Corpus de Referencia del Español Actual* (CREA), realizado por la Real Academia de la Lengua Española. El corpus, consultable en la página de la RAE (<http://www.rae.es>), cuenta más de ciento sesenta millones de formas en su última versión de 2008. Se trata de un conjunto de textos escritos y orales de procedencia de todos los países de habla hispana (50% de España, 50 % de Latinoamérica). Los textos recogidos entre 1975 y 2004 se dividen entre un 90% de fuentes escritas correspondientes a libros, periódicos y revistas que tratan más de cien materias diferentes, y textos orales que se componen de transcripciones de grabaciones procedentes de la radio y la televisión. En la versión de 2008 se agregan las listas de las 1.000, 5.000 y 10.000 formas más frecuentes, así como la lista de todas las formas ortográficas registradas en el CREA con su frecuencia absoluta y normalizada. En 2015 se ha publicado la versión anotada del CREA que permite la consulta por formas, lemas y categorías gramaticales pero no contiene aún los textos orales y se compone de fuentes fechadas hasta el 2000, año en el que limita con el **Corpus del**



**Español** del Siglo XXI, realizado a partir de textos escritos o producidos desde el año 2001 en adelante.

Como hemos apuntado anteriormente, los corpus proporcionan información ajustada a la realidad de la lengua, en especial modo, el corpus CREA es una herramienta útil y necesaria para la realización de los distintos diccionarios académicos. Además, desde su aparición, el ha sido el punto de partida para investigaciones sobre el español actual, sobre todo lingüísticas, pero también relacionadas a áreas diferentes como la publicidad, la terminología o la sociología. Asimismo, ha desempeñado un papel importante para la elaboración de diversos materiales derivados: gramáticas, diccionarios, tesauros, correctores ortográficos, métodos de didáctica del español y aplicaciones de traducción automática, entre otros.

# Capítulo II

## Estudios de disponibilidad léxica y cuestiones metodológicas

### 2.1 Estudios cuantitativos sobre léxico

Como ya puso de manifiesto Micheá (1953), hay algunas palabras que aunque sean perfectamente conocidas por los hablantes de una comunidad, solo se emplean cuando se habla de un determinado tema o en una cierta situación y por lo tanto no constituyen el léxico frecuente. La parcialidad de un léxico frecuente no podía ser la única manera de selección de unidades que componen el léxico fundamental de un idioma. Se hizo patente entonces que el justo complemento de ese tipo de léxico eran las palabras disponibles.

Para la elaboración del léxico disponible de una comunidad –no actualizado en un corpus o en un conjunto de ellos– los investigadores se sirven de unos centros de interés, estímulos temáticos amplios a partir de los cuales los sujetos de la muestra producen unas listas de unidades léxicas de las que disponen para uso inmediato. Como señala H. López Morales (1999:32) las pruebas en las cuales se actualizan estas listas son “las únicas que hacen posible en condiciones experimentales, que se realicen en la actuación lingüística las unidades léxicas con poca estabilidad estadística”.

Tras la realización de las primeras investigaciones sobre léxico disponible en Francia (Gougenheim et alii 1956, 1965) se empezaron a recoger otros conjuntos de listas. Cabe señalar entre ellos *Le vocabulaire disponible du français* (1971a, 1971b), trabajo canadiense encabezado por William Mackey, y también el sobre el inglés de Escocia de N. R. Dimitrijevic (1969) que introdujo algunos cambios metodológicos que fueron aceptados por proyectos posteriores.

### 2.2 Estudios de disponibilidad léxica en hablantes nativos de español

Las conclusiones obtenidas por las investigaciones sobre *El léxico fundamental* del francés, destinadas al francés como segunda lengua, impulsaron los primeros estudios sobre disponibilidad léxica que tuvieron mucho seguimiento en el ámbito del español estudiado en relación con hablantes nativos.

Los trabajos sobre disponibilidad léxica en español fueron iniciados por H. López Morales en 1973 mediante el estudio *Disponibilidad léxica de los escolares de San Juan* en Puerto Rico y se continuaron en 1978b, 1979, 1983 y 1986. Tras esta investigación, a partir de los años 90 otros estudios se llevaron a cabo en México (López Chávez, 1993), República Dominicana (Alba, 1995), Chile (Valencia y Echeverría, 1999). En España la primera investigación se realiza en la zona metropolitana de Madrid por mano de Benítez Pérez (1992). Desde entonces, a partir de la capital los estudios se han desarrollado en todas las provincias españolas hasta llegar a sus casi completas conclusiones y correspondientes publicaciones de resultados. Como señala Paredes García (2015), para obtener una visión de conjunto de la situación española se pueden consultar Carcedo González (1998), Samper Padilla, Bellón Fernández y Samper Hernández (2003), Samper Padilla y Samper Hernández (2006) y Gallego Gallego (2014). En estos momentos los equipos de investigación españoles están trabajando de manera conjunta a la realización del *Léxico Disponible de España*, a partir de una muestra ponderada de las pruebas llevadas a cabo en las distintas zonas del país (Bartol Hernández 2006).

### **2.3 El Proyecto Panhispánico**

Todos los estudios citados anteriormente han recibido un gran impulso gracias al ambicioso *Proyecto Panhispánico*, coordinado por H. López Morales, que, como se afirma en la página *DispoLex*<sup>4</sup>, tiene el objetivo de elaborar diccionarios de léxico disponible de las diferentes zonas pertenecientes al mundo hispánico para llegar a realizar comparaciones de tipo lingüístico, etnográfico y cultural, individuar áreas de difusión y, en general, para formular análisis posteriores.

En la base del proyecto hay un conjunto de normas que fueron establecidas en una reunión donde intervino la mayoría de los investigadores españoles, convocados por la profesora Maitena Etxebarria a finales de abril de 1999 en la Universidad de Bilbao. Los participantes llegaron a acuerdos sobre los criterios metodológicos comunes que se debían perseguir para garantizar la homogeneidad de los datos y permitir la

---

<sup>4</sup> *DispoLex* es una página electrónica ([www.dispolex.com](http://www.dispolex.com)) gestionada por los investigadores de la Universidad de Salamanca y destinada a los investigadores y otros interesados en los estudios de disponibilidad léxica.

comparación en investigaciones posteriores de los distintos léxicos disponibles. Las decisiones que se tomaron se presentan a continuación como encontramos en Samper Padilla, Bellón Fernández y Samper Hernández (2003).

#### **2.4 Los centros de interés**

A fin de obtener los datos necesarios a la elaboración de léxico disponible de las comunidades investigadas por los distintos equipos, se entrega a los encuestados una prueba de producción léxica de tipo asociativo controlado, que permite la actualización por parte de los informantes de respuestas relacionadas con el estímulo de partida, llamado centro de interés.

En el Proyecto Panhispánico los centros de interés establecidos son 16, los mismos investigados por los pioneros de estos tipos de estudios en Francia y deben ser exactamente los mismos tanto en todos los sujetos de la investigación, como en todos los proyectos de investigación para hacer posible la realización de comparaciones. Sin embargo, a lo largo de la historia de los estudios de disponibilidad léxica, no siempre se han utilizado todos los centros de interés o solo los 16 citados. Hay ejemplos donde la cantidad total ha sido reducida o aumentada adicionando centros de interés nuevos debido a propósitos de adaptación de los centros de interés en relación bien con peculiaridades dialectales, bien con las características extralingüísticas de los informantes.

Según Gougenheim (1967) los centros de interés cumplen “tanto la condición de la universalidad (y, por consiguiente, podrían interesar a los hablantes, independientemente de su estrato, situación o condición) como la de la coherencia interna” (Samper Padilla, Bellón Fernández y Samper Hernández, 2003: 48) y son los siguientes:

- 1) ‘Partes del cuerpo’
- 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) ‘Iluminación, calefacción y medios para airear un recinto’
- 10) ‘La ciudad’
- 11) ‘El campo’
- 12) ‘Medios de transporte’
- 13) ‘Trabajos del campo y del jardín’
- 14) ‘Los animales’
- 15) ‘Juegos y distracciones’
- 16) ‘Profesiones y oficios’

## **2.5 La prueba**

El material de la investigación se recoge mediante unas encuestas organizadas en listas abiertas que comúnmente suelen ser respondidas en un tiempo límite de dos minutos.

Como señala Samper Hernández (2002), al principio de los trabajos sobre disponibilidad léxica –en las investigaciones franceses y canadienses– el sistema utilizado era cerrado: los informantes debían escribir un total de 20 palabras por cada centro de interés en un tiempo máximo de 15 minutos. Sin embargo, este sistema permitía obtener sólo datos cualitativos; con el cambio hacia las listas abiertas se pudo averiguar además de las palabras disponibles, la cantidad de respuestas por cada informante. Se decidió también disminuir la limitación temporal hasta los dos minutos para que las actualizaciones fueran lo más natural posible evitando dejar demasiado tiempo para la producción de posibles palabras fruto de varios minutos de reflexión y, por lo tanto, menos disponibles.

## **2.6 El número de encuestados**

Otro criterio metodológico común que se emplea en las investigaciones del Proyecto Panhispánico es la limitación numérica de las personas que componen el total de la muestra por área de estudio. El número se ha fijado en 400 participantes y confiere según Samper Padilla, Bellón y Samper Hernández (2003) un mayor grado de fiabilidad a los datos que resultan de las comparaciones directas de investigaciones distintas,

evitando que las diferencias observadas no se deban al número de participantes en la encuesta.

## **2.7 Condicionantes extralingüísticos**

En las varias investigaciones sobre disponibilidad léxica que forman parte del Proyecto Panhispánico se han tomado en cuenta las siguientes variables: factor sexo, edad, nivel sociocultural, zona geográfica y tipo de centro.

### **2.7.1 Factor sexo**

Desde el inicio de los trabajos sobre disponibilidad, prácticamente todos los investigadores han considerado la condición masculina o femenina de los participantes a la encuesta a la hora de analizar los resultados. Como señalan Samper Padilla, Bellón y Samper Hernández (2003), este factor no suele presentar modificaciones entre los distintos estudios. Además, la muestra no se preestratifica ya que normalmente el proceso de selección de los informantes es aleatorio, lo cual explica las usuales diferencias numéricas entre hombres y mujeres. A pesar de que normalmente se presente un desequilibrio numérico donde destaca la superioridad numérica femenina, este factor no influencia en gran medida los resultados, que, con algunas excepciones, denotan una gran homogeneidad entre los sexos.

### **2.7.2 Edad**

Un criterio que comparten la mayoría de las investigaciones que integran el Proyecto Panhispánico es la variable edad que, a diferencia de la variable sexo, se suele seleccionar previamente. Normalmente los sujetos de las encuestas son estudiantes preuniversitarios del segundo año de Bachillerato, de edad que oscila entre los 17 y 19 años. Esta decisión se debe al carácter aún no contaminado por tecnicismos y rasgos típicos del lenguaje especializado propio de los estudiantes universitarios en sus respectivas licenciaturas. Además, el léxico que se recoge refleja el fruto de años de escolarización y se propone como muestra representativa del estado de la lengua de un hablante culto, homogénea y seleccionada con el fin de no alejarse de la norma léxica de la comunidad.

### **2.7.3 Nivel sociocultural**

Introducido en los estudios hispánicos por López Morales (1973, 1979), otro condicionante que no se suele preestratificar al lado de la variable sexo es el nivel sociocultural de los informantes, que se establece en la fase de postestratificación de la muestra a partir de la información aportada por los participantes en la primera fase de la encuesta de disponibilidad. La variable normalmente es el resultado de tres factores: el nivel de estudios de los padres, su profesión y la renta familiar. Debido a las dificultades en la obtención de datos sobre la situación económica de los padres de los informantes, en las investigaciones hispánicas se han seguido dos caminos: el empleo de tres indicadores (ingresos familiares, nivel educativo y ocupación profesional), según el ejemplo de López Morales (1999), o solo dos (escolaridad y profesión de los padres), como indicado en el trabajo de Samper y Hernández (1997)<sup>5</sup>. Tras la asignación de puntuaciones a cada uno de los padres tal como se explica la tabulación presente en el estudio recién citado, se establecen distintos niveles que resultan de la suma de los diferentes grados de puntuación y se dividen entre tres intervalos: alto (20-24), medio (9-19), bajo (4-8). En nuestro estudio hemos seguido los criterios utilizados y expuestos por Samper y Hernández (1997) para facilitar la aportación de información por parte de los informantes.

### **2.7.4 Zona geográfica y tipo de centro**

Los dos últimos condicionantes que se suelen tener en cuenta son las variables ‘zona geográfica’ y ‘tipo de centro’. La primera, ya tomada en consideración en el estudio de Gougenheim y sus colaboradores (1967), indica la zona espacial donde viven los informantes, considerada desde la perspectiva de oposición urbano/rural, lo cual permite analizar si existen diferencias entre los vocabularios producidos por los participantes según su lugar de residencia. La segunda variable se utiliza para definir el tipo del centro de estudios de los encuestados y pretende distinguir entre centros de enseñanza públicos y privados. Como explican Samper Padilla, Bellón y Samper Hernández (2003), se presenta como uno de los factores condicionantes que se han empezado a considerar solo a partir de los primeros trabajos de disponibilidad léxica en

---

<sup>5</sup> Las modalidades de asignación de puntuaciones se presentan en el trabajo de Samper y Hernández (1997) *Estudio de la disponibilidad léxica en Gran Canaria. Datos iniciales y variación sociolingüística*, págs. 237 – 238.

español, los cuales han destacado en distintos estudios la importancia que se le debe otorgar.

## **2.8 Estudios de disponibilidad léxica en ELE**

Como hemos visto, las aplicaciones de la léxico-estadística han dado lugar a valiosas aportaciones en el ámbito hispanohablante. Sin embargo, como afirman Rodríguez Muñoz y Muñoz Hernández (2009:9) “pocas veces se han sabido aprovechar las nuevas metodologías de análisis léxico y su aplicabilidad a la enseñanza de lenguas extranjeras”. De hecho, no encontramos muchas investigaciones que siguen esta línea. Distinguimos entre algunos investigadores que se han dedicado al ámbito del estudio de la disponibilidad léxica de hablantes nativos para la aplicación a la enseñanza de ELE (Benítez Pérez, 1994a; García Marcos y Mateo, 2000) y otros que han empleado los criterios comunes a estos trabajos para los estudios de léxico disponible de estudiantes no nativos de español. El iniciador de estos estudios es Carcedo González (2000a), el cual se ha centrado en la investigación de una comunidad de alumnos específica, es decir los alumnos de español en Finlandia, y ha comparado los resultados con los datos de disponibilidad léxica de los hablantes nativos de español. Su primer trabajo impulsó la realización de otras investigaciones que pretendían conocer la disponibilidad léxica de estudiantes alemanes (Hugo 2003), chinos (Lin 2006), eslovenos (Šifrar Kalan 2009), polacos (López González 2009), islandeses (Magnúsdóttir 2012), estadounidenses (Verdeses-Mirabal 2012, Ortiz Lacasa 2012), turcos (González Fernández 2013), marroquíes (Serfati y Aabidi 2013) y portugueses (Gamazo Carretero 2014, Fernández dos Santos 2014).

Además, se señalan los trabajos de estudiosos que han dirigido sus investigaciones a comunidades de inmigrantes hispanos en Estados Unidos (Moreno Fernández 2003, Sancho Sánchez 2006) o en España (Fernández-Merino Gutiérrez 2010, Jiménez Berrio 2012) y estudios que tienen como objetivo el de investigar la disponibilidad léxica de una muestra conjunta compuesta por estudiantes de ELE e inmigrantes en la Comunidad de Madrid (Gallego Gallego 2014).



## **2.9 Condicionantes extralingüísticos en los estudios de disponibilidad léxica en estudiantes de ELE**

Por el carácter propio de las investigaciones sobre disponibilidad léxica en estudiantes de ELE, a los factores condicionantes que normalmente se consideran en los trabajos pertenecientes al Proyecto Panhispánico, se añaden otros que son comunes en los varios estudios y que se analizan por su incidencia en los resultados: la variable lengua materna, conocimiento de otras lenguas, nivel de estudios.

La lengua materna de los participantes a la encuesta es un factor imprescindible para el análisis del vocabulario actualizado, tanto por la indicación del grado de dificultad de adquisición de la lengua según el grado de familiaridad entre las dos, como por la información que puede aportar para el estudio de los errores principales de aprendizaje.

Asimismo, como afirman Samper Padilla, Bellón y Samper Hernández (2003), el conocimiento de otras lenguas además de la materna influye positivamente a la hora de asimilar una nueva lengua, sobre todo si esta forma parte de la misma familia de la lengua madre.

Por último, es muy relevante la variable nivel de estudios en las pruebas de disponibilidad léxica en estudiantes de ELE. La razón de su importancia se debe a la hipótesis de partida que afirma la asociación fuerte entre este factor y el número de respuestas de los informantes y por lo tanto supone una más rica disponibilidad léxica en los participantes que presentan un mayor conocimiento y dominio de la lengua española

## **2.10 Gradación de la disponibilidad**

En los apartados anteriores hemos resumido los principales trabajos y pautas metodológicas que sirven para investigar sobre disponibilidad léxica. Sin embargo, no debemos olvidar el instrumento que está en la base de estos tipos de estudios, es decir el cálculo del índice de disponibilidad.

López Morales en su obra pionera ya puso de manifiesto la necesidad de encontrar una fórmula que tomara en consideración el grado de disponibilidad de las palabras en las listas para proporcionar una versión más refinada de la realidad señalando aquellas palabras que acudían primero a la memoria ante un estímulo determinado. En 1983,

López Morales y Lorán formulan una propuesta capaz de calcular la disponibilidad y ponderar la frecuencia y el lugar de aparición de las palabras en las listas. Poco después, en 1987, López Chávez y Strassburger, investigadores de la Universidad Nacional Autónoma de México, encuentran algunas irregularidades en la fórmula anterior y desarrollan otro artificio matemático que tiene en cuenta:

- la frecuencia absoluta de cada palabra en cada posición de la lista;
- la frecuencia absoluta de cada palabra resultante de la suma de las diferentes frecuencias alcanzadas en cada posición;
- el número de encuestados;
- el número de posiciones alcanzadas en los centros de interés analizados;
- la posición en que fue dicha cada palabra.

Como reconoce López Morales (1995:248), esta fórmula ha demostrado su superioridad y sus ventajas para la lingüística ya que “logra producir una adecuación descriptiva sumamente plausible, tanto en el grupo como en el individuo” sin perder su poder discriminador y su significado estadístico a medida que las palabras ordenadas en las listas se alejan de los primeros lugares, sin tener en consideración la extensión del listado.

Mediante los índices de disponibilidad es posible obtener una lista de palabras ordenadas de mayor a menor disponibilidad donde se expresa la correlación entre la frecuencia de un vocablo y el lugar de su aparición (orden de evocación de las palabras).

El análisis de los datos extraídos de las pruebas de disponibilidad léxica, señala Carcedo (1998) permite conocer el vocabulario de una comunidad de hablantes específica, examinar las distintas fases del proceso de adquisición aprendizaje de una lengua materna o extranjera, comparar los léxicos disponibles de comunidades hablantes la misma lengua y establecer cuáles unidades léxicas se deben enseñar gracias a una base estadística eficaz y estable.

### **2.11 Aplicaciones de los trabajos de disponibilidad léxica**

Los estudios de disponibilidad léxica pueden ofrecer materiales muy valiosos para sus aplicaciones a varias disciplinas además de revestir una importancia fundamental y primogenia dentro del ámbito pedagógico, que ha sido el primero en ver los resultados de sus aplicaciones.

A continuación presentamos los campos donde resulta especialmente útil la aplicación de la disponibilidad léxica.

### **2.11.1 Dialectología**

En este dominio la aplicación de la disponibilidad léxica permite destacar la aproximación o diferencia entre variedades diatópicas mediante el análisis cuantitativo de datos y su comparación.

Las primeras reales comparaciones de disponibilidad entre dialectos se llevan a cabo de manera científica gracias al desarrollo unas fórmulas elaboradas por parte de López Chavez (1992a, 1992b), quien las aplica para analizar las diferencias entre las disponibilidades arrojadas por los dialectos de Madrid, Gran Canaria, República Dominicana y Puerto Rico. Los resultados del estudio señalan unos grados de coincidencia relativamente bajos, datos que difieren de la tentativa sucesiva de Alba (1999), quien limita su cotejo a las primeras 50 palabras de cada lista de los tres centros de interés tomados en consideración hallando una compatibilidad muy alta entre los distintos dialectos encuestados.

Finalmente, la comparación realizada por Samper (1999) entre las disponibilidades de los dialectos de Puerto Rico y Gran Canaria, resulta en un 90% de compatibilidad y destaca la relevancia de incluir en el cotejo solo los vocablos que integran a la norma léxica de cada comunidad investigada. Sin embargo, el aspecto más interesante revelado por el investigador en su artículo de 1998 es la influencia de los criterios homogéneos de edición de los datos recogidos para el estudio comparativo de los aspectos cualitativos de disponibilidad léxica.

### **2.11.2 Sociolingüística**

Los estudios descriptivos sociolingüísticos permiten analizar el léxico disponible de una comunidad de habla en relación con las variables sociales de dicho grupo y establecer su correspondiente estratificación sociolingüística. Como ya hemos descrito anteriormente, las variables extralingüísticas que se toman en cuenta en los distintos trabajos atienden a diferencias de sexo, edad, nivel sociocultural, lugar de residencia y tipo de centro, a las cuales se añaden otros parámetros considerados en estudios de

disponibilidad léxica en alumnos de ELE como lengua materna, nivel de conocimiento del español y conocimiento de otras lenguas.

Como señala Samper Hernández (2002) el análisis de esta variable ha demostrado, entre otras cosas, la relación sistemática entre el nivel sociocultural bajo y una disponibilidad léxica menor a diferencia de los hablantes de otros estratos, lo cual coincide con los postulados de la “teoría del déficit” de Bernstein<sup>6</sup>.

Queda todavía por profundizar el análisis cualitativo de los datos en esta específica disciplina de estudio.

### **2.11.3 Psicolingüística**

Con respecto a la psicolingüística, las investigaciones sobre disponibilidad pueden aportar información muy relevante para el estudio de la estructura del lexicón mental. De hecho, ofrecen la posibilidad de establecer relaciones entre las diferentes palabras actualizadas por los informantes, característica que las distingue de las pruebas utilizadas en psicolingüística donde no se pide una única respuesta al estímulo que se proporciona al hablante (Aitchison, 1987; López, 1995) y por lo tanto, presentan un carácter más espontáneo a diferencia de las pruebas tradicionales.

Además, como cita Samper Hernández (2002), otra demostración de relación especial entre vocablos se puede obtener mediante los “índices de cohesión” ideados por M. S. Echeverría, los cuales indican el grado de cohesión entre palabras. El índice será mayor si los informantes escriben las mismas unidades léxicas dentro de un centro de interés, mientras que será menor si las respuestas son más variadas y difusas en relación con el estímulo dado.

### **2.11.4 Etnolingüística**

Las diferencias entre las distintas culturas y costumbres de las diversas comunidades analizadas, junto con su país, es otro aspecto que se puede relevar gracias a los trabajos de disponibilidad léxica. Samper Padilla (2003) señala el interés aportado

---

<sup>6</sup> En su teoría, Bernstein (1971, 1975) sostiene que la diferencia entre los sociolectos bajos y otros superiores por nivel sociocultural, no se debe solo al empleo de unidades lingüísticas distintas, sino también a la presencia/ausencia de elementos. En otras palabras, la menor cantidad de elementos lingüísticos en el estrato social bajo se diferencia de los estratos superiores por factores cuantitativos. Sin embargo, como destaca López Morales (1996) estas diferencias no implican dificultades cognitivas en el sociolecto que ha dado lugar a resultados más bajos como algunos han querido suponer.

en este caso por los resultados del estudio de Carcedo, donde se halla la estrecha relación entre algunas unidades del vocabulario de la comunidad encuestada y los factores culturales, geográficos y económicos de los grupos tomados en consideración. Además, el mismo estudioso (2002:12) afirma que para la etnolingüística puede ser muy productivo el estudio de disponibilidad léxica de hablantes no nativos de español ya que en ellos se puede reflejar “la influencia de las características culturales tanto del país de origen como del que ha constituido el lugar de aprendizaje de la LE”.

### **2.11.5 Didáctica de lenguas**

Entre los varios postulados de la enseñanza-aprendizaje de lenguas, una de las áreas más fecundas de la lingüística aplicada, llama la atención la gran importancia otorgada al léxico cuando se desea lograr un aprendizaje eficiente de una lengua. En las últimas décadas varias teorías lingüísticas como el análisis del discurso, la lingüística del texto, la pragmática, la psicolingüística y la lingüística cognitiva han influido notablemente en destacar la relevancia del papel que el vocabulario desempeña en el proceso de aprendizaje de una lengua por ser el componente que da sentido al sistema lingüístico y que es fundamental para obtener el establecimiento de la comunicación.

Los datos que se obtienen de las investigaciones sobre léxico disponible pueden resultar muy útiles para su aplicación práctica a la enseñanza, un aspecto suportado por Carcedo (1998: 53-54), quien menciona que “una de las ventajas más comúnmente admitidas para los estudios de disponibilidad es, sin duda, la posibilidad de aplicar sus resultados a la enseñanza de la lengua”. En general, los trabajos realizados en este ámbito son muy valiosos para la enseñanza-aprendizaje bien de la lengua materna, bien de la lengua extranjera.

Con respecto a la primera vertiente de aplicación, es indudable el valor que los estudios de disponibilidad tienen para la planificación del aprendizaje del vocabulario. Como destaca Samper Hernández (2002), por un lado, señalan las carencias o deficiencias en la competencia léxica de los estudiantes; por otro lado, nos ofrecen el reflejo fidedigno del vocabulario de la comunidad de habla encuestada. Por lo tanto, a partir de estas evidencias, es posible seleccionar los medios para mejorar aquellas áreas que presentan fallas y evitar caer en el error de incluir en los listados destinados a la enseñanza de léxico aquellos tecnicismos, arcaísmos, cultismos, regionalismos y otras

unidades poco o nada utilizadas por los informantes, pero presentes en los diccionarios generales de la lengua española.

Sin embargo, si la investigación objetiva sobre el léxico disponible de una comunidad nativa sirve para guiar el proceso de selección de léxico en las clases de enseñanza de lengua materna, como afirma Šifrar *Kalan* (2009) existe otra posible aplicación de la metodología de estudios sobre disponibilidad léxica a estudiantes no nativos de español, cuyos resultados ofrecen una panorámica del estado de adquisición de la lengua en cuestión.

En este sentido, algunos autores (Benítez Pérez, 1994; Zebrowski, 1993) han trabajado con el objetivo de comparar las palabras que aparecían en manuales de español para la enseñanza de ELE con las aportadas por los hablantes nativos en pruebas de disponibilidad léxica a fin de ofrecer una selección adecuada para la inclusión en diccionarios escolares o en manuales de ELE para profesores y alumnos que se apoyasen sobre bases científicas sólidas y reflejasen la utilidad de las palabras según el orden de aparición en las listas de disponibilidad.

Asimismo, otros investigadores (Mateo, 1994; García Marcos, 2000) han puesto de manifiesto la importancia de los trabajos de disponibilidad sobre nativos para la enseñanza a no nativos de español LE tanto por la mejor graduación de contenido léxico que permite, como por la base de estudio muy provechosa que ofrece a los autores de material didáctico de ELE.

Además, cabe señalar la aplicación a la didáctica de lenguas de aquellos estudios sobre disponibilidad léxica en estudiantes no nativos que permiten, como explica Carcedo (1998a) en Samper Hernández (2002:14) “examinar las distintas fases de aprendizaje del léxico español, descubrir tanto los errores más usuales como las tendencias de los hablantes no nativos a lo largo de este aprendizaje, planificar la enseñanza del léxico o realizar provechosas comparaciones con los listados de las distintas zonas hispanohablantes”. Como ya hemos señalado, estas investigaciones reducen también la arbitrariedad en la selección del léxico que debe conocer un estudiante de LE (Bartol Hernández, 2010). Así, ya que en el primer proceso de enseñanza-aprendizaje el alumno no sabe distinguir la importancia de una palabra respecto de otra (Benítez Pérez, 1992), es fundamental que se seleccione el léxico que los estudiantes deben aprender en cada fase del proceso educativo y con qué prioridad.

Nuestro trabajo se propone como una aproximación al estudio de disponibilidad léxica en estudiantes italianos de ELE y no quiere distanciarse de los objetivos y posibles aplicaciones educativas que hemos descrito. Por lo tanto, aprovechando también de las aportaciones de los trabajos anteriores, nos centraremos sobre la investigación de cuál es el vocabulario disponible de la muestra de los aprendices de ELE encuestados en los distintos centros de interés seleccionados para proceder con una evaluación del proceso educativo actual y añadir una nueva contribución dentro del panorama de la enseñanza de ELE a estudiantes no nativos, en nuestro caso específico a italianos.

# **BLOQUE II: PRESENTACIÓN Y ANÁLISIS DE LOS DATOS**

## **Capítulo III**

### **Investigación de disponibilidad léxica en alumnos de ELE en la Universidad de Padua**

La investigación sobre la disponibilidad léxica en alumnos de ELE en la Universidad de Padua, que se presenta a continuación, se propone como objetivo principal realizar un acercamiento al estudio del léxico disponible de los 112 alumnos participantes en la encuesta y ha sido llevada a cabo en una clase del curso de español como segunda lengua, asignatura de estudiantes del II año de dos licenciaturas diferentes pertenecientes al Departamento de Estudios Lingüísticos y Literarios (DiSLL).

Basándonos en diferentes trabajos sobre el tema de disponibilidad léxica en alumnos de ELE, a partir de la recogida de datos y de su sucesiva inserción en el específico banco de datos del programa de cálculo Dispoplex, plataforma informática gestionada por la Universidad de Salamanca, hemos procedido con el análisis de los resultados, donde se tendrán en cuenta tanto los aspectos cuantitativos, base inicial de la mayoría de los trabajos sobre disponibilidad léxica, como los aspectos cualitativos, imprescindibles para la interpretación de las conclusiones.

#### **3.1 Objetivos del trabajo**

Los objetivos específicos de esta investigación son los siguientes:

- 1) Aportar la cantidad absoluta y relativa de respuestas de los informantes en cada centro de interés.
- 2) Aportar la cantidad absoluta y relativa de vocablos en cada centro de interés.
- 3) Observar la influencia de las variables en cada centro de interés tanto en relación con el número de palabras como en la cantidad de vocablos.



- 4) Aportar los datos cuantitativos correspondientes a los índices de media de palabras por informante e índice de cohesión en cada centro de interés.
- 5) Enumerar los vocablos que ocupan los primeros lugares en cada centro de interés y realizar un análisis descriptivo de las unidades léxicas aportadas en cada campo léxico considerado.
- 6) Adjuntar en los anexos de este trabajo los vocabularios resultantes por cada centro de interés, según su índice de disponibilidad, aparición, frecuencia relativa y acumulada.
- 7) Debido al carácter introductorio de esta investigación, se optó por renunciar a establecer una segunda muestra para comparar los resultados ya que este objetivo se podrá perseguir en investigaciones sucesivas a la que se presenta en estas páginas.

A continuación explicamos de manera detallada la metodología que hemos seguido para el trabajo.

### **3.2 La muestra**

La recolección de los datos para este trabajo tuvo lugar a finales del mes de octubre de 2016, el día miércoles 26, y se hizo a través de unas encuestas realizadas a 112 alumnos de ELE de la Universidad de Padua, matriculados en su mayoría en las carreras en Mediación Lingüística y Cultural (MLC) o en Lenguas, Literaturas y Culturas Modernas (LLC). La prueba se ha llevado a cabo durante el horario de su clase de español como lengua extranjera, cuya profesora titular es Carmen Castillo Peña.

Es importante señalar que el número de participantes corresponde a los que estaban presentes en clase a la hora de realizar la encuesta y por lo tanto, se trata de una muestra de carácter no preestratificado. Asimismo, tenemos que mencionar que el universo muestral refleja un grupo heterogéneo, caracterizado por aspectos diferentes en cuanto a las variables consideradas, ya que es posible que algunos por su experiencia de estudios anteriores, tuviese un conocimiento de la lengua española mayor o inferior, variable que está debidamente analizada en las sucesivas páginas. Sin embargo, podemos considerar el grupo homogeneizado y cohesionado por el hecho de tener como

prerrequisito de acceso al curso haber aprobado el examen del curso anterior que reconocía un nivel final B1 de español (según el Marco Común Europeo de Referencia).

Además, el carácter preliminar y experimental de este estudio, que pretende ser un acercamiento a la disponibilidad léxica en español de un grupo de hablantes no nativos, así como los objetivos que se quieren perseguir, hicieron adecuada la selección de una muestra reducida con respecto al número de encuestados de las investigaciones estándar<sup>7</sup>.

### **3.3 Condicionantes extralingüísticos considerados**

Debido al carácter no Panhispánico de nuestro proyecto, hemos tenido que prescindir de algunas variables consideradas usualmente en estos tipos de investigaciones y, basándonos en otros trabajos realizados sobre hablantes no nativos, hemos decidido añadir otras.

Como se podrá comprobar en la ficha personal que se junta a continuación (ver Anexo I), a la hora de seleccionar las variables hemos considerado la situación de la comunidad en cuestión y todos los elementos que podían ser relevantes para el análisis de los resultados y las finalidades del trabajo.

En esta investigación se han considerado como variables de tipo extralingüístico las siguientes:

- a) Sexo
- b) Edad
- c) Carrera de estudios
- d) Zona geográfica de procedencia
- e) Nivel sociocultural
- f) Lengua materna
- g) Conocimiento de otras lenguas
- h) Años de estudio del español

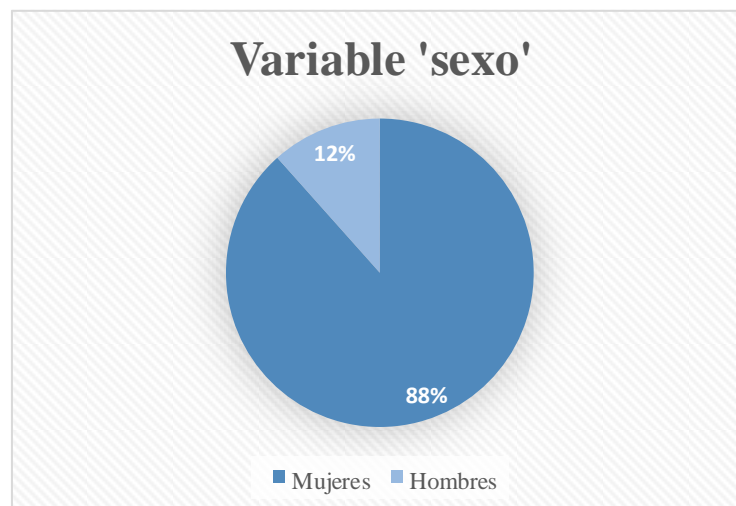
---

<sup>7</sup> Normalmente el número mínimo de informantes en una encuesta del Proyecto Panhispánico es de unos 400 participantes.

A diferencia de otros trabajos sobre no nativos, se considera en esta investigación un solo grupo de informantes, pertenecientes todos a un tipo de centro que, como hemos anticipado, es la Universidad de Padua, centro público de educación universitaria, ubicado en la zona urbana. Como ya hemos explicado, dado el carácter experimental de este trabajo se ha decidido tomar en cuenta una muestra limitada que se podrá comparar con otra diferente en un segundo momento en trabajos posteriores.

### 3.3.1 Variables 'sexo' y 'edad'

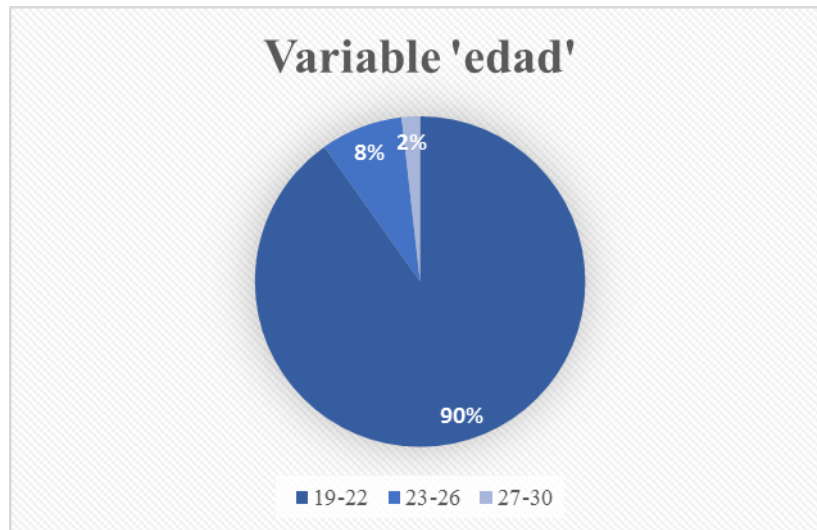
Empezamos a observar el número de informantes según la variable 'sexo' y destacamos una vez más, como en otras investigaciones, una superioridad cuantitativa femenina respecto de los sujetos de sexo masculino: 13 hombres frente a 99 mujeres, es decir, un 12% frente a un 88% aproximadamente (ver figura 1).



**Figura 1.** Distribución de los informantes según la variable 'sexo'.

Asimismo, la muestra se compone de sujetos que hemos decidido agrupar bajo otra variable, es decir, la edad de los informantes. Debido a la ligera heterogeneidad del grupo con respecto a este aspecto, hemos decidido estructurar la muestra en 3 intervalos de edad: 19-22, 23-26 y 27-30. El primer grupo, de 19 a 22 años, agrupa a 101 informantes (un 90% del universo muestral aproximadamente); el segundo, de 23 a 26 años, a 9 estudiantes (un 8% del total) y el último a dos alumnos, lo cual representa un 2% aproximadamente (ver figura 2).

Como es evidente, el grupo de edad predominante es el primero, lo cual se encuentra en línea con las expectativas debidas al haber realizado la encuesta en una clase del segundo año de estudios universitarios y considerando que normalmente los estudiantes se matriculan al primer año de estudios entre los 19-20 años de edad.



**Figura 2.** Distribución de los informantes según la variable ‘edad’.

### 3.3.2 Variable ‘carrera de estudios’ y ‘zona geográfica de procedencia’

Continuamos el análisis de la muestra tomando en cuenta otras dos variables, es decir la ‘carrera de estudios’ de los informantes y la ‘zona geográfica de procedencia’ de cada uno de ellos.

Como ya se ha anticipado en la introducción a este apartado (figura 3), es evidente que los encuestados se pueden colocar principalmente en dos carreras de estudios: un 76% de los participantes (85 personas sobre el total) estudia Mediación Lingüística y Cultural (MLC) y un 20% aproximadamente (23 informantes) se ha matriculado en el curso de Lenguas, Literaturas y Culturas Modernas (LLC). El resto de la muestra se compone de 4 personas que asisten al curso de español aunque su carrera de estudios sea diferente (3 personas aprovechan de una beca Erasmus y otra estudia Historia).



**Figura 3.** Distribución de los informantes según la variable 'carrera de estudios'.

Asimismo, es interesante considerar la variable que analiza la procedencia de los encuestados. Dada la ubicación del centro de estudios donde se ha realizado la encuesta, esperábamos que predominasen los informantes procedentes de la zona de Padua, los cuales sin embargo alcanzan solo un 38% sobre el total con 43 participantes. Por otro lado, destaca un 62% (69 encuestados) que proceden de otras ciudades, sean de la misma región (Veneto) o de otras partes de Italia, con la excepción de pocas personas procedentes de otros países (Reino Unido, Rumania, Croacia). (Ver figura 4).



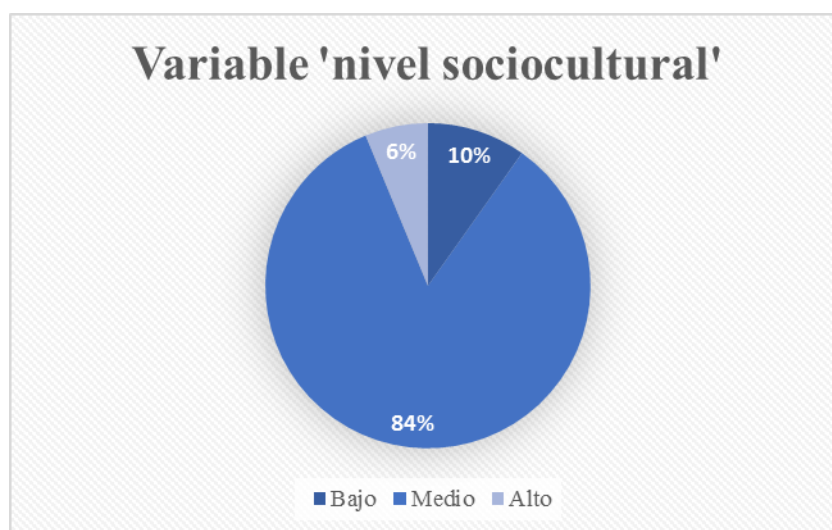
**Figura 4.** Distribución de los informantes según la variable 'carrera de estudios'.

### 3.3.3 Variable ‘nivel sociocultural’

Basándonos en el método que hemos descrito en las páginas anteriores, es decir la propuesta de tabulación de Samper Padilla y Hernández Cabrera (1997: 229-239), hemos establecido tres niveles socioculturales: 1) alto, 2) medio y 3) bajo, ponderando dos parámetros: el nivel de estudios y la profesión de los padres de cada informante.

El nivel sociocultural alto aglutina a 7 informantes (un 6% del total de los encuestados); el medio se compone de 94 personas (un 84% aproximadamente del universo muestral), y el bajo agrupa los restantes 11 estudiantes de la muestra objeto de estudio, resultando un 10% sobre el total.

Entre los datos relacionados con esta variable destaca el porcentaje del nivel sociocultural medio, que agrupa a la mayoría de los informantes dejando a los niveles bajo y alto una menor representatividad en la muestra total.



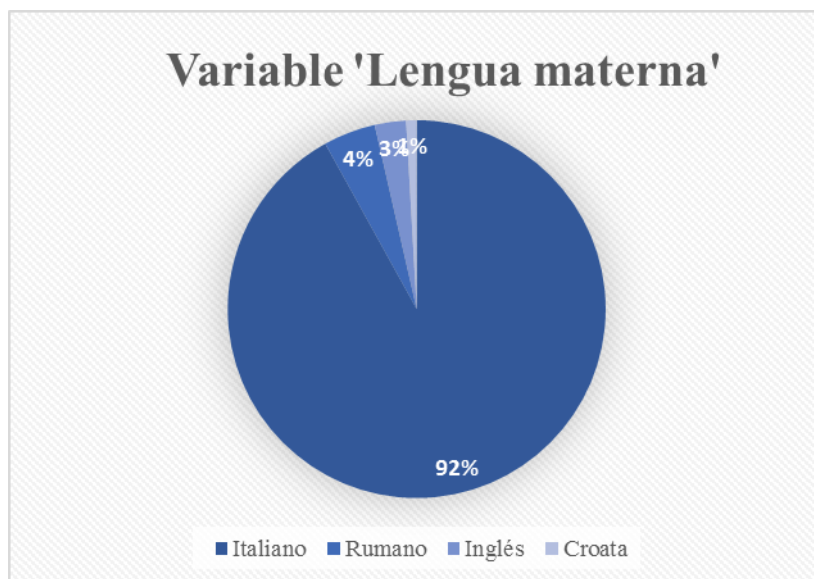
**Figura 5.** Distribución de los informantes según la variable ‘nivel sociocultural’.

### 3.3.4 Variables ‘lengua materna’ y ‘conocimiento de otras lenguas’

Otras variables que hemos investigado en este trabajo son ‘lengua materna’ y ‘conocimiento de otras lenguas’.

Como hemos señalado anteriormente, el universo muestral no ha sido preestratificado, por esta razón y a fin de obtener una visión de conjunto, hemos decidido agrupar a los informantes según el criterio de la lengua materna.

A continuación, la figura 6 muestra los diferentes grupos, cuatro en total, que componen la muestra. Destaca el dato que señala una mayoría de 103 hablantes nativos de italiano (un 92% sobre el total); los restantes informantes se subdividen entre 5 alumnos que hablan rumano como lengua materna (el 4% de los encuestados), 3 estudiantes de lengua materna inglesa (casi el 3% del universo muestral) y solo una persona de lengua croata que resulta aproximadamente en un 1% (ver figura 6).

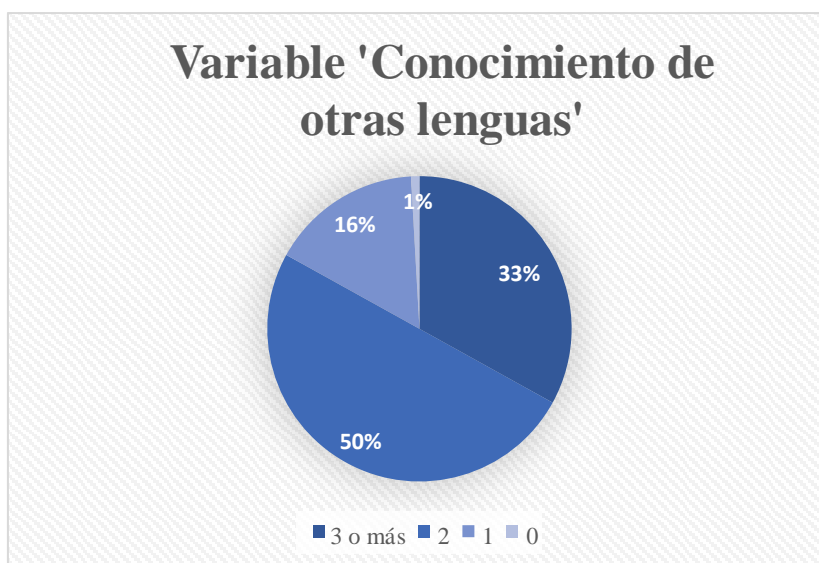


**Figura 6.** Distribución de los informantes según la variable 'lengua materna'.

Asimismo, hemos analizado el conocimiento de otras lenguas, es decir, una variable que se ha tomado en cuenta en otras investigaciones sobre hablantes no nativos y puede ser muy útil para proporcionar información sobre la influencia de otros idiomas en el aprendizaje de una nueva lengua.

Por necesidades técnicas de inserción de datos en el programa de cálculo estadístico, hemos decidido simplificar la introducción de las respuestas de los informantes según 4 criterios, sin explicitar cada lengua hablada por los encuestados. Prescindiendo por obvias razones de la lengua materna del informante y de la lengua española objeto de estudio de este trabajo, han resultado cuatro subdivisiones: 1) ningún conocimiento de otras lenguas (un único informante), 2) conocimiento de una lengua (18 alumnos – 16% de los encuestados), 3) conocimiento de 2 lenguas (56 personas – 50% del universo muestral), 4) conocimiento de tres o más lenguas (37 estudiantes – 33% sobre el total). (Ver figura 7).

Es interesante destacar el hecho de que la mitad exacta de los informantes domina dos lenguas además de la materna y del español y que solo una persona no conoce otras lenguas. Cabe señalar que no sorprende mucho el dato total sobre el conocimiento de lenguas ya que se trata de un grado en Lenguas Extranjeras, donde es obligatorio elegir al menos dos lenguas en el propio curriculum de estudios. Será interesante ver como la cantidad de idiomas hablados afecta al aprendizaje del vocabulario de ELE.



**Figura 7.** Distribución de los informantes según la variable ‘conocimiento de otras lenguas’.

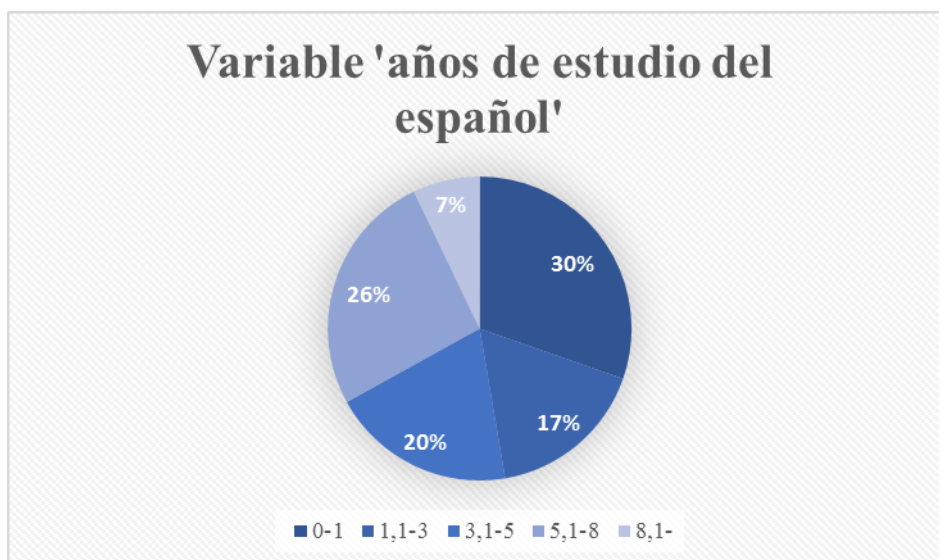
### 3.3.5 Variable ‘años de estudio del español’

La última variable que hemos tenido en cuenta en este trabajo es ‘años de estudio del español’. A la hora de insertar los datos en el programa informático, basándonos en las respuestas de los informantes, hemos decidido subdividir la variable en 5 intervalos de años de estudio que describimos a continuación.

Como destaca la figura 8, son dos los grupos que alcanzan la mayoría de las respuestas: en el momento de realización de la encuesta un 30% del total de estudiantes llevaba estudiando español de 0 a 1 año (34 informantes) y un 26% de 5,1 a 8 años (29 alumnos). Es interesante señalar como, a pesar de la diferencia de intervalo entre los dos agrupamientos, haya escasa diferencia en términos de número de informantes por cada grupo. El restante 44% sobre el total de informantes está repartido entre 19 personas (el 17% del universo muestral) que llevaban estudiando la lengua objeto de este estudio de



1,1 a 3 años; 22 alumnos (un 20%) de 3,1 a 5 años y por último, una minoría de 8 informantes se coloca en el intervalo más alto, es decir desde 8,1 años en adelante (ver figura 8).



**Figura 8.** Distribución de los informantes según la variable ‘años de estudio del español’.

### 3.4 La encuesta

La encuesta, como en la mayoría de las investigaciones realizadas sobre este tema, se proporcionó a cada alumno presente en la clase y se dividió en dos partes bien diferenciadas.

En primer lugar, después de explicar a los participantes los propósitos de la investigación y su modalidad, se entregó al estudiante una ficha preparada previamente en la cual al informante se le pedía que aportase informaciones personales que hemos utilizado para obtener los datos de las variables citadas en el apartado anterior, además de preguntar otros datos de carácter más o menos relevante (ver Anexo I). Se preguntó por ejemplo sobre los estudios anteriores y si el informante había vivido una experiencia de estudios o de residencia en España, criterio que sería vinculante solo en caso de una larga estancia, situación que no ha aparecido en el total de la muestra.

En segundo lugar, se continuó con la entrega a cada estudiante de un folio blanco con rayas, que una vez doblado por la mitad creó una división en ocho columnas de 31 líneas disponibles para las respuestas necesarias a realizar la parte que forma el cuerpo principal de la encuesta, es decir la prueba sobre disponibilidad léxica. Siguiendo el

procedimiento tipificado del Proyecto Panhispánico, el informante tenía a su disposición un tiempo limitado de dos minutos, cronometrado por el encuestador, a lo largo del cual debía enumerar en listas abiertas la mayor cantidad de vocablos que se le ocurriesen en relación con cada uno de los ocho centros de interés seleccionados para nuestro trabajo. Cada columna estaba asignada a un centro de interés; sin embargo, las indicaciones con el nombre de cada centro no aparecían para evitar que los estudiantes se distrajeran y que se adelantaran antes del tiempo establecido. Se pidió que pusiesen atención al escribir las respuestas empezando por la primera línea puesta en alto continuando hacía abajo, justificando la solicitud con los objetivos de la investigación, que quieren buscar y calcular el léxico disponible del estudiante, que acude a la mente del participante en la manera más rápida y espontánea. Además, se enfatizó que los informantes escribieran todas las palabras que se le vinieran a la mente, teniendo en cuenta que se podía tratar de sustantivos, adjetivos y verbos, incluso si los informantes no estaban seguros de su ortografía o existencia. A continuación, se revelaba a los alumnos el nombre del centro de interés encuestado, uno a uno, y se pedía que escribiesen al lado del centro en cuestión el número asignado como se enumera a continuación.

Dado el carácter experimental y descriptivo de la prueba, teniendo en cuenta los propósitos del Proyecto Panhispánico de estudio de disponibilidad léxica y algunas investigaciones en estudiantes de español como lengua extranjera, hemos optado para una elección de 8 de los 16 centros de interés del Proyecto Panhispánico, correspondientes a la mitad del número total. Para la selección de los centros de interés nos hemos servido como base de los temas del *Marco común europeo de referencia para el aprendizaje, la enseñanza y la evaluación de lenguas*. El MCER reconoce cuatro principales ámbitos en los que se produce la comunicación en relación a la enseñanza<sup>8</sup> y para nuestro trabajo hemos decidido abordar aquellas áreas donde se podían ubicar la mayoría de los centros de interés normalmente investigados<sup>9</sup>.

Así, los centros que se tomarán en cuenta en esta investigación son los siguientes:

- 1) ‘Partes del cuerpo’
- 2) ‘La ropa’

---

<sup>8</sup> El ámbito personal, que comprende las relaciones familiares; el ámbito público, que se refiere a todo lo relacionado con la actividad social habitual; el ámbito profesional, vinculado al ejercicio de la profesión; y el ámbito educativo, relacionado con el contexto de aprendizaje (Consejo de Europa 2001: 15, 49-50).

<sup>9</sup> Cf. J.A. Bartol Hernández (2010: 95) y Florentino Paredes García (2015: 8)

- 3) 'Partes de la casa'
- 4) 'Comidas y bebidas'
- 5) 'La escuela'
- 6) 'La ciudad'
- 7) 'Animales'
- 8) 'Profesiones y oficios'

Como se puede observar, hemos decidido omitir algunos centros de interés que coinciden con los ámbitos contemplados por el MCER motivando esta elección con la presumible concomitancia de respuestas en los centros que hemos seleccionado como 'Partes de la casa' con otros que comparten algunos aspectos como 'La cocina y sus utensilios' o 'Los muebles de la casa'. Asimismo, no hemos incluido los centros de interés 'El campo' o 'Medios de transporte' dada la posible coincidencia de asociaciones con el centro más general 'La ciudad'.

### **3.5 Edición de errores**

Tras haber completado la prueba, hemos dedicado gran parte del trabajo de investigación a la transcripción de los datos proporcionados por cada informante. Al término de este proceso, hemos pasado a la fase de edición del material, un pasaje que requiere mucha atención debido al gran número de problemas que el investigador encuentra a medida que limpia el material. A fin de poder proceder a introducir en el programa de cálculo una muestra homogeneizada que pudiese garantizar una mayor precisión en el cotejo de los resultados, nos hemos conformado a las normas comúnmente utilizadas en las investigaciones del Proyecto Panhispánico, señaladas en Samper Padilla (1998) y en Samper Padilla, Bellón Fernández y Samper Hernández (2003). A continuación se enumeran los criterios seguidos:

- 1) La eliminación de los términos repetidos: cuando un informante repite una respuesta dentro de un mismo centro de interés.
- 2) La corrección ortográfica: es muy común en investigaciones en estudiantes no nativos de español encontrar errores ortográficos, debidos a vacilaciones fonéticas o a deficiencias en el aprendizaje.

- 3) La unificación ortográfica: se aglutinan bajo una sola entrada las palabras que admiten dos grafías, manteniendo la que aparece mayormente entre los encuestados o en el diccionario de referencia (en nuestro caso el DRAE).
- 4) La neutralización de las variantes meramente flexivas: se reduce una palabra a la forma no marcada del paradigma. Los verbos se recogen al infinitivo, los adjetivos y los sustantivos al masculino, con la excepción de los sustantivos heterónimos o de aquellos cuyo femenino se caracteriza por tener un sufijo específico. Otras excepciones se señalan en los apartados a continuación.
- 5) La unificación de los derivados regulares: las formas no marcadas del diminutivo y del aumentativo, que no suponen una lexicalización o alteración del significado léxico, se reducen a la forma regular.
- 6) La unificación de formas plenas y acortamientos: en este caso se homogeneiza bajo una misma entrada las palabras susceptibles de ser acortadas.
- 7) El mantenimiento de marcas comerciales plenamente lexicalizadas: se admiten normalmente solo aquellas plenamente lexicalizadas en el uso de la lengua.

Asimismo se han seguido los criterios específicos de edición de errores en las investigaciones en estudiantes no nativos de español como aportado por Samper Hernández (2002):

- a) La corrección gráfica: se adaptan a la forma canónica castellana las palabras incorrectas desde del punto de vista gráfico.
- b) El tratamiento de los extranjerismos: no se aceptan palabras que no muestran cambio de la lengua materna; se mantienen entradas que presentan la ortografía de la lengua extranjera si tienen un uso frecuente en la lengua española; asimismo, se toman en cuenta los extranjerismos que aunque no aparezcan bajo la forma correcta del español, tienen una construcción similar a demostración del intento del alumno de adaptarse a la lengua de llegada.
- c) La unificación en la presentación de las unidades: las entradas siguen los mismos principios para la determinación de su forma final.
- d) La amplitud en las relaciones asociativas: la capacidad de los hablantes no nativos de generar asociaciones más o menos cercanas al centro de interés encuestado es típica de los hablantes no nativos de una lengua en las

investigaciones sobre disponibilidad léxica, debido al bajo conocimiento de la lengua. A continuación se podrá comprobar cuáles decisiones se han tomado en relación con este tema específico.

### 3.6 Toma de decisiones

#### 3.6.1 ‘Las partes del cuerpo’

A la hora de editar sus datos, el centro de interés en cuestión no causa demasiados problemas debido a la gran concreción de sus resultados y a su carácter cerrado, caso que parece compartir con varias investigaciones similares.

En nuestro trabajo, encontramos entradas asociativas que pertenecen a varias categorías gramaticales como adjetivos (*alto, ancho, azul, rubio*), verbos (*andar, arrodillarse, bailar, tocar*) y sustantivos (*oculista, ortopedia*).

Nuestro propósito de investigación nos ha empujado a eliminar la menor cantidad posible de palabras y hemos decidido prescindir solo de las inexplicables, inexistentes o indescifrables, como el caso de ‘*bocal*’, quizás un intento de asociación derivativa con la palabra más cercana y similar, es decir ‘*boca*’. Por la misma razón, pero con resultado contrario, hemos decidido incluir en el cotejo la expresión ‘*labio dental*’, asociación realizada junto a términos como ‘*labio*’ y ‘*boca*’ y convertida según los criterios de edición en ‘*labiodental*’.

Un aspecto que se ha mantenido para todos los centros de interés es la edición de todo el material en su forma singular, con las debidas excepciones que señalaremos. Además, hemos unificado los diminutivos bajo su forma regular (‘*bajito*’ > ‘*bajo*’).

Otro rasgo compartido con los demás centros de interés es la presencia de numerosos errores ortográficos, cuya causa principal es atribuible a la deficiencia del aprendizaje o a la distracción debida a la limitación temporal de la prueba. Sin enumerar todos los errores de acentuación, destacan por el gran número de repeticiones los siguientes: *naríz* > *nariz*, *estomago* > *estómago*, *pié* > *pie*. Además, entre los errores gráficos en relación con la fonética, señalamos los causados por una dificultad en la distinción en los fonemas /p/, /b/ como a continuación: ‘*bierna*’ > ‘*pierna*’, ‘*capeza*’ > ‘*cabeza*’. Asimismo, hemos completado el proceso de edición de datos en relación a

este centro de interés a través de la interpretación por analogía de palabras como ‘*esquina*’ > ‘*espalda*’, ‘*piéd*’ > ‘*pie*’, ‘*sobrecillas*’ > ‘*cejas*’.

### 3.6.2 ‘La ropa’

Por lo que se refiere a este centro de interés, cabe señalar que, como en el primer centro de interés, aquí también se encuentra un carácter concreto aunque menos cerrado que el anterior. De hecho, además de encontrar palabras relacionadas con las prendas de vestir, aparecen términos de accesorios y complementos como ‘*anillo*’, ‘*pendientes*’, ‘*pulsera*’ que hemos decidido dejar en línea con nuestro propósito principal de presentar todos los tipos de asociaciones semánticas formuladas por los informantes. Por lo tanto, hemos dejado también términos pertenecientes a diferentes categorías gramaticales como adjetivos que guardan relación con los colores o la descripción de la ropa (*ancho*, *amarillo*, *deportivo*, *nuevo*), verbos (*abrigarse*, *ir de compras*, *poner*, *vestir*, *lucirse*) y otros sustantivos que mantienen una relación asociativa con el centro de interés, como por ejemplo los tejidos (*seda*, *pana*) o los sitios (*tienda*, *probador*). Hemos tenido que prescindir solo de algunos términos no comprensibles o sin relación directa como *chantaje*, *chopela*, *luva*.

Uno de los criterios generales que hemos seguido en todo el proceso de edición de datos ha sido unificar las formas en masculino y singular tanto con los sustantivos como con los adjetivos que se presentaban en femenino o plural. Asimismo, cuando aparecían verbos tanto en forma transitiva/intransitiva como pronominal, se ha optado para la homogeneización bajo la forma no pronominal. Además, para la decisión de la forma final que hemos introducido en la herramienta informática nos hemos basado en su presencia en el DRAE; es por esta razón por la que aparece la forma singular *braga*, *calzoncillo* o *pantalón*, por ejemplo, aunque normalmente su uso se hace en plural. Excepciones se hicieron en los casos de los términos ‘*gafas*’ y ‘*tacones*’ que siempre se han encontrado en la forma plural y se ha decidido dejar así en la muestra ya que normalmente se utilizan en plural.

Además, hemos incluido varios compuestos como ‘*gafas de sol*’, ‘*camiseta de tirantes*’, ‘*ropa interior*’, ‘*pantalón corto*’ y hemos unificado los compuestos ‘*zapatilla deportivas*’ bajo ‘*zapatillas de deporte*’ y ‘*zapatos de tacos/de tacones/con tacón/zapatos altos*’ en ‘*zapato de tacón*’ que se mantiene junto con el término

‘*tacones*’ por ser también una variante del mismo núcleo lexemático y por aparecer varias veces. Hemos tenido que interpretar algunas palabras en este centro también, y hemos optado por modificar los términos que se presentaban como ‘*mantel*’ en ‘*capa*’. Aquí, se trata probablemente de una interferencia con la lengua materna que no hemos querido eliminar o dejar en su forma menos contextualizada para mantener la línea de nuestros propósitos iniciales.

Por último, señalamos que hemos optado también por incluir en nuestra lista los extranjerismos ‘*jeans*’, ‘*t-shirt*’ ya que aparecen en el DRAE y además, se han tenido en cuenta en otras investigaciones. Asimismo, se han incluido también aquellas variables utilizadas en algunos países hispanohablantes de América del Sur como ‘*bombacha*’, ‘*remera*’, ‘*pollera*’.

### 3.6.3 ‘Las partes de la casa’

Como ocurre en otras investigaciones, el centro de interés partes de la casa ha producido un gran número de asociaciones que se amplía con respecto de su núcleo principal. Este aspecto se evidencia por la presencia de numerosos datos que no representan solo específicas partes de la casa, sino también muebles (*armario, cama, mesa*), electrodomésticos (*microondas, frigorífico, lavavajillas*) utensilios específicos de la cocina (*cuchillo, cuchara, cucharón*) y otros elementos que se suelen utilizar para amueblar la casa y sus partes (*alfombra, cortina, lámpara*). Además, también en este centro de interés se repiten verbos y adjetivos que tienen relación con el tema en cuestión como ‘*alquilar*’, ‘*acostarse*’, ‘*almorzar*’ o ‘*maravilloso*’, ‘*luminoso*’, ‘*limpio*’. Aparecen también sustantivos que tienen una amplia relación asociativa con el centro, como los compuestos ‘*compañero de piso*’, ‘*reunión de familia*’ o el más general y abstracto ‘*libertad*’, que se han mantenido para el interés de la investigación y el objetivo de crear un vocabulario de cada centro encuestado.

Después de completar el proceso de unificación de los datos, siguiendo otras investigaciones decidimos mantener distintas variantes que pertenecen al mismo núcleo lexemático como *baño/cuarto de baño, cuarto/cuarto de estar, sala/sala de comer/sala de estar/sala de baño*.

Durante el proceso de edición de errores hemos homogeneizado los acortamientos ‘*tele*’ o ‘*tv*’ con la palabra correspondiente ‘*televisión*’ y nos hemos enfrentado con

algunas interesantes interferencias con el francés como ‘*verro*’ por ‘*vaso*’, ‘*fenetra*’ por ‘*ventana*’ y ‘*etaje*’ por ‘*piso*’ y otra con el portugués en ‘*ropeiro*’ en lugar de ‘*ropero*’. En fin, excluimos ‘*nevrelgico*’ de la lista de palabras por no haber encontrado correspondencia o interpretación alguna de la palabra.

#### 3.6.4 ‘Comidas y bebidas’

El centro de interés en cuestión, aunque se caracterice por ser uno de los más productivos, no ha creado demasiadas dificultades durante el proceso de corrección y edición de errores. Después de convertir los errores ortográficos en la forma regular, teniendo en cuenta también de los datos que se encontraban influidos de alguna manera por la lengua materna, lo que más destaca es la presencia de numerosos sustantivos estrechamente relacionados con el tema. Tampoco faltan verbos (*beber*, *comer*, *merendar*), adjetivos (*rico*, *vegetal*, *salado*) y sustantivos que designan lugares (*cervecería*, *chiringuito*, *pastelería*, *supermercado*). Es interesante además señalar que entre los distintos tipos de comida y bebidas enumeradas en la lista, la influencia de la cultura italiana es evidente: sustantivos como ‘*pizza*’, ‘*lasagna*’ y ‘*pasta*’ se encuentran con frecuencia. Sin embargo, la categoría de alimentos de la tradición hispana no es menor; al lado de comidas y bebidas típicas españolas como ‘*paella*’, ‘*gazpacho*’ y ‘*sangría*’ o de la tradición regional como ‘*fartón*’ o ‘*ensaimada*’, a numerosos informantes se le han ocurrido palabras como ‘*burrito*’, ‘*enchilada*’, ‘*empanada*’ o ‘*banana*’, términos procedentes de Latino América que se ha decidido dejar por la amplitud de asociaciones encontradas y por su aparición en el DRAE.

Asimismo, hemos mantenido como entradas los nombres de refrescos designados por su marca comercial como ‘*coca-cola*’ ‘*sprite*’, ‘*fanta*’ o ‘*pepsi*’ por ser bastante comunes entre los encuestados y por estar lexicalizados en la lengua española.

Por último, encontramos la presencia de entradas especiales como ‘*agua con gas/sin gas*’; diferentes tipos de café, zumo, jugo y jamón (*café cortado/con leche*; *zumo de naranja/de manzana*) y sustantivos de origen inglés que siguen las normas recomendadas por la RAE (*sándwich*, *kétchup*).



### 3.6.5 ‘La escuela’

La gran amplitud asociativa que ha caracterizado los centros de interés anteriores domina incluso en este caso específico. Al lado de palabras que se refieren a objetos y materiales de la escuela (*lápiz, cuaderno, libro*), hay muchas que distinguen estancias y muebles que se encuentran en este lugar (*silla, pupitre, aula, gimnasio, despacho*) y otras que designan relaciones con el exterior, por ejemplo los medios de transporte utilizados por los estudiantes para desplazarse (*autobús, bicicleta, tren*).

Una vez más aparecen verbos (*aprobar, asistir, rechazar, suspender*), adjetivos (*escolar, público, privado*), el adverbio ‘*temprano*’ o sustantivos de relación semántica más amplia (*amistad, cansancio, suerte*).

Tampoco faltan las ocurrencias de nombres de diferentes asignaturas: ‘*latín*’, ‘*física*’, ‘*literatura*’ o ‘*matemáticas*’ y ‘*ciencias*’, que hemos mantenido en plural por ser las formas más usadas según señala el DRAE, así como ‘*idiomas extranjeros*’. A lo largo del proceso de edición de datos ha aparecido el nombre de la marca comercial ‘*Scotch*’ que se ha convertido en ‘*cinta*’. Además, hemos desechado el apellido de la profesora encargada del curso asistido por la participantes en la encuesta y el nombre de la ciudad donde se encuentra el centro de estudios de los informantes.

Hemos modificado también aquellas palabras que aparecían en la forma acortada (*boli, cole, profe*) bajo su forma plena (*bolígrafo, colegio, profesor*). Además, hemos unificado bajo su forma regular ‘*papel*’ la forma derivada ‘*papelito*’ y hemos optado para mantener el anglicismo ‘*test*’ dado que la entrada aparece en el DRAE. Esta herramienta de comparación ha sido muy utilizada a lo largo del proceso de toma de decisiones y nos ha ayudado en la fase de edición de datos cuando, en lugar de convertir todas las palabras que se encontraban en plural hacia el singular, hemos decidido mantener las formas ‘*apuntes*’ y ‘*deberes*’ ya que en el DRAE se señala su mayor uso en plural.

Por último, hemos optado por simplificar algunas lexías que aparecían compuestas ya que se encontraban más bajo la forma simple. Algunos ejemplos son las entradas ‘*aula de estudio*’ y ‘*despacho del profesor*’. Como ya hemos afirmado, mantener todos los datos lo más posible cercanos a la realidad es nuestro objetivo. Sin embargo, para valorizar la muestra tal como ha sido producida y poder cuantificar las palabras, hemos pensado que en este caso unificar sería la solución mejor.

### 3.6.6 'La ciudad'

Como ya se puede imaginar dada la naturaleza asociativa de este centro de interés, el número de entradas generadas por los informantes ha sido bastante alto y amplio. Hemos decidido, como con los centros de interés pasados, eliminar los topónimos (*Barcelona, Valencia, Padua*) y palabras que fuimos incapaces de interpretar como 'quasques'.

Además, en este centro de interés se repite la tendencia a producir palabras de varios géneros gramaticales, como adjetivos generalmente destinados a describir una ciudad o sus elementos (*antiguo, innovador, moderno*), verbos (*descubrir, pasear, voltar*). Sin embargo, la mayoría de respuestas proporcionadas son sustantivos, muchos que tienen relación con lugares, edificios e incluso personas que desempeñan varias actividades u oficios en una ciudad.

Durante el proceso de edición de datos hemos encontrado un número alto de acortamientos que, siguiendo las normas de corrección de datos, hemos convertido en su lexía completa como 'bus' en 'autobús', 'bici' en 'bicicleta', 'moto' en 'motocicleta', 'metro' en 'metropolitano', 'auto' en 'automóvil'. Al lado de éstos, aparecen palabras coloquiales como 'disco' que se ha convertido en 'discoteca'. Asimismo, hemos modificado los verbos que se presentaban conjugados como 'gira' y 'sigue', normalmente utilizados para proporcionar información sobre las direcciones en una ciudad, unificándolos bajo el infinitivo 'girar' y 'seguir'.

Por fin, hemos actualizado lexías complejas como 'paso de peatones' y 'paso de cebra' o 'dar una vuelta', 'dar un paseo' o 'ir de compras'.

### 3.6.7 'Animales'

Por ser un centro bastante compacto de producción de entradas, el tema en cuestión no se caracteriza por ser uno de los más productivos, sino por la enumeración de muchas palabras que contienen intentos de acercarse a las españolas sin ser correctas.

Al lado de listados de animales, como ocurre en los otros centros de interés, destacan las ocurrencias de otras categorías gramaticales, que a veces aparecían junto con los sustantivos más comunes, sobre todo en aquellos casos donde el informante no producía muchas expresiones. Por esta razón encontramos verbos (*ladrar, cuidar, morder*) o adjetivos (*agresivo, cariñoso, doméstico*) o sustantivos que testimonian la

gran cantidad de asociaciones producidas por los encuestados como el interesante ‘*Caperucita Roja*’, los nombres abstractos ‘*felicidad*’, ‘*fidelidad*’, ‘*libertad*’ y muchos otros que están en relación con el mundo de los animales como ‘*safari*’, ‘*veterinario*’, ‘*zoológico*’.

En cuanto a la eliminación del material obtenido, hemos decidido dejar afuera los nombres propios ‘*África*’ o ‘*Taxi*’ (que indica el nombre de la mascota de un informante), palabras que no tienen asociación directa como ‘*departamento*’ o ‘*talones*’ y algunas palabras indescifrables como ‘*prado*’. Además, hemos optado por la unificación de los sustantivos femeninos ‘*gata*’, ‘*perra*’ bajo la forma masculina. Por otro lado, hemos decidido separar las variantes *león/leona*, *tigre/tigresa* aunque no se trate de sustantivos heterónimos como las entradas *caballo/yegua* ya que han aparecido muchas veces distintas y en la forma incorrecta. Su aparición en el vocabulario final, podrá llamar la atención de los alumnos e impulsar la memorización de las variantes correctas, evitando futuros errores debidos a interferencias con la lengua materna. Cabe señalar también la decisión de unificar bajo el nombre ‘*pez*’ la palabra ‘*pescado*’, sustantivo que se refiere a la comida.

### **3.6.8 ‘Profesiones y oficios’**

El último centro de interés encuestado se caracteriza, al igual que los pasados, por la presencia de entradas no solo estrictamente relacionadas con actividades laborales, sino también por la producción de términos que se pueden referir al mundo del mercado laboral en un sentido más amplio. Incluso en este centro destaca la ocurrencia de palabras correspondientes a distintas esferas: sustantivos que se refieren a sitios donde se realizan los trabajos (*oficina, banco, tienda*), nombres de herramientas que sirven para realizar un trabajo (*anteojos, gafas, ordenador*) verbos (*arreglar, cocinar, construir*) o adjetivos relativos a una profesión o al trabajo en general (*comercial, policial, puntual*).

El criterio de edición más utilizado junto a la corrección de la ortografía es la unificación de los datos bajo la forma no marcada del masculino, con algunas excepciones en los casos de ‘*ama de casa*’ y ‘*azafata*’ por ser profesiones fundamentalmente femeninas y que también el DRAE señala en la forma femenina; o

‘actriz’ y ‘reina’ por tener respectivamente un prefijo y un sufijo especiales y por aparecer bastante veces en lugar o al lado del masculino.

Hemos optado por mantener el anglicismo ‘*baby-sitter*’ aunque aparezca junto con el más coloquial ‘*canguro*’ y hemos prescindido solo de aquellas palabras que no tenían conexión directa con el tema central (*mama, mate, transpiración*). Asimismo, hemos unificado bajo la forma ‘*camarero*’ las ocurrencias ‘*barista*’ o ‘*camarero*’, eliminando las repeticiones si se actualizaban en la misma lista ya que la primera expresión no existe en la lengua española, en el intento de desecharnos de cuantas menos palabras posibles aún en este centro.

En general, podemos afirmar que la muestra se caracteriza por la gran cantidad de errores cometidos por los informantes en cada centro de interés, los cuales han llevado a un largo proceso de homogeneización en el cual hemos podido destacar también la gran heterogeneidad de respuestas y, por lo tanto, la extensa actividad de asociación semántica a la que nos hemos referido proporcionando varios ejemplos en las páginas anteriores. Sin embargo, siguiendo el objetivo principal de mantener casi todas las entradas actualizadas por los informantes, hemos prescindido solo de aquellos datos que no se podían descifrar o interpretar o que no tenían una relación estricta con el centro de interés encuestado. Además, dada la cercanía léxica de la principal lengua materna de los encuestados, es decir el italiano, hemos decidido no eliminar aquellas palabras que algunas veces aparecían poco diferentes de las italianas pero modificadas por los informantes en un intento de españolización. En fin, subrayamos una vez más el carácter experimental e introductorio de la investigación que nos ha dado la posibilidad de seguir las principales pautas metodológicas del Proyecto Panhispánico y a la misma vez de ajustar el método según otras investigaciones en estudiantes de ELE o según algunas consideraciones personales que han generado ligeras modificaciones respecto del método principal debido a razones evidentes.



## Capítulo IV

### Análisis y comentario de resultados cuantitativos

En este capítulo nos centraremos en la presentación de los resultados cuantitativos de la investigación, que se presentan a continuación.

#### 4.1 Resultados cuantitativos generales

Tras haber terminado la edición de los datos y la introducción de ellos en el programa de cálculo estadístico Dispolex, hemos procedido con el cálculo de los índices cuantitativos. La cantidad total de palabras obtenidas se eleva a 12.049, lo que ofrece una media de 107,59 palabras actualizadas por cada informante en las respectivas pruebas.

En la tabla 1 presentamos el número de palabras totales por centro de interés, el promedio de palabras por alumno en cada centro de interés y el rango o posición relativa que ocupan los campos léxicos según el número total de palabras.

Centros de interés	Palabras totales	Palabras por informante	Rango
01.Partes del cuerpo	1598	14,27	4
02.La ropa	1218	10,88	8
03.Partes de la casa	1460	13,04	5
04.Comidas y bebidas	1886	16,84	1
05.La escuela	1599	14,28	3
06.La ciudad	1655	14,78	2
07.Animales	1340	11,96	6
08.Profesiones y oficios	1293	11,54	7

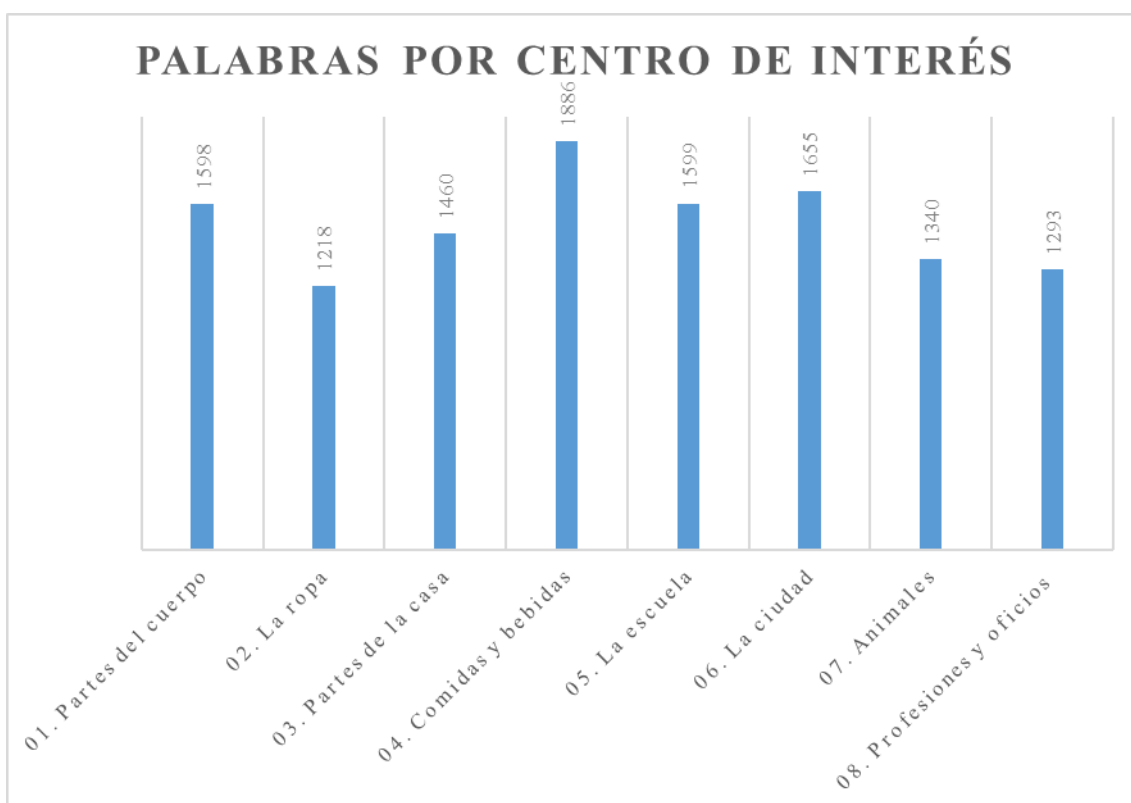
**Tabla 1.** Número de palabras y rango de cada centro de interés

A partir de estos datos es interesante observar el promedio de respuestas por cada sujeto en comparación con el promedio resultante en los trabajos hispánicos, que, como señala Samper Hernández (2002) normalmente superan casi siempre el número de 20 unidades por informante. En nuestra investigación la media total es de 13,45 respuestas, que resulta de dividir el número total de respuestas entre los 112 informantes. Por lo tanto, destaca un promedio menor a diferencia de los obtenidos en los estudios del Proyecto Panhispánico, lo cual deriva muy probablemente de la menor competencia

léxica en los participantes a nuestra encuesta, que, como ya hemos puesto de manifiesto en el capítulo anterior, son estudiantes de español como lengua extranjera, y por eso tienen un conocimiento de la lengua diferente de los nativos.

Además, en la tabla 1 se indica la enumeración del rango de cada centro de interés donde destaca al primer lugar el centro “Comidas y bebidas” (04), con un promedio de 16,84 palabras por informante, seguido por “La ciudad” (06) y “La escuela” (05). En oposición al primer lugar, al último puesto del listado encontramos el centro “La ropa” (02) con un promedio considerablemente inferior de 10,88 respuestas por alumno.

En la gráfica siguiente mostramos de manera más evidente la diferencia en el número de palabras totales aportadas en cada centro de interés:



**Figura 9.** Palabras totales por centro de interés

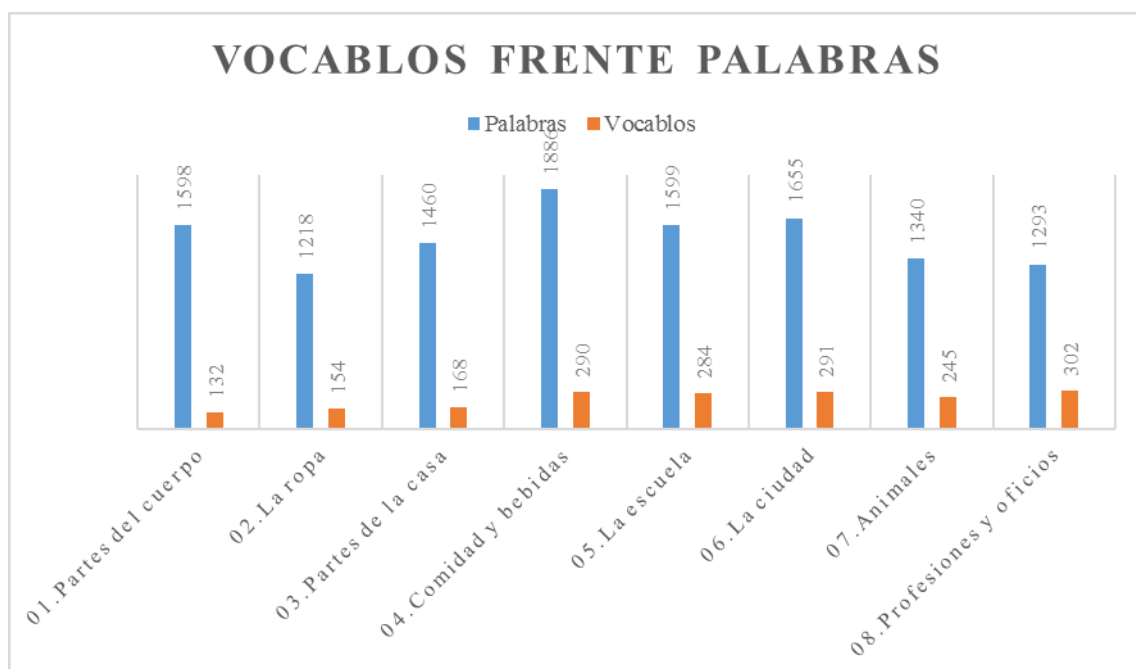
Como ya hemos señalado en el primer capítulo de esta investigación, existe una diferencia entre los términos “palabra” y “vocablo”, que normalmente se suelen utilizar como sinónimos. En los trabajos sobre disponibilidad léxica es importante distinguir los dos términos ya que designan dos conceptos diferentes. Todas las palabras respondidas por cada informante se denominan “palabras”(token, casos), mientras que las diferentes

palabras aportadas por los informantes y editadas según los criterios descritos en el tercer capítulo, se indican como “vocablos” (type, tipos). El número total de vocablos en nuestra investigación asciende a 1.882.

En la tabla 2 y en la figura 10 señalamos las diferencias numéricas en la producción de resultados según la distinción terminológica recién recordada y el rango en el cual se posiciona cada centro de interés.

Centros de interés	Palabras totales	Rango	Vocablos	Rango
01.Partes del cuerpo	1598	4	132	8
02.La ropa	1218	8	154	7
03.Partes de la casa	1460	5	168	6
04.Comidas y bebidas	1886	1	290	3
05.La escuela	1599	3	284	4
06.La ciudad	1655	2	291	2
07.Animales	1340	6	245	5
08.Profesiones y oficios	1293	7	302	1

**Tabla 2.** Producción de vocablos y rango de cada centro de interés



**Figura 10.** Número de vocablos frente a palabras por centro de interés

Como se observa, el orden de los distintos centros de interés ha sufrido algunas ligeras modificaciones. Algunos de ellos no varían mucho en relación con el rango, se



notan algunos que bajan de un punto u otros que suben de un punto. Sin embargo, llama la atención la gran diferencia que se produce en el campo léxico “Profesiones y oficios” (08), que asciende del antepenúltimo lugar en cuanto al número de respuestas totales al primero en el número de vocablos. Por el contrario, resalta el descenso del centro de interés “Partes del cuerpo” (01), cuya posición ha disminuido hasta llegar al último lugar en la nueva tabla.

Tras haber tenido en cuenta los datos que resultan sobre el número de palabras y vocablos, cabe señalar la relación que se establece entre ellos en cada centro de interés. A fin de medir esta relación, se emplea normalmente la fórmula desarrollada por Max S. Echeverría (1991), denominada índice de cohesión. Dividiendo el promedio de respuestas por el número de vocablos se mide la cohesión interna del centro de interés tomado en consideración y se pueden llevar a cabo comparaciones de índices. Si la mayoría de palabras respondidas por los informantes coinciden, el centro de interés se podrá definir compacto o cerrado. En el caso contrario, si el número de palabras diferentes ocurridas a los alumnos es alto, el campo se presentará como difuso o abierto. El índice coincidiría con la cifra 1 si todas las palabras actualizadas en el mismo centro de interés fueran las mismas, mientras que al variar de la cantidad de palabras distintas encontradas en cada centro de interés, el índice se aleja de 1.

En la tabla 3 se muestran los índices de cohesión de cada centro de interés junto con el respectivo rango según este criterio.

Centros de interés	Palabras totales	Promedio	IC	Rango
01.Partes del cuerpo	1598	14,27	0,11	1
02.La ropa	1218	10,88	0,07	3
03.Partes de la casa	1460	13,04	0,08	2
04.Comidas y bebidas	1886	16,84	0,06	4
05.La escuela	1599	14,28	0,05	5
06.La ciudad	1655	14,78	0,05	5
07.Animales	1340	11,96	0,05	5
08.Profesiones y oficios	1293	11,54	0,04	6

**Tabla 3.** Índice de cohesión por centro de interés

Es evidente que el centro de interés que presenta el índice de cohesión más alto y, por lo tanto un carácter más compacto, es “Partes del cuerpo” (01) que alcanza el 0.11, seguido por “Partes de la casa” (03) y “La ropa” (02), que no superan el 0.08. A partir de esta consideración se explica el porqué de las últimas posiciones en la ordenación por

vocablos de los tres campos léxicos citados. En el lado contrario, destaca la aparición del mismo índice de cohesión (0.05) en tres centros de carácter más abierto: “La escuela” (05), “La ciudad” (06), “Animales”. Por último, cierra el listado el centro más difuso “Profesiones y oficios”(08) que obtiene un índice de 0.04 y que en la ordenación por número de vocablos encabezaba el listado.

Finalmente, queremos terminar la presentación de los datos cuantitativos generales con la enumeración de las palabras que resultan más disponibles sobre el total de nuestro universo muestra. Para este propósito, hemos calculado cuáles palabras han alcanzado los índices de disponibilidad más altos y las presentamos a continuación en la tabla 4.

<b>Rango</b>	<b>Vocablo</b>	<b>Índice de disponibilidad</b>	<b>Centro de interés</b>
1	perro	0.93881	07.Animales
2	gato	0.85810	07.Animales
3	profesor	0.80357	08.Profesiones y oficios
4	ojo	0.73517	01.Partes del cuerpo
5	agua	0.70801	04.Comidas y bebidas
6	cocina	0.67884	03.Partes de la casa
7	camiseta	0.63352	02.La ropa
8	mano	0.60497	01.Partes del cuerpo
9	zapato	0.59360	02.La ropa
10	calle	0.52362	06.La ciudad
11	libro	0.51042	05.La escuela
12	cama	0.50027	03.Partes de la casa
13	boca	0.49513	01.Partes del cuerpo
14	pelo	0.48834	01.Partes del cuerpo
15	pantalón	0.47923	02.La ropa
16	brazo	0.46566	01.Partes del cuerpo
17	coche	0.45722	06. La ciudad
18	pierna	0.45333	01.Partes del cuerpo
19	falda	0.44572	02.La ropa
20	habitación	0.44431	.Partes de la casa

**Tabla 4.** Listado de los 20 vocablos más disponibles

Come se muestra, entre los veinte vocablos más disponibles, notamos el predominio del centro de interés “Partes del cuerpo” (01), que agrupa a seis unidades léxicas. También ponemos de relieve los cuatro vocablos del campo léxico “La ropa” (02) y los tres que forman parte del centro “Partes de la casa” (03). Además, es interesante señalar que los primeros lugares del listado están ocupados por dos vocablos

que designan “Animales” (07), donde “perro” domina, seguido por “gato”. En el tercer lugar se halla la presencia del único vocablo representativo del campo léxico “Profesiones y oficios” (08) que supera de 0.25 puntos la misma unidad perteneciente al centro de interés “La escuela”, y que habría debido aparecer en el listado al décimo puesto. Sin embargo, hemos decidido prescindir de ella debido a nuestro propósito de enumerar los 20 vocablos diferentes y más disponibles sobre el total de la muestra. Asimismo, a pesar de la gran cantidad de vocablos que los informantes han activado en relación con el centro “Comidas y bebidas” (04), destaca un solo vocablo de este campo léxico, es decir “agua”, que obtiene el cuarto puesto en el listado. Otras consideraciones nos llevan a señalar que ningún centro ha quedado fuera de este específico cotejo, aunque algunos no aparezcan muy representados en el listado.

#### **4.2 Resultados de la encuesta según las variables analizadas**

Tras haber presentado los resultados generales de nuestra investigación, en línea con nuestros objetivos de estudio, seguimos con la exposición de los datos resultantes en relación con las variables analizadas, tanto con respecto al número de palabras, como a la cantidad total de vocablos.

A fin de realizar una comparación de resultados según el método más fidedigno, que no se alejase de una realidad donde la diferencia numérica domina las distintas variables tomadas en cuenta, hemos seguido por la mayoría el ejemplo de Samper Hernández (2002), quien explica que atendiendo tanto a los promedios generales de palabras y vocablos, como a los particulares de cada centro de interés según la incidencia de los factores extralingüísticos considerados, se pueden obtener mejores cotejos en estudios donde la diferencia numérica afecta a las distintas variables y permite un mejor análisis de las semejanzas y contrastes en cada campo tomado en consideración.

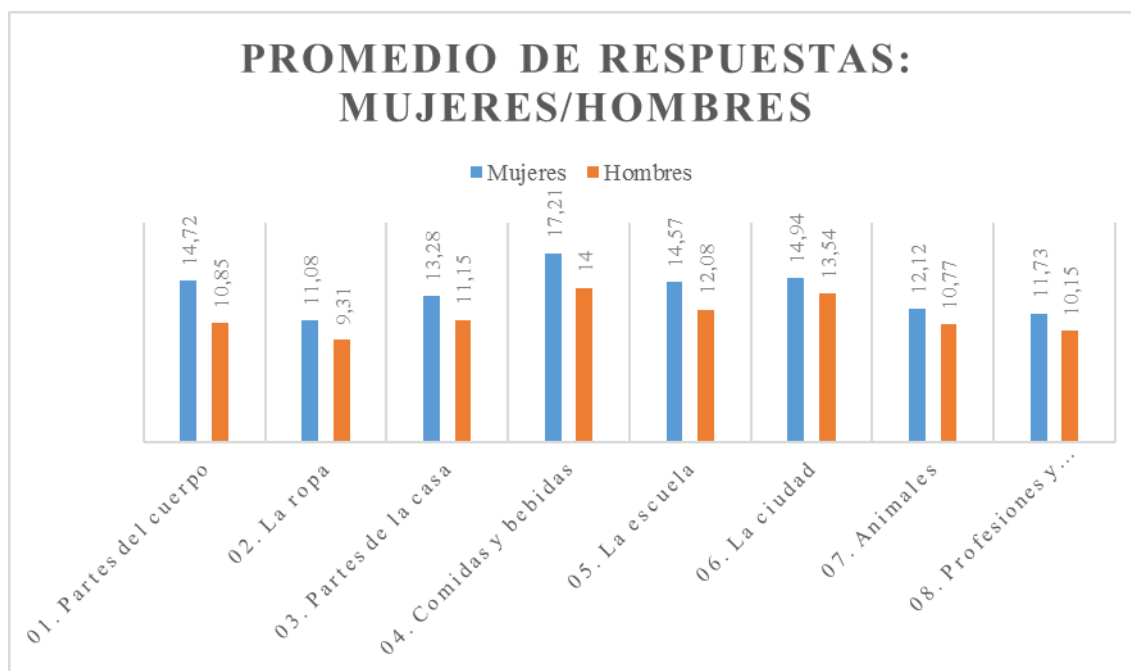
##### **4.2.1 Promedio de palabras según las variables analizadas**

Por lo que se refiere al factor sexo destaca la superioridad femenina en relación a las palabras aportadas por las mujeres frente a los hombres. La cantidad total de palabras respondidas por las mujeres es de 10.855, mientras que los hombres actualizan solo 1.194 respuestas.

A continuación mostramos los diferentes promedios de palabras por informante obtenidos en cada centro de interés.

Centro de interés	Rango	Promedio mujeres	Rango	Promedio hombres
01.Partes del cuerpo	3	14.72	5	10.85
02.La ropa	8	11.08	8	9.31
03.Partes de la casa	5	13.28	4	11.15
04.Comidas y bebidas	1	17.21	1	14.00
05.La escuela	4	14.57	3	12.08
06.La ciudad	2	14.94	2	13.54
07.Animales	6	12.12	6	10.77
08.Profesiones y oficios	7	11.73	7	10.15

**Tabla 5.** Promedio de palabras y rango según la variable ‘sexo’



**Figura 11.** Promedio de respuestas de las mujeres frente a las de los hombres

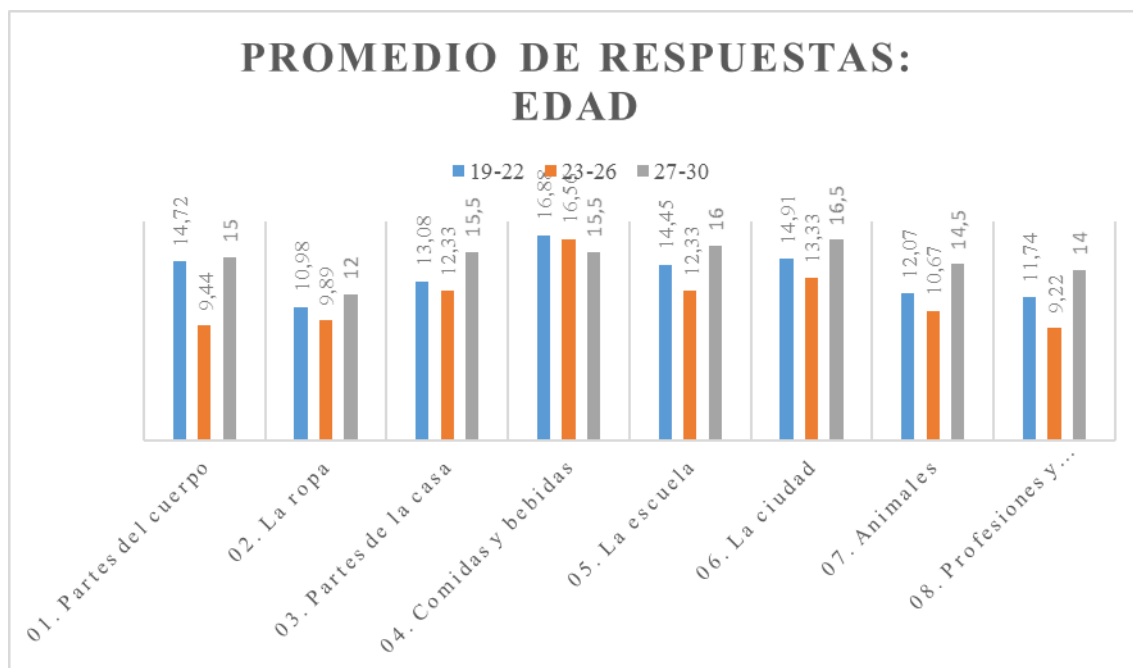
Como se observa, la ventaja de las mujeres es bastante alta en la mayoría de los centros de interés. Así, en los centros “Partes de la casa” (03) y “La escuela” (05) las mujeres superan de más de dos puntos los hombres, mientras que los centros “Partes del cuerpo” (01) y “Comidas y bebidas” (04) presentan una diferencia de promedio de más de tres puntos de distancia, en especial modo el centro “Partes del cuerpo” (01) que se aleja de 3.87 puntos del promedio de palabras de los hombres. Esto se refleja también en la diferencia entre los respectivos rangos, donde el centro en cuestión se encuentra al

tercer lugar en las mujeres y en quinta posición en los hombres. Sin embargo, el campo léxico donde señalamos la menor diferencia de puntos de promedio es el denominado “Animales” (07) donde la diferencia es de 1.35 puntos.

Otro factor considerado para este análisis de promedios de palabras por alumnos es la edad de los informantes. Podemos notar que a diferencia de lo que ocurre en algunas investigaciones donde, como cita Paredes García (2015:23), “a medida que aumenta la edad disminuye la producción léxica (López Rivero 2008, Gallego Gallego 2014)” o donde sucede la situación contraria, en nuestro caso los resultados no coinciden con ninguna consideración. De hecho, se ha verificado una tendencia poco clara (ver figura 12), en la cual se manifiesta para la mayoría de los centros de interés una curva que se da en ascenso desde la situación inicial hasta los alumnos de edad mayor, pero que sufre una caída en el intervalo intermedio (23-26). Por este motivo, nos resulta difícil encontrar similitudes con las tendencias descritas en relación a este factor.

Centro de interés	Rango	19-22	Rango	23-26	Rango	27-30
01.Partes del cuerpo	3	14.72	6	9.44	4	15.00
02.La ropa	8	10.98	5	9.89	7	12.00
03.Partes de la casa	5	13.08	3	12.33	3	15.50
04.Comidas y bebidas	1	16.88	1	16.56	3	15.50
05.La escuela	4	14.45	3	12.33	2	16.00
06.La ciudad	2	14.91	2	13.33	1	16.50
07.Animales	6	12.07	4	10.67	5	14.50
08.Profesiones y oficios	7	11.74	7	9.22	6	14.00

**Tabla 6.** Promedio de palabras y rango según la variable ‘edad’

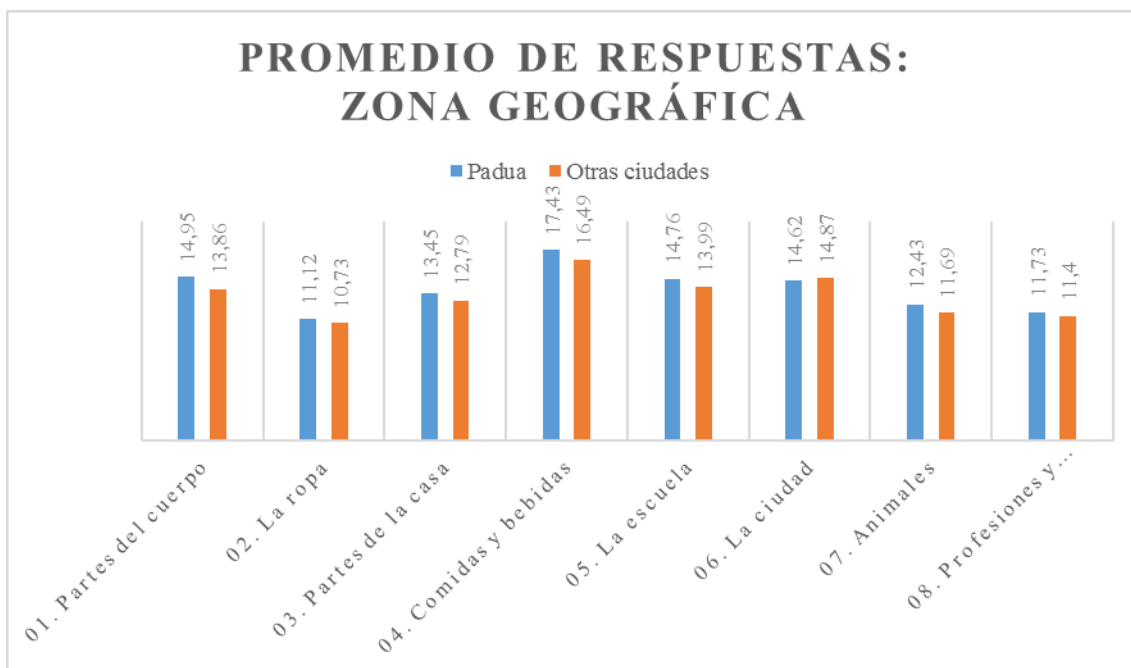


**Figura 12.** Promedio de respuestas según la edad de los informantes

En cuanto a la variable ‘zona geográfica’ de los informantes, indicamos en la tabla 6 y en el gráfico 13 las aportaciones de los datos obtenidos. Señalamos que debido a la dificultad de cotejar este factor con los estudios hispánicos, que proporcionan una distinción entre los encuestados de la zona urbana y rural, relevamos en general una proximidad de resultados en los promedios de los componentes de la muestra según sus diferentes zonas de procedencia. De hecho, la diferencia máxima de margen de promedio es de 1.09 puntos en el primer centro de interés “Partes del cuerpo” (01) que se evidencia también por ser el único campo junto con “La ciudad” (06) en no igualar el rango del otro grupo. Todo esto nos hace pensar en una posible y escasa capacidad discriminadora de esta variable.

Centro de interés	Rango	Padua	Rango	Otras ciudades
01.Partes del cuerpo	2	14.95	4	13.86
02.La ropa	8	11.12	8	10.73
03.Partes de la casa	5	13.45	5	12.79
04.Comidas y bebidas	1	17.43	1	16.49
05.La escuela	3	14.76	3	13.99
06.La ciudad	4	14.62	2	14.87
07.Animales	6	12.43	6	11.69
08.Profesiones y oficios	7	11.79	7	11.40

**Tabla 6.** Promedio de palabras y rango según la variable ‘zona geográfica’

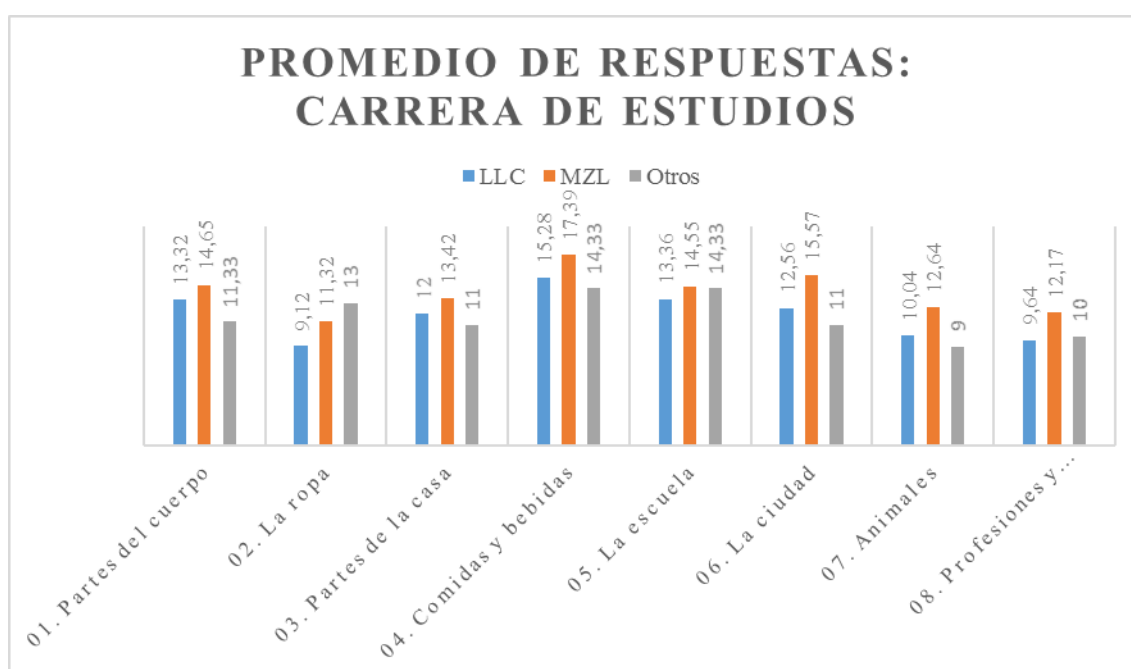


**Figura 12.** Promedio de respuestas de los informantes de la zona de Padua frente a otras ciudades

Continuamos el análisis de los promedios de palabras por informantes deteniéndonos en los resultados aportados según la variable ‘carrera de estudio’ de los alumnos. Principalmente destaca el predominio de los encuestados que estudian “Mediación Lingüística y Cultural” en todos los centros de interés con la excepción de uno, es decir el campo léxico “La ropa” (02) donde resalta el promedio de los estudiantes de otras carreras. A distancia de pocos puntos siguen los promedios de los alumnos del grado en “Lenguas, Literaturas y Culturas Modernas” que dejan el dominio en los centros “La escuela” (05) y “Profesiones y oficios” (08) a los estudiantes de otras licenciaturas. En suma, se evidencia una mayor competencia léxica en aquellos estudiantes cuyas asignaturas primarias son los idiomas extranjeros con enfoque a la mediación, lo cual podría hacernos suponer una mayor proximidad de sus estudios en relación con el desarrollo de aquellas competencias lingüísticas que sirven para manejar la lengua en las situaciones comunicativas que hemos seleccionado para nuestra investigación.

Centro de interés	Rango	LLC	Rango	MZL	Rango	Otros
01.Partes del cuerpo	3	13.32	3	14.65	3	11.33
02.La ropa	8	9.12	8	11.32	2	13.00
03.Partes de la casa	5	12.00	5	13.42	4	11.00
04.Comidas y bebidas	1	15.28	1	17.39	1	14.33
05.La escuela	2	13.36	4	14.55	1	14.33
06.La ciudad	4	12.56	2	15.57	4	11.00
07.Animales	6	10.04	6	12.64	6	9.00
08.Profesiones y oficios	7	9.64	7	12.17	5	10.00

**Tabla 7.** Promedio de palabras y rango según la variable ‘carrera de estudio’



**Figura 13.** Promedio de respuestas según la carrera de estudio de los informantes

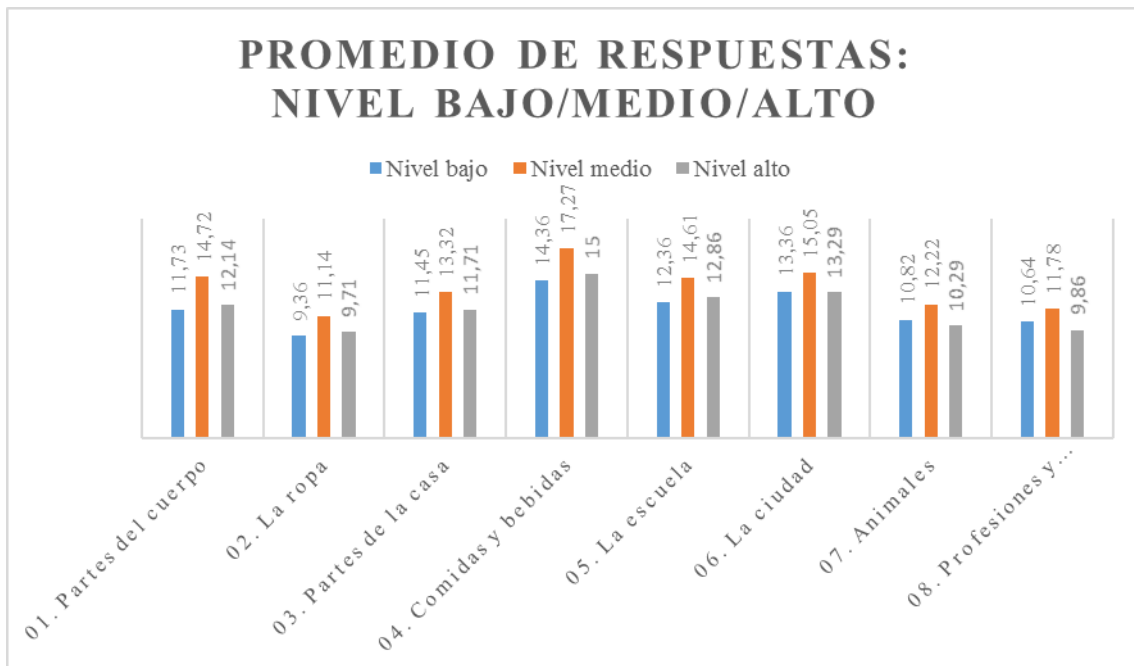
En lo que respecta al factor ‘nivel sociocultural’ (ver tabla 9 y figura 15), señalamos la casi total igualdad de resultados en relación con el rango en todas las variantes consideradas, con la excepción del tercer y cuarto puesto que ocupan respectivamente los centros “Partes del cuerpo” (01) y “La escuela” (05) en el nivel sociocultural medio en oposición a los restantes niveles donde los mismos centros intercambian su posición. Además, como se podía esperar, destacamos en todos los campos léxicos una inferioridad numérica de palabras respondidas por los informantes de nivel sociocultural bajo, que casi siempre resultan de un punto y medio más bajos que los promedios de sus compañeros de nivel medio. Sin, embargo, sorprendentemente



no se realiza una ascensión a medida que subimos al nivel sociocultural alto. El promedio de palabras actualizadas por los alumnos de este nivel, de hecho, no se caracteriza por dar lugar a puntuaciones mayores que los niveles menores, sino por tener muchas similitudes con los promedios de numerosos centros de interés del nivel bajo. La diferencia que ha resultado podría derivar de los pocos informantes que se agrupan al nivel alto, o para quedar en línea con la mayoría de las investigaciones sobre disponibilidad léxica, de la escasa capacidad discriminadora de esta variable.

Centro de interés	Rango	Nivel bajo	Rango	Nivel medio	Rango	Nivel alto
01.Partes del cuerpo	4	11.73	3	14.72	4	12.14
02.La ropa	8	9.36	8	11.14	8	9.71
03.Partes de la casa	5	11.45	5	13.32	5	11.71
04.Comidas y bebidas	1	14.36	1	17.27	1	15.00
05.La escuela	3	12.36	4	14.61	3	12.86
06.La ciudad	2	13.36	2	15.05	2	13.29
07.Animales	6	10.82	6	12.22	6	10.29
08.Profesiones y oficios	7	10.64	7	11.78	7	9.86

**Tabla 9.** Promedio de palabras y rango según la variable ‘nivel sociocultural’

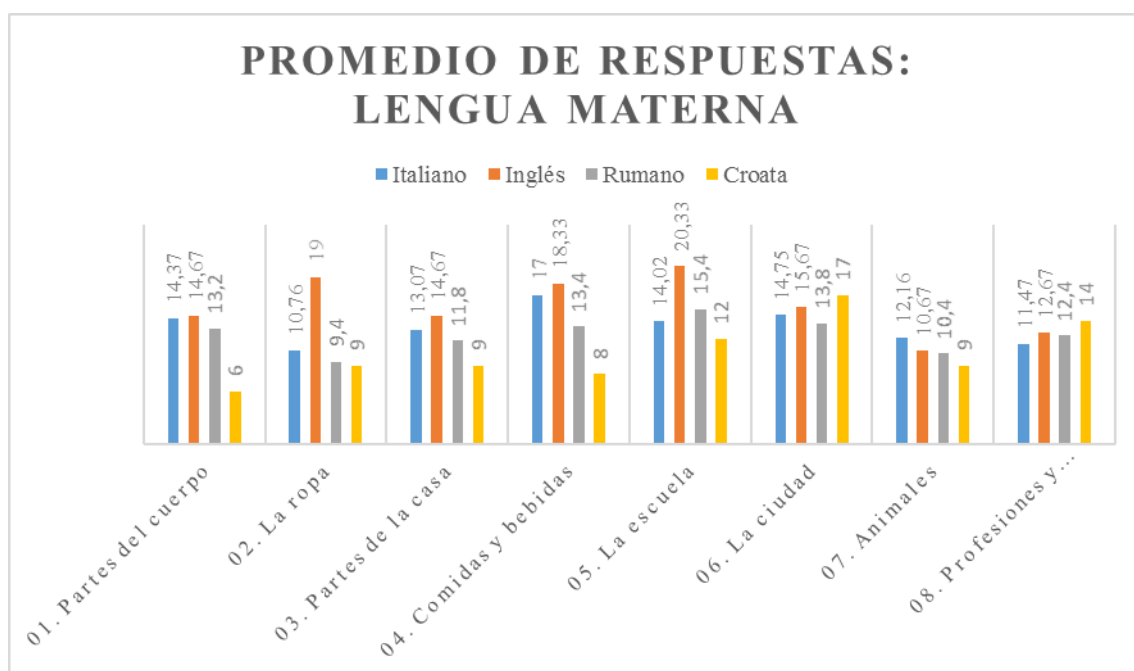


**Figura 15.** Promedio de respuestas en los niveles bajo, medio y alto

En la tabla 10 y en la sucesiva figura 16 aportamos los datos sobre la distribución de los promedios de palabras por informantes según la variable ‘lengua materna’.

Centro de interés	Rango	ITA	Rango	ING	Rango	RUM	Rango	CRO
01.Partes del cuerpo	3	14.37	5	14.67	4	13.20	6	6.00
02.La ropa	8	10.76	2	19.00	8	9.40	4	9.00
03.Partes de la casa	5	13.07	5	14.67	6	11.80	4	9.00
04.Comidas y bebidas	1	17.00	3	18.33	3	13.40	5	8.00
05.La escuela	4	14.02	1	20.33	1	15.40	3	12.00
06.La ciudad	2	14.75	4	15.67	2	13.80	1	17.00
07.Animales	6	12.16	7	10.67	7	10.40	4	9.00
08.Profesiones y oficios	7	11.47	6	12.67	5	12.40	2	14.00

**Tabla 10.** Promedio de palabras y rango según la variable ‘lengua materna’



**Figura 16.** Promedio de respuestas según la lengua materna de los informantes

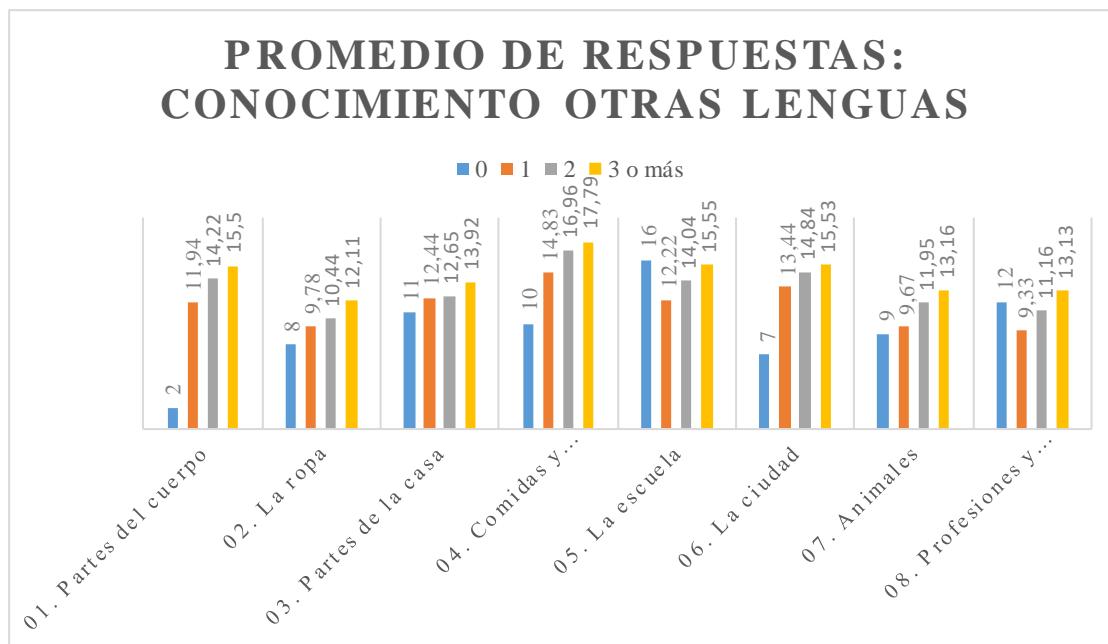
Como se observa claramente gracias al soporte gráfico de la figura 16, la superioridad casi general del grupo compuesto por hablantes nativos de lengua inglesa es evidente. Los hablantes nativos de italiano, por otro lado, demuestran una mayor competencia léxica en español que los informantes de las restantes lenguas maternas consideradas en nuestra investigación, con la excepción de los centros de interés “La escuela” (05) y “Profesiones y oficios” (08) donde se halla una diferencia de más de un

punto en favor de los participantes rumanos. Por fin, por lo que se refiere a la lengua croata en comparación con los promedios de las otras lenguas, destaca su dominio en los centros de interés “La ciudad” (06) y “Profesiones y oficios”(08).

En la siguiente tabla, se muestran los resultados en relación con el promedio de palabras por informantes según la variable “conocimiento de otras lenguas”. Como se evidencia de manera clara en la figura 17, cabe señalar la constante tendencia creciente en los datos que aportamos, donde destaca el predominio de los informantes que conocen tres o más lenguas aparte de la propia. Solo en el caso del centro de interés “La escuela” (05) el dominio pasa a quién no tiene conocimiento de otras lenguas, así como el centro “Profesiones y oficio” (08) en el mismo grupo supera de algunos puntos los promedios de los informantes que hablan 1 o 2 idiomas. Por lo que se refiere a nuestra investigación, entonces, las cifras indican que a medida que sube el número de lenguas conocidas por los estudiantes, aumenta su competencia léxica y su capacidad de retener una mayor cantidad de unidades léxicas en su memoria y de actualizarlas cuando se recibe un determinado estímulo.

<b>Centro de interés</b>	<b>Rango</b>	<b>0</b>	<b>Rango</b>	<b>1</b>	<b>Rango</b>	<b>2</b>	<b>Rango</b>	<b>3+</b>
01.Partes del cuerpo	8	2.00	5	11.94	3	14.22	4	15.50
02.La ropa	6	8.00	6	9.78	8	10.44	8	12.11
03.Partes de la casa	3	11.00	3	12.44	5	12.65	5	13.92
04.Comidas y bebidas	4	10.00	1	14.83	1	16.96	1	17.79
05.La escuela	1	16.00	4	12.22	4	14.04	2	15.55
06.La ciudad	7	7.00	2	13.44	2	14.84	3	15.53
07.Animales	5	9.00	7	9.67	6	11.95	6	13.16
08.Profesiones y oficios	2	12.00	8	9.33	7	11.16	7	13.13

**Tabla 11.** Promedio de palabras y rango según el factor ‘conocimiento de otras lenguas’



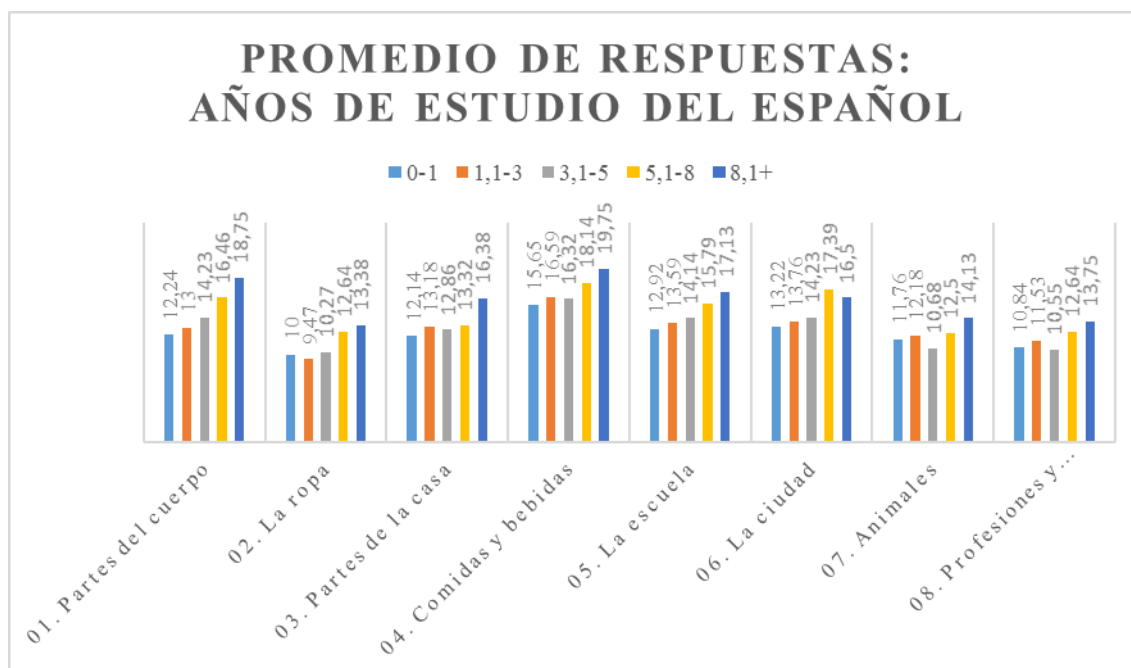
**Figura 17.** Promedio de respuestas según el conocimiento de otras lenguas

Seguimos nuestro análisis de los promedios de palabras por informante poniendo en evidencia los datos aportados por los encuestados según la variable ‘años de estudio del español’ que nos lleva a algunas interesantes consideraciones. Aquí también, como esperábamos, los resultados que mostramos ponen de manifiesto un general crecimiento de la cantidad de palabras activadas por los informantes a medida que aumenta su nivel de conocimiento de la lengua, que hemos medido mediante el número de años de estudio de la lengua española. En siete sobre ocho centros de interés, excepto en “La ciudad” (06), se encuentra un promedio mayor de los estudiantes de los niveles avanzados en comparación con los compañeros. Solo en el campo léxico recién citado los alumnos de nivel inmediatamente inferior –intervalo desde 5,1 hasta 8 años de estudio– dominan con casi un punto más que los compañeros. Por lo tanto, podemos concluir como Carcedo (1999a:87), citado por Samper Hernández (2002:58), que en nuestros datos se presenta “un desarrollo cuantitativo proporcionado en la incorporación de unidades léxicas a lo largo de este proceso de aprendizaje de la lengua española”. Es cierto que en algunos campos se hallan ligeras caídas de la curva (ver figura 15) por lo que se refiere a los datos de los centros de interés “Partes de la casa” (03), “Comidas y bebidas” (04), “Animales” (07), “Profesiones y oficios” (08) en los alumnos que se incluyen en el intervalo de estudio entre los 3,1 y 5 años. Sin embargo, se trata de pocos

casos sobre el total de la muestra, lo cual en general reconfirma la tendencia que hemos comentado.

Centro de interés	R	0-1	R	1,1-3	R	3,1-5	R	5,1-8	R	8,1+
01.Partes del cuerpo	4	12.24	5	13	2	14.23	3	16.46	2	18.75
02.La ropa	8	10	8	9.47	7	10.27	6	12.64	8	13.38
03.Partes de la casa	5	12.14	4	13.18	4	12.86	5	13.32	5	16.38
04.Comidas y bebidas	1	15.65	1	16.59	1	16.32	1	18.14	1	19.75
05.La escuela	3	12.92	3	13.59	3	14.14	4	15.79	3	17.13
06.La ciudad	2	13.22	2	13.76	2	14.23	2	17.39	4	16.50
07.Animales	6	11.76	6	12.18	5	10.68	7	12.50	6	14.13
08.Profesiones y oficios	7	10.84	7	11.53	6	10.55	6	12.64	7	13.75

**Tabla 12.** Promedio de palabras y rango según el factor ‘años de estudio del español’



**Figura 18.** Promedio de respuestas según los ‘años de estudio de español’

Tras haber realizado el análisis sobre el promedio de palabras según las variables tomadas en consideración en las páginas precedentes, pasamos ahora a la presentación

de los datos relativos a los promedios de vocablos por informante en los mismos factores.

#### **4.2.2 Producción de vocablos según las variables analizadas**

Como ya hemos puesto de manifiesto en la tabla 2, que mostraba la comparación del total de palabras con el total de vocablos y sus relativos rangos, no existe una correlación constante entre una alta cantidad de palabras aportadas en un centro de interés y el número de palabras diferentes o vocablos en el mismo centro. Por lo tanto, como cita Samper Hernández (2002:60) “no significa que el número de términos diferentes vaya a ser mayor ni viceversa”.

Nos habría interesado entonces, después haber expuesto los datos de los promedios de palabras por cada informante según las variables consideradas, llevar a cabo también un análisis sobre los promedios de vocablos según estos factores. Sin embargo, hemos encontrado algunas dificultades en realizar los cotejos de los promedios de manera fidedigna debido a la gran heterogeneidad numérica de los varios grupos en relación con los diferentes factores extralingüísticos, hecho que deriva del carácter no preestratificado de la muestra. Siguiendo el ejemplo de Magnúsdóttir (2012) hemos decidido presentar en las tablas sucesivas solo las cifras totales del número de vocablos por cada centro de interés según las distintas variables junto con la suma de todos los vocablos por cada variante (tabla 13). Además, aportamos la posición relativa de cada campo léxico en relación con los parámetros analizados a fin de comprobar cómo estos últimos afectan a la producción de vocablos (tabla 14).

Como hemos realizado en el apartado anterior, las variables que hemos tenido en cuenta son las siguientes: ‘sexo’ (hombres/mujeres), ‘edad’ (19-22/23-26/27-30), ‘zona geográfica’ (Padua/otras ciudades), ‘carrera de estudios’ (MLC/LLC/otros estudios), ‘nivel sociocultural’ (bajo/medio/alto), ‘lengua materna’ (italiano/inglés/rumano/croata), ‘conocimiento de otras lenguas’ (0/1/2/más de 3) y ‘años de estudio del español’ (0-1/1,1-3/3,1-5/5,1-8/8 o más).

Además, hemos utilizados algunas abreviaciones para designar los diferentes centros de interés tal como se muestra a continuación: 01.CUE (‘Partes del cuerpo’), 02. ROP (‘La ropa’), 03. CAS (‘Partes de la casa’), 04. COM (‘Comidas y bebidas’), 05.

ESC ('La escuela'), 06. CIU ('La ciudad'), 07. ANI ('Animales'), 08. PRO ('Profesiones y oficios').

En la tabla 13 se evidencia en general una correspondencia entre el número de personas que componen cada grupo de las variables y los vocablos que producen. Por lo tanto, no notamos que se supere el número total de vocablos aportados por aquellos grupos que cuentan menos informantes en comparación con los que presentan mayores componentes en ninguna variable.

Dicho esto, nos podrá ayudar a ahondar el tema el cotejo de los rangos de cada centro de interés según las variables consideradas. Como se observa en la tabla 14, el centro que ha resultado más productivo en general es "Comidas y bebidas" (04) que coincide once veces con el primer lugar de los rangos según las distintas variables, seguido por "La ciudad" (06) y "Profesiones y oficios" (08) que se clasifican iguales en cuanto al número de veces que encabezan las diferentes ordenaciones según sus posiciones relativas. Al lado opuesto, la incidencia general de las variables posiciona en el último puesto el centro de interés "Partes del cuerpo" (01) y en el penúltimo el campo léxico "La ropa" (02).

Hemos sido capaces de formular una clasificación del primer y último lugar en relación con el rango ya que en general no destacan grandes diferencias en la producción de vocablos según las variables consideradas. A pesar de que no siempre los resultados coincidan, las diferencias entre los varios rangos no superan muchas veces el margen de dos puestos salvo algunas excepciones.

Por ejemplo, en la variable 'sexo' el centro "Profesiones y oficios" (08) ocupa el primer puesto mientras que con las respuestas de los hombres se posiciona en el cuarto puesto.

Asimismo, en la variable 'edad' las mayores diferencias las producen los encuestados del intervalo de edad entre los 19 y 21 años en los centros "Partes de la casa" (03), "Comidas y bebidas" (04) y "Profesiones y oficios" (08).

Además, al estudiar el factor 'carrera de estudio' notamos que difieren de más de dos puestos los campos "La ropa" (02), "La escuela" (05) y "Profesiones y oficios" (08), debido principalmente a los datos de los informantes que se agrupan bajo la denominación 'otros estudios'.

En la variable ‘nivel sociocultural’ destaca el centro “Profesiones y oficios” (08) que obtiene tres posiciones relativas distantes en relación con las variantes tenidas en cuenta. En el nivel medio este campo léxico ocupa el primer puesto a diferencia del nivel bajo donde se encuentra en el sexto y del nivel alto donde llega a la cuarta posición.

Las menores coincidencias las encontramos al observar la variable ‘lengua materna’. En específico, en los centros “La ropa” (02), “Comidas y bebidas” (04) y “Profesiones y oficios” (08) se hallan considerables diferencias de rangos en todos los grupos considerados.

Sigue la variable ‘conocimiento de otras lenguas’ donde destaca la completa distancia de los rangos de los datos relacionados con la persona que no conoce otros idiomas aparte de la lengua materna y del español en comparación con los otros conjuntos considerados.

Finalmente, al estudiar la variable ‘años de estudio del español’ notamos una general coherencia de resultados con la ligera excepción que se da en los campos léxicos “Comidas y bebidas” (04) donde la producción de vocablos de los informantes que llevan estudiando español desde más de 5,1 años se posiciona en el cuarto puesto de rango a diferencia de los otros grupos que alcanzan los primeros lugares.



**Tabla 13.** Producción de vocablos según las variables extralingüísticas consideradas

C.I.	‘Sexo’		‘Edad’			‘Zona geográfica’		‘Carrera de estudios’		
	M (99)	V (13)	19-21 (101)	23-26 (9)	27-30 (2)	Padua (43)	Otras ciudades (69)	MLC (85)	LLC (23)	Otros (4)
01.CUE	122	48	125	39	26	95	100	123	60	22
02.ROP	148	47	144	40	22	108	125	132	73	30
03.CAS	160	66	157	55	29	108	146	149	78	23
04.COM	279	89	248	86	30	183	226	257	122	35
05.ESC	270	87	266	66	29	181	215	250	119	39
06.CIU	274	91	278	72	31	190	229	261	118	29
07.ANI	229	66	224	57	27	143	191	225	82	15
08.PRO	285	77	279	50	28	191	216	268	109	24
<b>Total</b>	<b>1.767</b>	<b>571</b>	<b>1721</b>	<b>465</b>	<b>222</b>	<b>1199</b>	<b>1448</b>	<b>1665</b>	<b>761</b>	<b>217</b>

C.I.	‘Nivel sociocultural’			‘Lengua materna’				‘Conocimiento de otras lenguas’				‘Años de estudio del español’				
	Bajo (11)	Medio (94)	Alto (7)	ITA (103)	ING (3)	RUM (5)	CRO (1)	0 (1)	1 (18)	2 (56)	3+ (37)	0-1 (34)	1,1-3 (19)	3,1-5 (22)	5,1-8 (29)	8+ (8)
01.CUE	46	126	39	128	26	32	6	12	60	97	85	71	57	58	78	45
02.ROP	103	143	35	1098	38	34	9	8	61	109	104	85	70	58	94	48
03.CAS	126	158	44	1333	30	37	9	11	72	119	109	100	76	73	92	53
04.COM	158	267	69	1734	38	49	8	10	113	207	179	189	117	120	137	81
05.ESC	136	266	58	1430	44	50	12	16	100	185	168	161	93	99	144	70
06.CIU	147	275	54	1504	37	49	17	7	110	206	173	154	100	125	165	76
07.ANI	119	230	46	1240	16	37	9	9	83	161	145	149	81	85	120	59
08.PRO	117	282	48	1170	26	51	14	12	85	198	172	149	115	93	145	70
<b>Total</b>	<b>952</b>	<b>1747</b>	<b>393</b>	<b>9637</b>	<b>255</b>	<b>339</b>	<b>84</b>	<b>85</b>	<b>684</b>	<b>1282</b>	<b>1135</b>	<b>1058</b>	<b>709</b>	<b>711</b>	<b>975</b>	<b>502</b>

**Tabla 14.** Rango de vocablos según las variables extralingüísticas consideradas

C.I.	‘Sexo’		‘Edad’			‘Zona geográfica’		‘Carrera de estudios’		
	M (99)	V (13)	19-21 (101)	23-26 (9)	27-30 (2)	Padua (43)	Otras ciudades (69)	MLC (85)	LLC (23)	Otros (4)
01.CUE	8	6	8	8	6	7	8	8	8	7
02.ROP	7	7	7	7	7	6	7	7	7	3
03.CAS	6	5	6	5	3	6	6	6	6	6
04.COM	2	2	4	1	2	3	2	3	1	2
05.ESC	4	3	3	3	3	4	4	4	2	1
06.CIU	3	1	2	2	1	2	1	2	3	4
07.ANI	5	5	5	4	5	5	5	5	5	8
08.PRO	1	4	1	6	4	1	3	1	4	5

C.I.	‘Nivel sociocultural’			‘Lengua materna’				‘Conocimiento de otras lenguas’				‘Años de estudio del español’				
	Bajo (11)	Medio (94)	Alto (7)	ITA (103)	ING (3)	RUM (5)	CRO (1)	0 (1)	1 (18)	2 (56)	3+ (37)	0-1 (34)	1,1-3 (19)	3,1-5 (22)	5,1-8 (29)	8+ (8)
01.CUE	8	8	7	8	5	6	6	2	8	8	8	7	8	7	8	7
02.ROP	7	7	8	7	2	5	4	6	7	7	7	6	7	7	6	6
03.CAS	4	6	6	4	4	4	4	3	6	6	6	5	6	6	7	5
04.COM	1	3	1	1	2	3	5	4	1	1	1	1	1	2	4	1
05.ESC	3	4	2	3	1	2	3	1	3	4	4	2	4	3	3	3
06.CIU	2	2	3	2	3	3	1	7	2	2	2	3	3	1	1	2
07.ANI	5	5	5	5	6	4	4	5	5	5	5	4	5	5	5	4
08.PRO	6	1	4	6	5	1	2	2	4	3	3	4	2	4	2	3



# Capítulo V

## **Análisis y comentario de resultados cualitativos**

Uno de los objetivos de nuestra investigación, además de aportar resultados cuantitativos en relación con los datos que hemos obtenido, es el de acercarnos también a un análisis complementario, de carácter cualitativo, de las respuestas aportadas por nuestros informantes en cada centro de interés, ya que el lenguaje no es solo número y estadística.

Los datos a nuestra disposición nos permitirían abarcar varios aspectos del comportamiento léxico de nuestros encuestados. Sin embargo, debido al tipo de investigación que estamos realizando, en los apartados sucesivos nos centraremos solo en una mínima parte de los análisis que es posible realizar. Por lo tanto, como conclusión de nuestro trabajo, presentaremos las primeras diez palabras más disponibles por cada centro de interés junto con la descripción general de las unidades léxicas recopiladas en cada campo léxico, que se añaden en forma de listados en el anexo II de este trabajo.

### **5.1 Palabras más disponibles por cada centro de interés**

#### **5.1.1 ‘Partes del cuerpo’**

En la base de lo que hemos descrito en el capítulo anterior sobre los datos generales referidas al índice de cohesión de los centros de interés, cabe recordar ahora el resultado obtenido por el primer centro “Partes del cuerpo”. Los datos que hemos expuesto han destacado la ventaja de este centro sobre los otros, un aspecto que se pone de manifiesto a la hora de presentar el conjunto de palabras que los informantes han actualizado para responder al estímulo proporcionado por este tema ya que las asociaciones realizadas están muy vinculadas con el campo léxico en cuestión.

Las primeras diez palabras listadas según el índice de disponibilidad (ver tabla 15) registran referencias muy estrechas con las partes del cuerpo en relación con la cabeza:

*ojo, boca, pelo, nariz, cabeza, cara*; las unidades que quedan, por el contrario, designan las extremidades del cuerpo con sus partes: *mano, brazo, pierna, pie*.

Nº	Palabra	Disponibilidad
1	ojo	0.73517
2	mano	0.60497
3	boca	0.49513
4	pelo	0.48834
5	brazo	0.46566
6	pierna	0.45333
7	nariz	0.44400
8	pie	0.43265
9	cabeza	0.41682
10	cara	0.32904

**Tabla. 15** Primeras diez palabras del centro de interés “Partes del cuerpo” (01) según el índice de disponibilidad

Si a partir de los primeros puestos de este listado queda patente que el conocimiento de este campo léxico ha generado relaciones semánticas muy cerradas y compartidas por la mayoría de los informantes hasta obtener la predominancia numérica incluso entre las veinte palabras más disponibles del cotejo global<sup>10</sup>; con todo, en el análisis del centro encontramos vocablos que hacen referencia a otras partes del cuerpo, así, como se podrá ver en el anexo II, se continua con la enumeración de unidades referidas a la cabeza y a las extremidades del cuerpo, y además, se añaden otras que designan órganos o partes internas como: *corazón, estómago, vientre, barriga, tripa<sup>11</sup>, músculo, pulmón, hueso, órgano*. Sin embargo, a partir de la posición 32 sobre las 133 palabras disponibles del centro en cuestión, encontramos el primer adjetivo del listado, es decir “*largo*”, el cual mantiene siempre una relación con el campo léxico, pero referida a la descripción del aspecto físico como las unidades que siguen: *largo, alto, bajo, gordo, delgado, redondo, corto, pequeño, grande, claro, ancho*. Asimismo, registramos algunas unidades usadas para designar colores: *azul, castaño, negro, rubio, rojo, color, verde*. Aparecen también sustantivos relacionados con la salud y la

<sup>10</sup> Ver capítulo 4, tabla 4.

<sup>11</sup> En el DRAE, con el término viene se designa tanto la “cavidad del cuerpo en la que se contienen los órganos del aparato digestivo y del genitourinario”, como “el conjunto de vísceras contenidas en el vientre” o “la región exterior del cuerpo, correspondiente al abdomen”. Las mismas definiciones aparecen al buscar la unidad “barriga”. Por lo tanto, se pueden referir tanto a un órgano, como a la parte exterior del cuerpo, pero debido al carácter de la prueba que lleva a la aportación de palabras aisladas de su contexto específico, se pueden crear ambigüedades.

medicina: *ortopedia, oculista*, y otros elementos que se alejan un poco más del tema a partir de asociaciones con algunas unidades que han generado sustantivos como *peluquería y peluquero*, elementos que se refieren a nombres de joyas como *brazalete, anillo y pendientes*, accesorios como *gafas y perfume*. También hay un gran número de verbos que se utilizan para denotar acciones realizadas por el ser humano como *oír, caminar, comer, tocar, pintar, escuchar, pensar, arrodillarse, dibujar, sentarse, echar un vistazo, andar, oler, peinar, ver, hablar, bailar, romper, lavarse, moverse, observar*.

Dentro del estudio de la variación léxica, destacan los disfemismos “*culo*” y su correspondiente eufemístico “*trasero*”.

En cuanto a los estudios de formación léxica, además, aparecen compuestos sintagmáticos con construcción sustantivo + preposición + sustantivo como *dedo de pie* o sintagmas verbales con estructura verbo transitivo + sintagma nominal como *echar un vistazo*.

Es interesante, por último, señalar que el centro se compone fundamentalmente de sustantivos, seguidos por un 18% de adjetivos y un 17% de verbos.

### 5.1.2 ‘La ropa’

El centro de interés dedicado a la ropa, al igual que el anterior, presenta una gran cohesión semántica en relación con el tema de las prendas de vestir. Las diez palabras más disponibles que encabezan el correspondiente listado (ver tabla 16) son: *camiseta, zapato, pantalón, falda, camisa, vaquero, jersey, bufanda, zapatilla y calcetín*. Como se puede deducir se refieren a ropa que de manera general puede ayudar a realizar una descripción de la indumentaria de una persona, sea hombre o mujer. Sin embargo, a medida que nos adentramos en el análisis de las unidades del listado, encontramos vocablos que hacen referencia a subconjuntos que podemos distinguir entre el de la ropa femenina (*vestido, minifalda, blusa, media, vestido de novia, camisión, bikini, mantilla*), ropa interior (*calzoncillo, sujetador, braga, lencería, camiseta interior*), elementos de vestuario masculino (*traje, corbata*), calzado (*bota, zapato de tacón, chancla, zapatilla de deporte, sandalia, espadrilla, playera, pantufla*) y prendas exteriores (*chaqueta, abrigo, chal, capa*). Figuran también elementos que no coinciden directamente con el estímulo principal como accesorios (*gorro, gorra, sombrero, pañuelo, gafas, gafas de sol, prismáticos, cinturón, reloj, tirantes*), joyas (*anillo, brazalete, collar, pulsera,*

pendiente) y otros complementos (*mochila, bolso, bolsa, maleta, cartera, saco*). Registramos también la presencia de materiales y formas relacionadas con las prendas de vestir (*algodón, pana, seda, de rayas, de flores, estampado*), unidades que designan la dimensión de la ropa (*largo, corto, estrecho, ancho*), colores (*verde, amarillo, rojo, negro, blanco, azul, gris*), así como otras características (*lindo, nuevo*). Asimismo, es interesante constatar la presencia de verbos vinculados con el tema central (*vestir, comprar, llevar, poner, ir de compras, planchar, abrigarse, probar, lavar, vender, estirar, quedar, pagar, sacarse, ir de escaparates, cambiar, quitarse, secar, gustarse, lucirse*).

Como vemos, aunque los verbos recién enumerados en apariencia son muchos, componen solo un 13% del total de vocablos producidos en relación con el centro de interés, seguidos por un 10% de adjetivos. Por lo tanto, se deduce que la mayoría de respuestas incluye los sustantivos, que constituyen un 77% del universo muestral.

A continuación de nuestro análisis señalamos la presencia de algunas unidades que se sitúan dentro de la variación diatópica. Registramos, de hecho, vocablos que se suelen utilizar en el español de América y para cuya correspondencia semántica nos ayudamos con el *Diccionario de americanismos*. Destacan vocablos como “*remera*” (camiseta de manga corta), “*campera*” (chaqueta de uso informal o deportivo), “*pollera*” (falda o parte de vestido que cae desde la cintura) y “*bombacha*” (prenda interior femenina).

Además, como ya hemos anticipado en el capítulo III, encontramos varios casos de préstamos; algunos de origen francés y adaptados (*blusa, chándal, pantufla*), algunos de origen inglés que se han adaptado a la ortografía española (*pulóver, suéter*) y otros que han mantenido la original (*jeans, t-shirt, jersey, cárdigan*).

Por último, se hallan compuestos sintagmáticos nominales formados por sustantivo+preposición+sustantivo (*zapato de tacón, vestido de novia, camiseta de tirantes, gafas de sol*) y compuestos verbales de construcción idiomática (*ir de escaparates, ir de compras*).

Nº	Palabra	Disponibilidad
1	camiseta	0.63352
2	zapato	0.59360
3	pantalón	0.47923
4	falda	0.44572

5	camisa	0.37186
6	vaquero	0.35704
7	jersey	0.34893
8	bufanda	0.29223
9	zapatilla	0.18666
10	calcetín	0.14846

**Tabla. 16** Primeras diez palabras del centro de interés “La ropa” (02) según el índice de disponibilidad

### 5.1.3 ‘Partes de la casa’

Al estudiar el centro de interés “Partes de la casa” según los resultados del índice de disponibilidad, notamos primero que las diez palabras más disponibles (ver tabla 17) incluyen no solo las distintas áreas que constituyen una vivienda (*cocina, habitación, baño, salón, cuarto de baño, cuarto*), sino también elementos, como los muebles, que se pueden encontrar en las diferentes partes de la casa (*cama, mesa, silla, sofá*). Así, ya se hace patente un alejamiento del tema central que funciona como estímulo y que ha llevado, como se verá en la descripción a continuación y en el anexo II, a la creación de una muestra variada que, de todas maneras, consigue mantener una conexión semántica con el centro primario.

Nº	Palabra	Disponibilidad
1	cocina	0.67884
2	cama	0.50027
3	habitación	0.44431
4	baño	0.38849
5	mesa	0.32339
6	salón	0.25896
7	cuarto de baño	0.25579
8	silla	0.24296
9	cuarto	0.23249
10	sofá	0.22725

**Tabla. 17** Primeras diez palabras del centro de interés “Partes de la casa” (03) según el índice de disponibilidad

Después de los primeros diez lugares, registramos una variedad de respuestas que se refieren a partes interiores y exteriores de la casa como *jardín, garaje, comedor, entrada, sala, dormitorio, terraza, trastero, sala de estar, pasillo, estudio, ático, piscina, sala de baño, aseo, living, cuarto de estar, pieza, biblioteca, corredor, azotea,*



*oficina, sala de comer, ascensor, tejado, salida, gimnasio, aparcamiento, césped.* Además, otros vocablos aportados están vinculados con otros subconjuntos, como los muebles (*armario, sillón, mesilla, mueble, mesilla de noche, mesita, estantería, escritorio, mesa de noche, mesita de noche, armario ropero, cómoda, perchero, zapatero, guardarropa*), los electrodomésticos (*nevera, frigorífico, horno, lavavajillas, microondas, refrigerador lavaplatos, heladera, aspirador, secadora, teléfono*), los aparatos electrónicos (*lámpara, televisión, luz, ordenador*) y las instalaciones sanitarias (*ducha, lavabo, inodoro, bidé*).

Asimismo, se incluyen partes de la construcción de la casa como: *puerta, ventana, techo, rincón, escalera, pared, pavimento, suelo* e incluso tipos de vivienda: *piso, casa, bloque, palacio, caseta*. Tampoco faltan utensilios que se colocan en la cocina como: *vaso, cuchara, cucharilla, cucharón, cuchillo*, accesorios de varios tipos: *alfombra, cortina, sábana, almohada, espejo, vela, flexo, mantel, toalla* y finalmente sustantivos que designan elementos que se alejan un poco más del centro aun manteniendo una cercanía semántica: *cocinero, libro, compañero de piso, familia, llave, calor, paraguas, almuerzo, tienda, reunión de familia*.

Lo que más destaca es el dominio de los sustantivos sobre los verbos, que componen un 10% del total, y los adjetivos (solo un 4% del universo muestral). Este dato nos hace pensar que los informantes han intentado mantenerse en el tema lo más posible, aportando una gran cantidad de sustantivos que, aunque en su mayoría están muy vinculados con el campo léxico no han centrado exactamente el tema dado. Para investigaciones futuras convendrá tener en cuenta en fase de definición de los centros de interés una denominación más específica que tenga en consideración este aspecto. Sin embargo, los datos reflejan una alta capacidad de conexión semántica de los estudiantes y por eso no podemos considerar totalmente equivocada la aportación de las respuestas obtenidas.

#### **5.1.4 ‘Comidas y bebidas’**

El siguiente apartado analiza el centro de interés que hace referencia al tema de la alimentación. Como ya hemos observado anteriormente, se trata del campo con el mayor número de palabras totales (291), cuyas diez primeras más disponibles se muestran a continuación en la tabla 18.

Nº	Palabra	Disponibilidad
1	agua	0.70801
2	tomate	0.37043
3	paella	0.34369
4	comer	0.30036
5	leche	0.27700
6	cerveza	0.26847
7	naranja	0.24561
8	jamón	0.22322
9	manzana	0.21076
10	beber	0.20726

**Tabla. 18** Primeras diez palabras del centro de interés “Comidas y bebidas” (04) según el índice de disponibilidad

Como destaca la tabla, ya entre las primeras palabras más disponibles se presenta una situación que se refleja en el total de palabras que componen la muestra del centro en cuestión. Encontramos en los primeros lugares tres palabras que se refieren a las bebidas (*agua, leche, cerveza*) y cinco que se relacionan con el tema de la comida incluyendo frutas y verduras (*tomate, naranja, manzana*), carne (*jamón*) y un plato típico de la cocina española (*paella*). Los dos vocablos restantes (*comer, beber*) son verbos que los informantes han derivado de las palabras que proporcionan el estímulo del centro de interés. Hay 38 vocablos que se refieren a bebidas y 174 unidades relacionadas con la comida. Podemos afirmar, por lo tanto, que se trata de un campo léxico que ha permitido activar muchas asociaciones entre los estudiantes y que no se ha alejado mucho del tema central. De hecho, además de los sustantivos que designan comidas o bebidas, registramos también 49 nombres que mantienen una cercanía con el campo léxico mediante la creación de subconjuntos como los establecimientos donde se consumen los alimentos (*restaurante, comedor, cocina, mesa, bar, cafetería, comedor universitario, chiringuito*) o se compran (*supermercado, pastelería, panadería, carnicería*). Se hallan incluso los utensilios que se ponen en la mesa (*plato, cuchillo, cuchara, toalla, servilleta, tenedor, copa, botella, cubierto*) o los electrodomésticos (*nevera, frigorífico*). Tampoco faltan sustantivos vinculados con las distintas comidas que se toman a lo largo del día (*desayuno, cena, almuerzo, merienda, aperitivo*), las sensaciones que se perciben en relación con el tema (*hambre, sed, frío*), o los efectos de acciones (*trago, mezcla*). Siguen algunos elementos que no están relacionados directamente con el centro de interés pero que hemos mantenido puesto que de alguna

manera se podrían conectar con la ayuda de un contexto mayormente estructurado (*menú, director, curso, dinero, aula*).

Dicho esto, consideramos que un 90% de la muestra está representada por los sustantivos, menor incidencia tienen los verbos (un 6% sobre el total) y aun menos son los adjetivos (4%).

Entre los verbos encontramos los que designan acciones en estrecha relación con la cocina (*cocinar, preparar, comprar, tragar, freír, hacer la compra*), con las comidas del día (*tomar, desayunar, almorzar, merendar*), con las sensaciones (*gustar*) y los efectos del beber mucho (*emborracharse*). Finalmente encontramos verbos que describen acciones que se suelen hacer a la hora de comer (*ordenar, pedir, elegir*). Además, destaca un elemento que se suele utilizar en el español de América: “*morfar*” que significa ‘ingerir alimentos’, según especifica el *Diccionario de americanismos*.

Entre los adjetivos registramos los que sirven para referirse a la comida en sentido estrecho (*bueno, frito, caliente, salado, rico, típico*) y a las personas (*vegetariano, comilón, gordo, borracho*).

### 5.1.5 ‘La escuela’

En la tabla 19 enumeramos los diez vocablos más disponibles del centro de interés relacionado con el mismo contexto donde se ha realizado la prueba. Probablemente es también por eso por lo que se posiciona, como hemos descrito en el capítulo anterior, en el tercer puesto por número total de palabras respondidas. Entre las diez primeras encontramos las principales categorías en las cuales se divide el campo léxico: los sujetos del centro de interés, la docencia y el material escolar.

Nº	Palabra	Disponibilidad
1	profesor	0.55427
2	libro	0.51042
3	bolígrafo	0.38701
4	lápiz	0.37787
5	estudiar	0.30569
6	examen	0.28100
7	pizarra	0.27702
8	cuaderno	0.27032
9	clase	0.24886
10	estudiante	0.24654

**Tabla. 19** Primeras diez palabras del centro de interés “La escuela” (05) según el índice de disponibilidad

Los sustantivos que constituyen el subconjunto de las personas vinculadas con el centro son: *profesor, estudiante, alumno, compañero, profesorado, amigo, estudiantado, alumnado, director, maestro, jefe, padres, empleado, rector, secretario, chico, dirigente, niño*, mientras que los objetos que se incluyen entre los materiales escolares son: *libro, bolígrafo, lápiz, pizarra, cuaderno, papel, mochila, ordenador, borrador, sacapuntas, tijeras, goma de borrar, tiza, hoja, fotocopia, pluma, computador, papelería, agenda, rotulador, lápiz de color, calculadora, bolsa, estuche, diario, marcador, cinta, grabador, maleta, hoja de papel, pantalla, corrector, mapa, bolso, cera, corrector blanco, encerado, afilalápices, micrófono, pegamento*). Un gran número de respuestas confluyen en la categoría de la docencia (*examen, deberes, curso, nota, apuntes, lección, solapamiento, horario, ejercicio, prueba, enseñanza, regla, prueba oral, prueba escrita, hora, tarea, lectura, pausa, semestre, texto, dictado, test, aprobación, evaluación, suspensión, tema, boletín de notas, recreo, soporte electrónico*) y una menor cantidad se refiere a los muebles del aula (*silla, mesa, pupitre, escritorio*). Asimismo, se añaden vocablos que se refieren al mundo de la escuela en un sentido más amplio, que recoge las diferentes áreas físicas en las cuales se puede dividir un establecimiento de enseñanza (*clase, aula, gimnasio, edificio, baño, jardín, laboratorio, cafetería, librería, despacho, oficina de los estudiantes, escalera, sala, cantina, campo de fútbol, piscina*), las diferentes áreas de estudio o asignaturas (*lengua, español, matemáticas, historia, inglés, ciencias, idiomas, italiano, francés, geografía, literatura, portugués, física, biología, lenguas extranjeras, latín, historia de la lengua, traducción, música, dibujo, estudios religiosos, arte, informática*) y las diferentes etapas de la enseñanza (*universidad, colegio, enseñanza primaria, enseñanza media, instituto, guardería, liceo, bachillerato*). También, encontramos varias respuestas que se relacionan de manera estrecha con el mundo de la universidad, hecho que se puede justificar por ser el entorno primario de los participantes a la encuesta. Los sustantivos que se pueden asignar a esta categoría, además de los anteriormente descritos, son: *universidad, facultad, tutoría, carrera, licenciatura, graduación, matriculación, beca Erasmus, institución, oficina de los estudiantes, correo electrónico, departamento, intercambio cultural*. Para terminar la descripción de los sustantivos, señalamos algunos de los varios que se refirieren a asuntos más generales si bien siempre contextualizables dentro del campo léxico de referencia: *autobús, gafas, tren, café, atención, suerte,*

*infancia, profesión, puesto, retraso, ansiedad, reposo, futuro, miedo, oportunidad, caso, huelga, bocadillo, madrugada, mañana, coche, amistad, sueño, vacación, fin de semana, vacaciones, cultura, divertimento, internet, experiencia, cansancio, agua, tiempo libre.*

Una vez más, en este centro de interés se repite la tendencia a actualizar sustantivos más que verbos o adjetivos, los cuales representan respectivamente solo el 16 % y 3 % de la muestra total. Los verbos describen en su mayoría acciones relacionadas con el mundo de los estudiantes y de los enseñantes como *estudiar, escuchar, aprobar, aprender, subrayar, hablar, asistir, suspender, repasar, repetir, ir a clase, enseñar, licenciar, examinar, saber, matricularse, explicar, corregir*. Los adjetivos, por el contrario, se pueden asociar a diferentes sujetos (*ausente, oral, primero, escolar, temprano, práctico, importante, público, privado, escrito*).

#### **5.1.6 ‘La ciudad’**

El siguiente centro de interés destaca por haber obtenido el segundo puesto en relación con el número total de repuestas proporcionadas por los estudiantes, lo cual no sorprende ya que las asociaciones léxico-semánticas que se han establecido con el tema de la ciudad son muy amplias.

Aquí también la presencia de sustantivos es abundante (86% del universo muestral) en relación con los verbos, que ocupan solo el 11%, y de los adjetivos, que alcanzan apenas el 3% sobre el total.

A partir de las primeras diez palabras (ver tabla 20) registramos diferentes tipos de categorías que han confluído en las respuestas de los informantes.

<b>Nº</b>	<b>Palabra</b>	<b>Disponibilidad</b>
1	calle	0.52362
2	coche	0.45722
3	tienda	0.30967
4	casa	0.30532
5	plaza	0.30073
6	parque	0.29076
7	autobús	0.26120
8	restaurante	0.24682
9	iglesia	0.24292
10	edificio	0.19744

**Tabla. 20** Primeras diez palabras del centro de interés “La ciudad” (06) según el índice de disponibilidad

El primer sustantivo que aparece, *calle*, hace referencia a las vías públicas y a la circulación como otros nombres: *plaza, carretera, semáforo, tráfico, avenida, cruce, vía, paso de cebra, paseo, paso de peatones, rotonda, señal, glorieta, ciclovía, camino, autopista, embotellamiento*. A continuación encontramos la palabra *coche*, que se incluye en la categoría de los medios de transporte como *autobús, bicicleta, tren, tranvía, metropolitano, taxi, carro, automóvil, avión, motocicleta, ciclomotor, scooter* y las palabra *tienda*, la cual se conecta con todos los tipos de locales comerciales como los que se refieren a la restauración o a la comida (*restaurante, bar, supermercado, panadería, mercado, cafetería, chiringuito, pastelería, pescadería, cervecería, pub, pizzería*), al sector hotelero (*hotel, albergue*), al ocio (*discoteca, cine, piscina, teatro, auditorio, complejo polideportivo, club*) y de otro tipo (*farmacia, banco, estanco, papelería, oficina, centro comercial, peluquería, tienda de ropa, librería, quiosco*).

Sin embargo, los sustantivos que representan la mayoría de respuestas actualizadas se refieren a las partes de la ciudad y, en un sentido más general, a elementos geográficos. Empezamos por la palabra *parque*, que se posiciona en el sexto lugar y continuamos registrando: *centro, barrio, árbol, municipio, jardín, centro histórico, playa, puente, pueblo, capital, cauce, río, plaza mayor, puerta de la ciudad, mar, paisaje, campo, casco antiguo, mirador, escalera, periferia, muro, país, suburbio, estado, bosque, sitio, naturaleza, fuente, esquina, jardín público, sitio turístico, lugar de interés, alrededor, aparcamiento, provincia, castillo, comunidad*. Asimismo, se nombran edificios y lugares públicos o de interés turístico: *iglesia, edificio, escuela, piso, monumento, estación, ayuntamiento, museo, hospital, palacio, rascacielos, catedral, aeropuerto, parada de autobús, biblioteca, estación de tren, estatua, estadio, residencia, apartamento, estación de autobús, bloque de pisos, banqueta, centro cultural, departamento, correos*.

Los restantes sustantivos se pueden distinguir entre los que designan elementos naturales (*luz, cielo, planta, arena, ambiente, flor, temporal, nieve, nube, sol*), personas (*turista, niño, ciudadano, amigo, peatón, transeúnte, habitante, inmigrante, ciclista*), varios objetos (*cartel, billete, rueda, dinero, cuadro, cruz, ventana*) y otros que

establecen asociaciones que destacan características negativas de la ciudad (*ruido, polución, contaminación, peligro, criminalidad, basura*) junto con otras referencias positivas y vitales (*música, historia, cultura, deporte, viaje, fiesta, oportunidad, diversión, movida*).

Finalmente, destacamos algunas respuestas que abarcan muchos de los temas considerados en esta investigación: el mundo laboral (*policía, alcalde, cliente, empleo, guía turística, bombero, carnicero, albañil, empleado, trabajador, doctor*), comidas y bebidas (*tapa, pizza, tequila, café, pan*), animales (*perro, paloma*), mundo de la escuela (*universidad, enseñanza media, liceo, colegio*).

Entre las otras categorías lingüísticas señalamos los verbos que indican acciones vinculadas estrictamente con el tema en cuestión (*cruzar, pasear, visitar, girar, seguir, andar, viajar, doblar, descubrir, ir, conducir, correr, voltar, mirar, encontrarse*), entre ellas algunos son compuestos sintagmáticos con núcleo verbal (*dar un paseo, ir de compras, dar una vuelta*).

Por lo que se refiere al bajo número de adjetivos encontrados, las asociaciones son varias y hacen referencia al tamaño (*grande, pequeño*), a las direcciones (*derecha, izquierda, recto*) y a las distintas características atribuibles al tema (*cosmopolita, multicultural, moderno, famoso, antiguo, principal, poblado, diferente*).

### 5.1.7 ‘Animales’

El penúltimo centro de interés encuestado es también el que presenta entre sus palabras más disponibles las dos palabras que encabezan el listado de disponibilidad general, es decir “*perro*” y “*gato*”, que designan animales domésticos muy comunes. Siguen animales salvajes (*león, tigre, elefante, jirafa*) y animales de granja (*caballo, vaca*), el sustantivo más común para referirse a los animales marinos (*pez*) y el más general en relación con las aves (*pájaro*). A continuación se muestra en la tabla 21 el orden de estas palabras dentro de este campo léxico.

Nº	Palabra	Disponibilidad
1	perro	0.93881
2	gato	0.85810
3	león	0.37422
4	caballo	0.37357
5	pez	0.26204

6	pájaro	0.26173
7	tigre	0.19722
8	elefante	0.18185
9	vaca	0.15561
10	jirafa	0.15167

**Tabla. 21** Primeras diez palabras del centro de interés “Los animales” (07) según el índice de disponibilidad

Entre las respuestas de la tabla 21, además de los dos primeros vocablos, no se encuentran otros animales domésticos como podríamos esperar. De hecho, los sustantivos que designan los animales de compañía los encontramos solo más adelante en el listado (*conejo, tortuga, loro, papagayo, hámster, mascota, conejillo de Indias, canario*). Muchos son los que pertenecen al subconjunto de los animales de granja (*pollo, cerdo, gallina, toro, burro, gallo, cabra, oveja, yegua, pato, puerco, cabrón, oca*), otros se incluyen entre los mamíferos silvestres (*ratón, tigresa, leona, mono, rato, cebra, lobo, zorro, camello, gacela, koala, oso, bisonte, alpaca, canguro, gorila, panda, pingüino, rinoceronte, armadillo, jabalí, ardilla, dingo, ciervo, coyote, ornitorrinco, bambi, hipopótamo, pantera, reno*). Se hallan también algunos pertenecientes a la clase de las aves (*mariposa, paloma, ave, búho, gaviota, mirlo, murciélago, golondrina, aramacao, cigüeña, flamenco, avestruz*) y otros pertenecientes a los reptiles (*serpiente, cucaracha, rana, cocodrilo, lagarto, camaleón, sierpe*). Se encuentran también los insectos y los invertebrados (*mariposa, hormiga, mosquito, caracol, abeja, insecto, mosca, araña, mariquita, escarabajo, chicharra, chapulín*) y finalmente los animales marinos (*delfín, tiburón, atún, sardina, mejillón, cangrejo, marisco, pulpo, ballena*).

Destaca, por lo tanto, un alto grado de cohesión con el tema central que ha dado lugar a una gran cantidad de respuestas muy cercanas al campo léxico encuestado. Sin embargo, no son pocos los sustantivos que forman parte de aquellos subconjuntos que se alejan un poco más del núcleo central. Entre ellos podemos distinguir entre algunos que se refieren a los productos de origen animal o a la comida para animales dependiendo del contexto (*comida, carne, queso, leche, croqueta, huevo, bebida, algodón, miel*), otros que están vinculados con las partes del cuerpo de los animales (*pelo, piel, cola, pata, ojo, pluma, diente*) y un buen número que designa el entorno donde se pueden encontrar los animales (*zoológico, naturaleza, parque, jardín, casa, bosque, árbol, finca, mar, tienda de animales, safari, agua, tienda, perrería, desierto,*



*circo*). Finalmente hay una serie de sustantivos que hacen referencia a personas que pueden tener relación con los animales (*veterinario, dueño, cuidador de perros, niño, cazador*), vocablos que proceden de relatos de fantasía (*dragón, Caperucita Roja*) y muchas asociaciones de distinto tipo (*amigo, amor, cariño, felicidad, paseo, caza, juego, beso, fidelidad, ladrido, techo, cucha, equitación, bestia, soledad, vida, ganadería, muerte, calle, vuelo, pelota, perfume, mordedura, huella, juguete, nombre, cielo, compañero, libertad, miedo, compañía, ruido, fusil, extinción, nostalgia, dulzura, radar, gallinero, torero, color*).

Como se puede observar, también en este centro la mayoría de los elementos está formada por sustantivos (un 75% sobre el total). Un 15% son verbos que tienen diferentes relaciones con el centro de interés: en su mayoría se trata de verbos de movimiento referidos a acciones cumplidas por los animales o por los seres humanos en relación con ellos (*pasear, cuidar, caminar, jugar, dar un paseo, nadar, cazar, volar, ladrar, beber, correr, dar de comer, amar, morder, querer, devorar, comer, curar, sacar, abrazar, morir, nutrir, acariciar, adiestrar, cabalgar, roer, gustar, comprar, limpiar, salir, dar una vuelta, matar, cantar, dormir, pastar, pescar*). Cierra la muestra un 10% de adjetivos que describen las dimensiones de los animales (*pequeño, grande, alto, bajo, largo, corto*), los colores (*negro, blanco, verde*), características atribuibles a ellos (*salvaje, doméstico, peligroso, tranquilo, fiel, malo, bonito, bueno, agresivo, sucio, limpio, cariñoso, feroz*). Por último, registramos dos adjetivos que se utilizan para hacer referencia a las conductas alimentarias de las personas (*vegetariano, vegano*) y que por eso tienen una relación bastante distante con el tema.

### **5.1.8 ‘Profesiones y oficios’**

Por lo que se refiere al último centro de interés “Profesiones y oficios” destaca aquí también el dominio de los sustantivos, que componen el 82 % sobre la muestra total. En la tabla 22 se ordenan las diez palabras más disponibles actualizadas por los informantes, donde señalamos la presencia de nueve sustantivos y un verbo muy genérico pero vinculado al tema, es decir *trabajar*. A continuación nos centraremos sobre los mecanismos de formación léxica que encontramos al distinguir las varias respuestas del campo léxico y presentaremos los vocablos según los diferentes sufijos utilizados para formar los sustantivos que designan profesiones.

Nº	Palabra	Disponibilidad
1	profesor	0.80357
2	médico	0.38388
3	estudiante	0.29056
4	abogado	0.27876
5	camarero	0.24678
6	cocinero	0.20414
7	doctor	0.17474
8	empleado	0.14050
9	policía	0.13688
10	trabajar	0.13407

**Tabla. 22** Primeras diez palabras del centro de interés “Profesiones y oficios” (08) según el índice de disponibilidad

La mayoría de los sustantivos que encontramos a partir de la primera palabra más disponible se derivan del prefijo *-tor/-dor/-or*, que, como define la RAE, indica un ‘agente’ (*profesor, doctor, escritor, director, traductor, trabajador, pintor, vendedor, cantor, nadador, jugador, jugador de futbol, autor, entrenador, director de cine, escultor, inventor, limpiador, preparador físico, entrevistador, monitor, buceador, tenor, embajador, constructor, controlador, administrador, lector*). En el segundo lugar encontramos un sustantivo que termina en *-ico*, sufijo que normalmente suele formar adjetivos que se sustantivan para obtener nombres de profesión (*médico, gráfico, físico, químico, técnico informático, músico, técnico, científico, mecánico, político*). Sigue como tercera palabra “estudiante”, cuyo sufijo *-nte* sirve para formar sustantivos a partir de verbos y significa ‘persona que realiza la acción’ (*cantante, presidente, ayudante, agente de policía, enseñante, negociante, agente, agente inmobiliario, comerciante, asistente, representante, dirigente*). Entre estas aparecen dos que sufren un cambio de significado: *dependiente* por ser un ‘empleado que tiene a su cargo atender a los clientes en las tiendas’ y *cliente*, que como define la RAE es una ‘persona que compra en una tienda o que utiliza con asiduidad los servicios de un profesional o empresa’. Sucesivamente registramos en la cuarta posición un sustantivo formado por el sufijo *-ado* que indica ‘cargo’ (*abogado, empleado, soldado*) y el sustantivo que hemos dejado tal como aparecía solo en femenino *criada*. Siguen dos vocablos que terminan con el sufijo *-ero*, uno de los mayormente utilizados en derivación, que aporta la connotación de ‘oficio, ocupación, profesión o cargo’ (*camarero, cocinero,*

*enfermero, peluquero, fontanero, bombero, obrero, ingeniero, jardinero, panadero, cajero, plomero, carnicero, florero, carpintero, banquero, niño, cartero, zapatero, verdulero, portero, camionero, carabinero, granjero, torero)* y el femenino *partera*. Además, en el listado registramos un gran número de sustantivos derivados del sufijo *-ista* que designa ‘a la persona que tiene determinada ocupación, profesión u oficio (*dentista, futbolista, periodista, electricista, taxista, recepcionista, artista, florista, deportista, esteticista, baloncestista, pianista, ecologista, maquillista, maquinista, motorista, tecladista, lingüista, oculista*), otros que tienen como morfema derivativo *-logo* que significa ‘persona versada o especialista’ en relación con lo que el primer elemento indica (*cosmetólogo, arqueólogo, biólogo, ginecólogo*). Tampoco faltan sustantivos que, formados por adjetivos, se han sustantivado mediante el morfema *-al* ‘pertenencia o relación’ y que hacen referencia a una profesión (*comercial, profesional, general*), así como *-ín/ina* (*bailarín, campesino*, con el interfijo *-es-*). El mismo significado lo aporta también el morfema *-ario* que encontramos en las palabras (*secretario, veterinario, empresario, voluntario, bancario*). Finalmente señalamos todas las unidades que no se incluyen en las distinciones formuladas, pero que, como sustantivos, se refieren al centro de interés en cuestión (*policía, canguro, jefe, albañil, cirujano, maestro, alumno, piloto, contable, guardia civil, poeta, pinchadiscos, alcalde, astrónomo, ladrón, modelo, rey, fotógrafo, gimnasta, atleta, lavaplatos, ministro, psiquiatra, filósofo, guía turística, consúl, juez*). Asimismo, resulta de gran interés, además, aquellos sustantivos que ya hemos analizado, la presencia de palabras que han aparecido muchas veces al lado del masculino y que destacan por su sufijo especial (*actriz, reina*) o que han aparecido solo en femenino como (*azafata, ama de casa, dueña de casa, mujer de negocios*). La última destaca por ser un préstamo adaptado del inglés y nos lleva a presentar otros extranjerismos, adaptados o no, que registramos en el listado: *baby-sitter, mánager, barman* (del inglés) y *chofer* (del francés). Sin embargo, no todos los sustantivos se refieren a profesiones u oficios en sentido estricto. Muchos otros están vinculados con lugares donde se trabaja (*oficina, trabajo, tienda, restaurante, escuela, hospital, peluquería, farmacia, empresa, universidad, despacho, estudio, panadería, bar, secretaría, banco, florería, fábrica, papelería, tienda de ropa, correos, colegio, casa, supermercado*).

Por último, señalamos compuestos sintagmáticos de tipo nominal compuestos por sustantivo+preposición+sustantivo (*ama de casa, jugador de futbol, agente de policía, director de cine, dueña de casa, mujer de negocios, jefe del gobierno, guía de turismo*) binominales (*guardia civil*), nombre+adjetivo (*técnico informático, preparador físico, agente inmobiliario, fuerzas armadas, guía turística*).

Entre los verbos (16%) aparecen muchos que están en relación con los nombres que se han actualizado (*enseñar, estudiar, jugar, vender, cantar, cocinar, conducir, comprar, aprender, ganar, correr, pintar*) y otros más generales. Los adjetivos son muy escasos y componen apenas el 2% de la muestra, lo cual manifiesta una vez más la inclinación general de los informantes a generar respuestas pertenecientes a la categoría de los sustantivos, que, como hemos visto, permiten mantener una cohesión más cercana al tema dado.



## Conclusiones

La presente investigación ha tenido como objetivo primario el estudio de la disponibilidad léxica en una muestra limitada de alumnos de español como lengua extranjera en la Universidad de Padua. Nuestro propósito es realizar una aproximación a estos tipos de estudios que pueda representar una contribución original dentro del panorama de la aplicación de la lingüística estadística a la enseñanza y aprendizaje de ELE en Italia.

Una vez terminados los análisis de los resultados cuantitativos calculados a partir de los 112 informantes que forman parte de una muestra no preestaticada, podemos exponer las conclusiones que hemos obtenido.

Primero, hemos destacado una menor cantidad en cuanto al número de unidades léxicas aportadas por cada informante en comparación con el promedio de palabras actualizadas por cada alumno en los trabajos hispánicos realizados con hablantes nativos, lo cual deriva de la menor competencia léxica en los participantes a nuestra encuesta.

Además, por lo que se refiere a las diferencias numéricas en la producción de resultados según la distinción entre palabras y vocablos, hemos observado que en general se mantiene un paralelismo entre la cantidad total de palabras y la cantidad total de vocablos en los ocho centros de interés, salvo algunas excepciones que se producen en algunos de ellos. Así, el campo léxico “Profesiones y oficios” (08) asciende de la antepenúltima posición en cuanto al número de palabras diferentes al primero en el número de vocablos. Por el contrario, el centro de interés “Partes del cuerpo” (01) desciende del cuarto puesto en la producción de palabras al último lugar en relación con los vocablos totales.

El análisis de los promedios de palabras según el factor ‘sexo’ ha demostrado la superioridad femenina en relación con los datos aportados por los hombres. Sin embargo, nos resulta difícil considerar discriminante este criterio ya que su representatividad es moderada. Una situación diferente y anómala se registra en los resultados de la variable ‘edad’, que da lugar a una tendencia que no encuentra similitudes en otras investigaciones similares, donde la curva que se da en ascenso a partir de los informantes de menor edad hasta los alumnos de edad mayor, sufre una caída en el intervalo intermedio. Asimismo, otro criterio que nos ha llevado al

establecimiento de su escaso grado de significación es la variable ‘zona geográfica’, la cual no ha generado diferencias relevantes entre los dos grupos considerados.

En la variable ‘carrera de estudios’ destaca el dato de los estudiantes en “Mediación Lingüística y Cultural”, los cuales han producido una mayor cantidad de palabras en siete centros de interés en comparación con sus compañeros matriculados en “Lenguas, Literaturas y Culturas Modernas” y otros grados. Podemos afirmar, por lo tanto, que se ha detectado un mayor repertorio léxico en los estudiantes cuyas asignaturas primarias son los idiomas extranjeros con enfoque a la mediación, lo cual podría hacernos suponer una mayor proximidad de sus estudios en relación con el desarrollo de aquellas competencias lingüísticas que sirven para manejar la lengua en las situaciones comunicativas consideradas en nuestro trabajo.

Al tomar en cuenta si existe una correspondencia entre el ‘nivel sociocultural’ de un informante y la cantidad de palabras respondidas, no notamos una relación directamente proporcional entre la ascensión del nivel sociocultural del alumno y el promedio de palabras. Si consideramos el nivel bajo y medio, se produce una tendencia que registra un aumento en la producción de palabras a medida que el nivel sube, pero esta progresión se interrumpe en relación con el nivel alto, lo cual podría reflejar la escasa capacidad discriminante de este factor, como se ha registrado en otras investigaciones, o podría deberse a la baja representatividad numérica que tienen los informantes de este nivel.

Por lo que respecta a la variable ‘lengua materna’, a pesar de la ventaja numérica de los participantes italianos en la muestra, no observamos su superioridad en relación con la producción de palabras. De hecho, no obstante el origen no románico de la lengua inglesa, los informantes que hablan este idioma son los que obtienen el mayor promedio de palabras en general, seguidos por los italianos y los estudiantes que hablan otros idiomas.

En la variable ‘conocimiento de otras lenguas’ se ha manifestado una tendencia general creciente en relación con el número de palabras producidas por los informantes a medida que aumenta el número de lenguas que ellos conocen, dato que refleja una fuerte relación asociativa entre el criterio considerado y la capacidad de memorizar una mayor cantidad de unidades léxicas y de actualizarlas en determinadas situaciones comunicativas.

Sin embargo, el dato que más destaca entre todas las variables consideradas se encuentra en relación con la última variable seleccionada, es decir ‘años de estudio del español’, donde se manifiesta una interesante correlación entre el número de palabras aportadas por cada informante y el nivel de conocimiento de la lengua, que suponemos aumente con la progresión de los años. La ventaja del nivel más avanzado se registra en siete sobre ocho centros de interés y nos lleva a considerar que a medida que el proceso de aprendizaje se desarrolla, hace falta introducir mayores y más complejas unidades léxicas que pueden ser retenidas por los informantes que ya tienen una mayor competencia léxica.

De manera diferente hemos cotejado los datos resultantes de la producción de vocablos por cada informante ya que hemos encontrado dificultades en la obtención de promedios de vocablos que nos ayudasen a realizar comparaciones fidedignas con los promedios de palabras. Debido a la gran heterogeneidad numérica de los grupos considerados por cada variable, hemos decidido considerar solo las cantidades totales de respuestas por cada centro de interés y cada factor extralingüístico. Los resultados generales han puesto de manifiesto una correspondencia directamente proporcional entre el número de vocablos totales y el respectivo número de componentes de los grupos pertenecientes a las distintas variables encuestadas.

En trabajos posteriores, a fin de obtener una mayor representatividad en cuanto al grado de significación de las distintas variables consideradas, será mejor determinar una muestra preestratificada, la cual podrá permitir llevar a cabo comparaciones más eficaces entre el promedio de palabras y el promedio de vocablos por informante. Sin embargo, los cotejos que hemos realizado no resultan de menor interés, aunque se trate de una muestra no preestratificada. De hecho, como se ha descrito anteriormente, hemos registrado en algunos casos resultados que consideramos bastante relevantes y satisfactorios en línea con nuestros objetivos de estudio.

Por lo que se refiere a los análisis de los resultados cualitativos, nos hemos centrado en la presentación de las primeras diez palabras más disponibles por cada centro de interés, a partir de las cuales hemos llevado a cabo una descripción general de las unidades léxicas que componen cada campo léxico y las hemos agrupado bajo diferentes subconjuntos semánticos para comprobar su cercanía al núcleo central. Estas



descripciones nos han permitido analizar también la relación de correspondencia de los resultados del índice de cohesión semántica de cada centro de interés con las diferentes aportaciones que se han registrado. En este sentido, destaca el dato del centro “Partes del cuerpo” (01) que muestra una relación asociativa muy vinculada con el tema en la mayoría de las respuestas obtenidas al igual que el campo léxico “La ropa” (02). Además, a diferencia de lo que se ha manifestado en relación con el índice de cohesión, hemos notado que el centro de interés “Comidas y bebidas” (04), que se encuentra en la cuarta posición, ha registrado una mayor relación semántica con el estímulo dado en comparación con “Partes de la casa”(03), su antecedente en el listado, que ha manifestado un general alejamiento del tipo de respuestas dadas en relación con el núcleo central. Los restantes campos léxicos muestran relaciones léxico-asociativas muy amplias, dato que se puede explicar también por el alto número de vocablos totales que se registran en los centros “Profesiones y oficios” (08), “La ciudad” (06) y “La escuela” (05).

Por lo que respecta al tipo de categoría gramatical que se encuentra mayormente en los distintos centros de interés, destaca la abundancia de los sustantivos en todos los centros de interés frente a la escasez de los verbos y adjetivos que no superan el 25% sobre el total en ningún campo léxico. Este dato pone de relieve la relación asociativa primaria que los informantes han establecido con los nombres, si bien en las explicaciones iniciales proporcionadas por el encuestador antes de empezar la prueba se hubiese aclarado la posibilidad de aportar palabras de diferentes categorías gramaticales.

Por otro lado, hemos querido analizar aquellos fenómenos lingüísticos que destacan las características cualitativas de las respuestas aportadas por nuestros informantes: las unidades relacionadas con la variación léxica, los cambios semánticos, casos de importación léxica y mecanismos de formación léxica. En general, hemos notado que el vocabulario de nuestros alumnos es de carácter principalmente estándar, con poca presencia de casos de variación diatópica –entre ellos se hallan unidades del español de América en el tema “La Ropa” (02)– y por lo que respecta a los tipos de cambios semánticos hemos detectado un número escaso de eufemismos –en el centro de interés “Las partes del cuerpo” (01). Sin embargo, encontramos una mayor cantidad de extranjerismos y prestamos procedentes en su mayoría del centro “Profesiones y

oficios” (08) donde aparecen también muchos mecanismos de formación y derivación de palabras. Además, en un buen número de centros de interés se registran algunos casos de compuestos sintagmáticos nominales y verbales.

Creemos que los estudios de disponibilidad léxica en alumnos de ELE representan un campo dentro del ámbito de la enseñanza-aprendizaje de lenguas que puede aportar interesantes aplicaciones. Además, las investigaciones sobre informantes que aprenden el idioma español en un país cuya lengua oficial es diferente, es un camino todavía poco explorado, como hemos puesto de relieve en el capítulo dedicado a los diferentes tipos de estudio realizados sobre este tema. Por eso, nuestra contribución, si bien de carácter introductorio, se puede colocar como una de las primeras de este tipo en el país donde se ha llevado a cabo y podrá dar lugar a investigaciones posteriores que profundicen aspectos que por razones de espacio no hemos podido tratar. Así, será muy interesante realizar un análisis de los errores cometidos por cada informante en la encuesta para aportar conocimiento sobre las diferentes lagunas registradas o llevar a cabo ulteriores análisis cualitativas que realicen comparaciones entre las palabras actualizadas según las distintas variables o entre datos aportados por muestras de informantes diferentes. Además, se ha manifestado la posibilidad de aplicar estos trabajos también a la comparación con manuales de estudio de ELE o a su elaboración, según las necesidades de aprendizaje de los alumnos.

Por último, en los anexos de esta investigación adjuntamos todo el material que hemos recogido y que, ordenado por cada centro de interés y según el índice de disponibilidad, se podrá utilizar como herramienta adecuada para la optimización del estudio de la lengua y la potenciación de las competencias léxicas de los alumnos mediante la realización de actividades en clase o para el estudio en casa, individuales o en grupo, que les ayuden a progresar a medida que aumenta su nivel de estudio.



# Anexo I

## Cuestionario DISPONIBILIDAD LÉXICA

### Datos de la encuesta:

Fecha: 26/10/2016

Centro escolar: Università degli Studi di Padova (DiSLL)

Ciudad: Padua

### Datos del informante:

1. **Sexo:**  Varón  Mujer

2. **Edad:** .....

### 3. Estudios actuales:

Lingue, Letterature e Culture Moderne ( I año  II año  III año)

Mediazione Linguistica e Culturale ( I año  II año  III año)

Otro curso .....

### 4. Estudios anteriores:

Título de educación secundaria: ..... (Centro:  Público  Privado)

Otro .....

5. **Localidad de residencia habitual:** ..... Provincia: .....

Si no ha vivido siempre en dicho lugar, indique dónde: .....

### 6. Lenguas habladas

Lengua materna: .....

Conocimiento de otras lenguas: ..... .....

Años de estudio del español: .....

Indique si ha vivido una experiencia de estudios en España o en un país de habla española:

Sí. ¿Dónde? ..... ¿Cuándo? ..... Duración de la estancia.....

No

### 7. ¿Cuál es la profesión u oficio de sus padres?

Padre: .....

Madre: .....

### 8. Estudios de sus padres:

	padre	madre
a) No ha ido a la escuela o muy poco	<input type="checkbox"/>	<input type="checkbox"/>
b) Enseñanza primaria (completa)	<input type="checkbox"/>	<input type="checkbox"/>
c) Enseñanza secundaria de primer grado	<input type="checkbox"/>	<input type="checkbox"/>

- d) Enseñanza secundaria de segundo grado
- e) Título universitario

## Anexo II

### Listados generales según el índice de disponibilidad por cada centro de interés

#### 01.Partes del cuerpo

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	ojo	0.73517	6.462 %	91.964 %	6.462 %
2	mano	0.60497	5.897 %	83.929 %	12.359 %
3	boca	0.49513	4.831 %	68.750 %	17.190 %
4	pelo	0.48834	5.144 %	73.214 %	22.334 %
5	brazo	0.46566	4.705 %	66.964 %	27.039 %
6	pierna	0.45333	4.642 %	66.071 %	31.681 %
7	nariz	0.44400	4.580 %	65.179 %	36.261 %
8	pie	0.43265	4.893 %	69.643 %	41.154 %
9	cabeza	0.41682	3.701 %	52.679 %	44.855 %
10	cara	0.32904	3.011 %	42.857 %	47.866 %
11	oreja	0.31540	3.639 %	51.786 %	51.505 %
12	dedo	0.30767	3.890 %	55.357 %	55.395 %
13	labio	0.25751	2.949 %	41.964 %	58.344 %
14	espalda	0.23998	3.325 %	47.321 %	61.669 %
15	rodilla	0.18428	2.384 %	33.929 %	64.053 %
16	cuello	0.18256	2.321 %	33.036 %	66.374 %
17	diente	0.17936	2.509 %	35.714 %	68.883 %
18	ceja	0.15446	1.882 %	26.786 %	70.765 %
19	uña	0.14533	2.008 %	28.571 %	72.773 %
20	pestaña	0.13267	1.819 %	25.893 %	74.592 %
21	lengua	0.10695	1.443 %	20.536 %	76.035 %
22	barriga	0.10634	1.506 %	21.429 %	77.541 %
23	codo	0.09621	1.129 %	16.071 %	78.670 %
24	pecho	0.09189	1.506 %	21.429 %	80.176 %
25	hombro	0.08864	1.255 %	17.857 %	81.431 %
26	cabello	0.07074	0.690 %	9.821 %	82.121 %
27	corazón	0.05220	0.690 %	9.821 %	82.811 %
28	estómago	0.04610	0.878 %	12.500 %	83.689 %
29	panza	0.03826	0.502 %	7.143 %	84.191 %
30	frente	0.03773	0.502 %	7.143 %	84.693 %
31	mejilla	0.03587	0.690 %	9.821 %	85.383 %
32	largo	0.03473	0.565 %	8.036 %	85.948 %
33	alto	0.03293	0.565 %	8.036 %	86.513 %
34	bajo	0.03085	0.502 %	7.143 %	87.015 %
35	piel	0.02883	0.439 %	6.250 %	87.454 %
36	garganta	0.02877	0.439 %	6.250 %	87.893 %
37	culo	0.02814	0.376 %	5.357 %	88.269 %
38	gordo	0.02769	0.502 %	7.143 %	88.771 %
39	oír	0.02767	0.314 %	4.464 %	89.085 %
40	delgado	0.02717	0.502 %	7.143 %	89.587 %
41	redondo	0.02713	0.251 %	3.571 %	89.838 %
42	caminar	0.02452	0.314 %	4.464 %	90.152 %
43	rubio	0.02433	0.439 %	6.250 %	90.591 %
44	comer	0.02337	0.251 %	3.571 %	90.842 %

45	corto	0.02195	0.376 %	5.357 %	91.218 %
46	pequeño	0.02144	0.314 %	4.464 %	91.532 %
47	grande	0.02120	0.251 %	3.571 %	91.783 %
48	entrecejo	0.02065	0.188 %	2.679 %	91.971 %
49	azul	0.01951	0.251 %	3.571 %	92.222 %
50	muñeca	0.01862	0.439 %	6.250 %	92.661 %
51	tobillo	0.01827	0.314 %	4.464 %	92.975 %
52	castaño	0.01584	0.188 %	2.679 %	93.163 %
53	moreno	0.01548	0.251 %	3.571 %	93.414 %
54	rostro	0.01539	0.125 %	1.786 %	93.539 %
55	sonrisa	0.01488	0.188 %	2.679 %	93.727 %
56	rizado	0.01216	0.251 %	3.571 %	93.978 %
57	cadera	0.01146	0.125 %	1.786 %	94.103 %
58	bigote	0.01144	0.125 %	1.786 %	94.228 %
59	negro	0.01081	0.125 %	1.786 %	94.353 %
60	muslo	0.01059	0.188 %	2.679 %	94.541 %
61	codazo	0.01055	0.125 %	1.786 %	94.666 %
62	cintura	0.01038	0.188 %	2.679 %	94.854 %
63	mentón	0.00967	0.125 %	1.786 %	94.979 %
64	ombbligo	0.00889	0.125 %	1.786 %	95.104 %
65	rojo	0.00882	0.125 %	1.786 %	95.229 %
66	barbilla	0.00853	0.125 %	1.786 %	95.354 %
67	feo	0.00829	0.063 %	0.893 %	95.480 %
68	gafas	0.00809	0.125 %	1.786 %	95.605 %
69	tocar	0.00787	0.125 %	1.786 %	95.730 %
70	cuerpo	0.00770	0.063 %	0.893 %	95.793 %
71	brazalete	0.00770	0.063 %	0.893 %	95.856 %
72	párpado	0.00749	0.125 %	1.786 %	95.981 %
73	pelirrojo	0.00733	0.188 %	2.679 %	96.169 %
74	vientre	0.00730	0.125 %	1.786 %	96.294 %
75	corporal	0.00715	0.063 %	0.893 %	96.357 %
76	señor	0.00715	0.063 %	0.893 %	96.420 %
77	labiodental	0.00715	0.063 %	0.893 %	96.483 %
78	pintar	0.00664	0.063 %	0.893 %	96.546 %
79	paseo	0.00616	0.063 %	0.893 %	96.609 %
80	oído	0.00613	0.125 %	1.786 %	96.734 %
81	escuchar	0.00572	0.063 %	0.893 %	96.797 %
82	claro	0.00572	0.063 %	0.893 %	96.860 %
83	tripa	0.00572	0.063 %	0.893 %	96.923 %
84	músculo	0.00531	0.063 %	0.893 %	96.986 %
85	pensar	0.00531	0.063 %	0.893 %	97.049 %
86	narizado	0.00531	0.063 %	0.893 %	97.112 %
87	brazada	0.00531	0.063 %	0.893 %	97.175 %
88	entraña	0.00531	0.063 %	0.893 %	97.238 %
89	hueso	0.00493	0.063 %	0.893 %	97.301 %
90	narizón	0.00493	0.063 %	0.893 %	97.364 %
91	beso	0.00493	0.063 %	0.893 %	97.427 %
92	arrodillarse	0.00493	0.063 %	0.893 %	97.490 %
93	anillo	0.00458	0.063 %	0.893 %	97.553 %
94	pulmón	0.00425	0.063 %	0.893 %	97.616 %
95	sobrecejo	0.00425	0.063 %	0.893 %	97.679 %
96	peluquero	0.00395	0.063 %	0.893 %	97.742 %
97	órgano	0.00395	0.063 %	0.893 %	97.805 %
98	culada	0.00395	0.063 %	0.893 %	97.868 %
99	peluquería	0.00367	0.063 %	0.893 %	97.931 %

100	dibujar	0.00367	0.063 %	0.893 %	97.994 %
101	dedo de pie	0.00367	0.063 %	0.893 %	98.057 %
102	ortopedia	0.00367	0.063 %	0.893 %	98.120 %
103	trasero	0.00367	0.063 %	0.893 %	98.183 %
104	sentarse	0.00367	0.063 %	0.893 %	98.246 %
105	manotazo	0.00367	0.063 %	0.893 %	98.309 %
106	echar un vistazo	0.00367	0.063 %	0.893 %	98.372 %
107	andar	0.00340	0.063 %	0.893 %	98.435 %
108	pulsera	0.00340	0.063 %	0.893 %	98.498 %
109	oculista	0.00340	0.063 %	0.893 %	98.561 %
110	color	0.00340	0.063 %	0.893 %	98.624 %
111	oler	0.00316	0.063 %	0.893 %	98.687 %
112	peinar	0.00316	0.063 %	0.893 %	98.750 %
113	ver	0.00316	0.063 %	0.893 %	98.813 %
114	lunar	0.00316	0.063 %	0.893 %	98.876 %
115	seno	0.00316	0.063 %	0.893 %	98.939 %
116	hablar	0.00293	0.063 %	0.893 %	99.002 %
117	cachete	0.00293	0.063 %	0.893 %	99.065 %
118	bailar	0.00293	0.063 %	0.893 %	99.128 %
119	romper	0.00272	0.063 %	0.893 %	99.191 %
120	mujer	0.00272	0.063 %	0.893 %	99.254 %
121	perfume	0.00272	0.063 %	0.893 %	99.317 %
122	lavarse	0.00272	0.063 %	0.893 %	99.380 %
123	verde	0.00272	0.063 %	0.893 %	99.443 %
124	ancho	0.00272	0.063 %	0.893 %	99.506 %
125	varón	0.00253	0.063 %	0.893 %	99.569 %
126	medir	0.00253	0.063 %	0.893 %	99.632 %
127	moverse	0.00253	0.063 %	0.893 %	99.695 %
128	observar	0.00253	0.063 %	0.893 %	99.758 %
129	pendientes	0.00235	0.063 %	0.893 %	99.821 %
130	tez	0.00235	0.063 %	0.893 %	99.884 %
131	pulso	0.00175	0.063 %	0.893 %	99.947 %
132	muela	0.00120	0.063 %	0.893 %	100.010 %

## 02.La ropa

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	camiseta	0.63352	7.331 %	79.464 %	7.331 %
2	zapato	0.59360	7.743 %	83.929 %	15.074 %
3	pantalón	0.47923	6.013 %	65.179 %	21.087 %
4	falda	0.44572	5.437 %	58.929 %	26.524 %
5	camisa	0.37186	4.613 %	50.000 %	31.137 %
6	vaquero	0.35704	4.201 %	45.536 %	35.338 %
7	jersey	0.34893	4.860 %	52.679 %	40.198 %
8	bufanda	0.29223	4.283 %	46.429 %	44.481 %
9	zapatilla	0.18666	2.718 %	29.464 %	47.199 %
10	calcetín	0.14846	2.554 %	27.679 %	49.753 %
11	vestido	0.13239	2.059 %	22.321 %	51.812 %
12	bota	0.13150	2.224 %	24.107 %	54.036 %
13	chaqueta	0.12625	2.224 %	24.107 %	56.260 %
14	abrigo	0.11435	1.812 %	19.643 %	58.072 %
15	traje	0.09693	1.565 %	16.964 %	59.637 %
16	sombrero	0.09478	1.565 %	16.964 %	61.202 %



17	minifalda	0.09014	1.318 %	14.286 %	62.520 %
18	guante	0.07825	1.565 %	16.964 %	64.085 %
19	vestir	0.07784	1.318 %	14.286 %	65.403 %
20	gafas	0.07684	1.236 %	13.393 %	66.639 %
21	tacones	0.07195	1.318 %	14.286 %	67.957 %
22	sudadera	0.05978	1.071 %	11.607 %	69.028 %
23	blusa	0.05832	0.906 %	9.821 %	69.934 %
24	zapato de tacón	0.05233	0.988 %	10.714 %	70.922 %
25	bolso	0.04669	0.906 %	9.821 %	71.828 %
26	bolsa	0.04499	0.988 %	10.714 %	72.816 %
27	calzoncillo	0.04480	0.659 %	7.143 %	73.475 %
28	tienda	0.04229	0.659 %	7.143 %	74.134 %
29	chancla	0.04173	0.824 %	8.929 %	74.958 %
30	cinturón	0.04107	0.988 %	10.714 %	75.946 %
31	zapatilla de deporte	0.03944	0.659 %	7.143 %	76.605 %
32	sujetador	0.03484	0.659 %	7.143 %	77.264 %
33	jeans	0.03465	0.494 %	5.357 %	77.758 %
34	comprar	0.03102	0.659 %	7.143 %	78.417 %
35	t-shirt	0.03002	0.577 %	6.250 %	78.994 %
36	armario	0.02827	0.329 %	3.571 %	79.323 %
37	largo	0.02761	0.577 %	6.250 %	79.900 %
38	sandalia	0.02728	0.577 %	6.250 %	80.477 %
39	media	0.02647	0.412 %	4.464 %	80.889 %
40	braga	0.02428	0.329 %	3.571 %	81.218 %
41	llevar	0.02336	0.329 %	3.571 %	81.547 %
42	gorro	0.02325	0.577 %	6.250 %	82.124 %
43	algodón	0.02198	0.412 %	4.464 %	82.536 %
44	reloj	0.02152	0.577 %	6.250 %	83.113 %
45	espadrilla	0.02130	0.329 %	3.571 %	83.442 %
46	amarillo	0.01929	0.329 %	3.571 %	83.771 %
47	corbata	0.01894	0.329 %	3.571 %	84.100 %
48	verde	0.01859	0.412 %	4.464 %	84.512 %
49	seda	0.01676	0.329 %	3.571 %	84.841 %
50	gafas de sol	0.01591	0.329 %	3.571 %	85.170 %
51	pañuelo	0.01573	0.247 %	2.679 %	85.417 %
52	capa	0.01553	0.247 %	2.679 %	85.664 %
53	rojo	0.01551	0.329 %	3.571 %	85.993 %
54	pijama	0.01535	0.412 %	4.464 %	86.405 %
55	poner	0.01489	0.329 %	3.571 %	86.734 %
56	negro	0.01449	0.247 %	2.679 %	86.981 %
57	collar	0.01439	0.329 %	3.571 %	87.310 %
58	remera	0.01406	0.165 %	1.786 %	87.475 %
59	bolsillo	0.01377	0.247 %	2.679 %	87.722 %
60	bañador	0.01319	0.412 %	4.464 %	88.134 %
61	corto	0.01301	0.329 %	3.571 %	88.463 %
62	tenis	0.01254	0.247 %	2.679 %	88.710 %
63	pulóver	0.01233	0.247 %	2.679 %	88.957 %
64	vestido elegante	0.01197	0.165 %	1.786 %	89.122 %
65	blanco	0.01173	0.247 %	2.679 %	89.369 %
66	anillo	0.01122	0.412 %	4.464 %	89.781 %
67	pulsera	0.01111	0.247 %	2.679 %	90.028 %
68	pantalón corto	0.01108	0.247 %	2.679 %	90.275 %
69	azul	0.01093	0.247 %	2.679 %	90.522 %

70	mochila	0.01092	0.165 %	1.786 %	90.687 %
71	vestido de novia	0.01083	0.165 %	1.786 %	90.852 %
72	ir de compras	0.01063	0.165 %	1.786 %	91.017 %
73	traje de baño	0.01062	0.247 %	2.679 %	91.264 %
74	brazalete	0.00990	0.165 %	1.786 %	91.511 %
75	cárdigan	0.00990	0.165 %	1.786 %	91.594 %
76	moda	0.00985	0.165 %	1.786 %	91.759 %
77	planchar	0.00980	0.165 %	1.786 %	91.924 %
78	manga	0.00980	0.165 %	1.786 %	92.089 %
79	color	0.00962	0.165 %	1.786 %	92.254 %
80	pendiente	0.00941	0.329 %	3.571 %	92.583 %
81	chica	0.00893	0.082 %	0.893 %	92.665 %
82	campera	0.00893	0.082 %	0.893 %	92.747 %
83	ropa	0.00893	0.082 %	0.893 %	92.829 %
84	de color	0.00891	0.165 %	1.786 %	92.994 %
85	bikini	0.00859	0.247 %	2.679 %	93.241 %
86	casco	0.00808	0.082 %	0.893 %	93.323 %
87	buso	0.00808	0.082 %	0.893 %	93.405 %
88	mujer	0.00808	0.082 %	0.893 %	93.487 %
89	lindo	0.00808	0.082 %	0.893 %	93.569 %
90	abrigarse	0.00797	0.247 %	2.679 %	93.816 %
91	gorra	0.00779	0.165 %	1.786 %	93.981 %
92	probar	0.00771	0.247 %	2.679 %	94.228 %
93	bombacha	0.00741	0.165 %	1.786 %	94.393 %
94	de manga larga	0.00733	0.165 %	1.786 %	94.558 %
95	camisón	0.00731	0.082 %	0.893 %	94.640 %
96	hombre	0.00731	0.082 %	0.893 %	94.722 %
97	de manga corta	0.00712	0.165 %	1.786 %	94.969 %
98	lavar	0.00712	0.165 %	1.786 %	95.134 %
99	chal	0.00691	0.165 %	1.786 %	95.299 %
100	playera	0.00661	0.082 %	0.893 %	95.381 %
101	de flores	0.00661	0.082 %	0.893 %	95.463 %
102	elegancia	0.00599	0.082 %	0.893 %	95.545 %
103	lencería	0.00599	0.082 %	0.893 %	95.627 %
104	cama	0.00599	0.082 %	0.893 %	95.709 %
105	maleta	0.00599	0.082 %	0.893 %	95.791 %
106	ropa interior	0.00599	0.082 %	0.893 %	95.873 %
107	vender	0.00542	0.082 %	0.893 %	95.955 %
108	estirar	0.00542	0.082 %	0.893 %	96.037 %
109	estilo vintage	0.00490	0.082 %	0.893 %	96.119 %
110	pantufla	0.00490	0.082 %	0.893 %	96.201 %
111	cuarto	0.00490	0.082 %	0.893 %	96.283 %
112	talla	0.00489	0.165 %	1.786 %	96.448 %
113	de rayas	0.00489	0.165 %	1.786 %	96.613 %
114	quedar	0.00443	0.082 %	0.893 %	96.695 %
115	cintura	0.00443	0.082 %	0.893 %	96.777 %
116	pagar	0.00443	0.082 %	0.893 %	96.859 %
117	manta	0.00443	0.082 %	0.893 %	96.941 %
118	estampado	0.00401	0.082 %	0.893 %	97.023 %
119	chándal	0.00401	0.082 %	0.893 %	97.105 %
120	deporte	0.00401	0.082 %	0.893 %	97.187 %
121	vestido de gala	0.00401	0.082 %	0.893 %	97.269 %
122	accesorio	0.00401	0.082 %	0.893 %	97.351 %

123	nuevo	0.00401	0.082 %	0.893 %	97.433 %
124	ropería	0.00363	0.082 %	0.893 %	97.515 %
125	camiseta interior	0.00363	0.082 %	0.893 %	97.597 %
126	pana	0.00363	0.082 %	0.893 %	97.679 %
127	camiseta de tirantes	0.00363	0.082 %	0.893 %	97.761 %
128	estrecho	0.00363	0.082 %	0.893 %	97.843 %
129	sacarse	0.00363	0.082 %	0.893 %	97.925 %
130	niño	0.00363	0.082 %	0.893 %	98.007 %
131	gris	0.00363	0.082 %	0.893 %	98.089 %
132	prismáticos	0.00328	0.082 %	0.893 %	98.171 %
133	ir de escaparates	0.00328	0.082 %	0.893 %	98.253 %
134	tirantes	0.00328	0.082 %	0.893 %	98.335 %
135	cambiar	0.00320	0.165 %	1.786 %	98.500 %
136	quitarse	0.00297	0.082 %	0.893 %	98.582 %
137	mantilla	0.00297	0.082 %	0.893 %	98.664 %
138	secar	0.00297	0.082 %	0.893 %	98.746 %
139	lana	0.00297	0.082 %	0.893 %	98.828 %
140	cartera	0.00269	0.082 %	0.893 %	98.910 %
141	codo	0.00269	0.082 %	0.893 %	98.992 %
142	saco	0.00269	0.082 %	0.893 %	99.074 %
143	ancho	0.00269	0.082 %	0.893 %	99.156 %
144	probador	0.00243	0.082 %	0.893 %	99.238 %
145	plancha	0.00243	0.082 %	0.893 %	99.320 %
146	alborno	0.00220	0.082 %	0.893 %	99.402 %
147	accesorio	0.00180	0.082 %	0.893 %	99.484 %
148	deportivo	0.00180	0.082 %	0.893 %	99.566 %
149	pollera	0.00180	0.082 %	0.893 %	99.648 %
150	gafas de sol	0.00163	0.082 %	0.893 %	99.730 %
151	gustarse	0.00148	0.082 %	0.893 %	99.812 %
152	suéter	0.00148	0.082 %	0.893 %	99.894 %
153	lucirse	0.00121	0.082 %	0.893 %	99.976 %

### 03.Partes de la casa

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	cocina	0.67884	6.658 %	86.607 %	6.658 %
2	cama	0.50027	5.491 %	71.429 %	12.149 %
3	habitación	0.44431	4.324 %	56.250 %	16.473 %
4	baño	0.38849	4.187 %	54.464 %	20.660 %
5	mesa	0.32339	4.324 %	56.250 %	24.984 %
6	salón	0.25896	3.020 %	39.286 %	28.004 %
7	cuarto de baño	0.25579	2.814 %	36.607 %	30.818 %
8	silla	0.24296	3.638 %	47.321 %	34.456 %
9	cuarto	0.23249	2.677 %	34.821 %	37.133 %
10	sofá	0.22725	2.951 %	38.393 %	40.084 %
11	jardín	0.20107	2.608 %	33.929 %	42.692 %
12	puerta	0.19325	2.883 %	37.500 %	45.575 %
13	ventana	0.19066	2.814 %	36.607 %	48.389 %
14	piso	0.18289	2.334 %	30.357 %	50.723 %
15	nevera	0.14538	1.922 %	25.000 %	52.645 %
16	armario	0.13744	2.265 %	29.464 %	54.910 %

17	lámpara	0.13112	1.922 %	25.000 %	56.832 %
18	sillón	0.12938	1.784 %	23.214 %	58.616 %
19	escalera	0.12595	1.990 %	25.893 %	60.606 %
20	ducha	0.12278	2.128 %	27.679 %	62.734 %
21	garaje	0.10806	1.579 %	20.536 %	64.313 %
22	alfombra	0.10363	1.441 %	18.750 %	65.754 %
23	mesilla	0.09040	1.373 %	17.857 %	67.127 %
24	comedor	0.08833	1.167 %	15.179 %	68.294 %
25	techo	0.07276	1.167 %	15.179 %	69.461 %
26	cocinar	0.07128	1.098 %	14.286 %	70.559 %
27	televisión	0.06666	1.304 %	16.964 %	71.863 %
28	frigorífico	0.06421	0.961 %	12.500 %	72.824 %
29	horno	0.06315	0.824 %	10.714 %	73.648 %
30	mueble	0.05427	0.755 %	9.821 %	74.403 %
31	lavadora	0.05236	0.892 %	11.607 %	75.295 %
32	grande	0.04844	0.618 %	8.036 %	75.913 %
33	mesilla de noche	0.04583	0.755 %	9.821 %	76.668 %
34	cortina	0.03947	0.686 %	8.929 %	77.354 %
35	lavavajillas	0.03822	0.686 %	8.929 %	78.040 %
36	entrada	0.03525	0.686 %	8.929 %	78.726 %
37	casa	0.03490	0.275 %	3.571 %	79.001 %
38	plato	0.03482	0.618 %	8.036 %	79.619 %
39	luz	0.03324	0.618 %	8.036 %	80.237 %
40	bañera	0.03204	0.549 %	7.143 %	80.786 %
41	sala	0.03197	0.412 %	5.357 %	81.198 %
42	balcón	0.03178	0.480 %	6.250 %	81.678 %
43	dormitorio	0.03157	0.343 %	4.464 %	82.021 %
44	limpiar	0.02899	0.480 %	6.250 %	82.501 %
45	mesita	0.02813	0.412 %	5.357 %	82.913 %
46	vaso	0.02773	0.412 %	5.357 %	83.325 %
47	terracea	0.02771	0.480 %	6.250 %	83.805 %
48	dormir	0.02640	0.480 %	6.250 %	84.285 %
49	pequeño	0.02605	0.412 %	5.357 %	84.697 %
50	estantería	0.02532	0.480 %	6.250 %	85.177 %
51	trastero	0.02507	0.275 %	3.571 %	85.452 %
52	ordenador	0.02499	0.549 %	7.143 %	86.001 %
53	cuchillo	0.02476	0.343 %	4.464 %	86.344 %
54	sala de estar	0.02259	0.275 %	3.571 %	86.619 %
55	espejo	0.02240	0.343 %	4.464 %	86.962 %
56	sábana	0.02187	0.343 %	4.464 %	87.305 %
57	rincón	0.02171	0.206 %	2.679 %	87.511 %
58	comer	0.01990	0.343 %	4.464 %	87.854 %
59	pasillo	0.01974	0.275 %	3.571 %	88.129 %
60	vivir	0.01893	0.275 %	3.571 %	88.404 %
61	estudio	0.01869	0.275 %	3.571 %	88.679 %
62	lavabo	0.01860	0.206 %	2.679 %	88.885 %
63	familia	0.01806	0.343 %	4.464 %	89.228 %
64	almohada	0.01709	0.275 %	3.571 %	89.503 %
65	escritorio	0.01607	0.412 %	5.357 %	89.915 %
66	ducharse	0.01461	0.275 %	3.571 %	90.190 %
67	mesa de noche	0.01364	0.206 %	2.679 %	90.396 %
68	planta	0.01249	0.275 %	3.571 %	90.671 %
69	lavarse	0.01235	0.206 %	2.679 %	90.877 %
70	pared	0.01230	0.275 %	3.571 %	91.152 %

71	ático	0.01223	0.137 %	1.786 %	91.289 %
72	piscina	0.01216	0.206 %	2.679 %	91.495 %
73	bloque	0.01188	0.137 %	1.786 %	91.632 %
74	sala de baño	0.01023	0.137 %	1.786 %	91.769 %
75	microondas	0.01009	0.206 %	2.679 %	91.975 %
76	aseo	0.01009	0.137 %	1.786 %	92.112 %
77	cuchara	0.01009	0.137 %	1.786 %	92.249 %
78	living	0.00956	0.137 %	1.786 %	92.386 %
79	refrigerador	0.00951	0.137 %	1.786 %	92.523 %
80	taberna	0.00917	0.137 %	1.786 %	92.660 %
81	centro	0.00893	0.069 %	0.893 %	92.729 %
82	limpio	0.00893	0.069 %	0.893 %	92.798 %
83	cuarto de estar	0.00893	0.069 %	0.893 %	92.867 %
84	alquilar	0.00865	0.137 %	1.786 %	93.004 %
85	lavaplatos	0.00864	0.137 %	1.786 %	93.141 %
86	mesita de noche	0.00864	0.137 %	1.786 %	93.278 %
87	bañarse	0.00811	0.069 %	0.893 %	93.347 %
88	aparcamiento	0.00811	0.069 %	0.893 %	93.416 %
89	cocinero	0.00811	0.069 %	0.893 %	93.485 %
90	cerradura	0.00811	0.069 %	0.893 %	93.554 %
91	compartido	0.00811	0.069 %	0.893 %	93.623 %
92	tejido	0.00792	0.137 %	1.786 %	93.760 %
93	heladera	0.00786	0.137 %	1.786 %	93.897 %
94	palacio	0.00737	0.069 %	0.893 %	93.966 %
95	ropero	0.00737	0.069 %	0.893 %	94.035 %
96	pieza	0.00737	0.069 %	0.893 %	94.104 %
97	jardinero	0.00737	0.069 %	0.893 %	94.173 %
98	biblioteca	0.00737	0.069 %	0.893 %	94.242 %
99	recoger	0.00736	0.137 %	1.786 %	94.379 %
100	libro	0.00735	0.206 %	2.679 %	94.585 %
101	compañero de piso	0.00670	0.069 %	0.893 %	94.654 %
102	compra	0.00670	0.069 %	0.893 %	94.723 %
103	poner la mesa	0.00670	0.069 %	0.893 %	94.792 %
104	pavimento	0.00670	0.069 %	0.893 %	94.861 %
105	luminoso	0.00670	0.069 %	0.893 %	94.930 %
106	grifo	0.00632	0.137 %	1.786 %	95.067 %
107	funda	0.00609	0.069 %	0.893 %	95.136 %
108	cucharilla	0.00609	0.069 %	0.893 %	95.205 %
109	buhardilla	0.00609	0.069 %	0.893 %	95.274 %
110	arreglar	0.00594	0.137 %	1.786 %	95.411 %
111	corredor	0.00576	0.137 %	1.786 %	95.548 %
112	champú	0.00568	0.137 %	1.786 %	95.685 %
113	cucharón	0.00553	0.069 %	0.893 %	95.754 %
114	tablero	0.00553	0.069 %	0.893 %	95.823 %
115	cepillarse los dientes	0.00514	0.137 %	1.786 %	95.960 %
116	armario ropero	0.00502	0.069 %	0.893 %	96.029 %
117	cocina	0.00502	0.069 %	0.893 %	96.098 %
118	cuchillo	0.00502	0.069 %	0.893 %	96.167 %
119	cuadro	0.00495	0.137 %	1.786 %	96.304 %
120	acostarse	0.00470	0.137 %	1.786 %	96.441 %
121	bello	0.00457	0.069 %	0.893 %	96.510 %
122	pañuelo	0.00457	0.069 %	0.893 %	96.579 %
123	azotea	0.00457	0.069 %	0.893 %	96.717 %

124	oficina	0.00415	0.069 %	0.893 %	96.786 %
125	habitar	0.00415	0.069 %	0.893 %	96.855 %
126	cómoda	0.00415	0.069 %	0.893 %	96.924 %
127	perchero	0.00408	0.137 %	1.786 %	97.061 %
128	vela	0.00401	0.137 %	1.786 %	97.198 %
129	suelo	0.00377	0.069 %	0.893 %	97.267 %
130	sala de comer	0.00377	0.069 %	0.893 %	97.336 %
131	flexo	0.00377	0.069 %	0.893 %	97.405 %
132	juguete	0.00377	0.069 %	0.893 %	97.474 %
133	hogar	0.00377	0.069 %	0.893 %	97.543 %
134	ascensor	0.00377	0.069 %	0.893 %	97.612 %
135	comida	0.00342	0.069 %	0.893 %	97.681 %
136	aspirador	0.00342	0.069 %	0.893 %	97.750 %
137	zapatero	0.00342	0.069 %	0.893 %	97.819 %
138	secadora	0.00342	0.069 %	0.893 %	97.888 %
139	inodoro	0.00342	0.069 %	0.893 %	97.957 %
140	maravilloso	0.00342	0.069 %	0.893 %	98.026 %
141	caseta	0.00311	0.069 %	0.893 %	98.095 %
142	libertad	0.00311	0.069 %	0.893 %	98.164 %
143	bidé	0.00311	0.069 %	0.893 %	98.233 %
144	paraguas	0.00311	0.069 %	0.893 %	98.302 %
145	construir	0.00283	0.069 %	0.893 %	98.371 %
146	planchar	0.00283	0.069 %	0.893 %	98.440 %
147	césped	0.00283	0.069 %	0.893 %	98.509 %
148	salida	0.00283	0.069 %	0.893 %	98.578 %
149	gimnasio	0.00283	0.069 %	0.893 %	98.647 %
150	guardarropa	0.00283	0.069 %	0.893 %	98.716 %
151	mantel	0.00283	0.069 %	0.893 %	98.785 %
152	enchufe	0.00257	0.069 %	0.893 %	98.854 %
153	llave	0.00257	0.069 %	0.893 %	98.923 %
154	calor	0.00257	0.069 %	0.893 %	98.992 %
155	toalla	0.00233	0.069 %	0.893 %	99.061 %
156	teléfono	0.00233	0.069 %	0.893 %	99.130 %
157	ropa	0.00233	0.069 %	0.893 %	99.199 %
158	carpeta	0.00233	0.069 %	0.893 %	99.268 %
159	pizarra	0.00233	0.069 %	0.893 %	99.406 %
160	tejado	0.00233	0.069 %	0.893 %	99.475 %
161	afeitarse	0.00212	0.069 %	0.893 %	99.544 %
162	levantarse	0.00193	0.069 %	0.893 %	99.613 %
163	estante	0.00193	0.069 %	0.893 %	99.682 %
164	estudiar	0.00193	0.069 %	0.893 %	99.751 %
165	fregar	0.00175	0.069 %	0.893 %	99.820 %
166	almuerzo	0.00175	0.069 %	0.893 %	99.889 %
167	reunión de familia	0.00159	0.069 %	0.893 %	99.958 %
168	tienda	0.00108	0.069 %	0.893 %	100.027 %

#### 04.Comidas y bebidas

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	agua	0.70801	5.168 %	86.607 %	5.168 %
2	tomate	0.37043	3.410 %	57.143 %	8.578 %
3	paella	0.34369	3.090 %	51.786 %	11.668 %
4	comer	0.30036	2.291 %	38.393 %	13.959 %

5	leche	0.27700	2.451 %	41.071 %	16.410 %
6	cerveza	0.26847	2.557 %	42.857 %	18.967 %
7	naranja	0.24561	2.504 %	41.964 %	21.471 %
8	jamón	0.22322	2.344 %	39.286 %	23.815 %
9	manzana	0.21076	2.291 %	38.393 %	26.106 %
10	beber	0.20726	1.705 %	28.571 %	27.811 %
11	vino	0.19725	2.025 %	33.929 %	29.836 %
12	pescado	0.18341	1.918 %	32.143 %	31.754 %
13	bocadillo	0.16351	1.438 %	24.107 %	33.192 %
14	café	0.15689	1.492 %	25.000 %	34.684 %
15	carne	0.14190	1.598 %	26.786 %	36.282 %
16	zumos	0.13827	1.225 %	20.536 %	37.507 %
17	arroz	0.13518	1.332 %	22.321 %	38.839 %
18	pizza	0.13100	1.225 %	20.536 %	40.064 %
19	queso	0.12922	1.492 %	25.000 %	41.556 %
20	tapa	0.12919	1.225 %	20.536 %	42.781 %
21	fruta	0.12874	1.332 %	22.321 %	44.113 %
22	azúcar	0.11772	1.332 %	22.321 %	45.445 %
23	tortilla	0.11699	1.066 %	17.857 %	46.511 %
24	aceite	0.11534	1.172 %	19.643 %	47.683 %
25	gazpacho	0.11338	1.172 %	19.643 %	48.855 %
26	pasta	0.11075	1.066 %	17.857 %	49.921 %
27	sal	0.11019	1.332 %	22.321 %	51.253 %
28	sangría	0.10678	1.012 %	16.964 %	52.265 %
29	chocolate	0.10440	1.012 %	16.964 %	53.277 %
30	pan	0.10334	1.225 %	20.536 %	54.502 %
31	limón	0.09411	1.066 %	17.857 %	55.568 %
32	plátano	0.09255	1.066 %	17.857 %	56.634 %
33	galleta	0.08888	0.906 %	15.179 %	57.540 %
34	plato	0.08613	0.799 %	13.393 %	58.339 %
35	verdura	0.08542	0.906 %	15.179 %	59.245 %
36	desayuno	0.08488	0.746 %	12.500 %	59.991 %
37	patata	0.08454	1.012 %	16.964 %	61.003 %
38	zumos de naranja	0.07705	0.586 %	9.821 %	61.589 %
39	churros	0.07175	0.693 %	11.607 %	62.282 %
40	pollo	0.07139	0.799 %	13.393 %	63.081 %
41	coca-cola	0.07022	0.639 %	10.714 %	63.720 %
42	fresa	0.07021	0.906 %	15.179 %	64.626 %
43	melocotón	0.06924	0.852 %	14.286 %	65.478 %
44	cocinar	0.06762	0.533 %	8.929 %	66.011 %
45	huevo	0.06501	0.693 %	11.607 %	66.704 %
46	té	0.06317	0.639 %	10.714 %	67.343 %
47	dulce	0.06203	0.799 %	13.393 %	68.142 %
48	calabacín	0.06144	0.639 %	10.714 %	68.781 %
49	cebolla	0.06062	0.586 %	9.821 %	69.367 %
50	cena	0.06018	0.639 %	10.714 %	70.006 %
51	banana	0.05810	0.746 %	12.500 %	70.752 %
52	almuerzo	0.05728	0.533 %	8.929 %	71.285 %
53	vinagre	0.05553	0.586 %	9.821 %	71.871 %
54	sopa	0.05479	0.639 %	10.714 %	72.510 %
55	ajo	0.05214	0.533 %	8.929 %	73.043 %
56	tarta	0.04939	0.479 %	8.036 %	73.522 %
57	berenjena	0.04908	0.426 %	7.143 %	73.948 %
58	romero	0.04854	0.426 %	7.143 %	74.374 %
59	hierbabuena	0.04726	0.426 %	7.143 %	74.800 %

60	zanahoria	0.04489	0.479 %	8.036 %	75.279 %
61	tequila	0.04372	0.373 %	6.250 %	75.652 %
62	cuchillo	0.04190	0.426 %	7.143 %	76.078 %
63	agua con gas	0.03976	0.320 %	5.357 %	76.398 %
64	restaurante	0.03789	0.479 %	8.036 %	76.877 %
65	desayunar	0.03760	0.426 %	7.143 %	77.303 %
66	postre	0.03457	0.373 %	6.250 %	77.676 %
67	mantequilla	0.03448	0.426 %	7.143 %	78.102 %
68	cuchara	0.03392	0.373 %	6.250 %	78.475 %
69	taco	0.03382	0.373 %	6.250 %	78.848 %
70	pera	0.03252	0.373 %	6.250 %	79.221 %
71	burrito	0.03247	0.266 %	4.464 %	79.487 %
72	chorizo	0.03229	0.373 %	6.250 %	79.860 %
73	patatas bravas	0.03171	0.373 %	6.250 %	80.233 %
74	refresco	0.03132	0.266 %	4.464 %	80.499 %
75	embutido	0.03077	0.373 %	6.250 %	80.872 %
76	cerdo	0.02815	0.373 %	6.250 %	81.245 %
77	agua sin gas	0.02728	0.213 %	3.571 %	81.458 %
78	helado	0.02581	0.213 %	3.571 %	81.671 %
79	legumbre	0.02477	0.266 %	4.464 %	81.937 %
80	nacho	0.02449	0.160 %	2.679 %	82.097 %
81	cenar	0.02445	0.266 %	4.464 %	82.363 %
82	alcohol	0.02410	0.213 %	3.571 %	82.576 %
83	vaso	0.02318	0.213 %	3.571 %	82.789 %
84	lenteja	0.02309	0.213 %	3.571 %	83.002 %
85	comida	0.02294	0.213 %	3.571 %	83.215 %
86	mojito	0.02110	0.213 %	3.571 %	83.428 %
87	pimienta	0.02086	0.266 %	4.464 %	83.694 %
88	bebida	0.02027	0.160 %	2.679 %	83.854 %
89	lechuga	0.01960	0.266 %	4.464 %	84.120 %
90	vino tinto	0.01870	0.213 %	3.571 %	84.333 %
91	hamburguesa	0.01833	0.213 %	3.571 %	84.546 %
92	comedor	0.01786	0.107 %	1.786 %	84.653 %
93	empanada	0.01767	0.266 %	4.464 %	84.919 %
94	marisco	0.01755	0.266 %	4.464 %	85.185 %
95	piña	0.01741	0.213 %	3.571 %	85.398 %
96	cocina	0.01733	0.160 %	2.679 %	85.558 %
97	bueno	0.01727	0.107 %	1.786 %	85.665 %
98	mesa	0.01697	0.213 %	3.571 %	85.878 %
99	bar	0.01692	0.213 %	3.571 %	86.091 %
100	maíz	0.01658	0.160 %	2.679 %	86.251 %
101	café con leche	0.01614	0.107 %	1.786 %	86.358 %
102	caña	0.01532	0.160 %	2.679 %	86.518 %
103	patatas fritas	0.01513	0.213 %	3.571 %	86.731 %
104	cafetería	0.01509	0.107 %	1.786 %	86.838 %
105	sandía	0.01502	0.266 %	4.464 %	87.104 %
106	pimiento	0.01482	0.213 %	3.571 %	87.317 %
107	almorzar	0.01438	0.160 %	2.679 %	87.477 %
108	bombón	0.01422	0.160 %	2.679 %	87.637 %
109	café cortado	0.01417	0.107 %	1.786 %	87.744 %
110	quesadilla	0.01379	0.107 %	1.786 %	87.851 %
111	tinto de verano	0.01297	0.160 %	2.679 %	88.011 %
112	comprar	0.01276	0.107 %	1.786 %	88.118 %
113	cóctel	0.01263	0.107 %	1.786 %	88.225 %
114	jugo	0.01249	0.107 %	1.786 %	88.332 %



115	uva	0.01232	0.107 %	1.786 %	88.439 %
116	bizcocho	0.01221	0.213 %	3.571 %	88.652 %
117	piña colada	0.01201	0.107 %	1.786 %	88.759 %
118	judía	0.01195	0.160 %	2.679 %	88.919 %
119	hambre	0.01190	0.107 %	1.786 %	89.026 %
120	calabaza	0.01183	0.107 %	1.786 %	89.133 %
121	perejil	0.01180	0.160 %	2.679 %	89.293 %
122	frío	0.01156	0.107 %	1.786 %	89.400 %
123	gustar	0.01151	0.107 %	1.786 %	89.507 %
124	chuchería	0.01150	0.107 %	1.786 %	89.614 %
125	pastelería	0.01147	0.160 %	2.679 %	89.774 %
126	jamón serrano	0.01135	0.107 %	1.786 %	89.881 %
127	vino blanco	0.01091	0.107 %	1.786 %	89.988 %
128	alcachofa	0.01076	0.107 %	1.786 %	90.095 %
129	ensalada	0.01066	0.160 %	2.679 %	90.255 %
130	atún	0.01065	0.107 %	1.786 %	90.362 %
131	cubierto	0.01050	0.160 %	2.679 %	90.522 %
132	pescado frito	0.01049	0.107 %	1.786 %	90.629 %
133	pez	0.01007	0.107 %	1.786 %	90.736 %
134	tostada	0.00991	0.107 %	1.786 %	90.843 %
135	aceituna	0.00980	0.107 %	1.786 %	90.950 %
136	seta	0.00960	0.107 %	1.786 %	91.057 %
137	sándwich	0.00953	0.107 %	1.786 %	91.164 %
138	pepino	0.00919	0.107 %	1.786 %	91.271 %
139	pepsi	0.00897	0.107 %	1.786 %	91.378 %
140	clase	0.00893	0.053 %	0.893 %	91.431 %
141	cervecería	0.00893	0.053 %	0.893 %	91.484 %
142	tragar	0.00893	0.053 %	0.893 %	91.537 %
143	universidad	0.00834	0.053 %	0.893 %	91.590 %
144	polenta	0.00834	0.053 %	0.893 %	91.696 %
145	chiringuito	0.00834	0.053 %	0.893 %	91.749 %
146	albóndiga	0.00825	0.107 %	1.786 %	91.856 %
147	servilleta	0.00820	0.107 %	1.786 %	91.963 %
148	cereal	0.00787	0.107 %	1.786 %	92.070 %
149	lima	0.00780	0.053 %	0.893 %	92.123 %
150	biblioteca	0.00780	0.053 %	0.893 %	92.176 %
151	calamar	0.00780	0.053 %	0.893 %	92.229 %
152	gamba	0.00780	0.053 %	0.893 %	92.282 %
153	morcilla	0.00780	0.053 %	0.893 %	92.335 %
154	trago	0.00780	0.053 %	0.893 %	92.388 %
155	bebida blanca	0.00780	0.053 %	0.893 %	92.441 %
156	fideo	0.00780	0.053 %	0.893 %	92.494 %
157	naranjada	0.00780	0.053 %	0.893 %	92.547 %
158	torta	0.00780	0.053 %	0.893 %	92.600 %
159	pomo	0.00748	0.107 %	1.786 %	92.707 %
160	frito	0.00743	0.107 %	1.786 %	92.814 %
161	ciudad	0.00729	0.053 %	0.893 %	92.867 %
162	vegetariano	0.00729	0.053 %	0.893 %	92.920 %
163	macarrón	0.00729	0.053 %	0.893 %	92.973 %
164	margarita	0.00729	0.053 %	0.893 %	93.026 %
165	profesor	0.00729	0.053 %	0.893 %	93.079 %
166	mezcla	0.00729	0.053 %	0.893 %	93.132 %
167	pasta fresca	0.00729	0.053 %	0.893 %	93.185 %
168	almendra	0.00701	0.107 %	1.786 %	93.292 %
169	amigo	0.00681	0.053 %	0.893 %	93.345 %

170	jamón	0.00681	0.053 %	0.893 %	93.398 %
171	caliente	0.00681	0.053 %	0.893 %	93.451 %
172	alumno	0.00681	0.053 %	0.893 %	93.504 %
173	merienda	0.00681	0.053 %	0.893 %	93.557 %
174	paella	0.00681	0.053 %	0.893 %	93.610 %
175	enchilada	0.00681	0.053 %	0.893 %	93.663 %
176	trozo de pizza	0.00681	0.053 %	0.893 %	93.716 %
177	emborracharse	0.00637	0.053 %	0.893 %	93.769 %
178	director	0.00637	0.053 %	0.893 %	93.822 %
179	champiñón	0.00637	0.053 %	0.893 %	93.875 %
180	torreznos	0.00637	0.053 %	0.893 %	93.928 %
181	morfar	0.00637	0.053 %	0.893 %	93.981 %
182	tomar	0.00637	0.053 %	0.893 %	94.034 %
183	melón	0.00606	0.107 %	1.786 %	94.141 %
184	picaña	0.00595	0.053 %	0.893 %	94.194 %
185	curso	0.00595	0.053 %	0.893 %	94.247 %
186	fartón	0.00595	0.053 %	0.893 %	94.300 %
187	pulpo	0.00595	0.053 %	0.893 %	94.353 %
188	montadito	0.00595	0.053 %	0.893 %	94.406 %
189	cacahuete	0.00595	0.053 %	0.893 %	94.459 %
190	jugo de naranja	0.00595	0.053 %	0.893 %	94.512 %
191	agua fresca	0.00595	0.053 %	0.893 %	94.565 %
192	fanta	0.00556	0.053 %	0.893 %	94.618 %
193	supermercado	0.00556	0.053 %	0.893 %	94.671 %
194	graduación	0.00556	0.053 %	0.893 %	94.724 %
195	perrito caliente	0.00556	0.053 %	0.893 %	94.777 %
196	sed	0.00556	0.053 %	0.893 %	94.830 %
197	borraja	0.00556	0.053 %	0.893 %	94.883 %
198	dulce de leche	0.00556	0.053 %	0.893 %	94.936 %
199	toalla	0.00556	0.053 %	0.893 %	94.989 %
200	sprite	0.00520	0.053 %	0.893 %	95.042 %
201	aula	0.00520	0.053 %	0.893 %	95.095 %
202	ojo	0.00520	0.053 %	0.893 %	95.148 %
203	comilón	0.00520	0.053 %	0.893 %	95.201 %
204	coliflor	0.00520	0.053 %	0.893 %	95.254 %
205	cuba libre	0.00520	0.053 %	0.893 %	95.307 %
206	mantel	0.00486	0.053 %	0.893 %	95.360 %
207	preparar	0.00486	0.053 %	0.893 %	95.413 %
208	carne de pollo	0.00486	0.053 %	0.893 %	95.466 %
209	tener hambre	0.00486	0.053 %	0.893 %	95.519 %
210	comedor universitario	0.00486	0.053 %	0.893 %	95.572 %
211	dinero	0.00454	0.053 %	0.893 %	95.625 %
212	zumos de fruta	0.00454	0.053 %	0.893 %	95.678 %
213	aperitivo	0.00454	0.053 %	0.893 %	95.731 %
214	pincho	0.00454	0.053 %	0.893 %	95.784 %
215	bacalao	0.00454	0.053 %	0.893 %	95.837 %
216	fuerte	0.00454	0.053 %	0.893 %	95.890 %
217	agua tónica	0.00454	0.053 %	0.893 %	95.943 %
218	mate	0.00454	0.053 %	0.893 %	95.996 %
219	borracho	0.00454	0.053 %	0.893 %	96.049 %
220	zumos de piña	0.00454	0.053 %	0.893 %	96.102 %
221	lechería	0.00454	0.053 %	0.893 %	96.155 %
222	kétchup	0.00454	0.053 %	0.893 %	96.208 %
223	nevera	0.00454	0.053 %	0.893 %	96.261 %

224	guacamole	0.00453	0.107 %	1.786 %	96.368 %
225	zum de manzana	0.00424	0.053 %	0.893 %	96.421 %
226	sin sal	0.00424	0.053 %	0.893 %	96.474 %
227	picante	0.00424	0.053 %	0.893 %	96.527 %
228	bife	0.00424	0.053 %	0.893 %	96.580 %
229	vegetal	0.00424	0.053 %	0.893 %	96.633 %
230	frigorífico	0.00424	0.053 %	0.893 %	96.686 %
231	freír	0.00396	0.053 %	0.893 %	96.739 %
232	nata	0.00396	0.053 %	0.893 %	96.792 %
233	espinaca	0.00396	0.053 %	0.893 %	96.845 %
234	hacer la compra	0.00396	0.053 %	0.893 %	96.898 %
235	batata	0.00396	0.053 %	0.893 %	96.951 %
236	papaya	0.00396	0.053 %	0.893 %	97.004 %
237	carne de vaca	0.00396	0.053 %	0.893 %	97.057 %
238	gordo	0.00396	0.053 %	0.893 %	97.110 %
239	ciruela	0.00396	0.053 %	0.893 %	97.163 %
240	panadería	0.00392	0.107 %	1.786 %	97.270 %
241	cordero	0.00371	0.053 %	0.893 %	97.323 %
242	aguacate	0.00371	0.053 %	0.893 %	97.376 %
243	copa	0.00371	0.053 %	0.893 %	97.429 %
244	lasaña	0.00371	0.053 %	0.893 %	97.482 %
245	lado	0.00371	0.053 %	0.893 %	97.535 %
246	familia	0.00371	0.053 %	0.893 %	97.588 %
247	merendar	0.00371	0.053 %	0.893 %	97.641 %
248	nachos	0.00346	0.053 %	0.893 %	97.694 %
249	bistec	0.00346	0.053 %	0.893 %	97.747 %
250	chontaduro	0.00346	0.053 %	0.893 %	97.800 %
251	ordenar	0.00346	0.053 %	0.893 %	97.853 %
252	harina	0.00324	0.053 %	0.893 %	97.906 %
253	pedir	0.00324	0.053 %	0.893 %	97.959 %
254	botella	0.00324	0.053 %	0.893 %	98.012 %
255	tirabeque	0.00324	0.053 %	0.893 %	98.065 %
256	tenedor	0.00302	0.053 %	0.893 %	98.118 %
257	legumbres	0.00302	0.053 %	0.893 %	98.171 %
258	calimocho	0.00302	0.053 %	0.893 %	98.224 %
259	elegir	0.00302	0.053 %	0.893 %	98.277 %
260	caldo	0.00302	0.053 %	0.893 %	98.330 %
261	miel	0.00283	0.053 %	0.893 %	98.383 %
262	salado	0.00283	0.053 %	0.893 %	98.436 %
263	rico	0.00283	0.053 %	0.893 %	98.489 %
264	menú	0.00283	0.053 %	0.893 %	98.542 %
265	menestra	0.00283	0.053 %	0.893 %	98.595 %
266	jengibre	0.00264	0.053 %	0.893 %	98.648 %
267	vodka	0.00264	0.053 %	0.893 %	98.701 %
268	puerro	0.00247	0.053 %	0.893 %	98.754 %
269	ensaimada	0.00247	0.053 %	0.893 %	98.807 %
270	sobremesa	0.00247	0.053 %	0.893 %	98.860 %
271	pastel	0.00247	0.053 %	0.893 %	98.913 %
272	medialuna	0.00247	0.053 %	0.893 %	98.966 %
273	salsa	0.00247	0.053 %	0.893 %	99.019 %
274	cebolleta	0.00231	0.053 %	0.893 %	99.072 %
275	harina	0.00231	0.053 %	0.893 %	99.125 %
276	avellana	0.00231	0.053 %	0.893 %	99.178 %
277	hielo	0.00231	0.053 %	0.893 %	99.231 %
278	comida de verano	0.00231	0.053 %	0.893 %	99.284 %

279	bollería	0.00231	0.053 %	0.893 %	99.337 %
280	nuez	0.00231	0.053 %	0.893 %	99.390 %
281	típico	0.00216	0.053 %	0.893 %	99.443 %
282	carnicería	0.00216	0.053 %	0.893 %	99.496 %
283	fruto	0.00216	0.053 %	0.893 %	99.549 %
284	garbanzo	0.00202	0.053 %	0.893 %	99.602 %
285	yogur	0.00202	0.053 %	0.893 %	99.655 %
286	pipa	0.00202	0.053 %	0.893 %	99.708 %
287	guisante	0.00188	0.053 %	0.893 %	99.761 %
288	caramelo	0.00165	0.053 %	0.893 %	99.814 %
289	galleta salada	0.00154	0.053 %	0.893 %	99.867 %
290	frutilla	0.00144	0.053 %	0.893 %	99.920 %
291	cereza	0.00134	0.053 %	0.893 %	99.973 %

## 05.La escuela

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	profesor	0.55427	5.517 %	78.571 %	5.517 %
2	libro	0.51042	5.016 %	71.429 %	10.533 %
3	bolígrafo	0.38701	4.326 %	61.607 %	14.859 %
4	lápiz	0.37787	3.950 %	56.250 %	18.809 %
5	estudiar	0.30569	3.197 %	45.536 %	22.006 %
6	examen	0.28100	3.260 %	46.429 %	25.266 %
7	pizarra	0.27702	2.884 %	41.071 %	28.150 %
8	cuaderno	0.27032	3.009 %	42.857 %	31.159 %
9	clase	0.24886	2.508 %	35.714 %	33.667 %
10	estudiante	0.24654	2.571 %	36.607 %	36.238 %
11	alumno	0.23356	2.382 %	33.929 %	38.620 %
12	universidad	0.22088	2.069 %	29.464 %	40.689 %
13	papel	0.21254	2.508 %	35.714 %	43.197 %
14	silla	0.17081	1.567 %	22.321 %	44.764 %
15	mesa	0.15835	1.442 %	20.536 %	46.206 %
16	mochila	0.13864	1.755 %	25.000 %	47.961 %
17	escribir	0.12220	1.818 %	25.893 %	49.779 %
18	deberes	0.11350	1.317 %	18.750 %	51.096 %
19	leer	0.10206	1.317 %	18.750 %	52.413 %
20	compañero	0.09575	1.317 %	18.750 %	53.730 %
21	aula	0.09121	1.129 %	16.071 %	54.859 %
22	asignatura	0.08910	1.129 %	16.071 %	55.988 %
23	curso	0.08575	1.129 %	16.071 %	57.117 %
24	ordenador	0.07464	0.940 %	13.393 %	58.057 %
25	nota	0.07452	1.003 %	14.286 %	59.060 %
26	lengua	0.07409	1.066 %	15.179 %	60.126 %
27	apuntes	0.06131	0.878 %	12.500 %	61.004 %
28	borrador	0.05946	0.690 %	9.821 %	61.694 %
29	español	0.05773	0.627 %	8.929 %	62.321 %
30	colegio	0.05731	0.564 %	8.036 %	62.885 %
31	matemáticas	0.05560	0.752 %	10.714 %	63.637 %
32	escuchar	0.05309	0.878 %	12.500 %	64.515 %
33	aprobar	0.05276	0.878 %	12.500 %	65.393 %
34	lección	0.04972	0.752 %	10.714 %	66.145 %
35	aprender	0.04714	0.564 %	8.036 %	66.709 %
36	sacapuntas	0.04223	0.502 %	7.143 %	67.211 %
37	tijeras	0.04059	0.502 %	7.143 %	67.713 %

38	profesorado	0.03814	0.376 %	5.357 %	68.089 %
39	escritorio	0.03806	0.376 %	5.357 %	68.465 %
40	gimnasio	0.03509	0.502 %	7.143 %	68.967 %
41	historia	0.03440	0.502 %	7.143 %	69.469 %
42	inglés	0.03338	0.439 %	6.250 %	69.908 %
43	solapamiento	0.03326	0.376 %	5.357 %	70.284 %
44	biblioteca	0.03316	0.376 %	5.357 %	70.660 %
45	horario	0.03244	0.439 %	6.250 %	71.099 %
46	amigo	0.03119	0.439 %	6.250 %	71.538 %
47	pupitre	0.03042	0.313 %	4.464 %	71.851 %
48	estudiantado	0.03011	0.313 %	4.464 %	72.164 %
49	ciencias	0.02965	0.439 %	6.250 %	72.603 %
50	goma de borrar	0.02951	0.376 %	5.357 %	72.979 %
51	ejercicio	0.02890	0.439 %	6.250 %	73.418 %
52	idiomas	0.02797	0.439 %	6.250 %	73.857 %
53	italiano	0.02589	0.251 %	3.571 %	74.108 %
54	escuela	0.02583	0.251 %	3.571 %	74.359 %
55	subrayar	0.02514	0.439 %	6.250 %	74.798 %
56	facultad	0.02511	0.313 %	4.464 %	75.111 %
57	estudio	0.02471	0.313 %	4.464 %	75.424 %
58	tiza	0.02348	0.439 %	6.250 %	75.863 %
59	hablar	0.02285	0.439 %	6.250 %	76.302 %
60	autobús	0.02246	0.313 %	4.464 %	76.615 %
61	hoja	0.02225	0.251 %	3.571 %	76.866 %
62	edificio	0.02210	0.251 %	3.571 %	77.117 %
63	asistir	0.02121	0.251 %	3.571 %	77.368 %
64	alumnado	0.02061	0.251 %	3.571 %	77.619 %
65	gafas	0.02057	0.188 %	2.679 %	77.807 %
66	suspender	0.01989	0.376 %	5.357 %	78.183 %
67	prueba	0.01953	0.376 %	5.357 %	78.559 %
68	enseñanza	0.01950	0.251 %	3.571 %	78.810 %
69	repasar	0.01928	0.188 %	2.679 %	78.998 %
70	repetir	0.01918	0.313 %	4.464 %	79.311 %
71	regla	0.01868	0.251 %	3.571 %	79.562 %
72	ir a clase	0.01865	0.188 %	2.679 %	79.750 %
73	baño	0.01863	0.188 %	2.679 %	79.938 %
74	fotocopia	0.01863	0.251 %	3.571 %	80.189 %
75	carrera	0.01786	0.125 %	1.786 %	80.314 %
76	director	0.01730	0.188 %	2.679 %	80.502 %
77	prueba oral	0.01725	0.251 %	3.571 %	80.753 %
78	borrar	0.01723	0.251 %	3.571 %	81.004 %
79	enseñanza primaria	0.01650	0.125 %	1.786 %	81.129 %
80	enseñar	0.01643	0.188 %	2.679 %	81.317 %
81	tren	0.01636	0.188 %	2.679 %	81.505 %
82	pluma	0.01622	0.188 %	2.679 %	81.693 %
83	francés	0.01577	0.188 %	2.679 %	81.881 %
84	licenciar	0.01554	0.313 %	4.464 %	82.194 %
85	computador	0.01536	0.125 %	1.786 %	82.319 %
86	tutoría	0.01522	0.188 %	2.679 %	82.507 %
87	prueba escrita	0.01493	0.251 %	3.571 %	82.758 %
88	geografía	0.01436	0.188 %	2.679 %	82.946 %
89	licenciatura	0.01401	0.188 %	2.679 %	83.134 %
90	maestro	0.01400	0.125 %	1.786 %	83.259 %
91	jefe	0.01286	0.125 %	1.786 %	83.384 %
92	merienda	0.01271	0.188 %	2.679 %	83.572 %

93	buena nota	0.01249	0.125 %	1.786 %	83.697 %
94	puerta	0.01245	0.188 %	2.679 %	83.885 %
95	papelera	0.01239	0.188 %	2.679 %	84.073 %
96	pintar	0.01228	0.188 %	2.679 %	84.261 %
97	agenda	0.01220	0.125 %	1.786 %	84.386 %
98	rotulador	0.01212	0.188 %	2.679 %	84.574 %
99	jardín	0.01189	0.188 %	2.679 %	84.762 %
100	número	0.01184	0.125 %	1.786 %	84.887 %
101	lápiz de color	0.01184	0.125 %	1.786 %	85.012 %
102	comer	0.01169	0.188 %	2.679 %	85.200 %
103	calculadora	0.01145	0.125 %	1.786 %	85.325 %
104	examinar	0.01106	0.125 %	1.786 %	85.450 %
105	laboratorio	0.01106	0.125 %	1.786 %	85.575 %
106	bolsa	0.01099	0.188 %	2.679 %	85.763 %
107	padres	0.01095	0.125 %	1.786 %	85.888 %
108	hora	0.01091	0.125 %	1.786 %	86.013 %
109	saber	0.01091	0.125 %	1.786 %	86.138 %
110	cafetería	0.01055	0.125 %	1.786 %	86.263 %
111	estuche	0.01050	0.251 %	3.571 %	86.514 %
112	tarea	0.01046	0.188 %	2.679 %	86.702 %
113	diario	0.01039	0.125 %	1.786 %	86.827 %
114	marcador	0.01018	0.125 %	1.786 %	86.952 %
115	café	0.00965	0.125 %	1.786 %	87.077 %
116	cinta	0.00947	0.188 %	2.679 %	87.265 %
117	lectura	0.00938	0.125 %	1.786 %	87.390 %
118	atención	0.00938	0.125 %	1.786 %	87.515 %
119	suerte	0.00929	0.125 %	1.786 %	87.640 %
120	grabador	0.00926	0.125 %	1.786 %	87.765 %
121	infancia	0.00893	0.063 %	0.893 %	87.828 %
122	salida	0.00893	0.063 %	0.893 %	87.891 %
123	dibujar	0.00879	0.125 %	1.786 %	88.016 %
124	color	0.00879	0.125 %	1.786 %	88.141 %
125	pausa	0.00871	0.125 %	1.786 %	88.266 %
126	organización	0.00864	0.125 %	1.786 %	88.391 %
127	maleta	0.00853	0.125 %	1.786 %	88.516 %
128	conocer	0.00836	0.125 %	1.786 %	88.641 %
129	librería	0.00836	0.125 %	1.786 %	88.766 %
130	semestre	0.00828	0.125 %	1.786 %	88.891 %
131	hoja de papel	0.00825	0.125 %	1.786 %	89.016 %
132	profesión	0.00822	0.063 %	0.893 %	89.079 %
133	puesto	0.00822	0.063 %	0.893 %	89.142 %
134	enseñanza media	0.00822	0.063 %	0.893 %	89.205 %
135	ausente	0.00822	0.063 %	0.893 %	89.268 %
136	despacho	0.00819	0.125 %	1.786 %	89.393 %
137	escalera	0.00819	0.125 %	1.786 %	89.518 %
138	apuntar	0.00809	0.125 %	1.786 %	89.643 %
139	página	0.00806	0.125 %	1.786 %	89.768 %
140	literatura	0.00796	0.125 %	1.786 %	89.893 %
141	matricularse	0.00788	0.125 %	1.786 %	90.018 %
142	graduación	0.00785	0.125 %	1.786 %	90.143 %
143	matriculación	0.00785	0.125 %	1.786 %	90.268 %
144	pantalla	0.00758	0.063 %	0.893 %	90.331 %
145	mala nota	0.00758	0.063 %	0.893 %	90.394 %
146	empleado	0.00758	0.063 %	0.893 %	90.457 %
147	retraso	0.00758	0.063 %	0.893 %	90.520 %

148	ansiedad	0.00698	0.063 %	0.893 %	90.583 %
149	reposo	0.00698	0.063 %	0.893 %	90.646 %
150	explicar	0.00698	0.063 %	0.893 %	90.709 %
151	futuro	0.00698	0.063 %	0.893 %	90.772 %
152	corrector	0.00698	0.063 %	0.893 %	90.898 %
153	instituto	0.00687	0.125 %	1.786 %	91.023 %
154	miedo	0.00669	0.125 %	1.786 %	91.148 %
155	rector	0.00653	0.125 %	1.786 %	91.273 %
156	oral	0.00644	0.125 %	1.786 %	91.398 %
157	mapa	0.00643	0.063 %	0.893 %	91.461 %
158	colores	0.00643	0.063 %	0.893 %	91.524 %
159	secretario	0.00643	0.063 %	0.893 %	91.587 %
160	beca Erasmus	0.00643	0.063 %	0.893 %	91.713 %
161	sala	0.00643	0.063 %	0.893 %	91.776 %
162	bolso	0.00643	0.063 %	0.893 %	91.839 %
163	oportunidad	0.00643	0.063 %	0.893 %	91.902 %
164	institución	0.00643	0.063 %	0.893 %	91.965 %
165	tomar notas	0.00633	0.125 %	1.786 %	92.090 %
166	dinero	0.00614	0.125 %	1.786 %	92.215 %
167	sacar notas	0.00594	0.125 %	1.786 %	92.340 %
168	caso	0.00592	0.063 %	0.893 %	92.403 %
169	cuadro	0.00592	0.063 %	0.893 %	92.466 %
170	oficina de los estudiantes	0.00592	0.063 %	0.893 %	92.529 %
171	texto	0.00592	0.063 %	0.893 %	92.592 %
172	huelga	0.00592	0.063 %	0.893 %	92.655 %
173	novela	0.00583	0.125 %	1.786 %	92.780 %
174	luz	0.00545	0.063 %	0.893 %	92.843 %
175	tiempo	0.00545	0.063 %	0.893 %	92.906 %
176	lenguas	0.00545	0.063 %	0.893 %	92.969 %
177	tomar apuntes	0.00545	0.063 %	0.893 %	93.032 %
178	trabajo	0.00545	0.063 %	0.893 %	93.095 %
179	portugués	0.00545	0.063 %	0.893 %	93.158 %
180	dar clases	0.00545	0.063 %	0.893 %	93.221 %
181	chico	0.00545	0.063 %	0.893 %	93.284 %
182	física	0.00545	0.063 %	0.893 %	93.347 %
183	bocadillo	0.00545	0.063 %	0.893 %	93.410 %
184	primero	0.00545	0.063 %	0.893 %	93.473 %
185	ir a la escuela	0.00545	0.063 %	0.893 %	93.536 %
186	correo electrónico	0.00545	0.063 %	0.893 %	93.599 %
187	pensar	0.00502	0.063 %	0.893 %	93.662 %
188	aulas	0.00502	0.063 %	0.893 %	93.725 %
189	corregir	0.00502	0.063 %	0.893 %	93.788 %
190	dirigente	0.00502	0.063 %	0.893 %	93.851 %
191	físico	0.00502	0.063 %	0.893 %	93.914 %
192	biología	0.00502	0.063 %	0.893 %	93.977 %
193	lenguas extranjeras	0.00502	0.063 %	0.893 %	94.040 %
194	dictado	0.00500	0.125 %	1.786 %	94.165 %
195	latín	0.00481	0.125 %	1.786 %	94.290 %
196	guardería	0.00463	0.063 %	0.893 %	94.353 %
197	departamento	0.00463	0.063 %	0.893 %	94.416 %
198	cera	0.00463	0.063 %	0.893 %	94.479 %
199	test	0.00463	0.063 %	0.893 %	94.542 %
200	horario académico	0.00463	0.063 %	0.893 %	94.605 %
201	madrugada	0.00463	0.063 %	0.893 %	94.668 %

202	piso	0.00463	0.063 %	0.893 %	94.731 %
203	aprobación	0.00463	0.063 %	0.893 %	94.794 %
204	grupo	0.00463	0.063 %	0.893 %	94.857 %
205	cuarto de baño	0.00426	0.063 %	0.893 %	94.920 %
206	idiomas extranjeros	0.00426	0.063 %	0.893 %	94.983 %
207	merendar	0.00426	0.063 %	0.893 %	95.046 %
208	deporte	0.00426	0.063 %	0.893 %	95.109 %
209	practicar	0.00426	0.063 %	0.893 %	95.172 %
210	sacar	0.00426	0.063 %	0.893 %	95.235 %
211	reloj	0.00426	0.063 %	0.893 %	95.298 %
212	sacar apuntes	0.00426	0.063 %	0.893 %	95.361 %
213	inscribirse	0.00426	0.063 %	0.893 %	95.424 %
214	niño	0.00426	0.063 %	0.893 %	95.487 %
215	alemán	0.00426	0.063 %	0.893 %	95.550 %
216	corrector blanco	0.00394	0.125 %	1.786 %	95.675 %
217	mañana	0.00393	0.063 %	0.893 %	95.738 %
218	aburrido	0.00393	0.063 %	0.893 %	95.801 %
219	cantina	0.00393	0.063 %	0.893 %	95.864 %
220	escolar	0.00393	0.063 %	0.893 %	95.927 %
221	poema	0.00393	0.063 %	0.893 %	95.990 %
222	encerado	0.00393	0.063 %	0.893 %	96.053 %
223	preparación	0.00393	0.063 %	0.893 %	96.116 %
224	tabla	0.00393	0.063 %	0.893 %	96.179 %
225	ver	0.00393	0.063 %	0.893 %	96.242 %
226	extranjero	0.00393	0.063 %	0.893 %	96.305 %
227	coche	0.00393	0.063 %	0.893 %	96.368 %
228	evaluación	0.00393	0.063 %	0.893 %	96.431 %
229	suspensión	0.00393	0.063 %	0.893 %	96.494 %
230	tema	0.00393	0.063 %	0.893 %	96.557 %
231	muy temprano	0.00362	0.063 %	0.893 %	96.620 %
232	viaje	0.00362	0.063 %	0.893 %	96.683 %
233	afilalápices	0.00362	0.063 %	0.893 %	96.746 %
234	amistad	0.00362	0.063 %	0.893 %	96.809 %
235	sueño	0.00362	0.063 %	0.893 %	96.872 %
236	imprimir	0.00362	0.063 %	0.893 %	96.935 %
237	intercambio cultural	0.00362	0.063 %	0.893 %	96.998 %
238	bicicleta	0.00362	0.063 %	0.893 %	97.061 %
239	historia de lengua	0.00362	0.063 %	0.893 %	97.124 %
240	comprender	0.00362	0.063 %	0.893 %	97.187 %
241	boletín de notas	0.00362	0.063 %	0.893 %	97.250 %
242	recreo	0.00362	0.063 %	0.893 %	97.313 %
243	vacación	0.00333	0.063 %	0.893 %	97.376 %
244	fin de semana	0.00333	0.063 %	0.893 %	97.439 %
245	papelería	0.00333	0.063 %	0.893 %	97.502 %
246	campo de fútbol	0.00333	0.063 %	0.893 %	97.565 %
247	ayudarse	0.00333	0.063 %	0.893 %	97.628 %
248	práctico	0.00333	0.063 %	0.893 %	97.691 %
249	importante	0.00333	0.063 %	0.893 %	97.754 %
250	cultura	0.00333	0.063 %	0.893 %	97.817 %
251	traducción	0.00333	0.063 %	0.893 %	97.880 %
252	divertimento	0.00307	0.063 %	0.893 %	97.943 %
253	grabación	0.00307	0.063 %	0.893 %	98.006 %
254	público	0.00307	0.063 %	0.893 %	98.069 %
255	vivir	0.00307	0.063 %	0.893 %	98.132 %



256	completar	0.00307	0.063 %	0.893 %	98.195 %
257	música	0.00283	0.063 %	0.893 %	98.258 %
258	graduarse	0.00283	0.063 %	0.893 %	98.321 %
259	piscina	0.00283	0.063 %	0.893 %	98.384 %
260	privado	0.00283	0.063 %	0.893 %	98.447 %
261	país	0.00283	0.063 %	0.893 %	98.510 %
262	tomar	0.00283	0.063 %	0.893 %	98.573 %
263	léxico	0.00283	0.063 %	0.893 %	98.636 %
264	enchufe	0.00283	0.063 %	0.893 %	98.699 %
265	escrito	0.00260	0.063 %	0.893 %	98.762 %
266	dibujo	0.00260	0.063 %	0.893 %	98.825 %
267	internet	0.00260	0.063 %	0.893 %	98.888 %
268	liceo	0.00260	0.063 %	0.893 %	99.014 %
269	septiembre	0.00240	0.063 %	0.893 %	99.077 %
270	bachillerato	0.00240	0.063 %	0.893 %	99.140 %
271	madrugar	0.00240	0.063 %	0.893 %	99.203 %
272	micrófono	0.00240	0.063 %	0.893 %	99.266 %
273	rechazar	0.00221	0.063 %	0.893 %	99.329 %
274	experiencia	0.00204	0.063 %	0.893 %	99.392 %
275	estudios religiosos	0.00204	0.063 %	0.893 %	99.455 %
276	cansancio	0.00204	0.063 %	0.893 %	99.518 %
277	agua	0.00187	0.063 %	0.893 %	99.581 %
278	arte	0.00187	0.063 %	0.893 %	99.644 %
279	tiempo libre	0.00187	0.063 %	0.893 %	99.707 %
280	informática	0.00173	0.063 %	0.893 %	99.770 %
281	salir	0.00159	0.063 %	0.893 %	99.833 %
282	soporte electrónico	0.00147	0.063 %	0.893 %	99.896 %
283	empezar	0.00147	0.063 %	0.893 %	99.959 %
284	pegamento	0.00124	0.063 %	0.893 %	100.022 %

## 06.La ciudad

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	calle	0.52362	4.532 %	66.964 %	4.532 %
2	coche	0.45722	4.834 %	71.429 %	9.366 %
3	tienda	0.30967	3.263 %	48.214 %	12.629 %
4	casa	0.30532	2.900 %	42.857 %	15.529 %
5	plaza	0.30073	2.719 %	40.179 %	18.248 %
6	parque	0.29076	3.021 %	44.643 %	21.269 %
7	autobús	0.26120	2.961 %	43.750 %	24.230 %
8	restaurante	0.24682	2.779 %	41.071 %	27.009 %
9	iglesia	0.24292	2.477 %	36.607 %	29.486 %
10	edificio	0.19744	1.752 %	25.893 %	31.238 %
11	bar	0.18310	2.417 %	35.714 %	33.655 %
12	escuela	0.17992	2.236 %	33.036 %	35.891 %
13	carretera	0.16380	1.511 %	22.321 %	37.402 %
14	piso	0.14017	1.329 %	19.643 %	38.731 %
15	gente	0.12811	1.631 %	24.107 %	40.362 %
16	monumento	0.12539	1.269 %	18.750 %	41.631 %
17	estación	0.12441	1.813 %	26.786 %	43.444 %
18	bicicleta	0.11007	1.390 %	20.536 %	44.834 %
19	centro	0.10257	0.846 %	12.500 %	45.680 %
20	tren	0.10230	1.208 %	17.857 %	46.888 %
21	cine	0.09985	1.208 %	17.857 %	48.096 %

22	farmacia	0.09826	1.027 %	15.179 %	49.123 %
23	supermercado	0.09590	1.088 %	16.071 %	50.211 %
24	ayuntamiento	0.09487	0.906 %	13.393 %	51.117 %
25	museo	0.09392	1.027 %	15.179 %	52.144 %
26	barrio	0.09249	0.785 %	11.607 %	52.929 %
27	árbol	0.09082	1.208 %	17.857 %	54.137 %
28	universidad	0.09070	1.269 %	18.750 %	55.406 %
29	semáforo	0.08627	0.906 %	13.393 %	56.312 %
30	jardín	0.08083	0.906 %	13.393 %	57.218 %
31	tráfico	0.07316	0.846 %	12.500 %	58.064 %
32	hospital	0.07287	1.088 %	16.071 %	59.152 %
33	cruzar	0.07248	0.665 %	9.821 %	59.817 %
34	tranvía	0.07061	0.906 %	13.393 %	60.723 %
35	banco	0.05932	0.785 %	11.607 %	61.508 %
36	palacio	0.05314	0.544 %	8.036 %	62.052 %
37	persona	0.05228	0.604 %	8.929 %	62.656 %
38	pasear	0.05062	0.785 %	11.607 %	63.441 %
39	panadería	0.04892	0.544 %	8.036 %	63.985 %
40	visitar	0.04831	0.604 %	8.929 %	64.589 %
41	rascacielos	0.04725	0.423 %	6.250 %	65.012 %
42	luz	0.04688	0.483 %	7.143 %	65.495 %
43	policía	0.04594	0.725 %	10.714 %	66.220 %
44	metropolitano	0.04587	0.604 %	8.929 %	66.824 %
45	avenida	0.04496	0.544 %	8.036 %	67.368 %
46	taxi	0.04458	0.544 %	8.036 %	67.912 %
47	cruce	0.03958	0.363 %	5.357 %	68.275 %
48	catedral	0.03955	0.363 %	5.357 %	68.638 %
49	aeropuerto	0.03881	0.604 %	8.929 %	69.242 %
50	trabajo	0.03692	0.423 %	6.250 %	69.665 %
51	parada de autobús	0.03588	0.483 %	7.143 %	70.148 %
52	biblioteca	0.03572	0.423 %	6.250 %	70.571 %
53	turista	0.03220	0.363 %	5.357 %	70.934 %
54	vía	0.03166	0.363 %	5.357 %	71.297 %
55	derecha	0.03166	0.302 %	4.464 %	71.599 %
56	izquierda	0.03134	0.302 %	4.464 %	71.901 %
57	niño	0.03122	0.363 %	5.357 %	72.264 %
58	girar	0.03104	0.302 %	4.464 %	72.566 %
59	municipio	0.03096	0.302 %	4.464 %	72.868 %
60	paso de cebra	0.03077	0.363 %	5.357 %	73.231 %
61	ciudadano	0.02984	0.423 %	6.250 %	73.654 %
62	mercado	0.02967	0.544 %	8.036 %	74.198 %
63	discoteca	0.02953	0.544 %	8.036 %	74.742 %
64	carro	0.02941	0.363 %	5.357 %	75.105 %
65	estación de tren	0.02682	0.363 %	5.357 %	75.468 %
66	trabajar	0.02608	0.363 %	5.357 %	75.831 %
67	estatua	0.02560	0.302 %	4.464 %	76.133 %
68	vivir	0.02506	0.363 %	5.357 %	76.496 %
69	medio de transporte	0.02493	0.363 %	5.357 %	76.859 %
70	cafetería	0.02454	0.242 %	3.571 %	77.101 %
71	estadio	0.02332	0.302 %	4.464 %	77.403 %
72	seguir	0.02329	0.242 %	3.571 %	77.645 %
73	negocio	0.02327	0.302 %	4.464 %	77.947 %
74	centro comercial	0.02309	0.242 %	3.571 %	78.189 %
75	paseo	0.02229	0.363 %	5.357 %	78.552 %
76	centro histórico	0.02198	0.242 %	3.571 %	78.794 %

77	parada	0.02182	0.302 %	4.464 %	79.096 %
78	cielo	0.02104	0.242 %	3.571 %	79.338 %
79	grande	0.02062	0.242 %	3.571 %	79.580 %
80	andar	0.02006	0.302 %	4.464 %	79.882 %
81	estanco	0.01952	0.302 %	4.464 %	80.184 %
82	compra	0.01892	0.181 %	2.679 %	80.365 %
83	papelería	0.01892	0.181 %	2.679 %	80.546 %
84	playa	0.01832	0.242 %	3.571 %	80.788 %
85	chiringuito	0.01805	0.181 %	2.679 %	80.969 %
86	hotel	0.01804	0.242 %	3.571 %	81.211 %
87	caminar	0.01790	0.242 %	3.571 %	81.453 %
88	oficina	0.01758	0.181 %	2.679 %	81.634 %
89	transporte	0.01707	0.181 %	2.679 %	81.815 %
90	punto	0.01672	0.242 %	3.571 %	82.057 %
91	viajar	0.01563	0.242 %	3.571 %	82.299 %
92	recto	0.01543	0.121 %	1.786 %	82.420 %
93	tapa	0.01529	0.181 %	2.679 %	82.601 %
94	paso de peatones	0.01529	0.121 %	1.786 %	82.722 %
95	doblar	0.01529	0.121 %	1.786 %	82.843 %
96	pueblo	0.01526	0.242 %	3.571 %	83.085 %
97	pequeño	0.01517	0.242 %	3.571 %	83.327 %
98	dar un paseo	0.01502	0.242 %	3.571 %	83.569 %
99	piscina	0.01461	0.181 %	2.679 %	83.750 %
100	ruido	0.01455	0.242 %	3.571 %	83.992 %
101	teatro	0.01409	0.242 %	3.571 %	84.234 %
102	capital	0.01384	0.181 %	2.679 %	84.415 %
103	música	0.01383	0.242 %	3.571 %	84.657 %
104	peluquería	0.01376	0.302 %	4.464 %	84.959 %
105	automóvil	0.01366	0.121 %	1.786 %	85.080 %
106	historia	0.01274	0.121 %	1.786 %	85.201 %
107	avión	0.01237	0.181 %	2.679 %	85.382 %
108	polución	0.01221	0.181 %	2.679 %	85.563 %
109	seguir recto	0.01198	0.121 %	1.786 %	85.684 %
110	rotonda	0.01198	0.121 %	1.786 %	85.805 %
111	motocicleta	0.01180	0.302 %	4.464 %	86.107 %
112	pastelería	0.01169	0.121 %	1.786 %	86.228 %
113	cauce	0.01107	0.121 %	1.786 %	86.349 %
114	tienda de ropa	0.01107	0.121 %	1.786 %	86.470 %
115	extranjero	0.01071	0.181 %	2.679 %	86.651 %
116	colegio	0.01056	0.121 %	1.786 %	86.772 %
117	planta	0.01034	0.121 %	1.786 %	86.893 %
118	señal	0.01028	0.121 %	1.786 %	87.014 %
119	vacación	0.01026	0.121 %	1.786 %	87.135 %
120	río	0.01025	0.242 %	3.571 %	87.377 %
121	cosmopolita	0.01023	0.121 %	1.786 %	87.498 %
122	residencia	0.00991	0.181 %	2.679 %	87.679 %
123	alcalde	0.00989	0.181 %	2.679 %	87.860 %
124	comprar	0.00974	0.121 %	1.786 %	87.981 %
125	lugar	0.00959	0.121 %	1.786 %	88.102 %
126	turismo	0.00950	0.121 %	1.786 %	88.223 %
127	librería	0.00936	0.121 %	1.786 %	88.344 %
128	apartamento	0.00928	0.121 %	1.786 %	88.465 %
129	plaza mayor	0.00893	0.060 %	0.893 %	88.525 %
130	arena	0.00893	0.060 %	0.893 %	88.585 %
131	ambiente	0.00893	0.060 %	0.893 %	88.645 %

132	puerta de la ciudad	0.00893	0.060 %	0.893 %	88.705 %
133	glorieta	0.00893	0.060 %	0.893 %	88.765 %
134	pescadería	0.00886	0.121 %	1.786 %	88.886 %
135	mar	0.00886	0.121 %	1.786 %	89.007 %
136	amigo	0.00873	0.121 %	1.786 %	89.128 %
137	cultura	0.00857	0.121 %	1.786 %	89.249 %
138	paisaje	0.00847	0.121 %	1.786 %	89.370 %
139	albergue	0.00841	0.121 %	1.786 %	89.491 %
140	deporte	0.00832	0.121 %	1.786 %	89.612 %
141	cartel	0.00832	0.121 %	1.786 %	89.733 %
142	campo	0.00831	0.121 %	1.786 %	89.854 %
143	ir a dar un paseo	0.00825	0.060 %	0.893 %	89.914 %
144	multicultural	0.00825	0.060 %	0.893 %	90.034 %
145	contaminación	0.00825	0.060 %	0.893 %	90.094 %
146	billete	0.00825	0.060 %	0.893 %	90.154 %
147	perro	0.00810	0.121 %	1.786 %	90.275 %
148	viaje	0.00810	0.121 %	1.786 %	90.396 %
149	casco antiguo	0.00807	0.121 %	1.786 %	90.517 %
150	país	0.00782	0.121 %	1.786 %	90.638 %
151	mirador	0.00762	0.060 %	0.893 %	90.698 %
152	escalera	0.00762	0.060 %	0.893 %	90.758 %
153	servicio	0.00762	0.060 %	0.893 %	90.818 %
154	periferia	0.00762	0.060 %	0.893 %	90.878 %
155	ciclovía	0.00762	0.060 %	0.893 %	90.938 %
156	muro	0.00704	0.060 %	0.893 %	90.998 %
157	pizza	0.00704	0.060 %	0.893 %	91.058 %
158	moderno	0.00704	0.060 %	0.893 %	91.118 %
159	patio	0.00704	0.060 %	0.893 %	91.178 %
160	ropa	0.00704	0.060 %	0.893 %	91.238 %
161	famoso	0.00704	0.060 %	0.893 %	91.298 %
162	suburbio	0.00704	0.060 %	0.893 %	91.358 %
163	camino	0.00704	0.060 %	0.893 %	91.418 %
164	noche	0.00704	0.060 %	0.893 %	91.478 %
165	estado	0.00704	0.060 %	0.893 %	91.538 %
166	aventón	0.00704	0.060 %	0.893 %	91.598 %
167	bosque	0.00704	0.060 %	0.893 %	91.658 %
168	estación de autobús	0.00698	0.121 %	1.786 %	91.779 %
169	ir de compras	0.00676	0.121 %	1.786 %	91.900 %
170	descubrir	0.00669	0.121 %	1.786 %	92.021 %
171	ir	0.00655	0.121 %	1.786 %	92.142 %
172	fin de semana	0.00650	0.060 %	0.893 %	92.202 %
173	sitio	0.00650	0.060 %	0.893 %	92.262 %
174	sueño	0.00650	0.060 %	0.893 %	92.322 %
175	empleo	0.00650	0.060 %	0.893 %	92.382 %
176	tequila	0.00650	0.060 %	0.893 %	92.442 %
177	bloque de pisos	0.00650	0.060 %	0.893 %	92.502 %
178	centro recreativo	0.00650	0.060 %	0.893 %	92.562 %
179	autopista	0.00650	0.060 %	0.893 %	92.622 %
180	club	0.00650	0.060 %	0.893 %	92.682 %
181	oficio	0.00650	0.060 %	0.893 %	92.742 %
182	rueda	0.00650	0.060 %	0.893 %	92.802 %
183	evento	0.00637	0.121 %	1.786 %	92.923 %
184	naturaleza	0.00636	0.121 %	1.786 %	93.044 %
185	fuelle	0.00618	0.121 %	1.786 %	93.165 %
186	dar una vuelta	0.00616	0.121 %	1.786 %	93.286 %

187	esquina	0.00601	0.060 %	0.893 %	93.346 %
188	cliente	0.00601	0.060 %	0.893 %	93.466 %
189	peligro	0.00601	0.060 %	0.893 %	93.526 %
190	fiesta	0.00601	0.060 %	0.893 %	93.586 %
191	cruzar la calle	0.00555	0.060 %	0.893 %	93.646 %
192	conducir	0.00555	0.060 %	0.893 %	93.706 %
193	fútbol	0.00555	0.060 %	0.893 %	93.766 %
194	dinero	0.00555	0.060 %	0.893 %	93.826 %
195	banqueta	0.00555	0.060 %	0.893 %	93.886 %
196	oportunidad	0.00555	0.060 %	0.893 %	93.946 %
197	criminalidad	0.00555	0.060 %	0.893 %	94.006 %
198	comedor	0.00555	0.060 %	0.893 %	94.066 %
199	antiguo	0.00512	0.060 %	0.893 %	94.126 %
200	parar	0.00512	0.060 %	0.893 %	94.186 %
201	habitación	0.00512	0.060 %	0.893 %	94.246 %
202	auditorio	0.00512	0.060 %	0.893 %	94.306 %
203	peatón	0.00512	0.060 %	0.893 %	94.366 %
204	jardín público	0.00512	0.060 %	0.893 %	94.426 %
205	correr	0.00508	0.121 %	1.786 %	94.547 %
206	cuadro	0.00473	0.060 %	0.893 %	94.607 %
207	cervecería	0.00473	0.060 %	0.893 %	94.667 %
208	pavimento	0.00473	0.060 %	0.893 %	94.727 %
209	voltar	0.00473	0.060 %	0.893 %	94.787 %
210	principal	0.00473	0.060 %	0.893 %	94.847 %
211	velocidad	0.00473	0.060 %	0.893 %	94.907 %
212	duende	0.00473	0.060 %	0.893 %	94.967 %
213	mochila	0.00473	0.060 %	0.893 %	95.027 %
214	política	0.00473	0.060 %	0.893 %	95.087 %
215	verano	0.00473	0.060 %	0.893 %	95.147 %
216	carrera	0.00473	0.060 %	0.893 %	95.207 %
217	habitante	0.00473	0.060 %	0.893 %	95.267 %
218	ciclomotor	0.00437	0.060 %	0.893 %	95.327 %
219	suelo	0.00437	0.060 %	0.893 %	95.387 %
220	sitio turístico	0.00437	0.060 %	0.893 %	95.447 %
221	complejo polideportivo	0.00437	0.060 %	0.893 %	95.507 %
222	perfume	0.00437	0.060 %	0.893 %	95.567 %
223	cruz	0.00437	0.060 %	0.893 %	95.627 %
224	centro cultural	0.00437	0.060 %	0.893 %	95.687 %
225	flor	0.00434	0.121 %	1.786 %	95.808 %
226	lugar de interés	0.00404	0.060 %	0.893 %	95.868 %
227	poblado	0.00404	0.060 %	0.893 %	95.928 %
228	café	0.00404	0.060 %	0.893 %	95.988 %
229	manifestación	0.00404	0.060 %	0.893 %	96.048 %
230	alrededor	0.00404	0.060 %	0.893 %	96.108 %
231	pan	0.00404	0.060 %	0.893 %	96.168 %
232	ciudad	0.00404	0.060 %	0.893 %	96.228 %
233	diferente	0.00373	0.060 %	0.893 %	96.348 %
234	enseñanza media	0.00373	0.060 %	0.893 %	96.408 %
235	aparcamiento	0.00352	0.121 %	1.786 %	96.529 %
236	pub	0.00345	0.060 %	0.893 %	96.589 %
237	provincia	0.00345	0.060 %	0.893 %	96.649 %
238	clima	0.00345	0.060 %	0.893 %	96.709 %
239	guía turística	0.00345	0.060 %	0.893 %	96.769 %
240	bombero	0.00345	0.060 %	0.893 %	96.829 %

241	diversión	0.00345	0.060 %	0.893 %	96.889 %
242	puerto	0.00345	0.060 %	0.893 %	96.949 %
243	innovador	0.00345	0.060 %	0.893 %	97.009 %
244	cantina	0.00345	0.060 %	0.893 %	97.069 %
245	animal	0.00345	0.060 %	0.893 %	97.129 %
246	estudiar	0.00345	0.060 %	0.893 %	97.189 %
247	elección	0.00345	0.060 %	0.893 %	97.249 %
248	liceo	0.00345	0.060 %	0.893 %	97.309 %
249	transeúnte	0.00318	0.060 %	0.893 %	97.369 %
250	castillo	0.00318	0.060 %	0.893 %	97.429 %
251	inmigrante	0.00318	0.060 %	0.893 %	97.489 %
252	paloma	0.00318	0.060 %	0.893 %	97.549 %
253	quiosco	0.00318	0.060 %	0.893 %	97.609 %
254	ecológico	0.00318	0.060 %	0.893 %	97.669 %
255	ciudadanía	0.00318	0.060 %	0.893 %	97.729 %
256	comunidad	0.00318	0.060 %	0.893 %	97.789 %
257	techo	0.00294	0.060 %	0.893 %	97.849 %
258	divertirse	0.00294	0.060 %	0.893 %	97.909 %
259	administración pública	0.00294	0.060 %	0.893 %	97.969 %
260	mirar	0.00294	0.060 %	0.893 %	98.029 %
261	departamento	0.00294	0.060 %	0.893 %	98.149 %
262	cenar	0.00294	0.060 %	0.893 %	98.209 %
263	reloj	0.00294	0.060 %	0.893 %	98.269 %
264	comida	0.00294	0.060 %	0.893 %	98.329 %
265	embotellamiento	0.00294	0.060 %	0.893 %	98.449 %
266	carnicero	0.00294	0.060 %	0.893 %	98.509 %
267	escaparate	0.00272	0.060 %	0.893 %	98.569 %
268	pizzería	0.00272	0.060 %	0.893 %	98.629 %
269	comer	0.00272	0.060 %	0.893 %	98.689 %
270	salir con los amigos	0.00251	0.060 %	0.893 %	98.749 %
271	albañil	0.00251	0.060 %	0.893 %	98.809 %
272	basura	0.00251	0.060 %	0.893 %	98.869 %
273	encontrarse	0.00251	0.060 %	0.893 %	98.929 %
274	correos	0.00251	0.060 %	0.893 %	98.989 %
275	movida	0.00232	0.060 %	0.893 %	99.049 %
276	empleado	0.00232	0.060 %	0.893 %	99.109 %
277	ventana	0.00232	0.060 %	0.893 %	99.169 %
278	ciclista	0.00232	0.060 %	0.893 %	99.229 %
279	trabajador	0.00214	0.060 %	0.893 %	99.289 %
280	familia	0.00214	0.060 %	0.893 %	99.349 %
281	servicio público	0.00214	0.060 %	0.893 %	99.409 %
282	población	0.00198	0.060 %	0.893 %	99.469 %
283	doctor	0.00198	0.060 %	0.893 %	99.529 %
284	chocar	0.00198	0.060 %	0.893 %	99.589 %
285	temblar	0.00183	0.060 %	0.893 %	99.649 %
286	temporal	0.00169	0.060 %	0.893 %	99.709 %
287	nieve	0.00169	0.060 %	0.893 %	99.769 %
288	nube	0.00156	0.060 %	0.893 %	99.829 %
289	sol	0.00144	0.060 %	0.893 %	99.889 %
290	scooter	0.00133	0.060 %	0.893 %	99.949 %

## 07. Animales

Nº	Palabra	Disponibilidad	Frecuencia	%	Frecuencia
----	---------	----------------	------------	---	------------

			<b>relativa</b>	<b>Aparición</b>	<b>acumulada</b>
1	perro	0.93881	8.358 %	100.000 %	8.358 %
2	gato	0.85810	8.134 %	97.321 %	16.492 %
3	león	0.37422	4.701 %	56.250 %	21.193 %
4	caballo	0.37357	4.851 %	58.036 %	26.044 %
5	pez	0.26204	3.507 %	41.964 %	29.551 %
6	pájaro	0.26173	3.358 %	40.179 %	32.909 %
7	tigre	0.19722	2.612 %	31.250 %	35.521 %
8	elefante	0.18185	2.537 %	30.357 %	38.058 %
9	vaca	0.15561	2.239 %	26.786 %	40.297 %
10	jirafa	0.15167	2.239 %	26.786 %	42.536 %
11	mariposa	0.14502	2.015 %	24.107 %	44.551 %
12	zoológico	0.14239	2.015 %	24.107 %	46.566 %
13	conejo	0.13714	2.015 %	24.107 %	48.581 %
14	ratón	0.13473	1.940 %	23.214 %	50.521 %
15	tortuga	0.12279	1.567 %	18.750 %	52.088 %
16	serpiente	0.10974	1.716 %	20.536 %	53.804 %
17	pollo	0.08863	1.194 %	14.286 %	54.998 %
18	cerdo	0.07318	0.970 %	11.607 %	55.968 %
19	pequeño	0.07262	1.045 %	12.500 %	57.013 %
20	tigresa	0.06741	0.896 %	10.714 %	57.909 %
21	comida	0.06635	0.821 %	9.821 %	58.730 %
22	gallina	0.06610	0.896 %	10.714 %	59.626 %
23	hormiga	0.06550	0.970 %	11.607 %	60.596 %
24	mosquito	0.06369	1.119 %	13.393 %	61.715 %
25	paloma	0.05007	0.597 %	7.143 %	63.058 %
26	mono	0.04912	0.672 %	8.036 %	63.730 %
27	toro	0.04811	0.821 %	9.821 %	64.551 %
28	burro	0.04784	0.597 %	7.143 %	65.148 %
29	pasear	0.04744	0.522 %	6.250 %	65.670 %
30	naturaleza	0.04492	0.672 %	8.036 %	66.342 %
31	gallo	0.04420	0.597 %	7.143 %	66.939 %
32	grande	0.04332	0.597 %	7.143 %	67.536 %
33	pelo	0.04285	0.522 %	6.250 %	68.058 %
34	rato	0.04274	0.522 %	6.250 %	68.580 %
35	parque	0.04106	0.522 %	6.250 %	69.102 %
36	cebra	0.04018	0.597 %	7.143 %	69.699 %
37	lobo	0.03876	0.448 %	5.357 %	70.147 %
38	amigo	0.03823	0.522 %	6.250 %	70.669 %
39	cabra	0.03788	0.522 %	6.250 %	71.191 %
40	cucaracha	0.03786	0.522 %	6.250 %	71.713 %
41	zorro	0.03514	0.522 %	6.250 %	72.235 %
42	amor	0.03219	0.448 %	5.357 %	72.683 %
43	oveja	0.03199	0.522 %	6.250 %	73.205 %
44	yegua	0.03039	0.448 %	5.357 %	73.653 %
45	jardín	0.03007	0.448 %	5.357 %	74.101 %
46	ave	0.02762	0.448 %	5.357 %	74.549 %
47	loro	0.02621	0.299 %	3.571 %	74.848 %
48	cuidar	0.02613	0.299 %	3.571 %	75.147 %
49	salvaje	0.02534	0.448 %	5.357 %	75.595 %
50	pato	0.02490	0.299 %	3.571 %	75.894 %
51	cariño	0.02466	0.373 %	4.464 %	76.267 %
52	delfín	0.02311	0.448 %	5.357 %	76.715 %
53	caminar	0.02280	0.448 %	5.357 %	77.163 %
54	puerco	0.02249	0.299 %	3.571 %	77.462 %

55	casa	0.02227	0.299 %	3.571 %	77.761 %
56	jugar	0.02220	0.299 %	3.571 %	78.060 %
57	caracol	0.02132	0.299 %	3.571 %	78.359 %
58	abeja	0.02005	0.299 %	3.571 %	78.658 %
59	rana	0.01993	0.299 %	3.571 %	78.957 %
60	bosque	0.01905	0.224 %	2.679 %	79.181 %
61	doméstico	0.01892	0.224 %	2.679 %	79.405 %
62	cocodrilo	0.01849	0.224 %	2.679 %	79.629 %
63	felicidad	0.01784	0.299 %	3.571 %	79.928 %
64	lagarto	0.01741	0.224 %	2.679 %	80.152 %
65	insecto	0.01668	0.224 %	2.679 %	80.376 %
66	árbol	0.01638	0.299 %	3.571 %	80.675 %
67	finca	0.01629	0.224 %	2.679 %	80.899 %
68	leona	0.01629	0.299 %	3.571 %	81.198 %
69	peligroso	0.01611	0.224 %	2.679 %	81.422 %
70	carne	0.01598	0.299 %	3.571 %	81.721 %
71	mar	0.01564	0.299 %	3.571 %	82.020 %
72	dar un paseo	0.01561	0.224 %	2.679 %	82.244 %
73	queso	0.01543	0.299 %	3.571 %	82.543 %
74	tiburón	0.01486	0.299 %	3.571 %	82.842 %
75	papagayo	0.01462	0.224 %	2.679 %	83.066 %
76	hámster	0.01449	0.224 %	2.679 %	83.290 %
77	paseo	0.01425	0.149 %	1.786 %	83.439 %
78	nadar	0.01393	0.299 %	3.571 %	83.738 %
79	cazar	0.01375	0.224 %	2.679 %	83.962 %
80	camello	0.01364	0.299 %	3.571 %	84.261 %
81	búho	0.01362	0.149 %	1.786 %	84.410 %
82	veterinario	0.01352	0.224 %	2.679 %	84.634 %
83	negro	0.01331	0.224 %	2.679 %	84.858 %
84	gaviota	0.01300	0.149 %	1.786 %	85.007 %
85	gacela	0.01297	0.149 %	1.786 %	85.156 %
86	volar	0.01285	0.299 %	3.571 %	85.455 %
87	ladrar	0.01261	0.224 %	2.679 %	85.679 %
88	piel	0.01259	0.149 %	1.786 %	85.828 %
89	mirlo	0.01259	0.149 %	1.786 %	85.977 %
90	croqueta	0.01251	0.149 %	1.786 %	86.126 %
91	murciélago	0.01251	0.149 %	1.786 %	86.275 %
92	dueño	0.01229	0.149 %	1.786 %	86.424 %
93	mascota	0.01199	0.149 %	1.786 %	86.573 %
94	beber	0.01177	0.149 %	1.786 %	86.722 %
95	correr	0.01150	0.224 %	2.679 %	86.946 %
96	caza	0.01144	0.149 %	1.786 %	87.095 %
97	juego	0.01087	0.149 %	1.786 %	87.244 %
98	cabrón	0.01067	0.149 %	1.786 %	87.393 %
99	leche	0.01038	0.224 %	2.679 %	87.617 %
100	koala	0.01013	0.149 %	1.786 %	87.766 %
101	tranquilo	0.00995	0.149 %	1.786 %	87.915 %
102	dragón	0.00947	0.149 %	1.786 %	88.064 %
103	oso	0.00944	0.149 %	1.786 %	88.213 %
104	pescado	0.00911	0.149 %	1.786 %	88.362 %
105	tienda de animales	0.00893	0.075 %	0.893 %	88.437 %
106	alto	0.00893	0.075 %	0.893 %	88.512 %
107	beso	0.00875	0.149 %	1.786 %	88.661 %
108	dar de comer	0.00849	0.149 %	1.786 %	88.810 %



109	cola	0.00825	0.075 %	0.893 %	88.885 %
110	pata	0.00825	0.075 %	0.893 %	88.960 %
111	atún	0.00825	0.075 %	0.893 %	89.035 %
112	amar	0.00825	0.075 %	0.893 %	89.110 %
113	bajo	0.00825	0.075 %	0.893 %	89.185 %
114	morder	0.00825	0.075 %	0.893 %	89.260 %
115	camaleón	0.00782	0.149 %	1.786 %	89.409 %
116	golondrina	0.00777	0.149 %	1.786 %	89.558 %
117	fidelidad	0.00762	0.075 %	0.893 %	89.633 %
118	ladrido	0.00762	0.075 %	0.893 %	89.708 %
119	querer	0.00762	0.075 %	0.893 %	89.783 %
120	devorar	0.00762	0.075 %	0.893 %	89.858 %
121	safari	0.00762	0.075 %	0.893 %	89.933 %
122	agua	0.00757	0.149 %	1.786 %	90.082 %
123	mosca	0.00744	0.149 %	1.786 %	90.231 %
124	comer	0.00727	0.149 %	1.786 %	90.380 %
125	ojo	0.00724	0.149 %	1.786 %	90.529 %
126	cuidador de perros	0.00704	0.075 %	0.893 %	90.604 %
127	techo	0.00704	0.075 %	0.893 %	90.679 %
128	conejillo de Indias	0.00704	0.075 %	0.893 %	90.754 %
129	curar	0.00704	0.075 %	0.893 %	90.829 %
130	sacar	0.00704	0.075 %	0.893 %	90.904 %
131	murciélago	0.00704	0.075 %	0.893 %	90.979 %
132	cucha	0.00704	0.075 %	0.893 %	91.054 %
133	abrazar	0.00704	0.075 %	0.893 %	91.129 %
134	morir	0.00704	0.075 %	0.893 %	91.204 %
135	ara macao	0.00704	0.075 %	0.893 %	91.279 %
136	pluma	0.00650	0.075 %	0.893 %	91.354 %
137	equitación	0.00650	0.075 %	0.893 %	91.429 %
138	nutrir	0.00650	0.075 %	0.893 %	91.504 %
139	tienda	0.00650	0.075 %	0.893 %	91.579 %
140	bestia	0.00650	0.075 %	0.893 %	91.654 %
141	soledad	0.00650	0.075 %	0.893 %	91.729 %
142	caballo	0.00650	0.075 %	0.893 %	91.804 %
143	vida	0.00650	0.075 %	0.893 %	91.879 %
144	bisonte	0.00650	0.075 %	0.893 %	91.954 %
145	alpaca	0.00629	0.149 %	1.786 %	92.103 %
146	huevo	0.00606	0.149 %	1.786 %	92.252 %
147	acariciar	0.00601	0.075 %	0.893 %	92.327 %
148	compañero	0.00601	0.075 %	0.893 %	92.402 %
149	adiestrar	0.00601	0.075 %	0.893 %	92.477 %
150	cabalgar	0.00601	0.075 %	0.893 %	92.552 %
151	roer	0.00601	0.075 %	0.893 %	92.627 %
152	gustar	0.00601	0.075 %	0.893 %	92.702 %
153	ganadería	0.00601	0.075 %	0.893 %	92.777 %
154	muerte	0.00601	0.075 %	0.893 %	92.852 %
155	sardina	0.00601	0.075 %	0.893 %	92.927 %
156	comprar	0.00555	0.075 %	0.893 %	93.002 %
157	calle	0.00555	0.075 %	0.893 %	93.077 %
158	vuelo	0.00555	0.075 %	0.893 %	93.152 %
159	limpiar	0.00555	0.075 %	0.893 %	93.227 %
160	canguro	0.00555	0.075 %	0.893 %	93.302 %
161	cigüeña	0.00555	0.075 %	0.893 %	93.377 %
162	sierpe	0.00555	0.075 %	0.893 %	93.452 %

163	pelota	0.00555	0.075 %	0.893 %	93.527 %
164	perfume	0.00555	0.075 %	0.893 %	93.602 %
165	blanco	0.00555	0.075 %	0.893 %	93.677 %
166	araña	0.00555	0.075 %	0.893 %	93.752 %
167	mejillón	0.00523	0.149 %	1.786 %	93.901 %
168	salir	0.00512	0.075 %	0.893 %	93.976 %
169	animal doméstico	0.00512	0.075 %	0.893 %	94.051 %
170	mariquita	0.00512	0.075 %	0.893 %	94.126 %
171	dar una vuelta	0.00512	0.075 %	0.893 %	94.201 %
172	mordedura	0.00512	0.075 %	0.893 %	94.276 %
173	oveja	0.00512	0.075 %	0.893 %	94.351 %
174	cigüeña	0.00512	0.075 %	0.893 %	94.426 %
175	gorila	0.00512	0.075 %	0.893 %	94.501 %
176	huella	0.00512	0.075 %	0.893 %	94.576 %
177	juguete	0.00512	0.075 %	0.893 %	94.651 %
178	panda	0.00512	0.075 %	0.893 %	94.726 %
179	nombre	0.00512	0.075 %	0.893 %	94.801 %
180	oca	0.00512	0.075 %	0.893 %	94.876 %
181	cangrejo	0.00492	0.149 %	1.786 %	95.025 %
182	cielo	0.00489	0.149 %	1.786 %	95.174 %
183	pingüino	0.00486	0.149 %	1.786 %	95.323 %
184	fiel	0.00473	0.075 %	0.893 %	95.398 %
185	marisco	0.00473	0.075 %	0.893 %	95.473 %
186	libertad	0.00473	0.075 %	0.893 %	95.548 %
187	miedo	0.00473	0.075 %	0.893 %	95.623 %
188	escarabajo	0.00473	0.075 %	0.893 %	95.698 %
189	vegetariano	0.00470	0.149 %	1.786 %	95.847 %
190	matar	0.00437	0.075 %	0.893 %	95.922 %
191	rinoceronte	0.00437	0.075 %	0.893 %	95.997 %
192	malo	0.00437	0.075 %	0.893 %	96.072 %
193	canario	0.00437	0.075 %	0.893 %	96.147 %
194	flamenco	0.00437	0.075 %	0.893 %	96.222 %
195	bonito	0.00437	0.075 %	0.893 %	96.297 %
196	armadillo	0.00437	0.075 %	0.893 %	96.372 %
197	niño	0.00437	0.075 %	0.893 %	96.447 %
198	jabalí	0.00437	0.075 %	0.893 %	96.522 %
199	ardilla	0.00437	0.075 %	0.893 %	96.597 %
200	dingo	0.00437	0.075 %	0.893 %	96.672 %
201	perrería	0.00437	0.075 %	0.893 %	96.747 %
202	diente	0.00437	0.075 %	0.893 %	96.822 %
203	cazador	0.00437	0.075 %	0.893 %	96.897 %
204	compañía	0.00404	0.075 %	0.893 %	96.972 %
205	bueno	0.00404	0.075 %	0.893 %	97.047 %
206	chicharra	0.00404	0.075 %	0.893 %	97.122 %
207	ruido	0.00404	0.075 %	0.893 %	97.197 %
208	ciervo	0.00404	0.075 %	0.893 %	97.272 %
209	desierto	0.00404	0.075 %	0.893 %	97.347 %
210	coyote	0.00404	0.075 %	0.893 %	97.422 %
211	fusil	0.00404	0.075 %	0.893 %	97.497 %
212	agresivo	0.00373	0.075 %	0.893 %	97.572 %
213	sucio	0.00373	0.075 %	0.893 %	97.647 %
214	circo	0.00373	0.075 %	0.893 %	97.722 %
215	chapulín	0.00373	0.075 %	0.893 %	97.797 %
216	cantar	0.00373	0.075 %	0.893 %	97.872 %

217	ornitorrinco	0.00373	0.075 %	0.893 %	97.947 %
218	limpio	0.00345	0.075 %	0.893 %	98.022 %
219	bambi	0.00345	0.075 %	0.893 %	98.097 %
220	avestruz	0.00345	0.075 %	0.893 %	98.172 %
221	extinción	0.00345	0.075 %	0.893 %	98.247 %
222	bebida	0.00345	0.075 %	0.893 %	98.322 %
223	cariñoso	0.00318	0.075 %	0.893 %	98.397 %
224	hipopótamo	0.00318	0.075 %	0.893 %	98.472 %
225	nostalgia	0.00318	0.075 %	0.893 %	98.547 %
226	pantera	0.00318	0.075 %	0.893 %	98.622 %
227	dulzura	0.00294	0.075 %	0.893 %	98.697 %
228	dormir	0.00294	0.075 %	0.893 %	98.772 %
229	pastar	0.00272	0.075 %	0.893 %	98.847 %
230	pulpo	0.00272	0.075 %	0.893 %	98.922 %
231	algodón	0.00251	0.075 %	0.893 %	98.997 %
232	pescar	0.00251	0.075 %	0.893 %	99.072 %
233	feroz	0.00232	0.075 %	0.893 %	99.147 %
234	reno	0.00214	0.075 %	0.893 %	99.222 %
235	caperucita roja	0.00214	0.075 %	0.893 %	99.297 %
236	radar	0.00198	0.075 %	0.893 %	99.372 %
237	ballena	0.00183	0.075 %	0.893 %	99.447 %
238	vegano	0.00183	0.075 %	0.893 %	99.522 %
239	miel	0.00169	0.075 %	0.893 %	99.597 %
240	gallinero	0.00156	0.075 %	0.893 %	99.672 %
241	torero	0.00156	0.075 %	0.893 %	99.747 %
242	verde	0.00123	0.075 %	0.893 %	99.822 %
243	color	0.00114	0.075 %	0.893 %	99.897 %
244	largo	0.00105	0.075 %	0.893 %	99.972 %
245	corto	0.00097	0.075 %	0.893 %	100.047 %

## 08. Profesiones y oficios

Nº	Palabra	Disponibilidad	Frecuencia relativa	% Aparición	Frecuencia acumulada
1	profesor	0.80357	7.913 %	91.071 %	7.913 %
2	médico	0.38388	4.888 %	56.250 %	12.801 %
3	estudiante	0.29056	3.413 %	39.286 %	16.214 %
4	abogado	0.27876	3.103 %	35.714 %	19.317 %
5	camarero	0.24678	3.336 %	38.393 %	22.653 %
6	cocinero	0.20414	2.870 %	33.036 %	25.523 %
7	doctor	0.17474	2.095 %	24.107 %	27.618 %
8	empleado	0.14050	1.862 %	21.429 %	29.480 %
9	policía	0.13688	2.172 %	25.000 %	31.652 %
10	trabajar	0.13407	2.095 %	24.107 %	33.747 %
11	cantante	0.12823	2.017 %	23.214 %	35.764 %
12	actor	0.11818	2.017 %	23.214 %	37.781 %
13	enfermero	0.11011	1.939 %	22.321 %	39.720 %
14	secretario	0.10000	1.319 %	15.179 %	41.039 %
15	peluquero	0.09731	1.164 %	13.393 %	42.203 %
16	oficina	0.09568	1.164 %	13.393 %	43.367 %
17	dentista	0.08648	1.241 %	14.286 %	44.608 %
18	trabajo	0.08078	1.086 %	12.500 %	45.694 %
19	jefe	0.07957	1.164 %	13.393 %	46.858 %
20	canguro	0.07859	0.931 %	10.714 %	47.789 %
21	escritor	0.06970	1.241 %	14.286 %	49.030 %

22	fontanero	0.06674	1.086 %	12.500 %	50.116 %
23	ama de casa	0.06458	0.931 %	10.714 %	51.047 %
24	albañil	0.06126	0.776 %	8.929 %	51.823 %
25	tienda	0.05587	0.853 %	9.821 %	52.676 %
26	enseñar	0.05485	0.776 %	8.929 %	53.452 %
27	bombero	0.05482	0.776 %	8.929 %	54.228 %
28	político	0.05467	0.776 %	8.929 %	55.004 %
29	obrero	0.05333	0.776 %	8.929 %	55.780 %
30	director	0.05328	0.776 %	8.929 %	56.556 %
31	futbolista	0.05119	0.776 %	8.929 %	57.332 %
32	arquitecto	0.05111	0.698 %	8.036 %	58.030 %
33	traductor	0.04853	0.853 %	9.821 %	58.883 %
34	trabajador	0.04417	0.543 %	6.250 %	59.426 %
35	actriz	0.04402	0.776 %	8.929 %	60.202 %
36	ingeniero	0.04252	0.621 %	7.143 %	60.823 %
37	psicólogo	0.04182	0.698 %	8.036 %	61.521 %
38	pintor	0.04036	0.621 %	7.143 %	62.142 %
39	restaurante	0.03996	0.543 %	6.250 %	62.685 %
40	jardinero	0.03944	0.543 %	6.250 %	63.228 %
41	mecánico	0.03739	0.698 %	8.036 %	63.926 %
42	vendedor	0.03629	0.543 %	6.250 %	64.469 %
43	dinero	0.03577	0.621 %	7.143 %	65.090 %
44	periodista	0.03552	0.543 %	6.250 %	65.633 %
45	estudiar	0.03504	0.388 %	4.464 %	66.021 %
46	dependiente	0.03318	0.465 %	5.357 %	66.486 %
47	cirujano	0.03153	0.388 %	4.464 %	66.874 %
48	panadero	0.02963	0.543 %	6.250 %	67.417 %
49	jugar	0.02924	0.465 %	5.357 %	67.882 %
50	cantor	0.02886	0.388 %	4.464 %	68.270 %
51	nadador	0.02780	0.465 %	5.357 %	68.735 %
52	escuela	0.02709	0.388 %	4.464 %	69.123 %
53	bailarín	0.02700	0.543 %	6.250 %	69.666 %
54	hospital	0.02685	0.465 %	5.357 %	70.131 %
55	maestro	0.02467	0.310 %	3.571 %	70.441 %
56	peluquería	0.02426	0.388 %	4.464 %	70.829 %
57	empresario	0.02318	0.388 %	4.464 %	71.217 %
58	músico	0.02138	0.388 %	4.464 %	71.605 %
59	vender	0.02070	0.388 %	4.464 %	71.993 %
60	chofer	0.02067	0.310 %	3.571 %	72.303 %
61	soldado	0.02008	0.388 %	4.464 %	72.691 %
62	azafata	0.02000	0.310 %	3.571 %	73.001 %
63	alumno	0.01991	0.233 %	2.679 %	73.234 %
64	cajero	0.01920	0.233 %	2.679 %	73.467 %
65	electricista	0.01898	0.388 %	4.464 %	73.855 %
66	plomero	0.01884	0.233 %	2.679 %	74.088 %
67	piloto	0.01834	0.388 %	4.464 %	74.476 %
68	jugador	0.01820	0.310 %	3.571 %	74.786 %
69	jugador de fútbol	0.01810	0.233 %	2.679 %	75.019 %
70	contable	0.01779	0.310 %	3.571 %	75.329 %
71	guardia civil	0.01764	0.233 %	2.679 %	75.562 %
72	deportivo	0.01764	0.233 %	2.679 %	75.795 %
73	deporte	0.01752	0.233 %	2.679 %	76.028 %
74	farmacia	0.01696	0.233 %	2.679 %	76.261 %
75	carnicero	0.01677	0.233 %	2.679 %	76.494 %

76	presidente	0.01657	0.233 %	2.679 %	76.727 %
77	leer	0.01636	0.155 %	1.786 %	76.882 %
78	empresa	0.01619	0.233 %	2.679 %	77.115 %
79	universidad	0.01557	0.155 %	1.786 %	77.270 %
80	florero	0.01557	0.155 %	1.786 %	77.425 %
81	enseñanza	0.01557	0.155 %	1.786 %	77.580 %
82	taxista	0.01553	0.388 %	4.464 %	77.968 %
83	poeta	0.01531	0.233 %	2.679 %	78.201 %
84	ayudante	0.01492	0.155 %	1.786 %	78.356 %
85	carpintero	0.01477	0.310 %	3.571 %	78.666 %
86	pinchadiscos	0.01420	0.155 %	1.786 %	78.821 %
87	repcionista	0.01419	0.233 %	2.679 %	79.054 %
88	despacho	0.01378	0.155 %	1.786 %	79.209 %
89	alcalde	0.01362	0.155 %	1.786 %	79.364 %
90	estudio	0.01329	0.233 %	2.679 %	79.597 %
91	cosmetólogo	0.01324	0.233 %	2.679 %	79.830 %
92	medicar	0.01295	0.155 %	1.786 %	79.985 %
93	banquero	0.01283	0.155 %	1.786 %	80.140 %
94	cantar	0.01257	0.155 %	1.786 %	80.295 %
95	astrónomo	0.01257	0.155 %	1.786 %	80.450 %
96	obligación	0.01241	0.155 %	1.786 %	80.605 %
97	panadería	0.01213	0.233 %	2.679 %	80.838 %
98	autor	0.01192	0.155 %	1.786 %	80.993 %
99	agente de policía	0.01182	0.155 %	1.786 %	81.148 %
100	veterinario	0.01164	0.233 %	2.679 %	81.381 %
101	campesino	0.01155	0.233 %	2.679 %	81.614 %
102	ladrón	0.01127	0.155 %	1.786 %	81.769 %
103	cocinar	0.01105	0.155 %	1.786 %	81.924 %
104	horario	0.01087	0.155 %	1.786 %	82.079 %
105	artista	0.01067	0.155 %	1.786 %	82.234 %
106	nadar	0.01061	0.155 %	1.786 %	82.389 %
107	ciudad	0.01033	0.155 %	1.786 %	82.544 %
108	negocio	0.01033	0.155 %	1.786 %	82.699 %
109	conducir	0.01008	0.155 %	1.786 %	82.854 %
110	niñero	0.00991	0.155 %	1.786 %	83.009 %
111	científico	0.00989	0.155 %	1.786 %	83.164 %
112	cita	0.00974	0.155 %	1.786 %	83.319 %
113	técnico	0.00919	0.155 %	1.786 %	83.474 %
114	profesional	0.00904	0.155 %	1.786 %	83.629 %
115	bar	0.00897	0.155 %	1.786 %	83.784 %
116	modelo	0.00888	0.155 %	1.786 %	83.939 %
117	florista	0.00882	0.155 %	1.786 %	84.094 %
118	comprar	0.00859	0.155 %	1.786 %	84.249 %
119	navegar	0.00814	0.078 %	0.893 %	84.327 %
120	viaje de negocios	0.00814	0.078 %	0.893 %	84.405 %
121	aprender	0.00814	0.078 %	0.893 %	84.483 %
122	rey	0.00814	0.078 %	0.893 %	84.561 %
123	secretaría	0.00814	0.078 %	0.893 %	84.639 %
124	extranjero	0.00814	0.078 %	0.893 %	84.717 %
125	banco	0.00814	0.078 %	0.893 %	84.795 %
126	enseñante	0.00814	0.078 %	0.893 %	84.873 %
127	natación	0.00814	0.078 %	0.893 %	84.951 %
128	deportista	0.00814	0.078 %	0.893 %	85.029 %
129	madre	0.00810	0.155 %	1.786 %	85.184 %

130	entrenador	0.00810	0.155 %	1.786 %	85.339 %
131	fotógrafo	0.00784	0.155 %	1.786 %	85.494 %
132	ordenador	0.00784	0.155 %	1.786 %	85.649 %
133	café	0.00784	0.155 %	1.786 %	85.804 %
134	ganar	0.00784	0.155 %	1.786 %	85.959 %
135	moda	0.00784	0.155 %	1.786 %	86.114 %
136	esteticista	0.00784	0.155 %	1.786 %	86.269 %
137	coche	0.00762	0.233 %	2.679 %	86.502 %
138	tiempo	0.00760	0.155 %	1.786 %	86.657 %
139	maquillaje	0.00746	0.155 %	1.786 %	86.812 %
140	sacar dinero	0.00743	0.078 %	0.893 %	86.890 %
141	automóvil	0.00743	0.078 %	0.893 %	86.968 %
142	director de cine	0.00743	0.078 %	0.893 %	87.046 %
143	técnico informático	0.00743	0.078 %	0.893 %	87.124 %
144	negociante	0.00743	0.078 %	0.893 %	87.202 %
145	baloncestista	0.00743	0.078 %	0.893 %	87.280 %
146	general	0.00743	0.078 %	0.893 %	87.358 %
147	reina	0.00743	0.078 %	0.893 %	87.436 %
148	deber	0.00743	0.078 %	0.893 %	87.514 %
149	escultor	0.00743	0.078 %	0.893 %	87.592 %
150	florería	0.00743	0.078 %	0.893 %	87.670 %
151	fábrica	0.00743	0.078 %	0.893 %	87.748 %
152	agente	0.00743	0.078 %	0.893 %	87.826 %
153	correr	0.00743	0.078 %	0.893 %	87.904 %
154	inventor	0.00715	0.155 %	1.786 %	88.059 %
155	dueña de casa	0.00678	0.078 %	0.893 %	88.137 %
156	mujer de negocios	0.00678	0.078 %	0.893 %	88.215 %
157	función	0.00678	0.078 %	0.893 %	88.293 %
158	químico	0.00678	0.078 %	0.893 %	88.371 %
159	papelería	0.00678	0.078 %	0.893 %	88.449 %
160	ejército	0.00678	0.078 %	0.893 %	88.527 %
161	barman	0.00678	0.078 %	0.893 %	88.605 %
162	limpiar las casas	0.00678	0.078 %	0.893 %	88.683 %
163	limpiador	0.00678	0.078 %	0.893 %	88.761 %
164	preparador físico	0.00678	0.078 %	0.893 %	88.839 %
165	guía	0.00678	0.078 %	0.893 %	88.917 %
166	agente inmobiliario	0.00678	0.078 %	0.893 %	88.995 %
167	cartero	0.00660	0.155 %	1.786 %	89.150 %
168	zapatero	0.00649	0.155 %	1.786 %	89.305 %
169	peluquera	0.00618	0.078 %	0.893 %	89.383 %
170	medicina	0.00618	0.078 %	0.893 %	89.461 %
171	atleta	0.00618	0.078 %	0.893 %	89.539 %
172	gimnasta	0.00618	0.078 %	0.893 %	89.617 %
173	escritorio	0.00618	0.078 %	0.893 %	89.695 %
174	físico	0.00618	0.078 %	0.893 %	89.773 %
175	fuerzas armadas	0.00618	0.078 %	0.893 %	89.851 %
176	limpiar las escuelas	0.00618	0.078 %	0.893 %	89.929 %
177	verdulero	0.00618	0.078 %	0.893 %	90.007 %
178	comerciante	0.00618	0.078 %	0.893 %	90.085 %

179	arqueólogo	0.00618	0.078 %	0.893 %	90.163 %
180	rector	0.00618	0.078 %	0.893 %	90.241 %
181	mujer	0.00564	0.078 %	0.893 %	90.319 %
182	mejorar	0.00564	0.078 %	0.893 %	90.397 %
183	asistente	0.00564	0.078 %	0.893 %	90.475 %
184	entrevistador	0.00564	0.078 %	0.893 %	90.553 %
185	avión	0.00564	0.078 %	0.893 %	90.631 %
186	portero	0.00564	0.078 %	0.893 %	90.709 %
187	medicinal	0.00564	0.078 %	0.893 %	90.787 %
188	lavaplatos	0.00564	0.078 %	0.893 %	90.865 %
189	hombre	0.00514	0.078 %	0.893 %	90.943 %
190	guía turística	0.00514	0.078 %	0.893 %	91.021 %
191	bailón	0.00514	0.078 %	0.893 %	91.099 %
192	gráfico	0.00514	0.078 %	0.893 %	91.177 %
193	enfermera	0.00514	0.078 %	0.893 %	91.255 %
194	foto	0.00514	0.078 %	0.893 %	91.333 %
195	política	0.00514	0.078 %	0.893 %	91.411 %
196	estanco	0.00514	0.078 %	0.893 %	91.489 %
197	tienda de ropa	0.00514	0.078 %	0.893 %	91.567 %
198	monitor	0.00514	0.078 %	0.893 %	91.645 %
199	camionero	0.00514	0.078 %	0.893 %	91.723 %
200	ministro	0.00514	0.078 %	0.893 %	91.801 %
201	plata	0.00514	0.078 %	0.893 %	91.879 %
202	representante	0.00514	0.078 %	0.893 %	91.957 %
203	pianista	0.00514	0.078 %	0.893 %	92.035 %
204	buceador	0.00514	0.078 %	0.893 %	92.113 %
205	vacación	0.00514	0.078 %	0.893 %	92.191 %
206	psiquiatra	0.00499	0.233 %	2.679 %	92.424 %
207	puntual	0.00469	0.078 %	0.893 %	92.502 %
208	deberes	0.00469	0.078 %	0.893 %	92.580 %
209	electrónico	0.00469	0.078 %	0.893 %	92.658 %
210	mánager	0.00469	0.078 %	0.893 %	92.736 %
211	cartera	0.00469	0.078 %	0.893 %	92.814 %
212	criada	0.00469	0.078 %	0.893 %	92.892 %
213	ganar dinero	0.00469	0.078 %	0.893 %	92.970 %
214	esperanza	0.00469	0.078 %	0.893 %	93.048 %
215	dueño	0.00469	0.078 %	0.893 %	93.126 %
216	ecologista	0.00469	0.078 %	0.893 %	93.204 %
217	jefe del gobierno	0.00469	0.078 %	0.893 %	93.282 %
218	traje	0.00469	0.078 %	0.893 %	93.360 %
219	autobús	0.00454	0.155 %	1.786 %	93.515 %
220	fácil	0.00428	0.078 %	0.893 %	93.593 %
221	carabinero	0.00428	0.078 %	0.893 %	93.671 %
222	guía de turismo	0.00428	0.078 %	0.893 %	93.749 %
223	estético	0.00428	0.078 %	0.893 %	93.827 %
224	abertura	0.00428	0.078 %	0.893 %	93.905 %
225	vivir	0.00428	0.078 %	0.893 %	93.983 %
226	filósofo	0.00428	0.078 %	0.893 %	94.061 %
227	pintar	0.00428	0.078 %	0.893 %	94.139 %
228	baloncesto	0.00428	0.078 %	0.893 %	94.217 %
229	practicar un deporte	0.00428	0.078 %	0.893 %	94.295 %
230	correos	0.00428	0.078 %	0.893 %	94.373 %
231	biólogo	0.00428	0.078 %	0.893 %	94.451 %
232	loto	0.00428	0.078 %	0.893 %	94.529 %

233	maquillista	0.00390	0.078 %	0.893 %	94.685 %
234	colegio	0.00390	0.078 %	0.893 %	94.763 %
235	cierre	0.00390	0.078 %	0.893 %	94.841 %
236	maquinista	0.00390	0.078 %	0.893 %	94.919 %
237	casa	0.00390	0.078 %	0.893 %	94.997 %
238	sueldo	0.00390	0.078 %	0.893 %	95.075 %
239	supermercado	0.00390	0.078 %	0.893 %	95.153 %
240	guía turística	0.00390	0.078 %	0.893 %	95.231 %
241	tenor	0.00390	0.078 %	0.893 %	95.309 %
242	tocar	0.00390	0.078 %	0.893 %	95.387 %
243	cónsul	0.00390	0.078 %	0.893 %	95.465 %
244	policial	0.00356	0.078 %	0.893 %	95.543 %
245	baby-sitter	0.00356	0.078 %	0.893 %	95.621 %
246	motorista	0.00356	0.078 %	0.893 %	95.699 %
247	comida	0.00356	0.078 %	0.893 %	95.777 %
248	huelga	0.00356	0.078 %	0.893 %	95.855 %
249	tren	0.00356	0.078 %	0.893 %	95.933 %
250	municipalidad	0.00356	0.078 %	0.893 %	96.011 %
251	pasión	0.00356	0.078 %	0.893 %	96.089 %
252	música	0.00356	0.078 %	0.893 %	96.167 %
253	embajador	0.00356	0.078 %	0.893 %	96.245 %
254	construir	0.00325	0.078 %	0.893 %	96.323 %
255	maquillar	0.00325	0.078 %	0.893 %	96.401 %
256	partera	0.00325	0.078 %	0.893 %	96.479 %
257	inteligente	0.00325	0.078 %	0.893 %	96.557 %
258	suportar la familia	0.00325	0.078 %	0.893 %	96.635 %
259	organizar	0.00325	0.078 %	0.893 %	96.713 %
260	tecladista	0.00325	0.078 %	0.893 %	96.791 %
261	paro	0.00325	0.078 %	0.893 %	96.869 %
262	documento	0.00325	0.078 %	0.893 %	96.947 %
263	guitarra	0.00325	0.078 %	0.893 %	97.025 %
264	constructor	0.00325	0.078 %	0.893 %	97.103 %
265	botón	0.00296	0.078 %	0.893 %	97.181 %
266	voluntario	0.00296	0.078 %	0.893 %	97.259 %
267	planear	0.00296	0.078 %	0.893 %	97.337 %
268	granjero	0.00296	0.078 %	0.893 %	97.415 %
269	utilizar	0.00296	0.078 %	0.893 %	97.493 %
270	pausa	0.00296	0.078 %	0.893 %	97.571 %
271	piano	0.00296	0.078 %	0.893 %	97.649 %
272	controlador	0.00296	0.078 %	0.893 %	97.727 %
273	jugar al voleibol	0.00296	0.078 %	0.893 %	97.805 %
274	ordenar	0.00270	0.078 %	0.893 %	97.883 %
275	padre	0.00270	0.078 %	0.893 %	97.961 %
276	arreglar	0.00270	0.078 %	0.893 %	98.039 %
277	dirigente	0.00270	0.078 %	0.893 %	98.117 %
278	jugar al baloncesto	0.00270	0.078 %	0.893 %	98.195 %
279	ir de vacaciones	0.00246	0.078 %	0.893 %	98.273 %
280	gafas	0.00246	0.078 %	0.893 %	98.351 %
281	administrador	0.00246	0.078 %	0.893 %	98.429 %
282	lingüista	0.00225	0.078 %	0.893 %	98.507 %
283	anteojos	0.00225	0.078 %	0.893 %	98.585 %
284	comer	0.00225	0.078 %	0.893 %	98.663 %



285	empleo	0.00225	0.078 %	0.893 %	98.741 %
286	entrenar	0.00225	0.078 %	0.893 %	98.819 %
287	bancario	0.00205	0.078 %	0.893 %	98.897 %
288	oculista	0.00205	0.078 %	0.893 %	98.975 %
289	ejercito	0.00205	0.078 %	0.893 %	99.053 %
290	traducir	0.00205	0.078 %	0.893 %	99.131 %
291	lector	0.00187	0.078 %	0.893 %	99.209 %
292	ginecólogo	0.00187	0.078 %	0.893 %	99.287 %
293	hablar	0.00187	0.078 %	0.893 %	99.365 %
294	seguridad	0.00187	0.078 %	0.893 %	99.443 %
295	traducción	0.00187	0.078 %	0.893 %	99.521 %
296	transporte público	0.00170	0.078 %	0.893 %	99.599 %
297	cliente	0.00170	0.078 %	0.893 %	99.677 %
298	carro	0.00155	0.078 %	0.893 %	99.755 %
299	sonrisa	0.00155	0.078 %	0.893 %	99.833 %
300	juez	0.00142	0.078 %	0.893 %	99.911 %
301	torero	0.00142	0.078 %	0.893 %	99.989 %
302	comercial	0.00098	0.078 %	0.893 %	100.067 %

## Bibliografía

- Almela, R. et al. (2005): *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Editorial Universitas.
- Alonso Ramos, M. (2012): “Explorando la frecuencia léxica para el Diccionario de colocaciones del español”, en Jiménez Juliá *et al.* (eds.): *Cum corde et in nova grammatica: estudios ofrecidos a Guillermo Rojo*, págs. 19-40.
- Alvar Ezquerro, M. (2003): *La enseñanza del léxico y el uso del diccionario*. Madrid: Arco/Libros.
- Alvar Ezquerro, M. (2004): “La frecuencia léxica y su utilidad en la enseñanza del español como lengua extranjera”, en M. A. Castillo Carballo *et al.* (eds.): *Las gramáticas y los diccionarios en la enseñanza de español como segunda lengua: deseo y realidad, Actas del XV Congreso Internacional de Asele*. Sevilla: Universidad de Sevilla, págs. 19-39.
- Ávila Martín, M<sup>a</sup> del Carmen (2010): “Estadística y lingüística de corpus: implicaciones pedagógicas en la enseñanza y el aprendizaje del léxico”, en *Cauce, Revista internacional de Filología, Comunicación y sus Didácticas*, n<sup>o</sup> 33, págs. 163-175.
- Baralo, M. (1997): “La organización del léxico en lengua extranjera”, *Revista de filología románica*, 14, pp. 59-72.
- Bartol Hernández, J. A. (2006) “La disponibilidad léxica”, *Revista Española de Lingüística*, 36, págs. 379-384.
- Bartol Hernández, J. A. (2010) “Disponibilidad léxica y selección del vocabulario”, en R. M. Castañer Martín y V. Lagüéns Gracia (eds.), *De moneda nunca usada: Estudios dedicados a José M<sup>a</sup> Enguita Utrilla*, Zaragoza: Institución Fernando El Católico, págs. 85-108.
- Benítez Pérez, P. y J. Zebrowski (1993): “El léxico español en los manuales polacos”, en S. Montesa y A. Garrido (eds.), *El español como lengua extranjera: de la teoría al aula, Actas del III Congreso Nacional de ASELE*, Málaga: ASELE, págs. 223-230.
- Benítez Pérez, P. (1994a): “Léxico real / léxico irreal en los manuales de español para extranjeros”, en S. Montesa Peydró y A. Garrido Moraga (eds.), *Español para extranjeros: didáctica e investigación. Actas del II Congreso Internacional de ASELE*, Málaga: ASELE, págs. 325-333.
- Boogards, P. (1994): *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris: Hatier-Didier.
- Carcedo González, A. (1998a): “Sobre las pruebas de disponibilidad léxica para estudiantes de español/LE», *RILCE*, 14,2 (número monográfico: “Español como lengua extranjera: investigación y docencia”), págs. 205-224.
- Carcedo González, A. (1999a): “Desarrollo de la competencia léxica en español LE: análisis de cuatro fases de disponibilidad”, *Pragmalingüística*, 5-6, págs. 75-94.

- Carcedo González, A. (2000a): *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Turku: Universidad de Turku.
- Carcedo González, A. (2000b): “Índices léxico-estadísticos y graduación del vocabulario en la enseñanza de E/LE (aspectos culturales)”, en M. Franco Figueroa et al. (eds.), *Nuevas perspectivas en la enseñanza del español como lengua extranjera*, Cádiz, Servicio de Publicaciones de la Universidad de Cádiz, págs. 174-183.
- Casadei, F. (2003): *Lessico e semantica*. Roma: Carocci.
- Cervero, M<sup>a</sup> Jesús y F. Pichardo (2000): *Aprender y enseñar vocabulario*. Madrid: Edelsa.
- Chomsky, N. (1995): “Some Notes on Economy of Derivation and Representation”, en *The Minimalist Program*, Massachusetts, The MIP Press, págs. 129-160.
- Consejo de Europa (2001): *Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación*. Traducción del Instituto Cervantes (2002), Madrid: MEC/Anaya.
- Davis, M. (2006): *A frequency dictionary of Spanish*. New York: Routledge.
- Dimitrijević, N. (1969): *Lexical Availability. A new aspect of the lexical availability of secondary school children*. Heidelberg: Julius Gross Verlag.
- Fajardo Aguirre, A. (1993): “Los vocabularios básicos”, en Montesa, S. y Garrido, A. (eds.): *El español como lengua extranjera: de la teoría al aula*. Actas del Tercer Congreso Nacional de Asele, Málaga, Universidad de Málaga, págs. 215-222.
- Gallego Gallego, D. J. (2014): *Léxico disponible de estudiantes de español como lengua extranjera en la Comunidad de Madrid*, Tesis doctoral inédita. Universidad de Alcalá.
- García Hoz, V. (1953): *Vocabulario usual, vocabulario común y vocabulario fundamental*. Madrid: C.S.I.C.
- Genouvrier, E. y J. Peytard (1970): *Linguistique et enseignement du français*. París: Larousse.
- Gómez Devís, B. (2004): *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: reflexión metodológica, análisis sociolingüístico y aplicaciones*. Valencia: Servicio de Publicaciones de la Universidad de Valencia.
- González Fernández, J. (2013): “La disponibilidad léxica de los estudiantes turcos de español como lengua extranjera”, *Marcoele. Revista de didáctica ELE*, 16, 14 pp.
- Gougenheim, G. et alii (1956), *L'élaboration du français élémentaire. Etude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier.
- Gougenheim, G. et alii (1964), *L'élaboration du français fundamental (1er degré). Étude sur l'élaboration d'un vocabulaire et d'une grammaire de base*. Paris: Didier.

- Gougenheim, G (1967): “La statistique de vocabulaire et son application dans l’enseignement des langues”, *Revue de l’Enseignement Supérieur*, 3, págs. 137-144.
- Haensch, G. et alii (1982): *La Lexicografía. De la lingüística teórica a la lingüística práctica*. Madrid: Gredos.
- Haensch, G. (1997): *Los diccionarios del español en el umbral del siglo XXI*. Salamanca: Ediciones Universidad de Salamanca.
- Haensch, G. y C. Omeñaca (2004): *Los diccionarios del español en el siglo XXI*. Salamanca: Ediciones Universidad Salamanca.
- Juilland, A. y Chang Rodríguez (1964): *Frequency Dictionary of Spanish Word*. The Hague: Mouton.
- Justicia, F. (1995): *El desarrollo del vocabulario. Diccionario de frecuencias*. Universidad de Granada.
- Lewis, M. (1993): *The Lexical Approach*. Londres: Language Teaching Publications.
- Lewis, M. (1997): *Implementing the Lexical Approach*. Londres: Language Teaching Publications.
- Lewis, M. (2000): *Teaching collocations. Further developments in the Lexical Approach*. Londres: Language Teaching Publications.
- López Morales, H. (1973): *Disponibilidad léxica en escolares de San Juan*. MS.
- López Morales, H. (1995-1996): “Los estudios de disponibilidad léxica: pasado y presente”, *Homenaje a Rodolfo Oroz, Boletín de Filología de la Universidad de Chile*, 35, págs. 245-259.
- López Morales, H. (1999): *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.
- López Rivero, E. (2008): *Estudio de disponibilidad léxica en 43 estudiantes de ELE*, Memoria de máster inédita. Universidad Antonio de Nebrija.
- Lorenzo, E. (1957): “Libros de texto y selección de vocabulario”. *El español y otras lenguas* (1980), Madrid: SGEL, págs. 109-115.
- Mackey, W. (1971) *Le vocabulaire disponible du français*, 2 vols. Paris-Bruxelles-Montreal: Didier.
- Magnúsdóttir, S. (2012): *Disponibilidad léxica en alumnos de español como lengua extranjera. Estudio sobre el léxico disponible en alumnos de ELE en la secundaria en Islandia*, Memoria de licenciatura. Universidad de Islandia.
- Marslen-Wilson W. (1989): *Lexical Representation and Process*. Cambridge: MIT Press.
- Mateo García, María Victoria (1994): “Disponibilidad léxica: posibles aplicaciones”, en J.D. Luque y A. Pamies (eds.), *Actas de las Primeras Jornadas sobre estudio y enseñanza del léxico*, Granada, Universidad de Granada, págs. 141-147.
- Michéa, R. (1953): “Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage”, *Les langues Modernes*, 47, págs. 338-344.

- Morales, A. (1986): *Léxico básico del español de Puerto Rico*. Academia Puertorriqueña de la Lengua Española.
- Müller, C. (1964): *Essai de statistique lexicale*. Paris: Bibliothèque française et romaine.
- Nattinger, J. R. y J. Decarrico (1992): *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Palapanidi, Kiriaki (2012): “La aplicación de la disponibilidad léxica a la didáctica del léxico de LE”, en *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, págs. 69-77.
- Paredes García, F. (2015): Disponibilidad léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula, en *Lingüística en la red*, pp. 1-32.
- Rodríguez Bou, I. (1952): *Recuento de vocabulario español*. San Juan de Puerto Rico: Organización de Estados Americanos.
- Rodríguez Bou, I. (1966): *Recuento de vocabulario de preescolares*. Rio Piedras: Consejo Superior de Enseñanza, Universidad de Puerto Rico.
- Rodríguez Muñoz, F. J. y Muñoz Hernández, I. O. (2009): “De la disponibilidad a la didáctica léxica”, en: *Tejuelo*, nº4. Almería: Universidad de Almería, págs. 8–18.
- 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.
- Samper Hernández, M. (2002): *Disponibilidad léxica en alumnos de español como lengua extranjera*. Málaga: ASELE.
- Samper Padilla, J. A. y Hernández Cabrera, C. E. (1997) “Estudio de la disponibilidad léxica en Gran Canaria: datos iniciales y variación sociolingüística.”, en *Contribuciones al estudio de la lingüística hispánica* (Homenaje a Ramón Trujillo Carreño), La Laguna: Universidad de la Laguna, vol. 2, págs. 229-239.
- Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible. *Lingüística*, 10, 311-333.
- Samper Padilla, J. A., J. J. Bellón Fernández y Samper Hernández, M. (2003): “El proyecto de estudio de la disponibilidad léxica en español”, R. Ávila, J. A. Samper y H. Ueda, *Pautas y pistas en el análisis del léxico hispano (americano)*, Frankfurt-Madrid: Vervuert Iberoamericana, págs. 27-140.
- Samper Padilla, J. A. y Samper Hernández, M. (2006): “Aportaciones recientes de los estudios de disponibilidad léxica”, *Lynx: Panorámica de estudios lingüísticos*, 5, págs. 5-95.
- Šifrar Kalan, M. (2009): “Disponibilidad léxica en español como lengua extranjera: el cotejo de las investigaciones en Eslovenia, Salamanca y Finlandia”, *Verba Hispanica*, 17, págs. 165-182.

## Sitios web consultados

<http://www.rae.es/> [Última consulta: 23/01/2017]

<http://www.dispalex.com/> [Última consulta: 23/01/2017]

<http://www.corpusdelespanol.org/> [Última consulta: 17/10/2016]

<http://www.rae.es/obras-academicas/diccionarios/diccionario-de-americanismos/>  
[Última consulta: 18/10/2016]

<http://clave.smdiccionarios.com/> [Última consulta: 27/11/2016]



## Riassunto

La disponibilità lessicale è un campo di indagine della linguistica che ha come obiettivo di studio la raccolta del lessico disponibile di una determinata comunità linguistica e la sua successiva analisi. Per lessico disponibile si intende un insieme di parole, organizzate nel lessico mentale di un individuo, che vengono attivate solo in relazione ad una situazione comunicativa concreta, grazie ad uno stimolo che ne va ad attivare le relative connessioni cerebrali. Basandosi su questi presupposti, le ricerche sulla disponibilità lessicale si propongono di studiare quali parole possono essere usate da un individuo in riferimento ad un determinato tema di comunicazione e si differenziano dagli studi sul lessico basilico, che è composto dalle parole più frequenti di una lingua indipendentemente dall'argomento trattato. In particolare, mentre il lessico basilico è formato principalmente da parole grammaticali, che costituiscono la struttura di supporto di una lingua, il lessico disponibile è caratterizzato da parole con una forte carica semantica, per lo più da un abbondante numero di sostantivi che si riferiscono a realtà concrete. L'utilità di queste parole è pertanto strettamente legata a dei temi, chiamati centri di interesse, che equivalgono in gran parte ai campi semantici stabiliti dal Quadro comune europeo di riferimento per le lingue (QCER).

I primi studi nell'ambito della disponibilità lessicale sono nati a metà del secolo scorso in Francia, dove un gruppo di ricercatori ha basato le proprie indagini sulla raccolta del lessico di una comunità di immigrati al fine di contribuire con i dati ottenuti al perfezionamento e alla selezione del vocabolario più adeguato da inserire nei programmi di insegnamento della lingua francese a stranieri. In ambito ispano-americano questa corrente ha avuto come suo pioniere il professore Humberto López Morales, il quale con il suo studio sul lessico disponibile di una comunità di studenti di San Juan, a Puerto Rico, ha dato l'avvio alle prime ricerche di questo genere sulla lingua spagnola e attualmente coordina il "Proyecto Panhispanico", un progetto di studi sulla disponibilità lessicale nel quale convergono congiuntamente le ricerche svolte su questo tema in Spagna e America Latina, con l'obiettivo di elaborare dizionari di disponibilità lessicale delle diverse zone del mondo ispano-americano per metterle a confronto a livello linguistico, etnografico e culturale, oltre che per fungere da punto di partenza per analisi posteriori.



Principalmente le ricerche condotte sulla lingua spagnola si sono concentrate sullo studio del lessico disponibile delle comunità madrelingua, eppure, negli ultimi decenni si è diffusa una nuova corrente di studi che ha come obiettivo la raccolta del lessico disponibile degli studenti di spagnolo come lingua straniera nei paesi in cui essi frequentano gli studi. A seguito del nostro interessamento su questo specifico campo di ricerca, non abbiamo rilevato la presenza di studi su questo argomento in Italia, motivazione che ci ha spinto ad avvicinarci a questo ambito di ricerca e a portare a termine questo lavoro sperimentale.

In particolare, la tesi qui presentata si basa sul materiale raccolto durante una prova di disponibilità lessicale svolta alla fine del mese di ottobre 2016, che ha coinvolto un campione non predefinito di studenti di spagnolo come lingua straniera dell'Università di Padova. La maggior parte di essi frequenta il secondo anno di studi in “Mediazione linguistica e culturale” e “Lingue, letterature e culture moderne”, mentre una bassissima percentuale proviene da altri percorsi universitari. Nel periodo in cui la prova è stata realizzata gli alunni seguivano il corso di spagnolo condotto dalla professoressa Carmen Castillo Peña, relatrice di questa tesi. In totale il campione di studenti ottenuto corrisponde a 112 intervistati, un numero limitato rispetto a quello delle ricerche standard svolte su madrelingua, ma che è in linea con il nostro proposito di fornire un contributo originale e di carattere introduttivo all'interno del panorama di ricerca oggetto di questa tesi, che ha il fine di studiare il lessico disponibile degli studenti che hanno partecipato alla prova.

Nello specifico, gli obiettivi di questa ricerca hanno tenuto in considerazione gli aspetti che elenchiamo di seguito:

1. Presentare il numero assoluto e relativo delle parole ottenute dagli informatori in ogni centro di interesse;
2. Presentare il numero assoluto e relativo dei vocaboli per ciascun centro di interesse;
3. Osservare l'influenza delle variabili extralinguistiche in ogni centro di interesse sia in relazione al numero di parole che alla quantità di vocaboli;
4. Fornire i risultati relativi ai dati quantitativi suddivisi per indici di media di parole per partecipante e indice di coesione per ogni centro di interesse;

5. Elencare le parole che si trovano tra i primi posti in ogni centro di interesse e realizzare un'analisi descrittiva delle unità lessicali presenti in ogni campo considerato;
6. Allegare in appendice al presente documento le liste di disponibilità lessicale raggruppate per centro di interesse, con l'indicazione degli indici di disponibilità, la frequenza percentuale, la frequenza relativa e assoluta dei singoli vocaboli;
7. Dato il carattere introduttivo di questa ricerca, abbiamo deciso di non stabilire un secondo campione con il quale confrontare i risultati ottenuti, dal momento che questo fine potrà essere perseguito in ricerche successive alla nostra.

Allo scopo di ottenere gli obiettivi che ci siamo proposti, abbiamo seguito la metodologia normalmente adottata per la ricerca in questo campo di studio, che impiega le tecniche della psicologia associativa per determinare il lessico mentale di una persona e stabilire il suo grado di competenza lessicale. In generale ci si avvale di una prova di disponibilità lessicale alla quale viene chiamato a partecipare un campione di persone a cui viene richiesto, con la guida di un somministratore, di elencare in una lista aperta tutte le parole che vengono in mente per prime riguardo ad uno specifico centro di interesse. in un tempo limite di due minuti. In questo modo, le parole più disponibili di un singolo individuo saranno quelle che verranno attivate più rapidamente e inserite nelle prime risposte, mentre dal punto di vista di una comunità linguistica, saranno quelle presenti tra le prime parole scritte dalla maggioranza degli intervistati e classificate in liste di disponibilità lessicale in base al calcolo dell'indice di disponibilità, che tiene conto della posizione relativa delle parole nella lista e della loro frequenza e che, in altre parole, si può definire come l'indice di possibilità d'uso di una parola in una comunità in relazione a un determinato argomento.

Al momento della prova, viene anche richiesto di fornire alcuni dati personali relativi al campione di partecipanti mediante un modulo in cui si specifica di apportare informazioni su ambiti che servono a determinare i criteri extralinguistici normalmente considerati nelle ricerche di disponibilità lessicale. Nel nostro caso, oltre a prendere in considerazione alcune variabili tipiche degli studi condotti su madrelingua (sesso, età, livello socioculturale, area geografica di residenza e tipo di studi seguiti) abbiamo

aggiunto, sulla base di lavori che hanno come soggetto gli alunni di spagnolo come lingua straniera, alcuni specifici elementi normalmente studiati nelle ricerche di questo genere, ovvero la lingua materna, il numero di eventuali altre lingue conosciute e la durata degli studi relativi alla lingua spagnola. Questi dati vengono poi elaborati ed impiegati per lo studio quantitativo e qualitativo delle risposte fornite durante la prova.

Dopo aver ottenuto il materiale, si è proceduto con il lavoro di trascrizione dei dati e la loro modifica, nel rispetto delle norme stabilite per le ricerche standard su studenti madrelingua e stranieri allo scopo di fornire una base valida di confronto per future ricerche. In seguito, ci siamo occupati dell'inserimento di tutti i dati raccolti all'interno del programma di calcolo statistico *Dispolex*, uno strumento informatico disponibile in rete, gestito da un'équipe di ricercatori dell'Università di Salamanca, che permette di ottenere numerosi indici statistici, tra cui il calcolo disponibilità lessicale. Anche se questo supporto offre numerose possibilità di confronto, nella nostra ricerca ci siamo limitati a presentare solo alcuni calcoli con le relative analisi; in futuro sarà possibile approfondire lo studio tramite lo studio parallelo di altri campioni.

A partire da queste premesse, abbiamo organizzato il presente elaborato in due parti: la prima tratta il quadro teorico generale in cui si colloca il nostro studio di disponibilità lessicale, mentre la seconda fornisce il contributo più innovativo della nostra ricerca ed è dedicata alla presentazione dei dati e alla loro analisi.

Nel primo capitolo di questa ricerca ci siamo soffermati su alcune nozioni di carattere concettuale in relazione al campo della lessicostatistica e abbiamo fornito le principali distinzioni terminologiche in riferimento a questo ambito. Prima di tutto, abbiamo definito che cosa si intende per *lessico* in contrapposizione a *vocabolario*. Il primo corrisponde all'insieme di tutte le unità virtuali (lessemi) che si possono creare in qualsiasi momento attraverso le regole di formazione delle parole e ha come base di riferimento la *parola*, mentre il secondo è la somma di tutte quelle unità che vengono utilizzate in una situazione comunicativa e che ricevono il nome di *vocaboli*. Nella lessicostatistica l'unità di base è la parola, ma il termine più adeguato per la definizione dei dati prodotti da un campione di partecipanti è *unità lessicale*. Successivamente, si stabilisce la differenza tra il vocabolario di un individuo, che è comune a una comunità linguistica poiché le sue unità sono conosciute e condivise dai parlanti, e il lessico

mentale, che si riferisce all'insieme delle parole conosciute da un singolo parlante e che possono essere create sulla base di regole e capacità innate. Questa capacità di capire e impiegare unità lessicali, ma anche di produrre unità mai percepite o utilizzate prima è sinonimo di competenza lessicale e si colloca all'interno delle basi psicolinguistiche alle quali si affida il nostro modello teorico. La digressione continua con la descrizione dei diversi tipi di lessico che costituiscono una lingua: il lessico frequente, che misura la frequenza di apparizione delle unità all'interno di un corpus; il lessico basico, che è composto dalle unità che un parlante deve conoscere e impiegare in qualsiasi situazione, le più generiche, comuni e importanti, che appaiono con molta frequenza in qualsiasi contesto comunicativo indipendentemente dal tema trattato a differenza del lessico disponibile che, come abbiamo già esposto, si attiva solo in relazione a degli argomenti concreti. Da qui viene generata la distinzione tra parole tematiche, legate a un contesto e caratterizzate da instabilità statistica, e quelle atematiche, più stabili e con una frequenza indipendente dal tema trattato. In linguistica, l'unione complementare del lessico basico con quello disponibile forma il lessico fondamentale. A chiusura del capitolo proseguiamo con una descrizione degli studi relativi alla statistica linguistica: i dizionari di frequenza, che raccolgono le parole più frequenti di una lingua a partire da un corpus eterogeneo, e i corpora linguistici dello spagnolo, strumenti di consultazione più recenti e rinnovati, accessibili in rete, che hanno in comune con i dizionari il grande interesse pedagogico che hanno suscitato per lo studio e l'insegnamento del lessico.

Nel secondo capitolo presentiamo l'origine e lo sviluppo degli studi quantitativi sul lessico, occupandoci delle ricerche che costituiscono la principale base teorica di questi tipi di lavoro. Inoltre, elenchiamo i principali studi svolti su comunità linguistiche di lingua materna spagnola e ci soffermiamo sui criteri che a partire dal 1999 hanno fornito importanti indicazioni circa la metodologia comune da adottare per indagare il lessico disponibile nel mondo ispano-americano. Attraverso queste linee guida sono stati stabiliti i sedici centri di interesse che possono essere impiegati come stimolo per l'esecuzione del test di disponibilità lessicale, ovvero: 1) 'Parti del corpo', 2) 'Abbigliamento', 3) 'Parti della casa (senza mobili)', 4) 'I mobili della casa', 5) 'Cibi e bevande', 6) 'Oggetti collocati in tavola durante i pasti', 7) 'La cucina e i suoi utensili', 8) 'La scuola: mobili e materiali', 9) 'Illuminazione, riscaldamento e metodi l'aerazione di un recinto', 10) 'La città', 11) 'La campagna', 12) 'Mezzi di trasporto', 13) 'Lavori

in campagna e in giardino’, 14) ‘Gli animali’, 15) ‘Giochi e distrazioni’, 16) ‘Professioni e mestieri’. Inoltre, sono state definite le norme per la preparazione e lo svolgimento della prova, il numero di partecipanti necessari, che ammonta a 400 per ogni comunità linguistica studiata, e le variabili extralinguistiche da tenere in considerazione. Il capitolo continua con l'esposizione dei contributi scientifici sulla disponibilità lessicale in relazione allo spagnolo come lingua straniera, alcuni dei quali hanno applicato i dati ottenuti tramite studenti madrelingua all'insegnamento dello spagnolo a studenti non nativi, mentre altri hanno impiegato i criteri comuni della ricerca ispano-americana per lo studio del lessico degli studenti di spagnolo come lingua straniera in diverse nazioni. L'iniziatore di questi studi è stato Alberto Carcedo González, che nel 2000 ha focalizzato le proprie ricerche sullo studio del lessico disponibile di una comunità di studenti in Finlandia e ne ha confrontato i risultati con altri studi condotti su parlanti madrelingua. Da quel momento, altre indagini simili sono state messe in campo, alcune in particolare su gruppi di immigrati di lingua materna spagnola negli Stati Uniti o in Spagna e altre su un campione congiunto composto da immigrati e studenti di spagnolo come lingua straniera in Spagna. Il capitolo continua con l'elenco dei criteri extralinguistici normalmente utilizzati nelle ricerche rivolte a parlanti non nativi (lingua materna, conoscenza di altre lingue e livello di studi dello spagnolo). Segue la descrizione del metodo di calcolo per ottenere l'indice di disponibilità lessicale, elaborato nel 1987 dai ricercatori López Chávez y Strassburger. Infine, si fornisce una breve descrizione delle diverse aree di applicazione dello studio di disponibilità lessicale: la dialettologia, che permette di segnalare le similitudini e le differenze diatopiche tra varianti linguistiche diverse; la sociolinguistica, che rende possibile l'analisi del lessico disponibile di una comunità linguistica e la determinazione della sua stratificazione sociolinguistica; la psicolinguistica, che come abbiamo già segnalato riceve un apporto importante dallo studio del lessico mentale di un parlante; la etnolinguistica, che può rilevare le differenze di cultura, usi e costumi delle diverse comunità considerate; e infine la didattica delle lingue, una delle aree più feconde della linguistica applicata, che pone grande interesse sull'importanza dello studio del lessico per ottenere un buon livello di apprendimento di una lingua e che può servirsi degli studi sulla disponibilità lessicale per ottenere informazioni sulle diverse fasi del processo di apprendimento di una

comunità, riconoscere gli errori più comuni e guidare il processo di pianificazione dell'insegnamento del lessico, oltre che permettere confronti tra gruppi di parlanti.

Segue nel terzo capitolo un'introduzione alla metodologia che abbiamo seguito nella nostra ricerca per lo svolgimento della prova in classe con i 112 partecipanti e l'esposizione delle decisioni prese in merito alla scelta delle variabili extralinguistiche considerate insieme alla presentazione dei dati ottenuti per ciascun fattore, presentati con il supporto di alcuni grafici. Nello specifico, il nostro campione si compone di 99 ragazze e 13 ragazzi, principalmente tra i 19 e i 22 anni. Per lo più gli alunni provengono dalle zone circostanti la città di Padova o da altre città italiane, ad eccezione di una minoranza che è originaria di paesi esteri. La maggioranza dei partecipanti si può raggruppare nel livello socioculturale intermedio, seguita da un numero esiguo di studenti di livello basso e alto. La lingua materna più comune è l'italiano, meno preponderante è il dato in percentuale relativo ad altri studenti madrelingua. Si evidenzia inoltre che metà del campione conosce due lingue e che un buon numero di persone ne conosce tre o più. In pochi ne padroneggiano meno di due. Infine, la distribuzione relativa agli anni di studio della lingua spagnola risulta abbastanza bilanciata, tranne nell'intervallo di anni di studio più alto dove si colloca la minoranza dei partecipanti. Inoltre, spieghiamo in dettaglio il tipo di prova che è stata svolta, come è stata sviluppata, mostriamo quali centri di interesse sono stato selezionati in base alle maggiori corrispondenze con i campi semantici del QCER, ovvero: 1) 'Parti del corpo', 2) 'Abbigliamento', 3) 'Parti della casa', 4) 'Cibi e bevande', 06) 'La città', 07) 'Gli animali' e 08) 'Professioni e mestieri'. Infine, dedichiamo una sezione alla descrizione delle decisioni che sono state prese durante il processo di inserimento del materiale nel database informatico per ciascun centro di interesse e vengono esposti i criteri seguiti nel lavoro di omogeneizzazione e pulizia dei dati raccolti in relazione ai singoli campi lessicali presi in considerazione, tra cui l'eliminazione delle ripetizioni, la correzione ortografica, l'unificazione ortografica, la neutralizzazione delle forme meramente flessive, l'unificazione delle forme derivate regolari, delle forme piene e accorciate, il mantenimento di marchi commerciali pienamente lessicalizzati. A queste si aggiungono altre norme proprie delle ricerche condotte su studenti non madrelingua che comprendono la correzione e l'adattamento grafico, l'eliminazione di prestiti

linguistici che non presentano cambi rispetto alla lingua materna, l'unificazione delle unità, l'ampiezza delle relazioni associative.

Il quarto capitolo fornisce inizialmente una presentazione dei dati quantitativi ottenuti nelle prove di disponibilità lessicale del nostro campione di studenti da un punto di vista generale, mostrando nel dettaglio i dati sul numero totale di parole, che corrisponde a 12.049 unità e che offre una media di 107,59 risposte totali per ciascun partecipante, i risultati relativi all'indice di coesione e l'elenco delle venti parole più disponibili. Per quanto riguarda i dati relativi alla media di parole per centro di interesse, si segnala un totale di 13,45 risposte per alunno, una quantità minore rispetto alla media degli studi ispano-americani, che si aggira attorno alle 20 unità. In relazione alla quantità totale di parole e vocaboli ottenuti, abbiamo osservato che si mantiene un parallelismo generale tra la quantità di parole totali e la quantità di vocaboli totali per ciascun centro di interesse, tranne in rare eccezioni, come accade nel centro 'Professioni e mestieri' (08) che dalla penultima posizione per numero di parole sale al primo posto per numero di vocaboli. Una differente sorte tocca al campo 'Parti del corpo' (01) che scende dal quarto posto in relazione al numero di parole al primo per numero di vocaboli. Le successive analisi relative alla media di parole a seconda del fattore extralinguistico in generale evidenziano uno scarso potere discriminante di alcune variabili dovuto ad una rappresentazione numerica sbilanciata, come ad esempio succede in riferimento al criterio sesso, dove il predominio viene ottenuto dalla quota femminile, e della variabile età, che presenta risultati anomali in quanto con l'aumento dell'età dei partecipanti la produzione di parole non segue una crescita regolare, ma subisce una diminuzione nell'intervallo di età intermedio. Allo stesso modo, scarsa importanza viene data al criterio relativo alla zona geografica dal momento che non ha generato differenze rilevanti. In riferimento alla variabile che considera il tipo di studi degli alunni, abbiamo riscontrato un maggiore repertorio lessicale da parte degli studenti di "Mediazione Linguistica e Culturale", al punto da farci supporre che una maggiore prossimità degli studi da loro seguiti in relazione alle competenze linguistiche necessarie a padroneggiare meglio la lingua nelle situazioni comunicative da noi considerate potrebbe aver inciso in positivo sul risultato. Per quanto riguarda il livello socioculturale, non si rileva una relazione di ascesa direttamente proporzionale tra la quantità di media di parole prodotte e l'aumento del livello. Infatti, se questo fenomeno

si presenta per i livelli basso e medio, in relazione al livello alto si interrompe a segnalare probabilmente il basso potere discriminante anche di questa variabile. Diverso è il comportamento del criterio relativo alla lingua materna, che mostra la superiorità nella produzione di parole dei pochi alunni inglesi rispetto a quelli madrelingua italiana, che sono presenti in maggior numero. La variabile conoscenza di altre lingue, dal canto suo, manifesta una tendenza generale crescente man mano che aumenta il numero di lingue conosciute dai partecipanti, fenomeno che fa presupporre un'alta relazione associativa tra il criterio considerato e la capacità di memorizzare un maggior numero di parole e di impiegarle in determinate situazioni comunicative. Tuttavia, il dato che risalta più di tutti è quello relativo al criterio che considera gli anni di studio dello spagnolo e che evidenzia una corrispondenza interessante tra il numero di parole risposte dagli studenti e il loro livello di conoscenza della lingua. Sette centri di interesse su otto registrano la predominanza del livello più avanzato rispetto agli altri e ci porta alla considerazione del fatto che con lo sviluppo del processo di apprendimento è necessario introdurre maggiori e più complesse unità lessicali, che gli studenti sono in grado di trattenere dal momento che posseggono una maggiore competenza lessicale. Infine, mediante il supporto di tabelle, vengono presentati i dati numerici relativi alla produzione del numero totale di vocaboli per ogni centro di interesse studiato in relazione alle differenti variabili extralinguistiche. I risultati generali evidenziano una corrispondenza direttamente proporzionale tra il numero dei vocaboli totali e il rispettivo numero di componenti per ciascun gruppo in cui sono suddivise le diverse variabili tenute in considerazione.

Il capitolo finale conclude la trattazione dell'argomento oggetto di questa tesi con l'analisi qualitativa dei centri di interesse considerati nella nostra ricerca, concentrandosi sulla presentazione delle prime dieci parole per ogni centro di interesse. A partire da esse è stato possibile determinare i diversi insiemi e sottoinsiemi di parole che compongono i pilastri del nostro lavoro, ovvero i centri di interesse, e determinarne la corrispondenza al nucleo centrale. Inoltre, abbiamo potuto verificare il grado di relazione dei campi lessicali con il corrispondente indice di coesione, che ne rappresenta il carattere diffuso/aperto o chiuso/compatto in base alla quantità di risposte uguali date. In questo senso, si evidenziano i dati relativi ai centri 'Parti del corpo' (01) e 'Abbigliamento' (02) che mostrano una relazione molto stretta tra centri di interesse e



indice di coesione. Inoltre, a differenza di quanto si è manifestato con il dato dell'indice di coesione, abbiamo notato che il centro di interesse 'Cibi e bevande'(04), al quarto posto, ha registrato una maggiore associazione semantica rispetto allo stimolo fornito, mentre 'Parti della casa' (03), suo antecedente nella lista relativa all'indice, ha mostrato un generale distanziamento dal nucleo centrale. I restanti campi lessicali presentano relazioni lessico-associative molto ampie, fenomeno che si rispecchia anche nell'alto numero di vocaboli totali ottenuti nei centri 'Professioni e mestieri' (08), 'La città' (06), 'La scuola' (05). Per quanto riguarda il tipo di categorie grammaticali nei vari centri di interesse, notiamo l'altissimo numero di sostantivi che abbonda in tutti i campi lessicali, rispetto al bassissimo numero di verbi e aggettivi che non superano mai il 25% del totale in nessun centro. Questi dati testimoniano il rapporto associativo primario che è stato stabilito con i nomi, nonostante la spiegazione iniziale fornita dal somministratore della prova avessero chiarito la possibilità di apportare parole di qualsiasi categoria grammaticale, qualora se ne fosse presentata l'eventualità. Inoltre, abbiamo voluto analizzare i fenomeni linguistici che evidenziano le caratteristiche qualitative dei vocaboli apportati dai partecipanti dal punto di vista della variazione lessicale, dei cambiamenti semantici, dei prestiti linguistici e dei meccanismi di formazione e derivazione lessicale. Il vocabolario del nostro campione di alunni in generale si è rivelato molto vicino allo spagnolo standard, con una presenza limitata di casi di variazione diatopica nel centro di interesse 'Abbigliamento' (02) e di eufemismi nel centro 'Parti del corpo' (01), per quanto riguarda i cambi semantici. Infine, troviamo una maggior quantità di prestiti linguistici nel centro di interesse 'Professioni e mestieri' (08), dove compaiono anche numerosi meccanismi di formazione e derivazione di parole. Infine, in un buon numero di centri di interesse si registrano alcuni casi di composizione sintagmatica di tipo nominale e verbale.

Per ultimo, ma non per importanza, abbiamo completato la ricerca con la presentazione del materiale da cui è stato possibile ricavare e svolgere questo contributo sperimentale. Nell'allegato I si mostra il modello del documento di raccolta dei dati personali dei singoli studenti, necessario per l'analisi dei risultati in base ai fattori extralinguistici. Nell'allegato II si elencano le liste generali di disponibilità lessicale, raggruppate per centro di interesse, secondo l'indice di disponibilità, la frequenza percentuale, la frequenza relativa e assoluta. Le unità lessicali che ne fanno parte

costituiscono i vocabolari di disponibilità, strumenti da cui partire per avviare ulteriori lavori di ricerca o da impiegare per le attività didattiche come sostegno all'insegnamento o all'apprendimento del vocabolario.