



## DISPONIBILIDAD LÉXICA Y ORTOGRAFÍA EN ELE: UN ESTUDIO PARA LA ENSEÑANZA DE LA LENGUA<sup>1</sup>

### LEXICAL AVAILABILITY AND ORTHOGRAPHY OF SFL LEARNERS: AN ANALYSIS FOR LANGUAGE TEACHING

Giulia Nalesso  
*Università degli Studi di Padova*  
[giulia.nalesso@unipd.it](mailto:giulia.nalesso@unipd.it)

#### RESUMEN

Es sabido que hasta la fecha la disponibilidad léxica en ELE ha proporcionado numerosos proyectos cuyo objetivo es el análisis del léxico disponible de los encuestados, pero las pruebas ofrecen materiales que permiten plantear otro tipo de estudios desde perspectivas diferentes. Esta investigación examina la situación de la ortografía de un grupo de estudiantes italianos mediante el análisis de los errores cometidos al realizar un test de disponibilidad léxica y propone un método de detección de errores ortográficos del que los docentes podrán servirse a partir de los resultados extraídos del léxico disponible de su propio estudiantado para la detección de los déficits, la averiguación de carencias de las propuestas didácticas y el planteamiento de actividades de refuerzo *ad hoc*.

**Palabras clave:** disponibilidad léxica, ELE, didáctica, ortografía, análisis de errores.

#### ABSTRACT

So far, lexical availability in SFL has provided many projects aimed at analysing the available lexicon of respondents. However, this kind of test offers materials that allow alternative studies, which have not yet been adequately explored. To this end, this work examines the orthographic proficiency of a group of Italian students by analysing the errors made in a lexical availability test. The survey can be used in language teaching, as professors could apply the results of their own students to detect their deficiencies, gaps in the didactic proposals and to design specific reinforcement activities.

**Keywords:** lexical availability, SFL, didactics, orthography, error analysis.

<sup>1</sup> Esta investigación es parte de la comunicación presentada en el I Congreso Virtual Internacional sobre investigaciones léxicas INLÉXICO 2021 (febrero de 2021), organizado por el equipo de investigación EL\_HUM4 de la Universidad de Jaén, España.



## 1. INTRODUCCIÓN Y OBJETIVOS

---

Tras la publicación de los trabajos pioneros de Carcedo (2000) y Samper Hernández (2002), se realizaron numerosos proyectos sobre disponibilidad léxica en ELE cuyo objetivo inmediato era, y sigue siendo, el análisis del léxico disponible de los encuestados (Nalesso, 2019). Sin embargo, esta tipología de prueba ofrece datos y materiales que permiten plantear estudios desde perspectivas diferentes, todavía no exploradas a fondo. Por ejemplo, no tenemos constancia de trabajos sobre estudiantes extranjeros que se centren en la situación de la ortografía, como se propuso para el español LM en Paredes (1999) o Trigo, Romero & Santos (2018, 2019, 2020).

En esta investigación llevamos a cabo un estudio que, a partir de la recopilación de un corpus cacográfico, examina los errores –categorizados bajo la propuesta de Vázquez (1999)– más frecuentes cometidos por un grupo de estudiantes universitarios italianos de nivel intermedio al realizar una prueba de disponibilidad léxica, diseñada según el modelo tradicional (Samper Padilla, 1998). Nuestro objetivo prioritario es proporcionar a los docentes datos útiles para la enseñanza de la lengua ayudando a solventar los problemas de ortografía del estudiantado. En concreto, damos cuenta de los resultados cuantitativos y cualitativos de un estudio basado en el léxico disponible de la muestra, realizado por medio del análisis de errores, ya que, en nuestra opinión, conocer la tipología, frecuencia y ámbito de aparición es fundamental para tomar decisiones y avanzar en el terreno de la didáctica de la ortografía española.

Tras la administración de la prueba, la recogida de datos y la recopilación del corpus, clasificamos los errores y establecemos cuáles son los más frecuentes de manera que el profesorado pueda intervenir en los puntos más problemáticos de sus propios alumnos. En efecto, confiamos en que los docentes podrán servirse de los resultados extraídos del léxico disponible para la detección de los déficits y de las eventuales carencias de las propuestas didácticas, así como el planteamiento de actividades de refuerzo *ad*

*boc* a la luz de los datos procedentes de los análisis. Asimismo, esperamos fomentar un nuevo camino de aplicación de la disponibilidad léxica en ELE abriendo una línea de investigación basada en esta metodología que proporcione datos aplicables en el ámbito más práctico de la didáctica, en este caso examinando la situación ortográfica del estudiantado reflejada en la prueba.

## 2. FUNDAMENTACIÓN TEÓRICA Y METODOLÓGICA

---

El marco teórico en el que se sitúa la presente investigación establece, en primera instancia, una diferencia entre los términos *palabra* y *vocablo*, ya que en el lenguaje común suelen utilizarse indistintamente para referirse a los lexemas de una lengua, o bien a los que un hablante utiliza en sus actos de habla (Casadei, 2003). Sin embargo, en léxico-estadística se distingue entre el cómputo de las unidades del texto y el de las unidades del léxico: en el primer caso, se entiende por *tokens* todas y cada una de las unidades léxicas que aparecen en un corpus (o en un texto); en el segundo caso, se denomina *types* a cada una de las piezas léxicas distintas que se registran en una muestra, sin repeticiones. Si consideramos los procesos empíricos de medición de la disponibilidad léxica, las *palabras* son todas las unidades aportadas por los informantes en un test, repeticiones incluidas (*tokens*); mientras que los *vocablos* son todos los ítems diferentes arrojados, sin tener en cuenta las repeticiones (*types*).

El *léxico disponible* se refiere a las palabras que se activan en una situación comunicativa concreta y comprende, en la mayoría de los casos, lexías de contenido nocional, esto es, *palabras temáticas* (Michéa, 1953) que son lexemas concretos relacionados con el tema de la comunicación<sup>2</sup>. El concepto de *léxico disponible* nació en la segunda mitad del siglo XX para superar las limitaciones del criterio de frecuencia utilizado hasta la fecha en los procesos de

---

<sup>2</sup> Michéa (1953) lo distinguía del concepto de *palabras atemáticas*, a saber, lexemas gramaticales que aparecen en cualquier texto independientemente del tema comunicativo.

selección del léxico destinado al aprendizaje<sup>3</sup>. Una palabra *disponible* es un ítem que, sin ser necesariamente frecuente, está siempre listo para ser usado según el asunto de la comunicación: se trata de *tokens* que acuden más fácil y rápidamente a la mente de un hablante en una situación concreta y se caracterizan por ser ítems de alto contenido nocional-informativo (Borrego Nieto, 2004). En definitiva, se trata del vocabulario “que se usa en una situación comunicativa específica, actualizándose sólo ante la necesidad de comunicar algo referente a esa situación concreta” (López Morales, 1999, p. 9).

La disponibilidad léxica tiene como objetivo la recogida y el análisis del léxico disponible de una comunidad de habla, activado en función del tema comunicativo. Se trata de un análisis que pretende descubrir qué y cuántas palabras sería capaz de usar un individuo en una situación determinada por un cierto tema. El estudio pionero se llevó a cabo por parte de un grupo de lingüistas franceses en la segunda mitad del siglo XX para solucionar las limitaciones de la selección léxica fundada exclusivamente en el criterio de la frecuencia y culminó con la publicación del *Français Fundamental* (Gougenheim *et al.*, 1964). El documento se presentaba como una lista de palabras esenciales para permitir un rápido acceso al vocabulario de la lengua y se

---

<sup>3</sup> A partir de este criterio se seleccionaba el *léxico frecuente*, a saber, el conjunto de palabras que presentan el mayor número de ocurrencias en el recuento de una muestra lingüística, por lo que era el patrón que se percibía como más provechoso para una selección pedagógica adecuada del léxico hasta la primera mitad del siglo XX, ya que permitía establecer cuáles eran los vocablos más estables de un idioma y que, sin duda, habrían aparecido en cualquier contexto comunicativo independientemente del tema, por lo que se debían enseñar y aprender. Sin embargo, se reconsideró la exclusiva aplicación de este índice porque muchas de estas unidades frecuentes eran palabras funcionales del discurso, de poco o nulo contenido semántico. Por ende, basarse solo en este criterio no resultaba tan provechoso como para ofrecer al discente el caudal y el tipo de léxias necesarias para el adecuado aprendizaje de la lengua objeto, por lo que dejó de considerarse el más proficuo para una apropiada elección del vocabulario, pues la literatura crítica subrayó muy pronto que el *léxico frecuente* presentaba palabras descontextualizadas, procedentes exclusivamente de textos escritos que dejaban de lado el vocabulario más común propio de la comunicación hablada espontánea que necesitaban los aprendientes de una lengua extranjera (Gougenheim *et al.*, 1964, pp. 31-35).

consideró la obra de referencia para el aprendizaje del léxico francés L2/LE, propiciando los ítems más apropiados para los aprendientes extranjeros.

Todo esto supuso un punto de inflexión en el marco de la léxico-estadística y un proceso de reajuste metodológico que culminó en la planificación de un nuevo sistema de recogida de datos para la elaboración del léxico disponible de una comunidad: se organizaron pruebas escritas de tipo asociativo (aún vigentes en este ámbito investigativo)<sup>4</sup> suministradas a grupos de informantes que recopilaban listas de palabras relacionadas con ciertos estímulos temáticos, definidos *centros de interés* (Michéa, 1950, p. 189)<sup>5</sup>. Estos estímulos siguen teniendo validez, pese a que hayan sido cuestionados por la supuesta universalidad de los temas. Efectivamente, están sujetos a arbitrariedad según los objetivos propuestos en cada trabajo: es posible reformularlos, reducirlos o sustituirlos en virtud de los intereses investigativos del autor<sup>6</sup>.

Tras los estudios franceses, los primeros proyectos sobre la disponibilidad léxica en español –la lengua que más ha contribuido al desarrollo de la disponibilidad léxica– fueron elaborados a raíz del *Proyecto Panhispánico de Disponibilidad Léxica*<sup>7</sup> a partir de los años setenta del siglo pasado y estaban centrados en el léxico disponible de los hispanohablantes nativos, pero,

---

<sup>4</sup> Las pautas metodológicas aplicadas para la realización de las encuestas de disponibilidad léxica han ido evolucionando a lo largo del tiempo, ajustándose a los cambios inducidos por nuevas finalidades investigativas. Sin embargo, se ha intentado mantener siempre ciertas constantes procedimentales, ya que es importante respetar unos mínimos criterios para no impedir los cotejos entre diferentes trabajos.

<sup>5</sup> Los originales son: 01. *Les parties du corps*; 02. *Les vêtements (homme et femme)*; 03. *La maison (sans les meubles)*; 04. *Les meubles de la maison*; 05. *Les aliments et boissons des repas*; 06. *Les objets placés sur la table pour les repas*; 07. *La cuisine, ses meubles et ses ustensiles*; 08. *L'école, ses meubles et son matériel scolaire*; 09. *Le chauffage et l'éclairage*; 10. *La ville*; 11. *Le village ou le bourg*; 12. *Les moyens du transport*; 13. *Les travaux des champs et du jardinage*; 14. *Les animaux*; 15. *Les jeux et distractions*; 16. *Les différents métiers* (Gougenheim *et al.*, 1964, p. 152-153).

<sup>6</sup> Mencionamos, entre otros, a Moreno Fernández (2012), González Fernández (2014), Paredes (2014), Tomé Cornejo (2015), Sánchez-Saus (2016), Satos Díaz (2017), Hidalgo (2019), quienes defendieron esta tesis.

<sup>7</sup> Visítese la página [www.dispoplex.com](http://www.dispoplex.com) para más información sobre el proyecto y la bibliografía relacionada.

rápidamente, junto a esta línea de investigación se inició otra, muy fructífera, dedicada a aprendientes no nativos en muchas zonas del mundo donde el aprendizaje de ELE alcanza cifras altas de estudiantes<sup>8</sup>.

Hasta hoy se aprecia la publicación de numerosos proyectos sobre disponibilidad léxica de estudiantes de español no nativos cuyo objetivo prioritario es el análisis del léxico disponible de los encuestados. Sin embargo, como ya hemos mencionado, las pruebas ofrecen materiales que permiten plantear análisis con finalidades distintas<sup>9</sup>: (i) análisis del léxico disponible de los aprendientes en un determinado momento; (ii) definición de las etapas de aprendizaje; (iii) estudio de los errores, interferencias de la L1 y de otras LE conocidas; (iv) cotejo del vocabulario de los participantes con otros grupos de estudiantes y con los nativos; (v) detección de las unidades poco rentables y déficits léxicos; (vi) comparaciones culturales; (vii) análisis de la adecuación de los manuales y de otros materiales pedagógicos; (viii) eliminación del léxico superfluo de los contenidos didácticos y proposición de contenidos más apropiados; (ix) ampliación de la cobertura de los estudios hacia la terminología, es decir, hacia el español para fines específicos (Paredes, 2015).

Este trabajo se centra en el análisis de la situación de la ortografía de estudiantes extranjeros, que a pesar de su indudable interés y utilidad para el español LM (Paredes, 1999; Trigo, Romero & Santos, 2018, 2019, 2020), no ha

---

<sup>8</sup> En Italia, teniendo en cuenta los importantes resultados de tales trabajos y su extensión geográfica (China, Estados Unidos, Eslovenia, Finlandia, Japón solo por citar algunos), se empieza a dedicarle atención desde 2013 gracias a las contribuciones de Caggiula (2013), Rubio Sánchez (2017), Del Barrio (2016, 2017a, 2017b, 2018), Del Barrio y Vann (2018) y Nalesso (2018a, 2018b, 2019, 2020).

<sup>9</sup> Aunque parece que el provecho mayor se refiere a finalidades educativo-pedagógicas relacionadas con la didáctica de las lenguas, materna y extranjera, en lo que atañe a: (i) planificación lingüística; (ii) elaboración de materiales; (iii) análisis de tipo cognitivo; (iv) comparación de sintopías o de comunidades bilingües; (v) averiguación de las carencias de las propuestas didácticas; por otro lado, la disponibilidad léxica no está limitada a los estudios de lingüística y didáctica, sino que hay otras disciplinas que pueden emplear esta metodología con resultados interesantes (Nalesso, 2019, pp. 105-107). En efecto, su aplicación abarca distintos campos, entre los cuales destacamos investigaciones (i) sociolingüísticas; (ii) dialectológicas; (iii) psicolingüísticas; (iv) etnolingüísticas.

recibido todavía la atención que merece en el ámbito de ELE. De hecho, no tenemos constancia de trabajos sobre discentes no nativos que se fijen detenidamente en este asunto, con excepción de algunos autores (Carcedo, 1999; Frey, 2008; Sánchez-Saus, 2016; Hidalgo, 2019), que si bien han tratado la disortografía a partir de datos obtenidos para el análisis de la disponibilidad léxica no tenían el primero como objetivo último.

Nuestra investigación consiste, sin embargo, en el análisis de un corpus cacográfico constituido a partir de del análisis del léxico disponible de un grupo de estudiantes italianos de ELE. Para su realización nos basamos en Nalesso (2019)<sup>10</sup>, trabajo desarrollado según las pautas de Carcedo (2000) y Samper Hernández (2002)<sup>11</sup>.

El análisis que presentamos se inscribe, además, en el marco metodológico del *Análisis de errores*, entendido como un método empírico de investigación surgido como alternativa al *Análisis contrastivo*, que en los años setenta del siglo pasado proponía el examen de los errores cometidos durante el aprendizaje lingüístico de no nativos con el propósito de detectar tanto las causas como las estrategias utilizadas en el proceso adquisitivo<sup>12</sup>. Corder (1967), uno de sus mayores teóricos, estableció que las pautas metodológicas del análisis consistían en analizar las actuaciones lingüísticas de los encuestados (sin basarse en la comparación de la lengua materna y la lengua meta como solía hacerse en el *Análisis contrastivo*) según los siguientes pasos: (i) identificación de los errores en su contexto; (ii) clasificación y descripción de los errores; (iii) explicación de su origen, buscando los mecanismos o estrategias psicolingüísticas y las fuentes de cada error: en este punto entraría la

---

<sup>10</sup> Véase §3.

<sup>11</sup> Retomadas más recientemente por otros autores, entre los que destacamos, a modo de ejemplo, Sánchez-Saus (2016) e Hidalgo (2019).

<sup>12</sup> De esta última propuesta deriva el concepto de *interlengua*, acuñado más tarde por Selinker (1969, 1972), que se refiere al sistema lingüístico del aprendiz extranjero detectable en las diferentes etapas del proceso de adquisición lingüística.

interferencia de la lengua materna, como una ulterior estrategia; (iv) evaluación de la gravedad del error y búsqueda de un posible tratamiento.

Lo más destacable de esta teoría coincide con el cambio de visión del error que deja de ser considerado negativo y empieza a tratarse como factor útil para la comprensión de la evolución del aprendizaje:

[...]. Una de las contribuciones indiscutibles del AE se concreta en la despenalización del error, ya que el mismo constituye una fuente valiosa de información y es un indicador de que el proceso de aprendizaje se está llevando a cabo por parte del aprendiente. (Alexopoulou, 2010)

Para el análisis de nuestro corpus cacográfico, hemos partido de los seis criterios propuestos por Vázquez (1999) para la clasificación de los errores<sup>13</sup>:

1. criterio lingüístico: errores de adición, omisión, yuxtaposición, falsa colocación, falsa selección;
2. criterio etiológico: errores interlingüales, intralingüales, de simplificación;
3. criterio comunicativo: errores de ambigüedad, irritantes, estigmatizantes, de falta de pertinencia;
4. criterio pedagógico: errores inducidos vs. creativos, transitorios vs. permanentes, fosilizados vs. fosilizables, individuales vs. colectivos, residuales vs. actuales, congruentes vs. idiosincrásicos, de producción oral vs. escrita, globales vs. locales;
5. criterio pragmático: errores de pertinencia (o discursivos);
6. criterio cultural: errores culturales.

### 3. MÉTODO DE INVESTIGACIÓN

---

Exponemos ahora, brevemente, la metodología empleada en la parte empírica de la investigación detallando las fases iniciales del trabajo que nos han permitido, en un primer momento, recopilar el corpus y, posteriormente, proceder a nuestros análisis.

Para alcanzar los objetivos, hemos dado los siguientes pasos:

1. selección de la muestra;
2. establecimiento de las técnicas de elicitación de datos, esto es, diseño de la prueba de disponibilidad léxica y de los procesos de recogida;
3. administración de la prueba;

---

<sup>13</sup> Véase §3.2.

4. fijación de los criterios de edición y tratamiento del material, tanto del léxico disponible como del análisis de los errores;
5. despojo de los datos.

Las etapas sucesivas han consistido en los análisis y la discusión de los resultados, de los que damos cuenta en los apartados correspondientes.

### 3.1 MUESTRA Y RECOGIDA DE DATOS

La población estudiada es la de los estudiantes procedentes de los cursos de Lengua Española del DISLL<sup>14</sup> de la Universidad de Padua, de la cual se extrajo una muestra representativa de cien sujetos y a la que se suministró la prueba de disponibilidad léxica: se trata de aprendientes de ELE de nivel intermedio, B1 y B2 según el MCER.

Diseñamos la encuesta según el modelo tradicional (Samper Padilla, 1998) partiendo de un cuestionario que recogiese algunos datos personales para analizar el grupo de participantes según tres variables sociolingüísticas<sup>15</sup> y de los dieciséis centros de interés clásicos, a pesar de que la bibliografía crítica y los estudios de campo hayan etiquetado algunos como poco rentables para cumplir con las necesidades y los intereses culturales y comunicativos de los discentes del siglo XXI<sup>16</sup>:

---

<sup>14</sup> *Dipartimento di Studi Linguistici e Letterari* (<https://www.disll.unipd.it/>).

<sup>15</sup> Se trata de (i) sexo; (ii) nivel de ELE; (iii) conocimiento de otras lenguas extranjeras, que no tendremos en cuenta ahora por estar fuera de los intereses del presente estudio.

<sup>16</sup> Es decir, incluimos en el test todos los estímulos temáticos originales mencionados arriba, propuestos en el *Français Fundamental* (Gougenheim *et al.*, 1964), sin eliminar los menos productivos, pese a que, corroborando las posturas defendidas por la literatura ya citada, los análisis realizados en Nalesso (2019) han puesto de manifiesto su escasa rentabilidad. En concreto, destacamos un grupo de tres campos que no alcanza las 400 unidades léxicas: C107, “la cocina y sus utensilios”; C109, “iluminación y calefacción”; C113, “trabajos del campo y del jardín”, frente a una mayor capacidad productiva de otros como, por ejemplo, C105, “alimentos y bebidas”; C101, “partes del cuerpo”; C110, “la ciudad”; C114, “los animales”, que superan los 1.000 lexemas, siendo los más productivos en absoluto para nuestros informantes. Sobre la cuestión, a modo de ejemplo, cabe recordar asimismo que Sánchez-Saus (2016) propone la aplicación de los contenidos que el MCER, *Marco común europeo de referencia para las lenguas*, (2002) recomienda para el nivel A1 en el proceso de selección de los centros de interés aplicables en las pruebas de disponibilidad léxica. Santos Díaz (2017) se basa, por su parte, en las nociones específicas del *Plan curricular del Instituto Cervantes* (2006) constituyendo el fundamento tanto para la elaboración de materiales como para la evaluación y certificación lingüística oficial de ELE.

- ci01, “partes del cuerpo” [CUE];
- ci02, “la ropa” [ROP];
- ci03, “partes de la casa (sin muebles)” [CAS];
- ci04, “los muebles de la casa” [MUE];
- ci05, “alimentos y bebidas” [ALI];
- ci06, “objetos colocados en la mesa para la comida” [MES];
- ci07, “la cocina y sus utensilios” [COC];
- ci08, “la escuela: muebles y materiales” [ESC];
- ci09, “iluminación y calefacción” [ILU];
- ci10, “la ciudad” [CIU];
- ci11, “el campo” [CAM];
- ci12, “medios de transporte” [TRA];
- ci13, “trabajos del campo y del jardín” [TRC];
- ci14, “los animales” [ANI];
- ci15, “juegos y distracciones” [JUE];
- ci16, “profesiones y oficios” [PRO].

Los folios de la encuesta que contenían el test asociativo se repartían en dieciséis columnas que los sujetos tenían que rellenar en relación con cada uno de los estímulos propuestos. Cada columna permitía un número máximo de aportaciones, pero para trabajar con listas abiertas explicamos que se podía utilizar el reverso de la hoja si la ficha no era suficiente. Los participantes respondieron a la encuesta de acuerdo con las pautas habituales en los trabajos de disponibilidad léxica, a saber, completando las fichas con las unidades léxicas que se les ocurrían según los centros (dos minutos para cada uno), empezando por la primera línea, de manera que los ítems más disponibles son los que aparecen en las primeras posiciones de las listas. Además, aconsejamos a los estudiantes que escribieran todas las palabras que les venían a la mente, aunque no estuviesen seguros de su grafía o de su significado. Como suele ser habitual, se recordó que la prueba era anónima y que no se habría tenido en cuenta para la evaluación del curso.

Se administró la encuesta durante una clase cualquiera sin avisar previamente al estudiantado para evitar ansiedades o eludir el absentismo y con el siguiente protocolo estándar:

1. presentación del investigador y del estudio;
2. explicación de los contenidos y de las instrucciones;

3. aclaración del anonimato;
4. administración de las hojas;
5. realización de la prueba.

### 3.2 EDICIÓN Y RECOPIACIÓN DEL CORPUS CACOGRÁFICO

Para el presente estudio no hemos seguido las pautas que suelen emplearse en disponibilidad léxica en lo referente la edición de los datos (Samper Hernández, 2002)<sup>17</sup>, cuyo objetivo es homogenizar el material recogido, eliminando las palabras inexplicables, inexistentes o indescifrables para facilitar el cálculo informático, ya que la finalidad perseguida era, precisamente, la constitución del corpus cacográfico.

Para la sistematización de los resultados hemos tenido en cuenta los siguientes criterios, basados, como ya se ha dicho, en Vázquez (1999): en primer lugar, el criterio lingüístico incluye los errores que podemos definir *de superficie*, a saber, las faltas de tipo fonológico, ortográfico, morfosintáctico, léxico, semántico y pragmático-discursivo que, como se verá, están a la base de nuestros análisis. Por su parte, el criterio etiológico se refiere a las causas del error: (i) la interferencia lingüística tanto de la lengua madre como de otros idiomas conocidos (errores interlingüales), (ii) la transferencia intralingüística debida a una generalización de las reglas de la lengua objeto, esto es, errores que derivan del conflicto interno de sus reglas por ser errores evolutivos, o de desarrollo, procedentes de estrategias universales que el alumno aplica en su proceso adquisitivo y por lo tanto independientes de la LM (errores intralingüales). Por último, (iii) la simplificación de estructuras, que supone la neutralización de los elementos que se alejan de la regla aprendida: según

---

<sup>17</sup> Para el análisis del léxico disponible, antes del tratamiento informático de los datos a través del software *Dispogen II* (Echeverría *et al.*, 2005), editamos el material conforme a los siguientes patrones: (i) eliminación de los términos repetidos y de los artículos; (ii) corrección de la ortografía; (iii) unificación ortográfica; (iv) unificación de variantes meramente flexivas; (v) unificación de derivados regulares; (vi) unificación de formas plenas y acortamientos; (vii) eliminación de los préstamos y de las interferencias de otras lenguas; (viii) eliminación de las marcas comerciales; (ix) amplitud de las relaciones asociativas.

la propuesta de Alexopoulou (2010) se trata de la reducción a un sistema simplificado (simplificación, neutralización, aplicación incompleta de las reglas, evasión) y de generalización (hipergeneralización, analogía, asociación cruzada, hipercorrección, ignorancia de las restricciones de las reglas). En efecto, los errores responden a varias causas que no consisten exclusivamente en la distracción, falta de esfuerzo, ignorancia de la norma, etc.

No hemos considerado los criterios comunicativo, pragmático y cultural porque no los hemos considerado relevantes para este tipo de investigación, si bien es notoria su importancia para cuestiones relacionadas con el desarrollo de una adecuada competencia comunicativa que posibilita al aprendiz realizar eductos (*output*) correcta y eficazmente, esto es, alcanzar la finalidad última del aprendizaje lingüístico (MCER, 2002).

En lo que se refiere al criterio pedagógico, conviene recordar que observamos solo errores de producción escrita, y no oral, debido a la manera de realización de la prueba a la que se sometió el muestreo.

Así pues, siguiendo la línea de investigación de los ya citados Paredes (1999), Trigo, Romero & Santos (2018, 2019, 2020) los análisis se han basado en la tipología de error y en su frecuencia de aparición en el léxico disponible de los encuestados, considerando (i) los errores ortográficos relacionados con el uso o la mala colocación de vocales, consonantes o tildes; (ii) los errores fonéticos ligados a cambios vocálicos o cambios consonánticos; (iii) los errores morfológicos nominales, verbales o pertenecientes a las locuciones preposicionales. Asimismo, tuvimos en cuenta los sinónimos, los extranjerismos y las falsas formaciones. Mención aparte merecen los italianismos que se han repartido en dos sub-categorías: (i) italianismos crudos, a saber, términos italianos sin modificación y (ii) calcos del italiano, esto es, palabras italianas adaptadas al español con alguna modificación. Por último, recogimos bajo la etiqueta *otros* las palabras inventadas y las unidades léxicas sin relación con el centro de interés en cuestión.

#### 4. DISCUSIÓN DE LOS RESULTADOS

Para poder procesar adecuadamente los resultados del estudio cuantitativo, presentamos en primer lugar los datos generales procedentes de los cómputos finalizados al análisis de la disponibilidad léxica de los que partimos para el desarrollo del presente trabajo. Empezamos por el número total de palabras y de vocablos activados, sus respectivos promedios por informante y por centro de interés, sin tener en cuenta los índices de cohesión y de densidad léxica por no ser útiles en esta ocasión.

Según los recuentos, las palabras activadas durante la prueba, esto es, el léxico disponible de los encuestados, son 12.579, lo cual supone un promedio de 126 *tokens* por informante y de 786 por centro. Por su parte, los vocablos rastreados son 1.490, lo que implica un promedio aproximativo de 15 *types* por sujeto y 93 por tema.

Ahora bien, si extraemos los datos derivados del análisis de los errores, es posible observar en la Tabla 1 que, en el total de *tokens* arrojados, 1.452 eran erróneos, lo cual representa un porcentaje de error del 11,54%. A partir de estos valores, es posible contabilizar una media de 14 unidades erróneas por cada sujeto y 91 por cada campo semántico.

Tipo de error	Palabras	
	Frecuencia de aparición	% del total
Errores ortográficos	480	33,06%
Errores fonéticos	171	11,78%
Errores morfológicos	106	7,30%
Sinónimos	39	2,69%
Italianismos	282	19,42%
Extranjerismos	112	7,71%
Falsas formaciones	163	11,23%
Otros	99	6,82%
Número total de errores	1.452	11,54%

TABLA 1. Número total de palabras con errores

El Gráfico 1 ilustra la situación recién descrita y pone de relieve cómo esta baja aparición de errores no perjudica el conocimiento del vocabulario de los informantes y en consecuencia su competencia léxica (considerando, además, el tipo de error que se analizará a continuación).

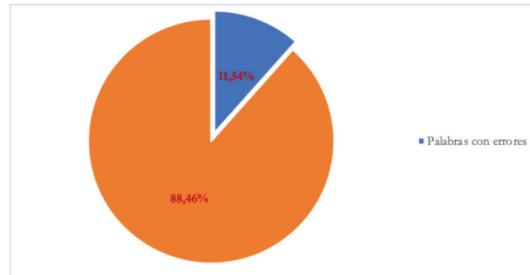


GRÁFICO 1. Porcentaje de palabras con errores sobre el total

Al desglosar este 11,54% en función de los tipos de error contemplados podemos analizar la distribución de los errores dentro del corpus a través del siguiente gráfico (Gráfico 2).

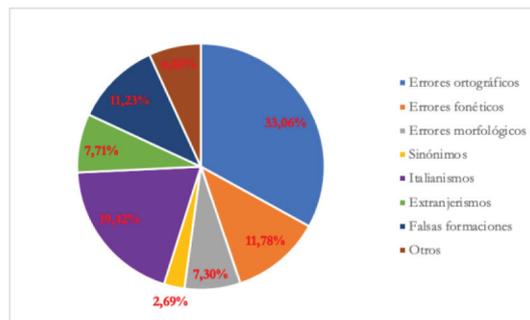


GRÁFICO 2. Porcentaje de errores según categoría

El mayor número de faltas forma parte de la categoría lingüística de Vázquez (1999). En esta categoría, las más numerosas son de tipo ortográfico, ya que constituyen un 33,06%, es decir, 480 en el total de 1.452. En segundo lugar, destaca el 19,42% de italianismos, atribuible a la afinidad lingüística entre la LM y la LE por la poca distancia léxica que caracteriza los lexemas (Calvi, 1995, pp. 86-90).

En cuanto a las faltas de tipo fonético y de falsa formación, representan respectivamente un 11,78% y un 11,23%, probablemente causados por una distracción: se han encontrado solo en algunas ocasiones mientras que en la mayoría de los casos las mismas palabras eran correctas. Los errores morfológicos (106 ítems), las palabras inventadas –errores creativos, según el criterio pedagógico– o las que no tenían relación con los centros de interés –errores de pertinencia, según el criterio pragmático– (99 lexemas) y los extranjerismos (112 unidades) aparecen unas 100 veces, llegando aproximadamente al 7% cada uno.

A este propósito, la presencia de errores interlinguales e intralinguales (extranjerismos y palabras inventadas, además de los calcos o préstamos del italiano) es causa de la interferencia de otras lenguas extranjeras conocidas, en particular del inglés y del francés, los dos idiomas más estudiados en el contexto académico de la muestra. Si sumamos las ocurrencias de las faltas causadas por las interferencias de la LM y de otras LE, que encajan tanto en los criterios pedagógico y etiológico como en el lingüístico<sup>18</sup>, podemos observar que presentan un alto porcentaje de error. De todo ello parece deducirse que las categorías de error más frecuentes son las lingüísticas e intralinguales o interlinguales a la vez. Según el criterio etiológico, esto depende de la organización del lexicón mental de un aprendiz extranjero, que favorecería la recuperación de material lingüístico del idioma nativo o de cualquier otro como estrategia de aprendizaje, denominada transferencia. Análogamente, la lingüística contrastivo-perceptiva considera que los procesos cognitivos del aprendiente sumados a su experiencia personal se reflejarían en el lexicón de la lengua objeto (Di Gesù, 2016).

Por último, los sinónimos componen el grupo más reducido siendo en total 39, que traducido en porcentaje se refiere al 2,96% del conjunto analizado.

---

<sup>18</sup> Como veremos en la exposición de las decisiones que hemos tomado durante la edición de las respuestas finalizada a la recopilación del corpus cacográfico por centro de interés.

Un desglose más fino de estos datos permite la distribución de estas categorías generales en subclases. Para ello, como muestra la Tabla 2, hemos eliminado los resultados correlacionados con los extranjerismos, las faltas causadas por falsa formación, los sinónimos y los errores agrupados en la categoría *otros*. De esta forma podemos observar los errores que cubren el 71,55% del corpus cacográfico.

Tipo de error		Palabras		
		Frecuencia de aparición	% respecto a su tipo	% del total
Errores ortográficos	Vocal	65	13,54%	4,48%
	Consonante	114	23,75%	7,85%
	Tilde	301	62,71%	20,73%
Errores fonéticos	Vocal	41	23,98%	2,82%
	Consonante	130	76,02%	8,95%
Errores morfológicos	Nominal	82	77,36%	5,64%
	Verbal	2	1,89%	0,14%
	Locución preposicional	22	20,75%	1,52%
Italianismos	Crudos	227	80,50%	15,63%
	Calcos	55	19,50%	3,79%

**TABLA 2.** *Desglose de los errores en subclases*

Con respecto al total de errores, los más frecuentes en absoluto con un 20,73% (301 apariciones) son los errores ortográficos relacionados con el uso de la tilde, a saber, por presencia o ausencia erróneas o porque está mal colocada. Siguen 227 italianismos crudos (15,63%), mientras que los calcos solo son 55 (3,79%). Asimismo, se han localizado más de 100 veces cada uno los errores de letra consonante, tanto fonéticos como ortográficos (8,95% y 7,85%). Por su parte, las vocales erróneas son 65 (4,48%) en la categoría ortográfica y 41 (2,82%) en la fonética, es decir, que los estudiantes se equivocan más en el uso de las consonantes posiblemente por analogía, hipercorrección o ignorancia de las restricciones de las reglas. En lo que se refiere a la morfología, los errores más difundidos atañen a la clase nominal (82 apariciones, 5,64%), seguidos por las locuciones pronominales (22 apariciones, 1,52%) y, por último, los verbos (2 apariciones, 0,14%), con muy pocos errores.

El Gráfico 3 ilustra la distribución de los errores relacionados con la grafía: más de la mitad de los errores afecta al uso de las tildes. En efecto, constituyen casi el 63% de su categoría, sobre todo en lo que atañe a la omisión y a la mala colocación: por ejemplo, *\*algodon*, *\*pantalon*, *\*habitacion*, *\*estanteria*, *\*oido*, *\*idraulico*, *\*naríz*, *\*pié*. Hemos encontrado en menor medida faltas relacionadas con las consonantes, que no llegan al 8%, como *\*garage*, *\*guchara*, y aún menos con las vocales, como *\*acetunas*, *\*menzana*, que por poco no alcanzan el 5%.

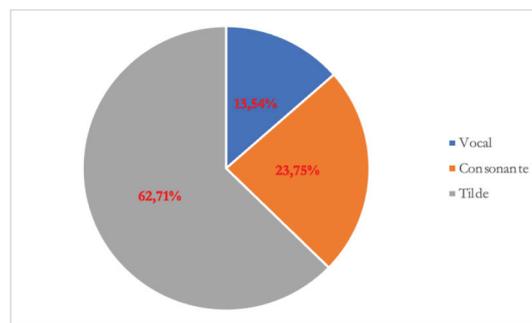


GRÁFICO 3. Porcentaje de errores ortográficos respecto a su tipo

Las dificultades mayores en lo que se refiere a los errores fonéticos afectan mucho más a las consonantes (76%) que a las vocales, (24%). En concreto, hemos detectado una gran confusión entre las letras /b/ y /v/ y entre /c/, /s/ y /z/: *\*tovillo*, *\*bever*, *\*carabela*, *\*camiceta*, *\*calzetín*, *\*gaspacho*, *\*caveza*, *\*cabeza*. Es posible que los errores respondan a varias causas relacionadas más con la ignorancia de la norma o la hipergeneralización que con la distracción o falta de esfuerzo (Gráfico 4).

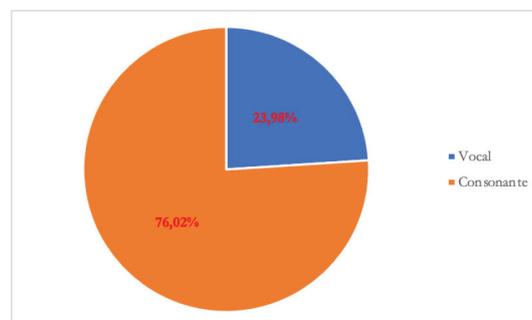


GRÁFICO 4. Porcentaje de errores fonéticos respecto a su tipo

Desde el punto de vista de la morfología (Gráfico 5), la clase nominal es la que más problemas causa a los aprendientes (77%) tanto en lo que se refiere al género (*\*pechugo de pollo*, *\*aire acondicionada*) como al número (*\*red sociales*, *\*dibujo animados*); las locuciones preposicionales no representan un número elevado de errores (21%), por ejemplo, *\*camiseta a mangas cortas*, *\*pantalón a campana*, *\*ir a compras*, y los verbos aún menos (2%), pero en este cómputo ha de ser relativizado porque el método de análisis del léxico disponible de los encuestados arroja una presencia abrumadora de sustantivos con respecto a las demás clases de palabras.

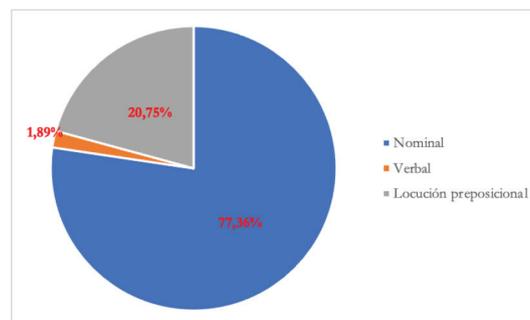


GRÁFICO 5. Porcentaje de errores morfológicos respecto a su tipo

Por último, con el 80,50% se ha manifestado significativa la aparición de italianismos crudos (Gráfico 6), a saber, la activación directa de términos italianos no modificados, con respecto a los calcos, es decir palabras italianas adaptadas a la lengua española (19,50%). A modo de ejemplo, aparecen las siguientes interferencias con la primera lengua de los participantes: *\*collo* (cuello), *\*frulador* (batidora), *\*impiegado* (empleado), *\*infermer* (enfermero), *\*anafiar* (regar), *\*seminar* (sembrar), *\*potar* (podar), *\*bonificar* (sanear)<sup>19</sup>.

<sup>19</sup> La mayoría de estas faltas se halla entre las respuestas aportadas en el C13, “trabajos del campo y del jardín”, confirmando la poca rentabilidad del estímulo para los estudiantes de ELE porque parece que han activado una cantidad elevada de palabras inventadas, a partir del italiano, a fin de escribir unidades léxicas relacionadas con este centro sin dejar vacía la lista.

Además, han aparecido falsos amigos como \**gamba* > *pierna*, \**autista* > *chofer*, \**tienda* > *cortina*, \**taberna* > *sótano*.

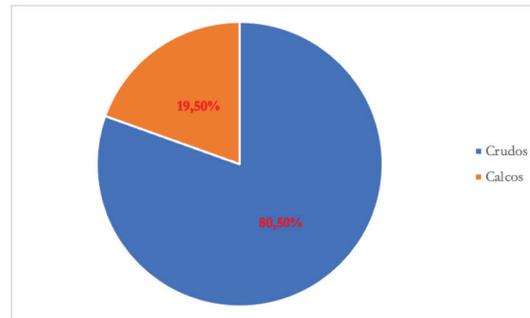


GRÁFICO 6. Porcentaje de italianismos respecto a su tipo

Teniendo en cuenta la capacidad productiva de los centros de interés (Tabla 3), analizamos los errores más frecuentes que han aparecido durante la edición de las respuestas, que (i) hemos corregido o eliminado a la hora de recopilar el diccionario de léxico disponible (Nalesso, 2019) y (ii) hemos utilizado para la creación del corpus cacográfico finalizado al desarrollo de la presente investigación. Con este ulterior análisis, basado en ejemplos representativos de cada uno de los centros de interés<sup>20</sup>, sin tener en cuenta la distribución exacta de errores, pretendemos averiguar si los campos semánticos menos productivos son, al mismo tiempo, los que presentan un porcentaje mayor de error. Es decir, examinamos si existe una correlación entre los que registran mayor disortografía y los más o menos productivos, como propusieron Trigo, Romero & Santos (2018).

---

<sup>20</sup> Por cuestiones de espacio no presentamos todos los errores lingüísticos (ortográficos, fonéticos, morfológicos) ni los derivados de la transferencia lingüística de la LM u otras LE que hemos señalado como los más recurrentes.

SECCIÓN: E/L2  
 DISPONIBILIDAD LÉXICA Y ORTOGRAFÍA EN ELE:  
 UN ESTUDIO PARA LA ENSEÑANZA DE LA LENGUA  
 Giulia Nalesso

Rango	CI	Palabras	Promedio por informante/CI
2	CI01.CUE	1.341	13,41
8	CI02.ROP	791	7,91
12	CI03.CAS	610	6,10
11	CI04.MUE	660	6,60
1	CI05.ALI	1.463	14,63
13	CI06.MES	530	5,30
14	CI07.COC	365	3,65
7	CI08.ESC	803	8,03
15	CI09.ILU	342	3,42
3	CI10.CIU	1.171	11,71
10	CI11.CAM	685	6,85
6	CI12.TRA	808	8,08
16	CI13.TRC	252	2,52
4	CI14.ANI	1.080	10,80
9	CI15.JUE	784	7,84
5	CI16.PRO	894	8,94
Total		12.579	7,86

**TABLA 3.** Productividad de los CI, número total de palabras y media por informante

Como ilustra la tabla y el apéndice con el comentario analítico de los errores de cada centro, parece que no hay una relación asociativa entre la mayor o menor rentabilidad de los centros de interés y la mayor o menor aparición de errores, pues se han cometido en todos indistintamente. En efecto, cada campo presenta cierta variedad de error y de los recuentos sobresale que el uno no supera de mucho el otro en lo referente la cantidad de palabras erróneas que contiene, a pesar del total de *tokens*.

## 5. CONCLUSIONES

En esta investigación hemos desarrollado un estudio que, a partir de la recopilación de un corpus cacográfico, ha examinado los errores cometidos por una muestra de estudiantes italianos de nivel intermedio (B1-B2) al realizar una prueba de disponibilidad léxica, diseñada según el modelo tradicional (Samper Padilla, 1998).

Hemos analizado la disortografía en el léxico disponible a través del *Análisis de errores* (Paredes, 1999; Trigo, Romero & Santos, 2018, 2019, 2020) porque la aplicación de la disponibilidad léxica en ELE ofrece materiales que permiten plantear estudios aplicables a un ámbito más práctico de la didáctica de la lengua. Para ello, después de la recogida del léxico disponible (Nalesso, 2019), hemos recopilado y analizado un corpus cacográfico que refleja la situación ortográfica de los participantes en la prueba y que ha llevado a averiguar que (i) los informantes no han activado un número muy elevado de palabras erróneas, el 11,54% sobre el total de *tokens* que componen su léxico disponible (1.452 errores frente a 12.579 palabras totales); (ii) los errores más frecuentes no perjudican el conocimiento del vocabulario.

En efecto, de las observaciones procedentes del análisis realizado por medio de la edición de cada centro de interés ha surgido que los errores colectivos más frecuentes abarcan los criterios (i) lingüístico (errores de adición, omisión, yuxtaposición, falsa colocación, falsa selección); (ii) etiológico (errores interlingüales, intralingüales); (iii) pedagógico (errores creativos, fosilizables, colectivos y obviamente de producción escrita). En detalle, las cantidades mayores forman parte de la categoría lingüística de Vázquez (1999), sobre todo en lo referente a las faltas ortográficas, que constituyen el 33,06% de la totalidad de errores cometidos. Se trata especialmente de errores ligados al uso de la tilde (por presencia o ausencia erróneas o porque está mal colocada). Las faltas de falsa formación que cubren el 11,23% del corpus cacográfico se deben probablemente a una distracción momentánea, ya que se han encontrado solo en algunas ocasiones, mientras que en las demás apariciones las mismas palabras eran correctas. Encaja en la categoría etiológica el 19,42% de los italianismos, en mayor medida crudos (80,50%) frente a los calcos (19,50%), posiblemente derivados de la cercanía lingüística que une la LM y la LE de los encuestados. Dentro de la misma categoría, los extranjerismos y las otras interferencias lingüísticas, a saber, los errores

interlingüales e intralingüales están igualmente representados. En particular hemos detectado interferencias del inglés y del francés al ser los dos idiomas más estudiados por los participantes.

En definitiva, las ocurrencias mayores de error están representadas por faltas causadas por transferencia y por los errores lingüísticos (ortográficos de tilde en particular, que constituyen casi el 63% respecto a su tipo). Por su parte, los errores fonéticos afectan más al uso de las consonantes que a las vocales. En concreto, hemos detectado una gran confusión entre las letras /b/ y /v/ y entre /c/, /s/ y /z/.

De todos modos, consideramos que no se trata de problemas graves, no afectan negativamente al conocimiento del vocabulario –no causan malentendidos comunicativos– aunque es imprescindible que el profesorado los tenga en cuenta para que no se fosilicen, por lo que sugerimos que sería necesaria una intervención didáctica específica. En este sentido, los docentes podrían servirse de los resultados de su propio estudiantado para plantear actividades de refuerzo *ad hoc* que solucionen los problemas detectados.

Pese a que el objetivo de la disponibilidad léxica no persiga el estudio de la disortografía, creemos que los datos aportados por encuestas de este tipo son de gran utilidad para la enseñanza de ELE, ya que permiten percatar las tipologías de los errores, su frecuencia y su ámbito de aparición para la detección de los déficits de los aprendientes y de carencias de las propuestas formativas. Todo esto permitiría avanzar en el terreno de la didáctica de la ortografía desde planteamientos innovadores abriendo el camino a nuevos estudios en el campo de la disponibilidad léxica en ELE. Esperamos poder despertar la atención de otros investigadores y fomentar una corriente de estudios basados en esta metodología que proporcione datos aplicables en el ámbito más práctico de la didáctica examinando la situación ortográfica reflejada en las pruebas de disponibilidad.

## APÉNDICE

---

### ci01, “partes del cuerpo”

Se trata de un centro muy productivo (rango 2) en el que se detectan algunos errores de acentuación o de elección falsa de consonante (\**oido*, \**musculo*, \**naríz*, \**tovillo*) y extranjerismos (*mustache*), palabras inventadas (\**cavil*, \**piesdes*) o adaptadas al español (\**orillas*, posiblemente del francés *oreille*; \**gamba* y \**unghia*, del italiano).

### ci02, “la ropa”

Este campo, que supera de poco el promedio de palabras (786) presentando en total 791 *tokens*, revela pocos casos de errores ortográficos (\**cinturon*, \**calzetines*, \**camiceta*, \**jersei*, \**falta*), algunas formas inventadas (\**jaqueta*, \**escarfa*, posiblemente del inglés *scarf*) y extranjerismos (*boxer*, *t-shirt*, *slip*, *collant*, *sweater*).

### ci03, “partes de la casa (sin muebles)”

En *partes de la casa* (rango 12 por su escasa productividad de palabras) cabe señalar algunos errores interlinguales ya citados arriba, \**tienda* > *cortina*, \**taberna* > *sótano*.

### ci04, “los muebles de la casa”

Estamos ante un centro ligeramente más rentable con respecto al anterior (rango 11) en el que la mayor parte de los errores lo constituyen las faltas de acentuación (\**estanteria*, \**lampara*, \**frigorifico*), lo cual apuntaría a su consideración como errores colectivos según el criterio pedagógico debido al elevado número de ocurrencias en el corpus.

### ci05, “alimentos y bebidas”

En este campo semántico, el más productivos en absoluto, los errores más frecuentes son faltas ortográficas que no perjudican el conocimiento del vocabulario: *\*lemón*, *\*gaspacho*, *\*jamon*, *\*salsicha*.

Hay errores ligados al uso de la tilde: por presencia o ausencia erróneas (*\*álcachofa*, *\*céreales*, *\*esparrago*, *\*platano*, *\*jamon*) o porque está mal colocada (*\*ázucar*, *\*sándia*, *\*júdias*) corroborando de nuevo el hecho de que estos son errores colectivos según el criterio pedagógico.

Asimismo, merece la pena subrayar la gran presencia de extranjerismos, como *riz*, *sushi*, *hamburger*, *cocktail*, *hot-dog*, *gin*, *spaghetti* (algunos tendrían su adaptación al español: *hamburguesa*, *cóctel*, *perrito caliente*, *ginebra*, *espagueti*). En lo que se refiere a los préstamos italianos *pizza* y *pasta*, hemos considerado error el primero por aparecer en cursiva en el DRAE, denotando su naturaleza de barbarismo, mientras que hemos aceptado como correcto el segundo al estar totalmente incorporado en castellano (no aparece en cursiva en el diccionario).

CI06, “objetos colocados en la mesa para la comida”

Los resultados de este centro, donde no se han rastreado muchos errores significativos, indican una escasa productividad de palabras (rango 13) y ningún dato que merece la pena comentar.

CI07, “la cocina y sus utensilios”

En *La cocina y sus utensilios* los errores más comunes encontrados dentro del conjunto de las 365 palabras que forman el léxico disponible de los informantes (rango 14) atañen a la interferencia lingüística con otras lenguas, pues hemos detectado tanto unidades léxicas extranjeras y calcos (*freezer*, *moka*, *pentola*, *mixer*) como formas inventadas (*\*cucino*), a saber, errores creativos.

CI08, “la escuela: muebles y materiales”

En este campo (posición 7 en función de su productividad), solo hemos hallado el extranjerismo inglés *post-it* y las interferencias *\*cantina* y *\*lavaña*, posiblemente del francés *cantine* y del italiano *lavagna*.

ci09, “iluminación y calefacción”

Este estímulo, siendo muy poco rentable para los aprendientes, se posiciona penúltimo en función de su producción de palabras. Los errores lingüísticos aparecidos se refieren a la ausencia de tilde *\*lampara*, *\*neon*, *\*calorífico*, o bien a la morfología nominal *\*aire acondicionada*, que coincidiría con la categoría de error de falsa selección o colocación.

ci10, “la ciudad”

El centro, rango 3 con 1.171 *tokens*, se caracteriza por la actualización de formas inventadas, es decir, errores creativos si tenemos en cuenta el criterio pedagógico, y extranjerismos, que pertenecen a la categoría del error interlingual en virtud del criterio etiológico. Sin embargo, hemos aceptado el anglicismo españolizado *esmog* (considerando error el barbarismo *smog*).

ci11, “el campo”

Se trata de una de las áreas léxicas menos productivas (rango 10), con un carácter asociativo muy disperso. En consonancia con los criterios expuestos, hemos clasificado como error los extranjerismos *contadino* y *prato* del italiano y *arbre* del francés.

ci12, “medios de transporte”

Otra vez, hemos encontrado varios extranjerismos: *skateboard*, *camper*, *roulotte*, *scooter* (tendría su versión adaptada, *escúter*), *jeep* (sería *todoterreno* en español). Además, según el criterio lingüístico hemos notado la presencia de errores ortográficos relacionados con la falsa colocación de la tilde y la falsa selección de vocal y consonante: *\*tandem*, *\*pié*, *\*pasejero*, *\*caravela*.

Nos parece oportuno destacar que no se ha revelado como uno de los centros más rentables (rango 6) a pesar de que creíamos que se trataba de un *input* mucho más provechoso.

CI13, “trabajos del campo y del jardín”

Como ya se ha subrayado, sobresale en este campo, que se ha revelado como el menos productivo en absoluto, una gran cantidad de palabras inventadas, calcos y falsos amigos procedentes del italiano, como *\*potar*, *\*potación*, *\*bonificar*, *\*anafiar*, *\*seminar*, que posiblemente los discentes han intentado proporcionar para no dejar en blanco la ficha correspondiente.

CI14, “los animales”

El área léxica *Los animales* se ha manifestado muy productiva (rango 4) y no ha supuesto grandes esfuerzos de edición: como en lo demás estímulos, hemos encontrado algunos extranjerismos de la LM y de otras LE (*aquila*, *pesce*, *elephant*).

CI15, “juegos y distracciones”

En este campo hemos percatado numerosas piezas léxicas del inglés consideradas errores en este trabajo (*baseball*, *football*, *volleyball*, *ping pong*, *ballet*, *rugby*, *computer*, *sport*, *social network*) y una del francés (*sortir*), invenciones y transferencias (*\*escondido*, *\*mosca ciega*, *\*luna parque*). Hemos admitido como forma correcta los vocablos adaptados al español *beisbol* y *voleibol* en la variante sin acento siendo lo más frecuente en el léxico disponible de los estudiantes.

En nuestra opinión habría podido ser un estímulo que permite la activación de un elevado número de palabras debido a la amplitud de asociaciones posibles, pero, al contrario, se ha colocado en el rango 9 por debajo del promedio computado con 784 *tokens* totales.

C116, “profesiones y trabajos”

El centro es bastante rentable, aunque no llega a los 1.000 lexemas (rango 5). En lo que se refiere a los errores más comunes, en primera instancia, según el criterio lingüístico, hemos rastreado faltas ortográficas relacionadas con el uso de la letra /h/ (\**idráulico*, \**befe*) y faltas de cambio consonántico con respecto a la unidad léxica *abogado* que apareció como \**avocado*, \**avogado*, \**abojado*, \**abocado*. Segundo, según el criterio etiológico hemos detectado extranjerismos, falsos amigos e invenciones: *hostess*, *influencer*, *autista*, \**judice*, \**regista*, \**impiegado*, \**infermer*.

## BIBLIOGRAFÍA

- Alexopoulou, A. (2010). Errores intralinguales e interlinguales en la interlengua escrita de aprendientes griegos de E/LE. En P. Civil & F. Crémoux. (coords.), *Nuevos caminos del hispanismo...: actas del XVI Congreso de la Asociación Internacional de Hispanistas. París, del 9 al 13 de julio de 2007 / Vol. 2, [(CD-ROM)]*.
- Borrego Nieto, J. (2004). Algunas preguntas en relación con el concepto de léxico disponible. En M.<sup>a</sup> V. Galloso Camacho & J. Prado Aragonés (coords.), *Diccionario, léxico y cultura* (pp. 59-69). Universidad de Huelva.
- Caggiula, S. (2013). *El español como lengua extranjera: un estudio de disponibilidad léxica y su aplicación a la enseñanza*. Lulu Press.
- Calvi, M.<sup>a</sup> V. (1995). *Didattica di lingue affini. Spagnolo e italiano*. Guerini scientifica.
- Carcedo González, A. (1999). Análisis de errores léxicos del español en la interlengua de los finlandeses. En T. Jiménez Juliá, M. C. Losada Aldrey, J. F. Márquez Caneda & S. Sotelo Docío (eds.), *Español como Lengua Extranjera: Enfoque Comunicativo y Gramática. Actas del IX Congreso Internacional de ASELE, Santiago de Compostela 23-26 de septiembre de 1998* (pp. 465-472). Universidad de Santiago de Compostela. [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/09/09\\_0468.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/09/09_0468.pdf).
- Carcedo González, A. (2000). *Disponibilidad léxica en español lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Turun Yliopisto.
- Casadei, F. (2003). *Lessico e semantica*. Carocci.
- Consejo de Europa (2002). *Marco común europeo de referencia para las lenguas: Aprendizaje, Enseñanza, Evaluación (MCER)*. Traducción del Instituto Cervantes. Ministerio de Educación, Cultura y Deporte.
- Corder, S. P. (1967). The significance of learners' errors. *IRAL*, 5(4), 161-170. <https://doi.org/10.1515/iral.1967.5.1-4.161>.

- Del Barrio de la Rosa, F. (2016). Algunas observaciones sobre la disponibilidad léxica en estudiantes itálofonos de español. En E. Sainz González, I. Solís García, F. Del Barrio de la Rosa & I. Arroyo Hernández (eds.), *Geométrica explosión. Estudios de lengua y literatura en homenaje a René Lenarduzzi* (pp. 127-143). Edizioni Ca' Foscari. <https://edizionicafoscari.unive.it/media/pdf/books/978-88-6969-080-8/978-88-6969-080-8-ch-08.pdf>.
- Del Barrio de la Rosa, F. (2017a). Los estudiantes itálofonos de ELE y los estudios de disponibilidad léxica en español. *Revista Nebrija de Lingüística Aplicada*, 22, 52-57. [https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo\\_58de02da8bc1d.pdf](https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_58de02da8bc1d.pdf).
- Del Barrio de la Rosa, F. (ed.). (2017b). *Palabras. Vocabulario. Léxico. La lexicología aplicada a la didáctica y a la diacronía*. Edizioni Ca' Foscari – Digital Publishing. [https://edizionicafoscari.unive.it/media/pdf/books/978-88-6969-170-6/978-88-6969-170-6\\_3GEHh4o.pdf](https://edizionicafoscari.unive.it/media/pdf/books/978-88-6969-170-6/978-88-6969-170-6_3GEHh4o.pdf).
- Del Barrio de la Rosa, F. (2018). Pares léxicos en el léxico disponible de estudiantes italianos. *Lingue e linguaggi*, 26, 173-196. [https://iris.unive.it/retrieve/handle/10278/3702889/136443/2018\\_L%26L\\_Pares.pdf](https://iris.unive.it/retrieve/handle/10278/3702889/136443/2018_L%26L_Pares.pdf).
- Del Barrio de la Rosa, F. & Mae Vann, M. (2018). *Disponibilidad léxica de los estudiantes de español en Italia. Estudio y diccionarios*. Aracne.
- Di Gesù, F. (2016). *Linguistica contrastivo-percettiva di lingue tipologicamente affini: italiano e spagnolo*. University press.
- Echeverría, M. S. et al. (2005). *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica*. Universidad de Concepción. <https://doi.org/10.4067/S0718-04622004049000010>
- Frey Pereyra, M. L. H. (2008). *El léxico disponible en los escritos de alumnos de español como lengua extranjera*. Tesis doctoral inédita. Universidad de Alcalá.
- González Fernández, J. (2014). Idoneidad de los centros de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera. *Revista Nebrija de Lingüística Aplicada*, 16, 41-53. <https://revistas.nebrija.com/revista-linguistica/article/view/223/192>.
- Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1964). *L'élaboration du français fondamental: étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Didier.
- Hidalgo, M. (2019). *Factores y fuentes que inciden en el léxico disponible de estudiantes sinohablantes de ELE*. Monografías ASELE.
- Instituto Cervantes (2006). *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Instituto Cervantes-Biblioteca Nueva. [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/plan\\_curricular/](http://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/).
- López Morales, H. (1999). *Léxico disponible de Puerto Rico*. Arco/Libros, S. L. <https://doi.org/10.31819/9783865278852-011>.
- Michéa, R. (1950). Vocabulaire et culture. *Les Langues Modernes*, 44, 188-189.

- Michéa, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les Langues Modernes*, 47, 338-344.
- Moreno Fernández, F. (2012). Disponibilidad léxica: cuestiones metodológicas. A propósito de disponibilidad léxica de los estudiantes hispanos de Redwood City, CA. *Revista Nebrija de Lingüística Aplicada*, 11, 53-61. <https://revistas.nebrija.com/revista-linguistica/article/view/174/145>.
- Nalesso, G. (2018a). La eficacia de actividades específicas para el aprendizaje del léxico en ELE. Estudio de caso. En M.<sup>a</sup> Bargalló Escrivá, E. Forgas Berdet & A. Nomdedeu Rull (eds.), *Léxico y cultura en LE/L2: corpus y diccionarios. Actas del XXVIII Congreso Internacional ASELE, Tarragona 6-9 de septiembre de 2017* (pp. 505-514). ASELE.
- Nalesso, G. (2018b). El desarrollo de la competencia léxica de estudiantes italianos universitarios de ELE. *Orillas. Rivista d'Italianistica*, 7, 381-394. [http://orillas.cab.unipd.it/orillas/articoli/numero\\_7/07Nalesso\\_Astilleros.pdf](http://orillas.cab.unipd.it/orillas/articoli/numero_7/07Nalesso_Astilleros.pdf).
- Nalesso, G. (2019). *Disponibilidad y riqueza léxica en un grupo de estudiantes universitarios italianos de ELE*. Tesis doctoral inédita. Università degli Studi di Padova. [http://padua-research.cab.unipd.it/12776/1/Giulia\\_Nalesso\\_tesi.pdf](http://padua-research.cab.unipd.it/12776/1/Giulia_Nalesso_tesi.pdf).
- Nalesso, G. (2020). Disponibilidad léxica terminológica en ELE: una propuesta de análisis. *Orillas. Rivista d'Italianistica*, 9, 609-631. [http://orillas.cab.unipd.it/orillas/wp-content/uploads/2020/11/2020\\_13Nalesso\\_Astilleros.pdf](http://orillas.cab.unipd.it/orillas/wp-content/uploads/2020/11/2020_13Nalesso_Astilleros.pdf).
- Paredes García, F. (1999). La ortografía en las encuestas de disponibilidad léxica. *Revista Estudio Adquisición de la Lengua Española (REALE)*, 11, 75-98.
- Paredes García, F. (2014). A vueltas con la selección de 'centros de interés' en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE. *Revista Nebrija de Lingüística Aplicada*, 16, 54-59. <https://revistas.nebrija.com/revista-linguistica/article/view/224/193>.
- Paredes García, F. (2015). Disponibilidad Léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula. *Linred*, 13. [http://www.linred.es/monograficos\\_pdf/LR\\_monografico13-articulo2.pdf](http://www.linred.es/monograficos_pdf/LR_monografico13-articulo2.pdf).
- Real Academia Española. (2014). *Diccionario de la lengua española (DRAE)*, 23.<sup>a</sup> ed., [versión 23.4 en línea]. <https://dle.rae.es> [24/05/2021]
- Rubio Sánchez, R. (2017). Acercamiento al léxico disponible de 173 estudiantes italianos preuniversitarios de Español como Lengua Extranjera. En F. Del Barrio de la Rosa (ed.), *Palabras Vocabulario Léxico. La lexicología aplicada a la didáctica y a la diacronía* (pp. 143-161). Edizioni Ca' Foscari. <http://edizionicafoscari.unive.it/media/pdf/chapter/978-88-6969-170-6/978-88-6969-170-6-ch-09.pdf>.
- Samper Hernández, M. (2002). *Disponibilidad léxica en alumnos de español como lengua extranjera*. ASELE.
- Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible: sugerencias. *Lingüística*, 10, 311-333.

- Sánchez-Saus Laserna, M. (2016). *Léxico disponible de los estudiantes de español como lengua extranjera en las universidades andaluzas*. Editorial Universidad de Sevilla.
- Santos Díaz, I. C. (2017). Selección del léxico disponible: propuesta metodológica con fines didácticos. *Porta Linguarum*, 27, 122-139. <https://doi.org/10.30827/Digibug.53965>.
- Selinker, L. (1969). Language transfer. *General Linguistics*, 9, 67-92.
- Selinker, L. (1972). Interlanguage. *IRAL*, X(2), 209-231. <https://doi.org/10.1515/iral.1972.10.1-4.209>.
- Tomé Cornejo, C. (2015). *Léxico disponible. Procesamiento y aplicación a la enseñanza de ELE*. Tesis doctoral inédita. Universidad de Salamanca. [https://gredos.usal.es/bitstream/handle/10366/128287/DLE\\_Tom%E9CornejoC\\_L%E9xicodisponible.pdf;jsessionid=755FBEF157497F5050AF3397A3B3BC86?sequence=1](https://gredos.usal.es/bitstream/handle/10366/128287/DLE_Tom%E9CornejoC_L%E9xicodisponible.pdf;jsessionid=755FBEF157497F5050AF3397A3B3BC86?sequence=1).
- Trigo, E., Romero, M. F. & Santos Díaz, I. C. (2018). Elaboración de un corpus cacográfico desde la disponibilidad léxica en estudiantes sevillanos: un análisis para la enseñanza de la lengua. *Revista de Lingüística y Lenguas Aplicadas*, 13, 119-131. <https://doi.org/10.4995/rlyla.2018.9176>.
- Trigo, E., Romero, M. F. & Santos Díaz, I. C. (2019). Aproximación empírica desde la disponibilidad léxica a la influencia de los factores sociolingüísticos en el dominio ortográfico. *Cultura y Educación*, 31(4), 814-844. <https://doi.org/10.1080/11356405.2019.1659007>.
- Trigo, E., Romero, M. F. & Santos Díaz, I. C. (2020). Disponibilidad léxica y dominio de la ortografía: un estudio empírico basado en la influencia de los factores sociales. *Onomazein*, 46, 27-45. <https://doi.org/10.7764/onomazein.47.02>.
- Vázquez, G. (1999). *¿Errores? ¡Sin falta!* Edelsa Didascalía.