

Diseño y construcción del CAELE2: Base para una planificación curricular

Design and construction of CAELE2: Base for curriculum planning

Diana Alejandra Hincapié Moreno*., Ruth Yanira Rubio López**

*Magister Enseñanza de Español como Lengua Extranjera de la Universidad Pontificia de Salamanca, Docente- investigadora de la Maestría en Enseñanza de Español como Lengua Extranjera y Segunda Lengua del Instituto Caro y Cuervo, E- mail: diana.hincapie@caroycuervo.gov.co

**Magister en Formación de profesores de Español como Lengua Extranjera, Docente- investigadora del Instituto Caro y Cuervo, E- mail: ruth.rubio@caroycuervo.gov.co

Citar: Hincapie, D.; Rubio, R. (2017). Diseño y construcción del CAELE2: Base para una planificación curricular. *Hechos y Proyecciones del Lenguaje*. 23(1), Págs.: (42 – 52).

Recibido: octubre 15, 2017

Aceptado: diciembre 3, 2017

Resumen

El presente documento tiene por objetivo la presentación de la fase inicial del proyecto *Corpus de Aprendientes de Español como Lengua Extranjera y Segunda Lengua (CAELE2)* y sus posibilidades de explotación como base para la planificación curricular. Para la construcción del corpus nos apoyamos en la metodología de la lingüística de corpus que sirve para la selección, organización y análisis de muestras auténticas de lengua. El CAELE2 recopila textos escritos producidos por aprendientes de español como lengua extranjera (ELE) y segunda lengua (L2) en el contexto colombiano, lo que permitirá en futuras fases el análisis de errores, el estudio de la interlengua, entre otros. Estos estudios y la colección de muestras sistematizadas servirán de soporte para el diseño y planeación de cursos y materiales centrados en las necesidades y el contexto de los aprendientes.

Palabras clave: Corpus, corpus de aprendientes, CAELE2, planificación curricular

Abstract

This document aims to present the initial phase of the project *Corpus de Aprendientes de Español como Lengua Extranjera y Segunda Lengua (CAELE2)* and its potential as a basis for curricular planning. For the construction of the corpus we use the methodology named corpus linguistics which serves for the selection, organization and analysis of authentic samples of language. CAELE2 collects written texts produced by learners of Spanish as a foreign and second language in the Colombian context, which will allow the analysis of errors, the study of interlanguage, among others in future phases. These studies and the collection of systematized samples of language will support the design and planning of courses and materials focused on the needs and context of learners.

Keywords: Corpus, learning corpus, CAELE2, curricular planning

Introducción

En los últimos años, la integración de corpus, herramientas tecnológicas y enseñanza de lenguas ha adquirido interés por la posibilidad de realizar análisis de grandes cantidades de datos y muestras producidas por los aprendientes de lenguas. Tradicionalmente, las dudas que emergen en la práctica docente se resuelven desde la subjetividad y la experiencia, lo que, aunque aporta a la solución de distintas dificultades, no lo hace de manera objetiva y sistemática. Los corpus permiten llevar a cabo investigaciones más rigurosas y experimentales sobre los procesos de enseñanza-aprendizaje de lenguas. En este sentido, se evidencia la importante necesidad de contar con corpus de aprendientes para la investigación y didáctica de las lenguas.

Penadés (2005) continúa con esta reflexión:

Para que el conocimiento sea científico debe derivarse no sólo de la experiencia del profesor/investigador —si acaso ésta debe ser el punto de partida—, sino del análisis de los datos existentes en materiales reales producidos por estudiantes de ELE. Pero para llevar a cabo estudios de ese tipo necesitamos disponer de corpus que recojan muestras orales y escritas de las producciones de los estudiantes de español. Y ahí nos encontramos con una seria dificultad metodológica, ya que no existe en la actualidad para el español no nativo ningún corpus extenso, de grandes dimensiones, y de fácil acceso para que los investigadores puedan llevar a cabo investigaciones sobre análisis de errores lingüísticos, sobre problemas de adquisición relacionados con la lengua materna, con el nivel de conocimientos, con el sexo o la edad de los estudiantes, sobre tipos, frecuencias, condicionamientos e implicaciones de las transferencias lingüísticas, sobre elaboración de materiales didácticos, entre otros tantos temas de estudio que se necesita abordar.
(p. 2)

Cestero et al. (2001) mencionan que los corpus de aprendientes extranjeros pueden ser utilizados para estudios de la interlingua; el análisis de errores y de transferencia lingüística; el diseño de materiales y el análisis de problemas relacionados con la L1, la edad u otros factores. Incluso, pueden servir para definir qué aspectos y temas trabajar en clase teniendo en cuenta las características, contexto y necesidades de los aprendientes y cómo trabajar los errores y la transferencia lingüística. Por otro lado, se debe tener en cuenta que el almacenamiento y sistematización de muestras con ayuda de herramientas tecnológicas pueden facilitar el acceso y explotación de numerosas cantidades de datos que, una vez organizadas en corpus podrán servir para investigaciones en diversas áreas del conocimiento.

El aprendizaje y enseñanza del español como lengua extranjera (ELE) está tomando cada vez más fuerza en el contexto colombiano, donde año tras año llega mayor número de extranjeros en aras de aprender español. Adicionalmente, en relación con la enseñanza aprendizaje de segunda lengua (L2) se debe mencionar que en Colombia hay pocas investigaciones en lo que se refiere a la enseñanza del español a hablantes cuya lengua materna son las indígenas y la lengua de señas. Por ende, contar con un repositorio de muestras de aprendientes de L2 facilitará el desarrollo de investigaciones en estos ámbitos y posibilitará el diseño de materiales relacionados con las necesidades de los alumnos.

En respuesta a estas necesidades el proyecto Corpus de Aprendientes de ELE y L2 (CAELE2) en su primera fase tiene como objetivo diseñar, recopilar, sistematizar y construir un corpus de aprendientes de ELE/L2 en el contexto colombiano, que incluya muestras escritas de estudiantes extranjeros, indígenas y sordos. Asimismo, pretende proporcionar muestras de lengua organizadas que faciliten y propicien investigaciones de carácter teórico o aplicado, incluyendo planeaciones curriculares y diseño de material didáctico.

En este documento se presentan inicialmente los objetivos y motivación del proyecto. Posteriormente, se explica el marco conceptual y metodológico para el diseño y elaboración del CAELE2. Luego se presentan los resultados y discusión en torno al estado actual del corpus y sus alcances para la planificación curricular. En último lugar, se exponen las conclusiones finales del texto.

Marco teórico

Un corpus computacional es definido por el *Expert Advisory Group on Language Engineering Standards -EAGLES* (2006) como “a corpus which is encoded in a standardized and homogenous way for open-ended retrieval tasks. Its constituent pieces of language are documented as to their origins and provenance”. Es decir, un corpus está constituido por piezas de lengua documentadas y sistematizadas de forma homogénea para la realización de diversas tareas de recuperación y búsqueda. En este sentido, hablamos de dos características fundamentales de los corpus: deben ser informatizado y estar compuestos por muestras de lengua organizadas y documentadas. De esta manera, un corpus lingüístico es un conjunto amplio de textos informatizados, producidos en situaciones reales, espontáneas o dirigidas con determinados fines y que se han seleccionado siguiendo una serie de criterios explícitos para garantizar que dicho corpus pueda ser usado como muestra representativa de la lengua que se quiere analizar.

Ahora bien, los corpus pueden ser clasificados desde distintos parámetros: como la modalidad de la lengua, el número de lenguas que contiene, el carácter general o especializado de los textos, el periodo cronológico que abarcan, el tamaño de las muestras y el tratamiento que se le aplica a los datos lingüísticos (Villayandre, 2008). Según la modalidad de la lengua, se habla de: corpus escrito, corpus oral, corpus multimodal, y corpus mixto (Torruela y Llisterri, 1999):

- Los corpus escritos o también llamados textuales están constituidos por textos o muestras de lengua escrita. Es uno de los tipos de corpus más comunes, ya que su recolección es más sencilla en comparación con los corpus orales o multimodales. Sus fuentes pueden provenir de libros, revistas, prensa, artículos, textos de internet, entre muchos otros. Este es el tipo de corpus que recolectaremos en la fase inicial del proyecto.
- Un corpus oral está formado por muestras de lengua oral, que corresponden a señales de voz, transcripciones y en algunos casos ambas. Este tipo de corpus hace referencia al que se organiza usualmente con propósitos netamente lingüísticos. Se

construye a partir de grabaciones de muestra oral o sus transcripciones, en un primer momento ortográficas. El objetivo de estos corpus es reflejar una lengua o variedad a partir de los usos de la lengua hablada ya sea discursos, conferencias, conversaciones, habla espontánea, etc.

- Un corpus mixto recoge muestras de lengua oral y lengua escrita, por tanto comparte las características anteriormente enunciadas.
- Los corpus multimodales combinan dos o más medios de producción, es decir pueden estar constituidos por texto, sonido, imagen o video. De esta manera los datos pueden contener información prosódica, kinésica, contextual, etc. Sus fuentes son usualmente documentales, lenguaje de señas, videoconferencias, entre otros.

Adicionalmente, los corpus se pueden catalogar y definir según los fines u objetivos de creación y explotación, es aquí donde aparecen los corpus de aprendientes. De acuerdo con Cruz (2012), estas herramientas lingüísticas reúnen grandes muestras de textos producidos por aprendientes de la lengua y permiten obtener información empírica sobre el proceso de aprendizaje, lo que facilita la posterior aplicación de estas investigaciones a la enseñanza diaria y también a la edición de materiales.

Por otro lado, el término diseño curricular hace referencia a una dimensión del currículo que da cuenta de la metodología, las acciones, la modelación, la estructuración y la organización de los proyectos curriculares. Esta organización busca solucionar problemas y satisfacer necesidades, para mejorar los procesos de enseñanza- aprendizaje (Fernández, 2015). El primer paso para la planificación curricular es la identificación y evaluación de las necesidades. Posteriormente, se deben definir las competencias que alcanzarán los estudiantes; precisar y organizar los contenidos; decidir las metodologías usadas por los docentes; y determinar el proceso de evaluación. Finalmente, se debe presentar el currículo y gestionar su desarrollo y revisión constante en un entorno educativo apropiado para su implantación (Palés, 2006).

Con el paso de los años la aparición de corpus lingüísticos en español aumenta de manera considerable, aunque los corpus de aprendientes no lo hacen con la misma velocidad. La mayoría de estos corpus son escritos u orales, con muestras de ELE más no de L2, y gran parte de estos han sido recolectados en Europa o Estados Unidos. Cruz (2012) enuncia algunos de estos corpus:

- *Corpus de aprendices de español (CAES)* del Instituto Cervantes: cuenta con aproximadamente 570.000 palabras provenientes de ensayos y escritos realizados por estudiantes de niveles A1- C1, cuya lengua materna es el árabe, portugués, inglés, francés, chino-mandarín o ruso.
- *Corpus Escrito del Español L2 (CEDEL2)*: corresponde a 510.000 palabras producidas en ensayos de aprendientes de nivel inicial hasta avanzado, con inglés como lengua materna.
- *CORELE*: contiene entrevistas semi-espontáneas tareas narrativas y descriptivas de estudiantes con lengua materna como el portugués, el italiano, el francés, el inglés,

el alemán, el holandés, el polaco, el chino y el japonés. Comprende un total de 52.759 unidades léxicas.

- *Spanish Corpus Proficiency Level Training*: contiene diálogos y transcripciones de 38 hablantes angloparlantes.
- *LANGSNAP*: está formado por entrevistas orales y áreas narrativas, de aprendientes angloparlantes de nivel intermedio.
- *Fono.Ele*: compuesto por conversaciones cortas, y lectura de textos, frases y palabras. Los 96 aprendientes provienen de 19 lenguas maternas distintas, y cuenta con aproximadamente 34.300 unidades.
- *Corpus oral de español como lengua extranjera*: recoge muestras de entrevistas de 40 aprendientes de nivel A2 y B1, provenientes de distintos países y estudiantes de la Universidad Autónoma de Madrid.

El CAELE2 no es el primer corpus de aprendientes de español, pero sí es el único que recoge muestras escritas de 3 poblaciones: estudiantes indígenas, sordos, y extranjeros. Este corpus busca suplir las necesidades de investigación en áreas como la lingüística aplicada, la enseñanza-aprendizaje de lenguas, la adquisición de segundas lenguas, entre otras. Adicionalmente, busca fomentar la investigación y uso de corpus para la enseñanza-aprendizaje de ELE y L2 en Colombia.

Metodología

Los corpus se elaboran siguiendo una metodología lingüística denominada Lingüística de corpus que tiene por objetivo la selección, organización y explotación de muestras auténticas de lengua. En el caso del CAELE2 el corpus busca tener muestras representativas de hablantes no nativos del español -aprendientes extranjeros, sordos e indígenas.

El primer paso de la investigación es la definición del tipo de corpus según los objetivos del proyecto. En términos generales, se apunta a la construcción de un corpus escrito de aprendientes de ELE y L2 en Colombia, abierto, y sin un número definido de informantes o palabras. Los aprendientes deben provenir de distintas lenguas maternas y tener diversos niveles de aprendizaje: básico, intermedio y avanzado. Además, las muestras son recolectadas mediante pruebas específicas diseñadas para el proyecto.

Tras la definición del corpus meta y la construcción de una base de datos con información sobre las instituciones y características generales de los aprendientes en Colombia, se realiza una propuesta y selección de metadatos que acompañan el corpus. Estos metadatos corresponden tanto al informante como al archivo. Las etiquetas se pueden observar en las Tablas 1 y 2.

Tabla 1. *Metadatos de archivos escritos*

Archivos
Id (código de identificación del archivo)
Ubicación Id.txt
Tipo de texto
Temas
Descripción de la tarea escrita
¿Tuvo ayuda en la tarea? / ¿Qué tipo de ayuda?
Subcorpus (ELE, L2/Indígenas, L2/Sordos)
Fecha de recolección
Encuestador/ Transcriptor
Revisor

Tabla 2. *Metadatos del aprendiente de español como lengua extranjera*

Informante/Aprendiente
Id Aprendiente (Código de identificación del aprendiente)
Edad
Sexo
Nacionalidad
Lengua materna
Segunda lengua
Nivel educativo
Profesión/ Ocupación
Lenguas extranjeras que maneja
Nivel actual de español
Tiempo de aprendizaje
Países hispanohablantes que ha visitado
Contacto con el español
Ciudad en la que más tiempo ha permanecido en Colombia ¹
Tiempo en Colombia

Los metadatos del informante, corresponden a la información externa general. Para los subcorpus de L2 se han agregado nuevos campos como tipo de dificultad auditiva, padre o/ madre sorda, edad de adquisición de LSC, comunidad indígena, ciudad de residencia.

Como se enunciaba anteriormente, las muestras se recolectan mediante pruebas específicas diseñadas para la investigación. Después de revisar los criterios de evaluación

por nivel del Marco Común Europeo de referencia- MCER (2002), y basados en el análisis de criterios que realizan Gozalo y Martín (2009) se diseñaron siete pruebas divididas en los niveles A1, A2, B1, B2, y C, las cuales corresponden a su vez a tres tipologías textuales: textos descriptivos, textos narrativos, y textos argumentativosⁱⁱ.

La primera fase de recolección se llevó a cabo mediante la aplicación de la prueba a 25 aprendientes extranjeros. Esta fase se tomó como el pilotaje de los instrumentos de recolección y, a partir de los resultados se ajustaron los metadatos y las instrucciones de las tareas. Terminado el pilotaje se continúa con la recolección de las muestras de ELE obteniendo como resultado parcial a 2017ⁱⁱⁱ 78 muestras de español escrito.

En la fase de sistematización se genera una base de datos con los metadatos de cada archivo y cada informante, marcados de tal manera que coincidan en las búsquedas. A la vez que se construye este instrumento, las pruebas deben ser digitalizadas y almacenadas en formato *pdf*, transcritas y almacenadas en formato *.docx*, y a su vez se guardan en formato *.txt*, que a largo plazo será el archivo que podrá ser tratado computacionalmente. Cada una de las muestras está acompañada por un consentimiento informado almacenado en *pdf* o *jpg* y marcado con el metadato *IdAprendiente*. Para evitar algún tipo de error en este proceso, se finaliza con la revisión de la sistematización efectuada por un integrante del grupo distinto al encuestador o transcriptor.

Resultados y discusión

El resultado de la recolección y sistematización de muestras en su componente escrito ELE del corpus CAELE2 se presenta a continuación en la Tabla 3.

Tabla 3 Muestras escritas CAELE2 componente aprendientes extranjeros

Datos	Resultados
Número de muestras	73
Número de palabras	15.000 aprox.
Número de informantes	42
Sexo de los informantes	23 hombres 19 mujeres
Edad de los informantes	Entre los 20 y 40 años
Nivel de español	A1: 3 aprendientes A2: 4 aprendientes B1: 8 aprendientes B2: 7 aprendientes C: 17 aprendientes
Lengua materna de los aprendientes	Inglés Portugués Alemán Francés Creole Lituano

Nivel educativo de los aprendientes	Pregrado: 23 aprendientes Posgrado: 14 aprendientes Doctorado: 3 aprendientes
-------------------------------------	---

Como se enuncio anteriormente en el marco teórico, los corpus escritos de aprendientes de español como lengua extranjera más reconocidos hasta el momento son el Corpus de aprendientes de español CAES (2014) creado por el Instituto Cervantes y el Corpus escrito del español L2 CEDEL2 (2006) de la Universidad Autónoma de Madrid y la Universidad de Granada. Estos corpus superan las 500.000 palabras, cifra que aún no es posible comparar con el CAELE2 debido a que se encuentra en su fase inicial de recolección. En cuanto al número de informantes- CAES contó con 1423, y el CEDEL2 con más de 2000, número que aspira a alcanzar nuestro corpus en fases posteriores. Sobre los niveles de lengua el CAELE2 y el CAES manejan los propuestos por el MCER, mientras que el CEDEL2 utiliza las categorías: *beginner*, *intermediate*, y *advanced*. Finalmente, las pruebas utilizadas en estos dos corpus corresponden a ensayos y tareas diseñadas según el nivel de lengua. En el caso del CAELE2, las tareas se centraron en la función: narrativa, descriptiva y argumentativa.

Por otro lado, los metadatos de los aprendientes usados en los corpus de este tipo concuerdan en su mayoría: sexo, edad, nivel educativo, tiempo de estudio de español, países hispanohablantes visitados, y conocimiento sobre otras lenguas. El CAELE2 presenta las etiquetas de profesión y contacto con el español como información nueva, además de la información extra presentada para los aprendientes sordos e indígenas.

El CAELE2 y sus posibilidades para el diseño curricular

Las características del CAELE2 permiten su explotación y uso para la investigación y para diversas tareas de enseñanza aprendizaje de ELE y L2. Granger (2013) considera que una de las principales cualidades que la investigación con corpus trae al campo de la enseñanza de ELE es la amplia base de información empírica con la que no contábamos anteriormente. Asimismo, López (2008) menciona que “Los corpus son valiosos en la enseñanza de lenguas por su condición de conjuntos de textos reales de una lengua, aplicables, por ejemplo, a la práctica con ejercicios extraídos directamente de ellos” (p. 116). El CAELE2 pretende funcionar como una herramienta pedagógica y de investigación para los docentes de ELE y L2 en Colombia, por ejemplo para el diseño de ejercicios, actividades didácticas, y materiales extraídos del corpus mismo.

Agregado a lo anterior, el corpus puede servir de insumo para la planeación curricular, al respecto el Spanish Learner Language Oral Corpora (SPLLOC) (2008) enuncia:

Research such as this enables us to better understand the processes involved in learning a second language in a classroom setting, and thus supports curriculum design for L2 education, as well as exploring fundamental theoretical questions in second language acquisition.

Como lo mencionamos en el marco teórico, el proceso inicial que se debe llevar a cabo para la planificación curricular es el diagnóstico de problemas y necesidades, que conduzca a la modelación, estructuración y organización para la puesta en práctica. Es en este punto donde un corpus de aprendientes puede servir como base para la planeación curricular, ya que la explotación de los datos arroja las dificultades más frecuentes según las características sociodemográficas de los estudiantes, permite un análisis de contenidos en cada nivel de aprendizaje, y por tanto permite que los docentes-investigadores determinen necesidades a partir de datos objetivos.

Los corpus de aprendientes dan la posibilidad de estudiar la interlengua y analizar los errores de los aprendices de español (Cabrera y Gomez, 2017). Gracias a ellos se puede revisar e identificar las dificultades de aprendizaje y los errores más frecuentes. A partir de esta revisión se pueden planear actividades y contenidos para los currículos; comparar los datos con corpus de referencia para así definir la secuencia de contenidos; y realizar investigaciones sobre el proceso de aprendizaje que ayuden en la determinación de técnicas de enseñanza y la creación de materiales basados en las necesidades de los estudiantes.

Como Cruz (2012) lo describe:

Los corpus de aprendices, tanto textuales como orales, son especialmente útiles para los profesionales de la enseñanza de lenguas interesados en investigar sobre el proceso de aprendizaje de sus alumnos. Constituyen, pues, una fuente empírica de datos para trabajar en la línea de análisis de errores y del estudio de la interlengua y, a partir de los resultados, diseñar cursos y materiales más adecuados a las necesidades reales de los aprendices. (p. 137)

En este sentido, el corpus CAELE2 busca convertirse en una herramienta para docentes, estudiantes e investigadores interesados en trabajar sobre la adquisición de la lengua, el análisis de errores, el estudio de la interlengua, ente otros.

Conclusiones

Los corpus de aprendientes recogen conjuntos de textos (orales o escritos) producidos por estudiantes de lenguas. En las últimas décadas ha habido un creciente interés por este tipo de corpus, ya que proporcionan información empírica para la realización de investigaciones en torno a la enseñanza aprendizaje de lenguas. El Corpus de Aprendientes de Español como Lengua Extranjera y Segunda Lengua (CAELE2) busca recopilar en su fase inicial muestras escritas de aprendientes de español extranjeros, indígenas y sordos en el contexto colombiano. Este corpus pretende servir como base para profesores e investigadores que deseen realizar estudios y usar el corpus para enseñanza aprendizaje del ELE y L2.

Actualmente, el corpus cuenta con 73 textos escritos producidos por estudiantes extranjeros hombres y mujeres con distintos niveles de lengua (A1, A2, B1, B2, C1) y de distintas nacionalidades (Francia, Trinidad y Tobago, Estados Unidos, etc.). En muchos casos los cursos y materiales usados en clase de ELE/L2 están alejados del contexto del

estudiante, lo que hace que no se aborden temas de interés para ellos. A futuro, el uso de este corpus permitirá realizar investigaciones relacionadas con el análisis de errores, el estudio de la interlengua, la adquisición del español, entre otros. Estos estudios servirán de base para el diseño de materiales y contenidos relacionados con las dificultades de aprendizaje y las necesidades de los estudiantes que son base esencial para la planificación curricular.

Adicionalmente, se podrá dar respuesta a distintas preguntas en la enseñanza aprendizaje de ELE y L2 que hasta ahora se han respondido desde la intuición y la experiencia y que aunque han proporcionado algunas soluciones se deben trabajar con mayor rigor. Con el fin de facilitar y mejorar el trabajo que se hace dentro y fuera del aula en la enseñanza de español es fundamental fomentar la investigación rigurosa y empírica en estas áreas y proporcionar recursos como los corpus para facilitar esta tarea.

Referencias Bibliográficas

- Baralo, M. (2010). La investigación del español como lengua segunda y extranjera. *Actas Congreso internacional de la lengua española Valparaíso*. Recuperado de: http://congresosdelalengua.es/valparaiso/ponencias/lengua_educacion/baralo_marta.htm
- Cabrera, A. y Gómez, J. (2017). Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE. *Revista Brasileña de Lingüística Aplicada*, 17, 509-538.
- Cestero, A., Penadés, I., Blanco, A., Camargo, L. y Granda J. (2001). Corpus para el análisis de errores de aprendices de E/LE (CORANE). *ASELE. Actas XII*. Recuperado de: http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/12/12_0527.pdf
- Cruz, M. (2012). *Lingüística de corpus y enseñanza del español como 2/L*. Madrid: Arco Libros.
- Fernández, A. (2015). El diseño curricular. La práctica curricular y la evaluación curricular. Recuperado de: http://sistemas2.dti.uaem.mx/evadocente/programa2/Psic009_13/documentos/06%20DISENO%20Y%20EVALUACION%20CURRICULAR.pdf
- Ferreira, A., Elejalde, J., y Vine, A. (2014). Análisis de Errores Asistido por Computador basado en un Corpus de Aprendientes de Español como Lengua Extranjera. *Revista signos*, 47(86), 385-411.
- Gozalo, P. y Martín, M. (2009). *Pruebas de nivel. Modelos de examen de clasificación*. Madrid: SGEL.
- Granger, S. (2013). The contribution of learner corpora to second language acquisition and foreign language teaching. A critical evaluation. En Aijmer, K. *Corpora and Language Teaching*. (pp. 13- 32). Amsterdam: John Benjamins Publishing Company.
- Hincapié y Bernal (2015). *Lingüística de Corpus*. (Libro producto de investigación, documento no publicado. Bogotá.
- Instituto Cervantes. (2014). *Corpus de aprendices de español- CAES*. Recuperado de <http://galvan.usc.es/caes>
- López, V. (2008). Integración de los corpus como herramienta de apoyo en la enseñanza de ESP. *Porta Linguarum*, 10, 115- 136.
- Martínez, I. y Llorián, S. (2016). Aplicaciones de la lingüística de corpus a las tareas del aula, al currículo y a la evaluación del español de los negocios. Recuperado de <http://mdl.univ->

poitiers.fr/images/medias/fichier/amphi-espe-martinez-martinez-llorian-gonzales-0900-945_1466670441337-pdf?INLINE=FALSE

Palés, Jorge L.. (2006). Planificar un currículum o un programa formativo. *Educación Médica*, 9(2), 59-65. Recuperado de http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-8132006000200004&lng=es&tlng=es

Penadés, I. (2005). Corpus para el análisis de errores en el aprendizaje de E/LE: Presentación. *LINRED: Revista electrónica de lingüística Informaciones sobre cuestiones lingüísticas*, 3, 1-5.

Torrueja, J., y Llisterra, J. (1999). Diseño de corpus textuales y orales. En Bleca, J., Clavería, G., Sánchez, C., y Torruela, J. (Eds.) *Filología e informática: Nuevas tecnologías en los estudios filológicos*. (pp.45– 77). Barcelona: Milenio.

Universidad Autónoma de Madrid, y Universidad de Granada. (2006). *Corpus Escrito del Español L2-CEDEL2*. Recuperado de <https://www.uam.es/proyectosinv/woslac/cedel2.htm>

University of Southampton, University of Newcastle, y University of York (2006). Spanish Learner Language Oral Corpora- SPLLOC. Recuperado de <http://www.sploc.soton.ac.uk/index.html>

Villayandre, M. (2008). Lingüística con corpus (I). *E.H. Filología*. 30, 329- 349.

ⁱ Metadato agregado después del pilotaje de los instrumentos.

ⁱⁱ El diseño de las pruebas y la recolección de las muestras de L2 se encuentra en proceso, por tal motivo el texto presenta únicamente el avance de la investigación con los datos de ELE.

ⁱⁱⁱ El proyecto continúa su curso en 2018 con el fin de alimentar el corpus con más muestras de ELE, y recolectar las muestras de L2.