

doi: <https://doi.org/10.15446/fyf.v33n1.84182>

PROPUESTA DE UNA TAXONOMÍA ETIOLÓGICA PARA ETIQUETAR ERRORES DE INTERLENGUA EN EL CONTEXTO DE UN CORPUS ESCRITO DE APRENDIENTES DE ELE *

Anita Ferreira Cabrera **

Jessica Elejalde Gómez ***

Universidad de Concepción, Chile

Resumen

La identificación, clasificación y anotación de errores en la producción escrita de una segunda lengua o lengua extranjera (L2-LE), se basa principalmente en la taxonomía con la cual se implementa el análisis de los errores en el estudio de la lengua. Por esta razón, las taxonomías deben responder con la mayor precisión a la categorización de los errores, entregando al investigador la posibilidad de indagar en análisis posteriores sobre fenómenos que subyacen a la conformación de la interlengua. El propósito principal de esta investigación es proponer una taxonomía que pueda adaptarse a diferentes contextos de aprendizaje y análisis de la lengua en estudio, lo cual supone un avance relevante en materia de investigación a partir de clasificaciones más acertadas de la interlengua de un aprendiente. En este sentido, el presente artículo aborda la importancia de la problemática de categorización y etiquetado de errores y expone una propuesta hacia una taxonomía estandarizada para la clasificación de errores de interlengua con un criterio etiológico. Esta propuesta surge del avance en la investigación sobre errores en el marco de proyectos Fondecyt-Chile del área de la lingüística aplicada a la enseñanza y aprendizaje de ELE.

Palabras clave: *taxonomía de errores etiológicos; análisis de errores asistido por computador; español como lengua extranjera; errores interlingüísticos.*

Cómo citar este artículo:


Ferreira Cabrera, A., & Elejalde Gómez, J. (2020). Propuesta de una taxonomía etiológica para etiquetar errores de interlengua en el contexto de un corpus escrito de aprendientes de ELE.

Forma y Función, 33(1), 115-146. <https://doi.org/10.15446/fyf.v33n1.84182>

Artículo de investigación: Recibido: 31-01-2019, aceptado: 18-06-2019

* Este artículo se ha desarrollado en el contexto del Proyecto Fondecyt 1180974 «Diseño e implementación de un corpus escrito de aprendientes de Español como Lengua Extranjera (ELE) para el análisis de la interlengua», 2018-2021.

**  ORCID <https://orcid.org/0000-0001-7979-6467>. aferreir@udec.cl

***  ORCID <https://orcid.org/0000-0002-6024-2381>. jelejalde@udec.cl

PROPOSAL FOR AN ETIOLOGICAL TAXONOMY TO LABEL INTERLANGUAGE ERRORS IN THE CONTEXT OF A WRITTEN CORPUS OF SFL LEARNERS

Resumen

The identification, classification, and annotation of errors in the written production of a second language or foreign language (L2-FL) is mainly based on the taxonomy used to implement the analysis of errors in the study of the language. For this reason, taxonomies should respond more precisely to the categorization of errors, thus allowing the researcher, in subsequent analyses, to inquire into the phenomena underlying the formation of the interlanguage. The main objective of this research project is to propose a taxonomy that can adapt to different contexts of learning and analysis of the studied language, and which marks a relevant step forward in research since it provides more accurate classifications of a learner's interlanguage. In this sense, the article addresses the importance of the issue of categorization and labeling of errors, and proposes a standardized taxonomy for the classification of interlanguage errors with an etiological criterion. This proposal arose as a result of progress in error research, within the framework of Fondecyt-Chile projects in the area of linguistics applied to SFL teaching and learning.

Keywords: *taxonomy of etiological errors; computer-assisted error analysis; Spanish as a foreign language; interlanguage errors.*

PROPOSTA DE UMA TAXONOMIA ETIOLÓGICA PARA ETIQUETAR ERROS DE INTERLÍNGUA NO CONTEXTO DE UM CORPUS ESCRITO DE APRENDIZES DE ELE

Resumo

A identificação, classificação e anotação de erros na produção escrita de uma segunda língua ou língua estrangeira (L2-LE) está baseada, sobretudo, na taxonomia com a qual a análise dos erros foi implantada no estudo da língua. Por essa razão, as taxonomias devem responder, com maior precisão, à categorização dos erros, propiciando ao pesquisador a possibilidade de questionar em análises posteriores sobre fenômenos que subjazem a conformação da interlíngua. O propósito principal desta pesquisa é propor uma taxonomia que possa adaptar-se a diferentes contextos de aprendizagem e de análise da língua em estudo, o que supõe um avanço relevante em matéria de pesquisa a partir de classificações mais adequadas da interlíngua de um aprendiz. Nesse sentido, este artigo aborda a importância da problemática de categorização e etiquetado de erros, e expõe uma proposta por uma taxonomia padronizada para classificar erros de interlíngua com um critério etiológico. Essa proposta surge do avanço na pesquisa sobre o tema no âmbito de projetos Fondecyt-Chile da área da linguística aplicada ao ensino e aprendizagem de ELE.

Palavras-chave: *taxonomia de erros etiológicos; análise de erros assistido por computador; espanhol como língua estrangeira; erros interlinguísticos.*

INTRODUCCIÓN

El análisis de errores asistido por computador (del inglés *computer aided-error-analysis* —CEA—) es un enfoque de investigación basado en corpus electrónicos de textos de aprendientes con el objeto de identificar, clasificar, describir y explicar los errores en el proceso de adquisición de una L2. Junto con dichos procedimientos, se realiza un etiquetado semiautomático a través de software especializados para tareas de etiquetamiento que se basan en un sistema de anotación previo. Una vez concluido dicho proceso, se pueden procesar los datos y obtener los resultados según las variables del estudio. El concepto base de error se sustenta en el clásico planteamiento de Corder (1967), quien entiende por error la desviación sistemática y consistente que caracteriza el sistema lingüístico de un aprendiente en un nivel determinado del desarrollo de su L2.

El análisis de errores asistido por computador conlleva la identificación y clasificación de los errores encontrados para realizar posteriores análisis en los cuales se pueda: (1) determinar la frecuencia de los errores acorde con las etiquetas, (2) identificar las tendencias de uso erróneo y (3) evaluar el estado de la interlengua del aprendiente en la L2 o LE. Por estas razones, es fundamental contar con un sistema de anotación que abarque criterios de identificación, para así determinar con mayor precisión los fenómenos que atañen al sistema lingüístico. Por tanto, una de las etapas importantes en la investigación basada en el CEA es diseñar un sistema de etiquetas de anotación de errores para reconocer y etiquetar los errores de un corpus de aprendientes en diversos niveles de categorización.

El objetivo en este artículo es proponer una taxonomía para la anotación de los errores etiológicos sistemáticos causados por transferencia lingüística de la L1. Para ello, nos basaremos en la revisión de la literatura de la especialidad y en los avances y desarrollos previos realizados en el contexto de proyectos de investigación (Ferreira, 2011, 2014).

El artículo se organiza en las siguientes secciones: primero nos referimos a los principales fundamentos teóricos en materia de tipologías de errores y taxonomías con criterio etiológico. Después abordamos el estudio sobre una propuesta de taxonomía para errores con criterio etiológico. En la siguiente sección, presentamos la descripción de la taxonomía para clasificar errores etiológicos. Por último, presentamos algunas conclusiones sobre los avances y logros obtenidos en esta investigación.

FUNDAMENTOS TEÓRICOS

Si bien no existe un consenso en cuanto a la clasificación y taxonomías de errores, los criterios en la literatura especializada (Corder, 1967; Alexopoulou, 2006, 2007; De Alba

Quiñones, 2009) sientan las bases para la construcción de cualquier tipo de taxonomía. Así, como lo menciona Díaz-Negrillo (2007), la taxonomía de errores describe la observación de los datos de un corpus, minimizando la interpretación subjetiva de los errores y la categorización de estos. Sin embargo, es necesario delimitar una taxonomía estandarizada para designar los errores de aprendices en un corpus computarizado (Granger, 2004, 2015; Díaz-Negrillo, 2007), puesto que los resultados permitirán indagar con mayor precisión fenómenos inherentes a la interlengua en estudio. En la medida en que se disponga de corpus en otras lenguas con el mismo tipo de anotación, se van a poder realizar estudios contrastivos muy valiosos para contribuir en otras áreas de investigación. De acuerdo con la literatura mencionada, el diseño de las taxonomías se ha basado en tres criterios:

- *El criterio descriptivo o de modificación de la lengua objeto de estudio (LO):* este criterio busca identificar las formas y estructuras presentes en la conformación de la interlengua. Las categorías para su clasificación corresponden a la adición, omisión, falsa elección de elementos y algunas más específicas para determinar otro tipo de características como, por ejemplo, la elección errónea, la sustitución o forma errónea.
- *El criterio lingüístico o gramatical:* este criterio delimita el nivel o forma en que se encuentra el error, determinando así su localización, ya sea a nivel de la palabra, la oración o el párrafo. Se relaciona con las características del error en sus formas gramaticales, ortográficas, sintácticas, semánticas, pragmáticas o discursivas de la lengua.
- *El criterio etiológico:* en este caso se trata de inferir las causas de los errores sobre la base de la transferencia negativa de la lengua materna del aprendiente. Se formulan hipótesis sobre las posibles fuentes que producen el error. Desde esta perspectiva, se intenta explicar la desviación del aprendiente desde un punto de vista psicolingüístico para determinar la forma en la que opera el proceso de aprendizaje de la LO.

Además, para el proceso de identificación y etiquetado de los errores, es necesario acordar con el evaluador de los textos la forma de identificar y proceder a la descripción, la cual se sugiere que sea desde lo que se observa y está incorrecto, evitando así interpretar lo que debería ser correcto o lo que debería ir.

En el ámbito de los corpus de aprendientes de ELE, existen corpus orales y escritos recolectados en formato computacional. La mayoría de estos corpus consideran variables como el nivel de competencia, la lengua materna y años de estudio de los aprendientes con el objeto de identificar aspectos diferenciadores del proceso de adquisición. Uno de los más representativos por su rigor metodológico y tamaño es el Corpus CEDEL2. Este surge en el seno del proyecto Word Order in Second Language Acquisition Corpora

(WOSLAC), dirigido por la doctora Amaya Mendikoetxea. El objetivo del programa de investigación viene marcado por interesantes hallazgos en L2 para determinar el papel que tienen las interfaces en el desarrollo de la interlengua (Lozano & Mendikoetxea, 2007, 2013). CEDEL2 es un corpus escrito de aprendientes anglófonos de español de niveles de competencia inicial, intermedio y avanzado, según los resultados obtenidos en un test de diagnóstico estandarizado.

Por su parte, Cestero Mancera y Penadés (2009), de la Universidad de Alcalá, han dado origen al corpus Corane (Corpus para el Análisis de Errores de Aprendices de E/LE). Este es uno de los pocos corpus de corte longitudinal en ELE. El corpus está integrado por 1091 composiciones escritas por 321 aprendientes de diversas lenguas (entre ellas: inglés, alemán, francés, japonés, pakistání y húngaro); tiene niveles de competencia elemental, intermedio, avanzado y superior. La recolección del corpus se llevó a cabo en el ambiente académico universitario.

Otro corpus interesante es el Corpus de Aprendices de Español como Lengua Extranjera (CAES) (Rojo & Palacios Martínez, 2015), el cual ha sido recientemente compilado. Corresponde a un conjunto de textos producidos por estudiantes de ELE de todos los niveles de competencia. Los textos han sido generados por aprendientes de diversas lenguas maternas: árabe, chino, mandarín, francés, inglés, portugués y ruso. El total de textos recolectados es de 3878 producidos por 1423 estudiantes que escribieron dos o tres textos cada uno.

Taxonomías de errores etiológicos en el ámbito de ELE

La aplicación de una taxonomía para la identificación de errores etiológicos permite indagar el estado de la interlengua. El criterio etiológico es muy relevante en cuanto al estudio de la interlengua, dado que aporta conocimiento sobre los rasgos y características provenientes, ya sea de la relación interlingüística entre la lengua materna, la lengua extranjera o segunda que se aprende, y otras terceras lenguas aprendidas. Así, las propuestas de taxonomías de diferentes autores (Alexopoulou, 2005, 2006; De Alba Quiñones, 2009; Ferreira, 2014; Ferreira, Elejalde & Vine, 2014; Elejalde & Ferreira, 2016) buscan un acceso hacia los fenómenos descriptibles y, de esta manera, entregar mayor información que posibilite una posterior corrección o tratamiento del error.

Algunas taxonomías con criterio etiológico consideran problemáticas como las que se describen en la Tabla 1. En esta se ilustran las diferentes formas de clasificación inter e intralingüística propuestas por varios autores a lo largo de las últimas décadas. En esta tabla, se observa una división entre: (1) los errores interlingüísticos, los cuales son producto de la transferencia directa negativa de la L1 u otras lenguas previamente

aprendidas; y (2) los errores intralingüísticos que se asocian a la evolución o desarrollo (*developmental errors*), cuya fuente se origina del uso de mecanismos y estrategias de tipo universal independiente de la L1. En el caso de los errores producto de la transferencia de otras lenguas, existen clasificaciones propuestas como la *transferencia directa* entendida como el traspaso directo tanto de estructuras gramaticales como sintácticas, la traducción literal o falsas analogías provocadas por la L1 (Vázquez, 1991, 1999; Alexopoulou, 2006; De Alba Quiñones, 2009).

En cuanto a la producción de la lengua meta basada en las hipótesis que hace el aprendiente, los errores intralingüísticos se han clasificado de acuerdo con las diferentes estrategias que usa el aprendiente. No obstante, dichas clasificaciones en algunos criterios podrían superponerse por tratarse de errores que provienen del desarrollo de la interlengua. Por ejemplo, la simplificación (Tabla 1), es un tipo de clasificación que permite identificar errores producto de la reducción del sistema a reglas mínimas que asocia el aprendiente. Esto ocurre cuando el aprendiente tiene una competencia de nivel principiante en la L2, donde puede utilizar estructuras o formas lingüísticas más simples, según lo que haya aprendido en dicha lengua. Otro ejemplo de esta clasificación es la evasión de estructuras lingüísticas de mayor complejidad, como en el caso del español el uso del subjuntivo. Así, el estudiante prefiere usar estructuras que ya conoce y son simples frente aquellas que le generan inseguridad por el grado de dificultad o de comprensión de las mismas.

Tabla 1. Tipología de errores proveniente de diferentes estudios (De Alba Quiñones, 2009; Alexopoulou, 2005, 2006; Ferreira, Elejalde & Vine, 2014; Elejalde & Ferreira, 2016)

Categoría	Profundidad	Nivel
Interlingüístico	Transferencia directa L1/L3	Uso erróneo del sistema gramatical, léxico, semántico o sintáctico de la L1 a la L2/LE.
	Traducción literal de la L1/L3	Traspaso palabra por palabra de la lengua materna u otras lenguas aprendidas a la L2/LE.
	Analogías provocadas por el sistema lingüístico de la L1	Gramatical, léxico, semántico, sintáctico.
	Falsos cognados	Semántico y léxico.
	Calcos semánticos	Provenientes de estructuras semánticas apropiados para una estructura de la LO, pero que no son acertados en su contexto.
	Cambio de código	Uso de la escritura y código alfabético de la L1/L3 hacia la LO.

Categoría	Profundidad	Nivel
Intralingüístico	Desarrollo de la interlengua	Confusiones por semejanza morfológica u otro aspecto de la lengua.
	Simplificación	Omisión o evasión de la regla.
	Sobregeneralización	Uso de la misma regla en distintos aspectos de la lengua.
	Superproducción	Uso reiterativo de la regla por desconocimiento de otras o inseguridad frente a las hipótesis no comprobadas.
	Falsa correspondencia gramatical	Cuando se desconoce o se usa erróneamente el significado en donde la LO tiene una o más correspondencias.
	Errores producto del uso de otras estrategias de aprendizaje	De producción, comunicativas, compensatorias o recuperatorias.

Por otra parte, otras clasificaciones como la sobregeneralización y la superproducción son estrategias que usa el aprendiz para equilibrar o solucionar vacíos comunicacionales.

En relación con la estrategia de la sobregeneralización, el estudiante, a partir de una regla global, aplica la misma norma en toda su producción lingüística sin considerar las posibles excepciones. Un ejemplo de esto es el uso de la categoría de género gramatical en español. En estos casos, los aprendientes suelen generalizar la regla de que todos los sustantivos terminados en «-a» son femeninos. Sin embargo, algunas lenguas que caracterizan el rasgo de género, como el alemán, suelen transferir el género de los sustantivos. Por ejemplo, en la oración «...vamos a hacer una *tour** en barca...» (Sujeto 10, alemán) el sustantivo *tour* es femenino en alemán, por lo que el sujeto utiliza el artículo en femenino, interfiriendo negativamente en la estructura morfológica de género en el sustantivo. De esta manera, el sujeto transfiere los rasgos morfológicos de su lengua materna, considerando similitudes en ambas lenguas.

Con respecto a la superproducción, refiere al uso deliberado y reiterativo de la regla para suplir vacíos comunicacionales. Es decir, cuando un estudiante se enfrenta a situaciones comunicativas que exigen estructuras complejas, usa aquellas que tiene internalizadas y con las cuales puede dialogar fácilmente. En este escenario, podría reconocerse una superposición con la sobregeneralización-superproducción y en contraste con la simplificación. En un sentido crítico, ambas clasificaciones operan dentro de la identificación objetiva en la sistematización de un error, conduciendo así a la posibilidad de separar lo que corresponde al uso de la regla, sujeta al conocimiento adquirido por el estudiante, del uso deliberado por desconocimiento o inseguridad frente a problemáticas lingüísticas complejas.

PROPUESTA DE UNA TAXONOMÍA PARA ERRORES ETIOLÓGICOS

La propuesta de taxonomía que se presenta en este artículo para identificar los errores de interlengua con criterio etiológico, proviene del desarrollo de proyectos de investigación previos (Ferreira, 2011, 2014) centrados en la habilidad escrita en el español como lengua extranjera. En su origen, el desarrollo de la taxonomía, si bien incluye el criterio lingüístico y etiológico para identificar y clasificar los errores, estaba más orientada a delimitar problemáticas centradas en aspectos gramaticales (Ferreira, Elejalde & Vine, 2014; Elejalde & Ferreira, 2016). En el proyecto actual de investigación (Ferreira, 2018), el propósito es identificar errores de transferencia negativa. Por este motivo, la perspectiva más bien es etiológica, y para ello se ha revisado y adaptado la taxonomía inicial, para que pueda dar cuenta de manera fehaciente de las problemáticas de la interlengua de aprendientes de ELE. Una vez que se ha organizado la estructura principal para la identificación y el etiquetado de errores, se han ido explicitando las etiquetas para clasificar los errores producto de la conformación de la interlengua. A continuación, se explica la constitución y evolución de la propuesta taxonómica de errores de escritura de aprendientes de ELE.

Contexto investigativo de la taxonomía de errores etiológicos

La primera aproximación al análisis de errores respondió a la necesidad de determinar problemáticas de la escritura de los estudiantes extranjeros que tomaban cursos de ELE en la universidad (Ferreira, 2011). En este caso, se recolectó un pequeño corpus para identificar errores y definir la línea de trabajo. De esta forma, los errores detectados, inicialmente, fueron organizados de la manera como se presenta en la Tabla 2.

Tabla 2. Primera organización de la identificación de errores

Nivel	Tipo de error	Clasificación
Palabra	Gramatical, semántico	Lingüística
Oración	Concordancia sintáctica	Lingüística
Párrafo	Coherencia textual	Lingüística

En un primer esquema, fue necesario establecer los criterios principales para reconocer los diferentes tipos de errores y, así, construir la taxonomía base. Como se observa en la Tabla 2, la clasificación se organizó en función de cómo los errores se iban a identificar de acuerdo con la profundidad del análisis, esto es, a nivel de palabra, oración y párrafo. Luego de esta clasificación, se procedió a buscar la forma de dar cuenta del error, para lo cual se estableció una anotación simple en Word de lo observado por

el identificador/etiquetador. En este paso, se determinó que el etiquetador debía señalar lo incorrecto y no la estructura correcta. Es decir, si había una palabra que faltaba, el etiquetador señalaría el error y escribiría una pequeña descripción del error sin entrar a interpretar. Por ejemplo, cuando se identificaban errores donde faltaba un elemento, se procedía de la siguiente manera:

Yo voy a *visitar** *mi amiga* en USA (Sujeto 4, LI inglés).

Omisión de la preposición *a* para introducir complemento directo de persona.

En otros casos, cuando se identificaron errores en los cuales el aprendiente usaba elementos de la LI y creaba por derivación las palabras, se definieron características de tipo morfológico, producto de la utilización de las reglas de la LI, L2 y la LO. Un ejemplo de este tipo es el siguiente:

«El *cazabando** fue al bosque» (Sujeto 9, LI inglés).

Creación del sustantivo a partir de la derivación del español
y la combinación del sustantivo en inglés «*hunting*»

Una vez establecido este criterio y cómo se definiría la jerarquía, se hizo la revisión de propuestas ya referenciadas en la sección anterior, como la de Alexopoulou (2005, 2006) y la de De Alba Quiñones (2009). Con dichos modelos, se procedió a clasificar las problemáticas de acuerdo con los criterios de modificación de la LO.

El segundo aspecto que se consideró en la delimitación de la taxonomía corresponde al *criterio descriptivo o de modificación de la LO*. Este se refiere a las alteraciones que usa el aprendiente presentes en el error. Para este tipo de clasificación existen varias propuestas de anotación. A continuación, se presenta una definición de las categorías de este criterio de acuerdo con estudios anteriores sobre análisis de errores en ELE (Tabla 3).

Tabla 3. Descripción del error (Alexopoulou, 2005, 2006; Vázquez, 1991, 2009; Fernández, 1991; Tono, 2003; Rakaseder & Schmidhofer, 2014; Ferreira, Elejalde & Vine, 2014; Elejalde & Ferreira, 2016; Ferreira & Elejalde, 2017)

Clasificación	Descripción
Adición	Se adhieren uno o varios elementos innecesarios a la producción de la LO.
Omisión	Se omite un elemento o varios dentro de la producción de la LO.
Elección falsa/ falsa selección	Se elige una forma gramatical o semántica de acuerdo con la fuente de la LM o la L2.

Clasificación	Descripción
Elección errónea	Elección ortográfica inadecuada por desconocimiento de la regla o asociada a las reglas de acentuación o puntuación de la LM.
Forma errónea	Elección incorrecta del significado de expresiones o palabras homógrafas/homófonas por desconocimiento o confusión del uso de estas.
Colocación falsa	Alteración del orden de los elementos dentro de un sintagma, oración o párrafo en la producción de la LO.

Con respecto a la profundidad y el nivel de análisis de los errores y su localización dentro del sistema lingüístico, es necesario establecer las categorías sobre las que operan las etiquetas (Tabla 3). En ese aspecto, las propuestas de estudios previos (Ferreira, Elejalde & Vine, 2014; Rakaseder & Schmidhofer, 2014) buscan desglosar en qué categoría del sistema lingüístico se ha cometido el error. Obsérvese en la Tabla 4 los tipos de categoría y profundidad de análisis.

Tabla 4. Criterio para clasificar errores lingüísticos

Categorías	Profundidad	Nivel de análisis
Gramática	Categorías gramaticales, morfosintaxis	Palabra, sintagma, oración
Ortografía	Acentuación y literal	Palabra
Concordancia gramatical	Concordancia sintáctica	Sintagma, oración y párrafo
Coherencia textual	Uso discursivo de los conectores, conjunciones y puntuación	Párrafo, composición completa de la escritura

Como se observa en la Tabla 4, las categorías y profundidad se organizan de la siguiente manera: (1) gramática (categorías gramaticales y morfosintaxis de la conformación de la palabra), cuyo análisis opera al nivel de la palabra, el sintagma o la oración; por ejemplo, la omisión o adición de una categoría gramatical en un sintagma, entre otros; (2) ortografía, respecto del uso de la acentuación y la ortografía literal de los grafemas, cuya identificación se ubica en el nivel de análisis de palabra; (3) concordancia gramatical, que corresponde a las normas lingüísticas de concordancia entre los elementos de la oración, y cuyo análisis se localiza en el nivel de la oración, sintagma o párrafo; y (4) coherencia textual, en relación con el uso de los conectores y marcadores discursivos, correspondientes al nivel de análisis en el párrafo y la composición completa de la escritura.

Evolución de la taxonomía en estudio del Corpus Especializado de Aprendientes de Español como Lengua Extranjera (Caele) y la importancia de errores etiológicos

La primera aplicación de la taxonomía arrojó resultados de frecuencias de errores en un nivel lingüístico (Ferreira, Elejalde & Vine, 2014). No obstante, dichos errores carecían de una posible interpretación sobre su origen o fuente. Para ello, se propuso ampliar la taxonomía en la descripción de errores a partir del criterio etiológico. Este criterio clasifica los errores en relación con la influencia de la lengua materna y las estrategias utilizadas por el aprendiz. Estos errores se dividen en dos categorías: (1) interlingual (transferencia y problemas relacionados con la L1 u otras lenguas aprendidas), y (2) intralingual (problemáticas relacionadas con la independencia y desarrollo de la L2: sobregeneralización, neutralización, desconocimiento de la regla) (Tabla 5).

En relación con la propuesta, se puede señalar que, desde el origen del esquema, los errores se identificaban a nivel de la palabra y la oración. Los rangos de identificación fueron más simples y se posicionaron en el reconocimiento de errores más generales, como en el caso de la clasificación interlingual, por ejemplo, *uso de la L1*, cuyo objetivo fue identificar y detectar cualquier posible interferencia en general de la L1 del estudiante. Principalmente, esta etiqueta operó en inglés y alemán, dado que los estudiantes provenían de los Estados Unidos, Inglaterra, Alemania y Canadá. Otro ejemplo fue el uso de la etiqueta *traducción literal*, para detectar aquellas estructuras que fueron traspasadas de la traducción de la L1 hacia la L0. Como puede evidenciarse, el uso de estas etiquetas fue general, sin considerar el criterio descriptivo que reconoce la modificación de la L0 a partir de la L1.

El Corpus Caele para la realización de los estudios de interlengua

El Caele (Ferreira, 2014-2018) es una colección de 780 textos escritos producidos por 135 aprendientes de español como lengua extranjera (ELE), recogidos entre los años 2014 y 2018. El Caele ha sido desarrollado en el marco de los proyectos de investigación Fondecyt n.º 1140651 (Ferreira, 2014) y Fondecyt n.º 1180974 «Diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua» (Ferreira, 2018). Los estudiantes provienen de los programas de intercambio que tiene la Universidad de Concepción en el marco de la movilidad estudiantil. Los países de origen son, principalmente, Francia, Bélgica, Alemania, Sudáfrica, Estados Unidos, Inglaterra, Brasil, Portugal, entre otros. Las lenguas maternas de los estudiantes son el inglés, francés, alemán, y portugués. La recolección de los textos se realiza durante las clases de ELE de los niveles A2 avanzado y B1, donde asisten los estudiantes durante su semestre de intercambio. La secuencia textual es predominantemente narrativa con

variadas temáticas sobre cultura, gastronomía, música y problemáticas juveniles. El nivel de competencia de los alumnos extranjeros fue determinado a través de la Prueba de Multinivel con Fines Específicos Académicos (Ferreira, 2016), instrumento alineado con los niveles del Marco Común Europeo de Referencia (Consejo de Europa, 2002).

Tabla 5. Taxonomía para clasificación de errores proyecto Fondecyt n.º 1110812 y n.º 1180974¹

Criterio	Nivel	Categoría	Subcategoría	Adición	Falsa selección	Omisión	
Lingüístico	Palabra/oración	Gramática	Categorías gramaticales	ADI-PREP	FS-PREP	OMI-PREP	
			Coherencia textual	Concordancia sintáctica	FS-CG-GEN		
				Puntuación	ADI-coma	FS-coma	OMI-coma
				Conectores	ADI-CONNECT	FS-CONNECT	OMI-CONNECT
	Palabra	Léxico	Léxico creado por derivación	FS-LEX-creado-DERIV			
			Léxico morfología	FS-LEX-MORF			
			Léxico innecesario	ADI-LEX-INNE			
		Ortografía	Acentual	ADI-ORTO-grave	FS-ORTO-grave	OMI-ORTO-grave	
			Literal	ADI-ORTO-LIT	FS-ORTO-LIT	OMI-ORTO-LIT	
			Dierética	ADI-ORTO-hiato	FS-ORTO-hiato	OMI-ORTO-hiato	
	Etiológico	Palabra/Oración	Uso L1	uso-L1			
			Uso L2	uso-L2			
			Orden de palabras	orden de palabras			
Analogía semántica			analogía semántica				
Traducción literal			traducción literal				
Sobregeneralización			sobregeneralización				
Simplificación			simplificación				
Error de tipeo			error tipeo				

¹ En cuanto a la equivalencia de las etiquetas expuestas en esta tabla, convendría señalar que, específicamente en el criterio lingüístico, responden a las tres identificaciones de la taxonomía en función de la modificación de la LO. Es decir, las abreviaturas ADI-FS-OMI, corresponden a la adición, falsa selección y omisión de una categoría gramatical, ortográfica o léxica. Por ejemplo, en la subcategoría *categorías gramaticales* se encuentra como ejemplo etiquetas para identificar problemáticas en las preposiciones. Para esto se usa ADI-PREP (adición de la preposición), FS-PREP (falsa selección de la preposición) y OMI-PREP (omisión de la preposición). De esta manera, el etiquetador podrá clasificar con mayor precisión la problemática que observa en un nivel más profundo.

Con respecto al tiempo de escritura, los textos fueron recolectados durante tres meses con un intervalo de un texto cada 15 días. La extensión de los textos fue de 150-200 (A2) y 250-400 (B1) palabras y fueron escritos en el bloc de notas para cautelar el uso de correctores ortográficos que intervinieran en el uso real por parte del sujeto.

El objetivo principal de este corpus es estudiar diferentes fenómenos que atañen a la conformación de la interlengua de aprendientes de ELE, de acuerdo con la producción más auténtica y real de muestras escritas. El formato digital que se utiliza es en *txt*, almacenado en Sketch Engine y procesado con el programa UAM Corpus Tool para la identificación de tendencias y usos erróneos en la escritura.

Etiquetamiento de errores etiológicos, propuesta actualizada

El criterio etiológico referido a la clasificación interlingüística, se organiza considerando los aspectos provenientes de la LI del aprendiente hacia la LO. De esta forma, cuando la taxonomía opera, se clasifica de acuerdo con el error identificado, considerando además si tiene uno o más errores y su posible explicación. Los errores interlingüales se distribuyen principalmente en transferencias directas o indirectas de la LI u otras lenguas aprendidas hacia la LO, uso de traducción literal, analogías falsas, falsos cognados, calcos semánticos y cambio de código. Todas estas reflejan cómo el aprendiente usa una estrategia determinada para solucionar un problema de comunicación en la LO, en la cual desconoce la norma o correspondencia de esta.

En esta propuesta, se busca cubrir todos los aspectos posibles, integrando de manera adecuada la descripción de la modificación de la LO, el nivel del sistema lingüístico en el que se encuentra el error, si es de tipo oracional, por palabra o de párrafo y su posible explicación (criterio lingüístico, descriptivo, explicativo). Por ejemplo, el criterio de omisión, adición, falsa selección, describe el error que se transfiere directa o indirectamente de la lengua materna o de otras lenguas aprendidas. Asimismo, otras categorías buscan describir otros procesos del desarrollo de la interlengua, como en el caso de la neutralización, las transferencias basadas en la LI, la sobregeneralización, entre otras. En la Figura 1, se puede ver el desglose de una etiqueta a modo de ejemplo para ilustrar la utilización de la taxonomía y su sistema de anotación.

La estructura de la etiqueta permite identificar el error de acuerdo con la descripción de la modificación de la LO, en este caso *omisión* (OMI) de la categoría gramatical, seguido de la localización y profundidad de la categoría gramatical en el *artículo definido* (ART-DEF). De esta forma, la etiqueta refleja que el error está localizado en las categorías gramaticales en un nivel de profundidad de análisis por palabra.

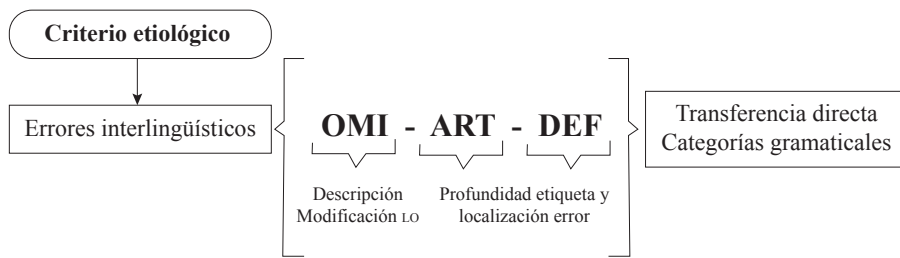


Figura 1. Etiqueta para el uso de la taxonomía en errores de transferencia directa de categorías gramaticales

En la Figura 2, la etiqueta «FS-CG-GEN-atributo» permite dar cuenta de un nivel de mayor profundidad y localización del error en el sintagma oracional. Asimismo, permite etiquetar sobre una estructura completa para determinar el tipo de error en la oración. Al igual que la etiqueta anterior, la descripción refiere a la *falsa selección* (FS) de la *concordancia gramatical* (CG) del *género* (GEN) en el *atributo* de las oraciones copulativas con ser y estar.

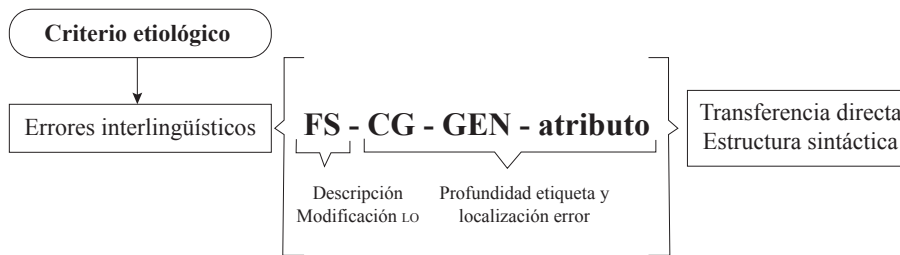


Figura 2. Etiqueta para el uso de la taxonomía en errores de transferencia directa de estructuras sintácticas

Ahora bien, considerando las problemáticas que surgen durante el etiquetado y la clasificación de los errores, también pueden proponerse diferentes metodologías de análisis para un etiquetado objetivo y con mayor precisión. En este punto, cabe señalar dos modalidades de etiquetamiento: (1) etiquetado objetivo y preciso de lo que se observa y se identifica, cautelando que lo observado no corresponda a una interpretación subjetiva de lo que debería encontrarse; y (2) identificación del problema y su solución o forma esperada, cuya etiqueta debería reflejar la interpretación del etiquetador. Esta segunda

modalidad, considerando el cambio constante durante la conformación de una interlengua, podría representar un problema dado su carácter subjetivo. No así con la modalidad más objetiva, donde el etiquetador identifica el problema y refleja lo observado.

Sin embargo, esta modalidad no está exenta de dificultades para el etiquetador, dado que en muchas ocasiones algunos errores podrían representar más de una problemática o podrían tener ambigüedad en su significado o la estructura con la que construye la oración. El etiquetador con criterio etiológico debe identificar cuáles de los errores podrían atribuirse a una posible interferencia de la LI u otras lenguas aprendidas, o al desarrollo propio de la interlengua. En este último punto, surgen cuestiones interesantes sobre cómo etiquetar problemas relacionados con estructuras o con categorías gramaticales, puesto que en ambos casos se identifica el error, pero aparece la dificultad para dirimir sobre etiquetar la alteración de acuerdo con la categoría gramatical o la función. Por ejemplo, cuando el aprendiente comete el error dentro de una perífrasis verbal, omitiendo la preposición «a», en «voy comer con mi amiga». En este ejemplo, se puede etiquetar la omisión a partir de la categoría gramatical o identificar la problemática en la perífrasis de acuerdo con la caracterización del error y la problemática que se pretende estudiar.

Considerando la presentación de la composición de las etiquetas y posibles procedimientos metodológicos para su uso, a continuación, se explica la taxonomía propuesta acorde con la identificación de errores producidos por la transferencia (interlingüísticos) y los errores referidos a los niveles de desarrollo que tiene la interlengua (intra lingüísticos).

DESCRIPCIÓN DE LA TAXONOMÍA PARA CLASIFICAR ERRORES ETIOLÓGICOS

El criterio etiológico permite determinar la fuente del error y entregar una posible explicación de acuerdo con el contraste entre los sistemas lingüísticos en estudio. Así, la fuente podría originarse desde la interferencia de la LI u otras lenguas aprendidas o producto de la evolución misma de la interlengua del aprendiente.

De esta manera, la clasificación de este criterio se distribuye en dos grandes categorías: errores interlingüísticos y errores intra lingüísticos. En este contexto, la propuesta que se presenta en este artículo responde a la evolución taxonómica tanto de estudios previos como del proyecto Fondecyt n.º 1180974 (Ferreira, 2018), dando como resultado la clasificación taxonómica para identificar y clasificar errores de ELE. En la Tabla 6, se presenta la taxonomía propuesta.

Tabla 6. Propuesta de taxonomía con criterio etiológico

Clasificación	Tipo	Profundidad	Nivel	Descripción LO
Interlingüístico	Transferencia directa	Categorías gramaticales	Palabra	Omisión, adición, falsa selección y forma errónea
		Concordancia sintáctica	Oración	
		Estructura morfológica	Palabra	
		Léxico	Palabra	
		Coherencia textual	Párrafo/texto	
		Cambio de código	Palabra	
		Falsos cognados	Palabra	
		Traducción literal	Oración	
		Interferencia de otras lenguas aprendidas	Palabra	

Intralingüístico	Neutralización	
	Sobregeneralización	
	Hipercorrección	
	Simplificación	Aplicación incompleta de la regla
		Aplicación incorrecta de la regla
		Desconocimiento de la regla
Léxico creado por derivación		

La Tabla 6 muestra el detalle de la clasificación propuesta, donde se observan dos grandes fenómenos, el interlingüístico y el intralingüístico. Los errores interlingüísticos refieren al traspaso directo de la LI u otras lenguas aprendidas, distribuyéndose la clasificación en nueve categorías: (1) transferencia directa del uso de las categorías gramaticales al nivel de la palabra, específicamente en el sintagma oracional; (2) transferencia del uso de la estructura y concordancia sintáctica de la oración, considerando las diferencias en la concordancia sintáctica del sujeto y del predicado, y la concordancia entre los modificadores de los sintagmas de la oración; (3) transferencia de rasgos morfológicos, especialmente en el género y número de los sustantivos; (4) transferencia de léxico a nivel de significado, referida a la expansión de una palabra con varias acepciones; (5) transferencia de la coherencia textual relacionada con el uso de conectores y el hilo conductor del texto; (6) cambio de código respecto de palabras escritas en la LI dentro de una oración; (7) falsos cognados, palabras que por similitud gráfica se confunden con el significado real en la LO; (8) traducción literal considerando el traspaso palabra por palabra de una estructura; y (9) interferencia de otras lenguas aprendidas diferentes a la LO.

Con respecto a los errores intralingüísticos, relacionados con el desarrollo y evolución de la interlengua de la LO, se distribuyen en cinco categorías: (1) neutralización referida a la correspondencia dual entre uno o más usos de diferentes categorías gramaticales o significados de una palabra, (2) sobregeneralización en el uso de ciertas reglas del sistema lingüístico, (3) hipercorrección de palabras de acuerdo con la norma de la LO, (4) simplificación del sistema lingüístico y (5) léxico creado por derivación relacionado con la invención de palabras a partir de la estructura de la LO y la LI. En el siguiente apartado se profundiza en cada nivel y tipo de error de acuerdo con la taxonomía propuesta.

Errores interlingüísticos producto de la transferencia directa negativa de la L1 o L2 hacia la LO

Las semejanzas o diferencias de la LI pueden contribuir a transferencias positivas o interferir negativamente induciendo a un error en el uso, omisión o falsa selección de cualquier estructura del sistema lingüístico de la LO. El fenómeno de la transferencia surge de la necesidad, por parte del aprendiente de lenguas, para solucionar alguna situación comunicativa (de la cual no es consciente en algunas ocasiones), para lo cual recurre a su lengua materna con el objeto de suplir la falencia. No obstante, existen otras estrategias que usa el aprendiz, donde el grado de conocimiento o de monitorización (conciencia de los errores) juega un papel importante cuando se comunica en la LO. Por consiguiente, los errores interlingüales son aquellos que provienen de la influencia directa o indirecta de la lengua materna u otras lenguas aprendidas por el aprendiente. En esta categoría, se clasifican los errores acordes con la naturaleza de la influencia, el tipo de estrategia utilizada por el aprendiente y el tipo de error cometido. Las clasificaciones existentes, como la propuesta por Alexopoulou (2005, 2006, 2011) o Vázquez (1991) son de tipo general y describen los errores según la taxonomía de la modificación de la lengua meta (Bustos, 1998; Santos Gargallo, 2004; Alexopoulou, 2005; Ferreira, Elejalde & Vine, 2014; Ferreira & Elejalde, 2017). A continuación, se presentan las categorías de acuerdo con el nivel de profundidad en los errores de transferencia directa de la LI hacia la LO.

Errores de transferencia directa de categorías gramaticales

La subcategoría de la transferencia directa de una estructura gramatical se entiende como el traspaso directo del uso de las categorías gramaticales de la LM hacia la L2, lo cual induce a errores en la L2 o LE. Este tipo de etiquetamiento busca clasificar y entregar información sobre cómo afecta el uso erróneo y cuáles son las tendencias en las

categorías gramaticales. En este aspecto, se consideran todas las categorías gramaticales. Para ello, se determinó una estructura de etiquetamiento que se ilustra en la Tabla 7.

Los errores de transferencia directa corresponden al traspaso de una estructura gramatical, léxica o sintáctica hacia la LO que hace el estudiante desde su lengua materna. Así, la taxonomía que clasifica este tipo de errores (Tabla 7) muestra una subdivisión en las categorías gramaticales del sistema lingüístico. En este aspecto, se propone identificar errores en siete categorías principales: (1) la preposición, considerando el uso erróneo afectado por la influencia de las reglas gramaticales en la LI; (2) el sustantivo en la misma categorización, es decir, cuando el aprendiente elige erróneamente un adjetivo en vez de un sustantivo o la omisión de este en el sujeto; (3) el adjetivo, considerando la falsa selección (FS) entre adjetivos, sustantivos y adverbios, o la adición (ADI) u omisión (OMI) de estos; (4) falsa selección, omisión o adición de los pronombres (personal, relativo, demostrativos); (5) falsa selección u omisión en los cinco aspectos del verbo (modo, tiempo, voz, perfectivo y formas no personales); y (6) artículos definidos e indefinidos, en relación con la identificación del sujeto, persona o cosa.

Tabla 7. Errores de transferencia de estructuras gramaticales

Profundidad	Categoría	Etiqueta ejemplo
Uso incorrecto de las categorías gramaticales a nivel de palabra u oración	Preposición	OMI-PREP-A
		FS-PREP-CON
		ADI-PREP-DE
		OMI-SUST
	Sustantivo	FS-SUST
		ADI-SUST
	Adjetivo	OMI-ADJ
		FS-ADJ
		ADI-ADJ
	Pronombre	FS-PRON-IMPER
		FS-PRON-relativo
	Verbo	FS-modo-SUBJ/INDIC/IMPER
		FS-tiempo verbal
		FS-formas-personales
		FS-verbo-AUX
	Artículos	OMI-ART-DEF
		FS-ART-INDEF
		ADI-ART-DEF

Todas estas categorías presentan normas lingüísticas que, dependiendo de la lengua en estudio, tienen similitudes o diferencias que inducen a error, como en el caso del inglés, donde claramente ciertas estructuras erróneas son atribuibles a reglas del sistema en dicha lengua. Obsérvense las divisiones en algunos ejemplos extraídos del Corpus Caele (Ferreira 2014; 2018) (Tablas 8, 9 y 10).

Tabla 8. Ejemplos lingüísticos de etiquetamiento en la transferencia directa en el uso de preposiciones

Ejemplo	Explicación
«...Es muy probable que va a comer algo adentro () * una cosa de greda, como pastel de choclo...» (Sujeto 11, A2 inglés). <i>Etiqueta: OMI-PREP-DE</i>	En este caso, hay una omisión de la preposición «de», proveniente de la estructura «inside () * a thing», cuya conformación muestra claramente la preposición en inglés sin otra adicional, seguido del elemento regido por la preposición. Este tipo de casos es muy recurrente, dado que son estructuras complejas y que funcionan en conjunto, es decir, la tipificación del uso de las preposiciones corresponde a su uso por estructura y caso de la situación comunicativa.
«...En Inglaterra la música típica es muy <i>diferente a</i> * la música típica...» (Sujeto 23, B1 inglés). <i>Etiqueta: FS-PREP-A</i>	En este ejemplo, se observa una falsa selección de la preposición «a» proveniente de la estructura «different to» (diferente de) traducida por la preposición «a» como posible correspondencia múltiple en español. En este tipo de oraciones, el estudiante hace una falsa analogía del uso homólogo entre las preposiciones, usando la misma acepción que identifica como equivalente en su LI.
«...una acción concreta para ayudar () * los músico chileno es de hacer...» (Sujeto 3, francés) <i>Etiqueta: OMI-PREP-A</i>	En francés, <i>ayudar (aider)</i> es un verbo transitivo en este contexto y no introduce con la preposición «a» para complemento de persona. Solo se indica la persona con ausencia de una preposición.

Tabla 9. Ejemplos lingüísticos de etiquetamiento en la transferencia directa en el uso del artículo

Ejemplo	Explicación
Otro ejemplo de esta estructura típica errónea, corresponde a: «... () * Personas pasan tiempo con amigos y familia...» (Sujeto 53, B1 inglés). <i>Etiqueta: OMI-ART-DEF</i>	En esta oración, se observa nuevamente la omisión del artículo en generalizaciones del sustantivo colectivo, en este caso del sustantivo «people», el cual se utiliza en inglés de forma singular.

Tabla 10. Ejemplos lingüísticos de etiquetamiento en la transferencia directa en el uso de verbos

Ejemplo	Explicación
«...Me encanta que <i>hay</i> * tanta diversidad en mi país...» (Sujeto 13, A2+ inglés). <i>Etiqueta: FS-modo-INDIC</i>	En este ejemplo, se puede observar cómo las reglas del sistema verbal del inglés afectan la estructura en el uso del subjuntivo en oraciones que determinan los gustos e intereses subordinadas por la conjunción «que». Es decir, el aprendiente hace una falsa selección del modo indicativo proveniente del inglés, cuyas estructuras se usan en presente de indicativo.
«...Fuimos en el teleférico y <i>pode-mos</i> * ver la ciudad...» (Sujeto 39, A2+, portugués). <i>Etiqueta: FS-tiempo-PRET-simple</i>	Otro ejemplo que ilustra este tipo de error en la falsa selección del presente de indicativo en español, donde nuevamente se observa el traspaso del indicativo, producto del vacío lingüístico entre el pretérito imperfecto del subjuntivo y el presente de indicativo en sus usos más comunes en el español.

Errores de transferencia directa en la concordancia sintáctica

En la Tabla 11, se presentan los errores de transferencia directa en las categorías de género y número en la concordancia gramatical a nivel del sintagma nominal (modificadores directos del sustantivo y el sujeto), verbal y la concordancia gramatical entre sujeto y predicado en la oración. Se evidencia la clasificación de dos aspectos principales en la concordancia sintáctica: la primera referida al uso incorrecto de estructuras concordantes en la LI transferidas de forma negativa a la LO; y la segunda relacionada con el orden canónico de la oración del español, muchas veces influenciada la producción de esta como ELE por el orden de la LI, como en el caso del inglés y el alemán.

Tabla 11. Errores de transferencia en la concordancia sintáctica de la oración

Profundidad	Tipo de error	Etiqueta
Uso incorrecto de la estructura sintáctica en la concordancia gramatical debido al traspaso de la LI a la LO	Concordancia gramatical de género	FS-CG-GEN
	Concordancia gramatical de número	FS-CG-NUM
	Concordancia gramatical de atributo	FS-CG-ATRIB
	Concordancia gramatical sujeto predicado	FS-CG-SUJ-PRED
	Estructuras pasivas/impersonales	FS-ESTRUCT-TIP

Profundidad	Tipo de error	Etiqueta
Orden canónico de la LI hacia la LO. Falsa colocación de los elementos dentro del sintagma o en la oración.	Orden sintáctico de la oración	FC-ELEM-SN (ADJ-SUST)
		FC-ELEM-SV (V-ADV)
		FC-PRON

Con respecto a la primera clasificación, se encuentran las etiquetas correspondientes a problemáticas en las categorías morfosintácticas de género y número a nivel de los sintagmas (FS-CG-NUM/GEN). Es decir, cuando no existe concordancia entre el sustantivo y los modificadores dentro de los sintagmas. La tercera y cuarta etiqueta se refieren a errores detectados a nivel de la concordancia entre sujeto-cópula-atributo y sujeto-predicado en oraciones de predicado verbal. La quinta etiqueta corresponde a errores en las estructuras pasivas e impersonales de la oración, donde la mayoría de los casos se transfieren en inglés o alemán el uso de la segunda persona del singular como oraciones impersonales. Por último, los errores relacionados con el orden canónico de la oración responden a la estructura típica de las oraciones en español, las cuales en la producción de ELE se ven afectadas por la transferencia del orden en los sintagmas nominales y verbales, antecediendo los modificadores directos al núcleo de los sintagmas. Muchas veces, este cambio del orden no afecta el significado de la oración, pero, refiriéndose al estilo de la escritura, se consideraría un error irritable dentro de la comunicación. Obsérvense algunos ejemplos (Tablas 12 y 13).

Tabla 12. Ejemplos lingüísticos de etiquetamiento en la transferencia directa en la concordancia sintáctica

Ejemplo	Explicación
«...Unos de las ferias mas <i>popular*</i> es lo de Sante Fe, Nuevo México...». (Sujeto 3, A2+, inglés). <i>Etiqueta: FS-CG-NUM</i>	En esta oración puede revisarse la falsa selección en la concordancia gramatical del adjetivo con el sujeto, proveniente del inglés cuyo rasgo no posee número o género ni concuerda con el sujeto en cualquier modificador del núcleo; así, el traspaso proviene de «One of the most* popular holiday in New Mexico», donde el superlativo «the most*» no posee el rasgo referido anteriormente.
«... <i>Dicen la gente*</i> que la comida cocida en grueda es mejor...». (Sujeto 11, A2 inglés). <i>Etiqueta: FS-CG-SUJ-PRED</i>	En este error, se ve una falsa selección de la concordancia en el número gramatical entre el sujeto y el verbo por influencia del inglés, donde el sustantivo colectivo está en plural «people are», cuya correspondencia en español es singular.
«...Los mariscos <i>es exquisito*</i> ...». (Sujeto 12, A2+ inglés). <i>Etiqueta: FS-CG-NUM-TRIB</i>	En este ejemplo se ilustra la interferencia del inglés debido al sustantivo singular e invariable «Seadfood is exquisite», dado que las oraciones copulativas en inglés carecen de atributos concordantes en género y número con el sujeto de la oración.

Tabla 13. Ejemplos lingüísticos de etiquetamiento en la transferencia directa en el orden sintáctico de la oración

Ejemplo	Explicación
«...Por ejemplo, es un <i>muy quieto lugar*</i> y conozco mucha gente aquí...» (Sujeto 14, alemán). <i>Etiqueta: FC-orden-ELEM</i>	En esta oración, se puede ver el adjetivo antepuesto al sustantivo, típica estructura del alemán en el sintagma nominal de cualquier parte de la oración. El error no afecta la comunicación o significado del mensaje, pero no responde a las normas de estilo y concordancia en el orden sintáctico del español.
«...La protagonista se llama Katniss y tiene que ser 10 días <i>con 22 otros jóvenes*</i> en una arena...» (Sujeto 10, B I alemán). <i>Etiqueta: FC-orden-ELEM</i>	Como se evidencia en la oración, este error proviene de la estructura canónica del alemán, donde se transfiere el orden del adjetivo antepuesto al nombre y determinante « <i>muss mit 22 anderen Jugendlichen</i> ». Este orden altera la concordancia en el orden de los elementos, provocando falta de sentido y coherencia en el mensaje que quiere transmitir el aprendiente.

Errores de transferencia directa de rasgos morfológicos, léxicos, interferencia gráfica y coherencia textual

En este último apartado, los errores, por lo general, corresponden más a la forma superficial de la oración y del mensaje que quiere transmitir el aprendiente de lenguas. Es decir, no afecta el significado de la situación comunicativa que se pretende dar, aun cuando altera las normas de estilo o de concordancia en general del español. Para etiquetar este tipo de errores, se fijan parámetros como rasgos genéricos de la morfología que modifican la estructura original de la LO, expansión del vocabulario, pequeñas interferencias en la escritura gráfica de la palabra y uso de conectores como organizadores discursivos del texto. A continuación, se explican las etiquetas y se ejemplifican (Tabla 14).

Tabla 14. Errores de transferencia morfológica

Profundidad		Tipo de error	Etiqueta
Trasaso de los rasgos morfológicos de la LI hacia la LO	Rasgos morfológicos del sustantivo	Género	FS-GEN
		Número	FS-NUM

Como se muestra en la Tabla 14, con respecto a los errores de tipo morfológico, los aprendientes traspasan rasgos morfológicos de género y número. Este tipo de error se clasifica especialmente respecto de lenguas con reglas de género y número similares, como en el caso de las lenguas romances, germánicas u otras con sistemas lingüísticos parecidos. Las etiquetas para clasificar este tipo de errores corresponden a FS-GEN (falsa selección de género) y FS-NUM (falsa selección de número). Obsérvense algunos ejemplos en la Tabla 15.

Tabla 15. Ejemplos lingüísticos de etiquetamiento en la transferencia directa de rasgos morfológicos

Ejemplo	Explicación
«...Voy <i>un vez</i> a semana <i>al clase*</i> ...» (Sujeto 10, B I alemán). <i>Etiqueta: FS-GEN</i>	En este error, se observa la transferencia del rasgo morfológico del alemán al sustantivo «clase», el cual tiene género masculino en alemán «Unterricht». Muchos sustantivos alemanes o franceses no concuerdan con el género de español, produciendo errores que afectan la relación entre el género del sustantivo y el uso de modificadores directos, como en el uso de los determinantes.
«...No tengo un objeto favorito pero <i>un cosa*</i> que es muy importante para mi es mi piano...» (Sujeto 15, alemán B I). <i>Etiqueta: FS-GEN</i>	Como se observa en la oración, este error proviene del rasgo de género del sustantivo «Ding» (cosa), el cual tiene artículo neutro, más cercano al masculino en alemán. El uso en español, el aprendiente lo determina por el género masculino, sin considerar que el género es femenino.
«...Chile habla a menudo de temas como <i>los problemas*</i> en el país: los inegales...» (Sujeto 3 B I, francés) <i>Etiqueta: FS-GEN</i>	En francés, el sustantivo «problema» es masculino. Por esta razón, un aprendiente de L I francesa asigna a todos los nombres masculinos que terminan en «o».

Por otra parte, así como se ilustra en la Tabla 16, el uso del léxico de forma expansiva corresponde al traspaso del significado de las palabras. Para este tipo de error léxico, se determinó el uso de la etiqueta forma errónea o FS-LEX, debido al alcance que podría tener una palabra de la LO con múltiples correspondencias.

Tabla 16. Errores léxicos provenientes del significado y la escritura de la palabra

Profundidad	Tipo de error	Etiqueta
Léxico	Falsa elección del significado	FS-LEX/forma errónea
	Interferencia gráfica de la L I u otra lengua	INTERF-GRAF

Algunos ejemplos de este tipo de errores se evidencian en el uso de palabras como *por qué, porque, el porqué*; u otras como *rato, momento, ahora, después*, en relación con el tiempo. Por otra parte, con los errores de tipo gráfico como la interferencia gráfica dentro de la categoría léxico, corresponde a aquellas palabras escritas en la LO, pero que, en medio o al final, usan alguna tipificación de la L I (Tabla 17).

Tabla 17. Ejemplos lingüísticos de etiquetamiento en la transferencia directa de léxico e interferencia gráfica

Ejemplo	Explicación
«...A la fin* de mes vamos a viajar al Valparaíso...» (Sujeto 25, B1 inglés). <i>Etiqueta: forma errónea</i>	En este error, se observa el uso de final, proveniente del inglés «at the end* of the month», cuya correspondencia en español es doble a diferencia del inglés con una sola.
«...los “Korrigans” son pequeñas criaturas como nanos que poseen muchos <i>tresores*</i> » (Sujeto 4, B1 francés). <i>Etiqueta: INTERF-GRAF</i>	Interferencia de la escritura de la palabra «trésor» (tesoro) en francés hacia el español. En este caso, se observa una combinación entre la palabra francesa y el español.
«...LOS EE. UU. describa las modas de vida <i>deifférentes*</i> de las gentes estadounidenses...» (Sujeto 13, A2+ inglés). <i>Etiqueta: INTERF-GRAF</i>	Error inducido por interferencia gráfica del inglés proveniente de «different». En este error, se observa cómo el aprendiente, por transferencia, combina la escritura de la LO y la LI, resultando en un error que, si bien no afecta el significado, sí afecta la atención sobre la escritura en ELE.
«...Voy a visitar <i>Neuva*</i> York...» (S10, B1 alemán). <i>Etiqueta: INTERF-GRAF</i>	La escritura del adjetivo «nueva» de forma errónea proviene de la escritura en alemán «neu» produciéndose una combinación en la palabra «Neuva*».
«...tambien alusion a <i>sentimentos*</i> de lo día a día, como por ejemplo amor y sentimientos normales...» (sujeto 6, A2+ portugués) <i>Etiqueta: INTERF-GRAF</i>	Este sustantivo fue escrito con la interferencia del portugués, «sentimentos», donde claramente se observa el cambio de la sílaba «mientos» por «mentos», derivación del portugués.

Por último, los errores referidos a la coherencia textual (Tabla 18) corresponden al uso directo de los conectores provenientes de la LI hacia la LO, cuyo uso puede afectar el mensaje en cuanto a la coherencia, o no producir ningún efecto negativo.

Tabla 18. Errores de coherencia textual referido al uso de conectores

Profundidad		Tipo de error	Etiqueta
Conectores	Uso inadecuado de los conectores	Copulativos	FS-CONNECT-copulativo
		Disyuntivos	FS-CONNECT-disyuntivo
		Adversativos	FS-CONNECT-ADVER
		Causa efecto	FS-CONNECT-causa-EFEC
		Numeración	FS-CONNECT-numeración

A continuación, se pueden revisar algunos ejemplos en el uso erróneo de los conectores en español como LE (Tabla 19).

Tabla 19. Ejemplos lingüísticos de etiquetamiento en la transferencia directa de conectores

Ejemplo	Explicación
«... <i>Al primero</i> *, no sabía que mi teléfono no estaba en mis manos...» (Sujeto 56 B1, inglés). <i>Etiqueta: FS-CONNECT-numeración</i>	Uso del conector «at first», proveniente del inglés, cuya equivalencia corresponde «primero».
«... <i>Mientras la guerra</i> *, he tenido algunos muy lacrimógenos momentos y nunca realicé que un libro me podría hacerlo...» (Sujeto 19, B1 alemán). <i>Etiqueta: FS-CONNECT-temporal</i>	En alemán se utiliza la preposición («während»), la cual puede ser traducida al español como «durante» o como el adverbio «mientras» dependiendo del contexto.

Otros errores de transferencia

Además de los errores de transferencia directa explicados en la sección anterior, se consideraron como errores etiológicos aquellos provenientes de diferentes estudios previos (Dulay & Burt, 1974; Alexopoulou, 2006, 2011) que buscan determinar la similitud en términos de la morfología y la semántica del sistema lingüístico de la L0 y la L1, cuando ocurren diferentes tipos de transferencia.

Tabla 20. Errores por influencia de la L1

Clasificación	Profundidad	Etiqueta
Traducción literal	Traducción palabra por palabra de la L1	Traducción literal
Interferencia L2	Uso de palabras de otra L2	INTERF-L2
Cambio de código	Escritura de una palabra de la L1	Cambio código
Falso cognado	Traducciones incorrectas de la L1 a la L0	Falso cognado

En la taxonomía que se propone (Tabla 20) se considera etiquetar los errores correspondientes a las siguientes problemáticas generadas por influencia de la L1:

- *Traducción literal*: este error corresponde a la traducción palabra por palabra de una oración completa o de alguno de los sintagmas oracionales; por lo general, se observa en los niveles iniciales de la L0.
- *Interferencia de otra L2*: esta etiqueta clasifica errores de otras lenguas aprendidas por parte del aprendiente diferentes a la L0. Muchas veces los aprendientes durante el *code switching* utilizan tanto palabras de la L1 como de otras L2, L3 existentes.
- *Cambio de código*: esta etiqueta corresponde a la escritura de una palabra en la L1 por parte del estudiante sin alterar el resto de la estructura de la oración que está en español.
- *Falso cognado*: referido a los falsos amigos provenientes de la confusión y asimilación de significados erróneos de la L1 hacia la L0.

Obsérvense en la Tabla 21 algunos ejemplos de las etiquetas de acuerdo con las problemáticas identificadas de tipo intralingüístico.

Tabla 21. Ejemplos lingüísticos de etiquetamiento de otros errores de transferencia

Ejemplo	Explicación
«...Me parece, las ferias con <i>supor-te*</i> local son buenas lugares para la gente...» (Sujeto 11, A2 inglés). <i>Etiqueta: falso cognado</i>	En esta oración se evidencia un falso cognado del inglés «support» (apoyo), cuyo uso lo modifica, considerando tanto las reglas del español como de la L1.
«...Deje entonces las canciones <i>quedaran*</i> muy famosas en todo lo país...» (Sujeto 6 A2+, portugués) <i>Etiqueta: traducción literal</i>	Traducción literal del verbo «ficar*», <i>quedarse</i> en español, cuyo significado no se corresponde con el traspaso a la L2. En este caso convendría utilizar «se volvieron o tornaron» muy famosas.
«...Nuestras <i>classes*</i> tienen lugar una vez de la semana...». (Sujeto 21, B1 alemán) <i>Etiqueta: cambio de código</i>	En este error, se evidencia el cambio de código de una sola palabra en la oración sin afectar el significado de esta. Muchas veces ocurre esto por descuido o por la similitud de la palabra en la L1 y la L0.
«...A las una y <i>quatro*</i> ...» (Sujeto 10, A2 alemán) <i>Etiqueta: INTERF-L2</i>	Interferencia del inglés en la producción escrita de un estudiante alemán, donde la escritura de <i>cuarto</i> corresponde en inglés «a quarter». Este error se clasificaría como una interferencia de otra L2, aprendida previamente, en el español.

Errores intralingüísticos producto del desarrollo de la interlengua de la L0

Finalmente, la taxonomía para clasificar los errores etiológicos contempla también los errores intralingüísticos, los cuales corresponden a la aplicación de las reglas del sistema lingüístico de la L0 independiente de la L1, manifestándose principalmente en estrategias basadas en la simplificación y sobregeneralización de las estructuras de la L0. Sin embargo, algunos de estos errores también poseen rasgos de la interferencia negativa de la LM hacia la L2, de los cuales puede destacarse la estrategia de la neutralización. Los errores producto de dicha estrategia corresponden a la interferencia por bifurcación tanto morfosintáctica como léxico-semántica, donde el sistema lingüístico de la L0 posee dos realizaciones. A continuación, se presentan la clasificación y los ejemplos de este tipo de errores (Tabla 22).

Tabla 22. Errores intralinguales producto de la conformación de la L0

Nivel	Explicación del criterio	Tipo de error	Etiqueta ejemplo
Neutralización	Cuando en la L0 existen más correspondencias a diferencia de la L1	Confusión de <i>ser</i> y <i>estar</i> , <i>por</i> y <i>para</i> , otros	FS-NEUTRA
Sobregeneralización	El aprendiente sobregeneraliza las reglas aplicando las formas lingüísticas sin conocimiento de las excepciones	Morfema de género (-a/-o) y número	Etiqueta lingüística+INTRA
Hipercorrección	El aprendiente corrige estructuras dadas las normas de la L0	Corrección de nombres, adjetivos, estructuras sintácticas	FS-hipercorrección
Simplificación	Léxico por derivación de la L2	Usa reglas de la L1 y la L0 para crear palabras	FS-LEX-DERIV
	Aplicación incompleta de la regla	Usa la regla incompleta	Etiqueta lingüística+INTRA
	Aplicación incorrecta de la regla	No aplica adecuadamente la regla	
	Desconocimiento de la regla	Por desconocimiento comete error	

Como se especifica en la Tabla 22, las etiquetas buscan con su nomenclatura evidenciar el fenómeno de una forma concisa y concreta. En este caso, los errores intralinguales se dividen en las siguientes categorías:

- *Neutralización*: en este tipo de errores se evidencia la confusión presente en dos realizaciones de una misma palabra en español, a diferencia de la L1. Son comunes los errores de las preposiciones *por*, *para*, el uso de los verbos *ser* y *estar* y *haber*, donde el estudiante los usa de manera intercambiable en las oraciones. Es decir, como no identifica los usos de cada realización, muestra a lo largo de su escritura uso tanto de una como de otra realización.
- *Sobregeneralización*: los errores por sobregeneralización corresponden específicamente a la utilización de una misma regla sobre otras, desconociéndose el uso de excepciones a la regla. Es común identificar errores del género y número provenientes del inglés.
- *Hipercorrección*: corresponde a la corrección errónea que hace el aprendiente, considerando las reglas del sistema lingüístico. Este tipo de error no afecta el sentido de la oración, más que la forma estilística de la palabra corregida en el español.
- *Simplificación*: este es un error bastante común, dado que el aprendiente, considerando el nivel de dificultad o por desconocimiento de las reglas, simplifica el sistema

lingüístico de la LO. En esta categoría se consideran errores provenientes de la aplicación incorrecta de la regla, desconocimiento de la regla, léxico creado por derivación y aplicación incompleta de la regla.

Véanse algunos ejemplos que permiten dilucidar la aplicación de la taxonomía y el uso de las etiquetas intralingüísticas (Tabla 23).

Tabla 23. Ejemplos lingüísticos de etiquetamiento de errores intralingüísticos

Ejemplo	Explicación
<p>«...Todos los <i>chileanos</i>* adoran sus cosas de grueda y es probable que todas las familias tienen alguna en su cocina...». (Sujeto 11 A2, inglés). <i>Etiqueta: FS-LEX-DERIV</i></p>	<p>En esta oración, se puede observar cómo el aprendiente crea una nueva palabra por derivación, proveniente de la mezcla entre «Chilean» y la terminación del morfema derivativo en español «-os».</p>
<p>«...Sistema de micro para llegar a <i>nuestra hostel</i>*...». (Sujeto 25 B1, alemán). <i>Etiqueta: SOBREGEN-GEN</i></p>	<p>En este ejemplo, se evidencia la sobregeneralización del género femenino usando la forma del determinante femenino. Este error ocurre por la falsa asociación de las reglas morfológicas de género relacionadas con la terminación «-a/-o».</p>
<p>«...Debo ir a terminal Collado...» (Sujeto 14, inglés). <i>Etiqueta: FS-hipercorrección</i></p>	<p>En este ejemplo, se observa que la palabra «Collado», nombre de un lugar de un país, se corrige, desconociendo su forma original («Collao»), lo que podría suponer una corrección a la simplificación mal usada, como en el ejemplo «cansado/cansao*»; Bilbao/Bilbado» u otros parecidos. En este caso, esta hipercorrección correspondería al uso incorrecto de la regla de la terminación «-ado» en algunos nombres propios del español.</p>
<p>«...Si yo <i>estuviera</i>* una poeta o una narradora, me gustaría escribir sobre mi vida...». (Sujeto 3 A2, inglés). <i>Etiqueta: FS-NEUTRA</i></p>	<p>En esta oración, se puede identificar el uso del verbo <i>estar</i> de acuerdo con la neutralización que hace el sujeto, al desconocer las reglas de uso tanto del verbo <i>estar</i> como de <i>ser</i>. No solo se observa neutralización en <i>estar</i> y <i>ser</i>, sino también en el uso del subjuntivo y el indicativo, el uso de las preposiciones <i>por</i> y <i>para</i> y algunos pronombres interrogativos, como «por qué/porque*». En estos casos, el aprendiente de lenguas debe identificar con mayor precisión la realización y uso de los casos con doble correspondencia.</p>

CONCLUSIONES

En este artículo, se ha abordado la problemática de categorización y etiquetado de errores y se ha delimitado una propuesta para una taxonomía hacia la estandarización en la clasificación de errores de interlengua con un criterio etiológico.

Como se ha planteado en el desarrollo de dicha propuesta taxonómica, las problemáticas frecuentes para etiquetar un corpus surgen principalmente cuando se delimita el tipo de análisis o procesamiento de los textos en estudio. Es decir, el etiquetado debe responder, principalmente, a los objetivos de la investigación y, segundo, a la forma de recopilación de los datos para su posterior procesamiento. En este punto, cabe señalar que pueden presentarse dos modalidades para etiquetar, la primera relacionada con el etiquetado de lo que se observa y se identifica, cautelando que el problema observado no corresponda a una interpretación subjetiva de lo que debería encontrarse. La segunda modalidad corresponde a la identificación del problema y su solución o forma esperada, cuya etiqueta debería reflejar la interpretación del etiquetador. Esta segunda modalidad, considerando la dinamicidad y conformación de una interlengua, podría representar un problema dado su carácter subjetivo. No ocurre lo mismo con la modalidad más objetiva, donde el etiquetador identifica el problema y refleja lo observado. Sin embargo, esta modalidad no está exenta de interpretaciones de parte del etiquetador, dado que en muchas ocasiones algunos errores podrían representar más de una problemática o podrían tener ambigüedad en su significado o la estructura con la que construye la oración.

De acuerdo con lo sugerido en la literatura (Dagneaux, Denness & Granger, 1998; Díaz-Negrillo & Fernández, 2006), los investigadores de corpus en el ámbito del ELE debemos hacer nuestros esfuerzos por ponernos de acuerdo sobre un esquema general de anotación de error, sobre la base de las tendencias semejantes. La ventaja de la propuesta aquí sugerida es que se basa en la revisión de varias taxonomías previas y aprovecha sus principales aportes, adaptándolos a la realidad de los errores que se quieren identificar y clasificar. La diversidad de sistemas de etiquetado de errores evidencia una limitación para la anotación de errores. Las taxonomías de error tienden a explicar las diversas dimensiones de la clasificación de errores, las convenciones de codificación y los modelos de anotación, lo que demuestra que no existe un estándar de error. Surgen temas interesantes de abordar, como etiquetar problemas relacionados con estructuras o con categorías gramaticales, puesto que en ambos casos se identifica el error, pero ocurre la dicotomía sobre etiquetar la alteración de acuerdo con la categoría gramatical o la función. Por ejemplo, cuando el aprendiente comete el error dentro de una perífrasis verbal, omitiendo la preposición «a», en «voy comer con mi amiga»; se puede etiquetar la omisión a partir de la categoría gramatical o identificar la problemática en la perífrasis.

En la configuración de un sistema de anotación, la propuesta de los criterios señalados por Granger (2015) y Payrató (1995), si bien se ajusta a cualquier objetivo

de investigación, requiere un desarrollo de etiquetas para identificar errores que den cuenta de manera más exhaustiva de las problemáticas en estudio. En el caso de estos criterios, el sistema de taxonomía y anotación que se ha propuesto en este artículo, responde a las siguientes categorías: (1) *informativas*, considerando que las etiquetas tienen tres partes en la taxonomía propuesta, entregan la información necesaria para detectar rápidamente el nivel de profundidad, la localización del error y el criterio que se aplica; (2) *reutilizables*, en relación con la posibilidad de utilizar las etiquetas en otras lenguas, para este caso lenguas romances e indoeuropeas como la alemana; (3) *flexibles* para permitir la adición o eliminación de etiquetas considerando nuevas problemáticas identificadas; (4) *neutras*, evitando la modalidad de revisión subjetiva, para así determinar con mayor precisión la problemática; y (5) *universales*, lo que involucra la compatibilidad con otros sistemas de anotación, como el etiquetario del Corpus CAES (Instituto Cervantes, 2018), propuesto por el proyecto del Instituto Cervantes, u otros etiquetarios que permitan dar cuenta de errores de interlengua.

REFERENCIAS

- Alexopoulou, A. (2005). Aproximación al tratamiento del error en la clase de E/LE desde la perspectiva del análisis de errores. *Estudios de Lingüística Aplicada*, 23(41), 101-125.
- Alexopoulou, A. (2006). Los criterios descriptivo y etiológico en la clasificación de los errores del hablante no nativo: una nueva perspectiva. *Porta Linguarum*, 5, 17-35.
- Alexopoulou, A. (2007). Errores intralingüales e interlingüales en la interlengua escrita de aprendientes griegos de ELE. Trabajo presentado en el *Congreso de la Asociación Internacional de Hispanistas XVI: Nuevos caminos del hispanismo*. Madrid, España. Recuperado de https://cvc.cervantes.es/literatura/aih/pdf/16/aih_16_2_004.pdf
- Alexopoulou, A. (2011). El papel de la transferencia en los errores léxicos. *Revista Nebrija de Lingüística Aplicada*, 9(5), 27-36.
- Bustos, G. (1998). Análisis de errores: problemas de categorización. *Cuadernos de Filología Hispánica*, 16, 11-40.
- Cestero Mancera, A., & Penadés, I. (2009). *Corpus de textos escritos para el análisis de errores de aprendices de E/LE (Corane)*. Alcalá de Henares: Universidad de Alcalá.
- Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas*. Strasburgo: Consejo de Europa, Ministerio de Educación, Cultura y Deporte/Instituto Cervantes.
- Corder, P. (1967). The significance of learner's errors. *International review of Applied Linguistics in Language Teaching*, 5, 1-4.
- Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided error analysis. *System*, 26(2), 163-174.

- De Alba Quiñones, A. (2009). El análisis de errores en el campo de ELE. Algunas cuestiones metodológicas. *Revista Nebrija de Lingüística aplicada a la enseñanza de Lenguas*, 5. Recuperado de <https://www.nebrija.com/revista-linguistica/el-analisis-de-errores-en-el-campo-del-esp%C3%B1ol-como-lengua-extranjera.html>
- Díaz-Negrillo, A. (2007). *A fine-grained error tagger for learner corpora* (tesis de doctorado). Universidad de Jaén, España.
- Díaz-Negrillo, A., & Fernández, J. (2006). Error tagging systems for learner corpora. *Revista española de lingüística aplicada*, 19, 83-102.
- Dulay, H., & Burt, M. (1974). You Can't Learn without Goofing: An Analysis of Children's Second Language Errors. En J. C. Richards (ed.), *Error Analysis* (pp. 95-123). Londres: Longman.
- Elejalde, J., & Ferreira, A. (2016). Errores de transferencia en comunidades de aprendizaje en línea por aprendientes de español como lengua extranjera (ELE). *Trabalhos em Linguística Aplicada*, 55(3), 619-650.
- Fernández, S. (1991). *Interlengua y análisis de errores en el aprendizaje del español como lengua extranjera* (tesis de doctorado). Universidad Complutense de Madrid, España.
- Ferreira, A. (2011). *Proyecto de investigación Fondecyt n.º 1110812: un sistema tutorial inteligente para la focalización en la forma en la enseñanza del español como lengua extranjera*. Universidad de Concepción, Chile.
- Ferreira, A. (2014). *Proyecto de investigación Fondecyt n.º 1140651: el feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera*. Universidad de Concepción, Chile.
- Ferreira, A. (2014-2018). *Corpus de Aprendientes de Español Lengua Extranjera, Caele*. Universidad de Concepción, Chile.
- Ferreira, A. (2016). *Prueba de multinivel con fines específicos académicos*. Universidad de Concepción, Chile.
- Ferreira, A. (2018). *Proyecto de investigación Fondecyt n.º 1180974: diseño e implementación de un corpus escrito de aprendientes de ELE en formato computacional para el análisis de la interlengua*. Universidad de Concepción, Chile.
- Ferreira, A., & Elejalde, J. (2017). Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, Caele. *Revista Brasileira de Linguística Aplicada*, 17(3), 509-538.
- Ferreira, A., Elejalde, J., & 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.
- Granger, S. (2004). Computer learner corpus research: current status and future prospects. *Language and Computers*, 52(1), 123-145.

- Granger, S. (2015). Contrastive Interlanguage Analysis: A Reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7-24.
- Instituto Cervantes. (2018). Corpus de aprendices de español como lengua extranjera (CAES). Recuperado de <http://galvan.usc.es/caes>
- Lozano, C., & Mendikoetxea, A. (2007). Learner Corpora and the Acquisition of Word Order: A Study of the Production of Verb-Subject Structures in L2 English. En M. Davies, P. Rayson, S. Hunston & P. Danielsson (eds.), *Proceedings of the Corpus Linguistics Conference*. Birmingham: University of Birmingham.
- Lozano, C., & Mendikoetxea, A. (2013). Learner corpora and Second Language Acquisition: The Design and Collection of CEDEL2. En A. Díaz-Negrillo, N. Ballier & P. Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 65-100). Amsterdam: John Benjamins.
- Payrató, L. (1995). Transcripción del discurso coloquial. En L. Cortés Rodríguez (ed.), *El español coloquial. Actas del I Simposio sobre Análisis del Discurso Oral*. Almería, España.
- Rakaseder, J., & Schmidhofer, A. (2014). Errores de producción de textos por parte de germanoparlantes en las primeras etapas de aprendizaje de ele. *MarcoELE Revista de Didáctica*, 19, 1-17.
- Rojo, G., & Palacios Martínez, I. M. (2015) Corpus de aprendices de español (CAES). *Journal of Spanish Language Teaching*, 2(2), 194-200.
- Santos Gargallo, I. (2004). El análisis de los errores en la interlingua del hablante no nativo. En J. Sánchez Lobato & I. Santos Gargallo (eds.), *Vademécum para la formación de profesores* (pp. 391-410). Madrid: SGEL.
- Tono, Y. (2003). Learner corpora: Design, development and applications. En D. Archer, P. Rayson, A. Wilson & T. McEnery (eds.), *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 800-809). Lancaster: University Centre for Computer Corpus Research on Language.
- Vázquez, G. (1991). *Análisis de errores y aprendizaje de español/lengua extranjera*. New York: Peterlang.
- Vázquez, G. (1999). ¿Errores? ¡Sin falta! Madrid: Edelsa.
- Vázquez, G. (2009). El concepto de error: estado de la cuestión y posibles investigaciones. *Revista Nebrija de Lingüística aplicada a la enseñanza de Lenguas*, 5.