

PALABRAS PROBLEMÁTICAS Y FRASES INCORRECTAS: UNA SOLUCIÓN AUTÓNOMA PARA DETECTAR LO INDETECTABLE.

FULGENCIO HERNÁNDEZ GARCÍA
CENTRO DE ENSEÑANZA INFANTIL Y BÁSICA ARTEAGA

Resumen: El presente estudio da cuenta de la investigación que se ha llevado a cabo para evaluar un corrector gramatical que da respuesta de forma efectiva a dos de los objetivos del currículo de la Educación Secundaria Obligatoria (ESO) y de Bachillerato: el aprendizaje autónomo y el uso de las tecnologías de la información y la comunicación. El corrector gramatical e-gramm puede engranar las actividades de escritura del inglés como lengua extranjera con el aprendizaje autónomo a través de sus dos herramientas, Palabras Problemáticas y Frases Incorrectas. Este corrector permite no sólo detectar muchos de los errores sintácticos y léxicosemánticos que suelen pasar desapercibidos en otros correctores, sino que parece que también se le puede enseñar a que corrija el 100% de los errores encontrados en corpus de aprendices.

Palabras clave: *aprendizaje autónomo, corrector gramatical, e-gramm, escritura del inglés como lengua extranjera, errores, corpus de aprendices.*

Abstract: This article will report the research carried out to evaluate a grammar checker that fulfils efficiently two of the objectives in the curriculum for ESO (lower secondary education) and Bachillerato (higher secondary education): autonomous learning and the use of information and communication technologies. The grammar checker “e-gramm” can marry English writing as a foreign language activities with autonomous learning through its tools Problem Words and Incorrect Sequences. This checker can detect not only many of the syntactic and lexicosemantic errors that are usually undetected by other checkers, but can also, it appears, be trained to detect 100% of the errors found in learner corpora.

Keywords: *autonomous learning, grammatical checker, e-gramm, English writing as a foreign language, errors, learner corpus.*

1. Introducción

En el currículo de la Educación Secundaria Obligatoria (en adelante ESO) (BORM 24.09.07) y de Bachillerato (BORM 10.09.08) de la Región de Murcia, se destaca la importancia del aprendizaje autónomo y de las nuevas tecnologías en los objetivos generales de la lengua extranjera.

6. Desarrollar la autonomía en el aprendizaje, reflexionar sobre el mismo y transferir al estudio de la lengua extranjera los conocimientos adquiridos en la lengua materna o en el aprendizaje de otras lenguas.
7. Utilizar los recursos didácticos a su alcance (diccionarios, libros de consulta, materiales multimedia) incluidas las TIC, para la obtención, selección y presentación de la información oral y escrita de forma autónoma. (BORM 24.09.07: 27254)
7. Adquirir y desarrollar estrategias de aprendizaje autónomo empleando todos los medios a su alcance, incluidas las tecnologías de la información y la comunicación con el fin de utilizar la lengua extranjera de forma autónoma y para seguir progresando en su aprendizaje.
10. Afianzar estrategias de autoevaluación y de planificación en la adquisición de la competencia comunicativa en la lengua extranjera, desarrollando actitudes de iniciativa, confianza y responsabilidad en este proceso. (BORM 10.09.08: 28047)

El objetivo de desarrollar entre los alumnos el aprendizaje autónomo aparece tanto en el ámbito de las lenguas extranjeras como en el resto de materias del currículo. Sin embargo,

según Benson (2001), existe una tendencia a interpretar que aprendizaje autónomo es igual a aprender por uno mismo sin la ayuda de un profesor.

[...] it is often assumed that autonomy implies learning in isolation, learning without a teacher or learning outside the classroom, such that the relevance of the concept to language teaching is unclear. [...] These misconceptions are, at least in part, a result of terminological and conceptual confusion within the field itself. (Benson 2001: 1)

La definición de autonomía en el aprendizaje apunta más bien al desarrollo de un enfoque general en el proceso de aprendizaje y no a la adopción de un modo de aprendizaje particular. Se pretende que el alumno tome el control de su propio aprendizaje y decida sobre los recursos, contenidos, evaluación, etc.

Autonomy can be broadly defined as the capacity to take control over one's own learning. [...] is not a method of learning, but an attribute of the learner's approach to the learning process. [...] Learners who lack autonomy are capable of developing it given appropriate conditions and preparation. The conditions for the development of autonomy include the opportunity to exercise control over learning. The way in which we organize the practice of teaching and learning therefore have an important influence on the development of autonomy among our learners. (Benson 2001: 2)

Si aplicamos esta definición al ámbito de la destreza de la escritura en una segunda lengua, tendríamos que permitir que los aprendices pudieran tomar decisiones sobre cuándo, cómo y qué errores corregir en sus producciones escritas.

Cabe esperar que los alumnos cometan errores como parte implícita de un proceso de aprendizaje en el que nada puede ser asimilado totalmente hasta que todas las partes sean aprendidas y por ello, no deberíamos considerarlos un problema sino una señal de que los aprendices están adquiriendo la lengua extranjera con éxito. Están dando los pasos necesarios para que se produzca el aprendizaje (Corder 1981).

Los errores, por lo tanto, representan una posibilidad de aprendizaje muy importante.

[...] we have to change our attitude towards mistakes. We must not think of them as something negative which needs some kind of punishment. Think of correction as a way of giving information, or feedback, to your students, just when it will support their learning. (Edge 1989: 17)
El error como signo de progreso ha de interpretarse a la luz de la madurez de la persona. Un error sin importancia a los 6 años puede tenerla a los 8 años. De todo esto concluimos con rotundidad que carece de sentido utilizar el error con sentido punitivo. Quien se equivoca en el proceso de aprendizaje no merece castigo, sino explicación y aclaración (...) El error nos brinda un excelente instrumento diagnóstico del desarrollo mental del sujeto, de su nivel de ejecución, de destrezas, de atención, de estilo cognitivo, etc. Podríamos hablar de 'un buen error' si permite mejorar en la realización de otras tareas. (Torre, de la 1993: 78-84)

Consecuentemente, el error es algo necesario que el profesor tiene que permitir en el aula creando las condiciones necesarias para que sea el propio alumno el que los corrija y participe de forma activa en la comprensión de una hipótesis falsa y en la búsqueda de otra que fuese correcta. Esto es posible a través de la retroalimentación como una vía crucial para promover y consolidar el aprendizaje (Vygotsky 1978; Brophy 1981; Anderson 1982).

Sin embargo, la retroalimentación en la Enseñanza de Segundas Lenguas ha cambiado mucho en los últimos 20 años. Haylad y Hyland (2006) aseguran que la retroalimentación del profesor ha sido sustituida por la retroalimentación de los compañeros, los talleres de escritura, las conferencias, y la retroalimentación a través del ordenador. "Over the past 20 years, changes in writing pedagogy and research have transformed feedback practices, with teacher comments often supplemented with peer feedback, writing workshops, conferences, and computer-delivered feedback." (Hyland y Hyland 2006: 1)

Este último tipo de retroalimentación, es decir, la retroalimentación a través del ordenador, ha sido objeto de nuestra investigación.¹

Una retroalimentación indirecta que fuerce a los alumnos a reflexionar, analizar y decidir sobre sus errores está más en el enfoque del aprendizaje autónomo que una retroalimentación directa en la que el alumno recibe la corrección de sus errores sin ningún esfuerzo.

Los beneficios de la retroalimentación indirecta sobre la directa han sido ampliamente demostrados por los investigadores.

[...] researchers have argued for the superiority of indirect feedback: indicating an error through circling, underlining, highlighting, or otherwise marking it at its location in a sentence, with or without a verbal rule reminder or an error code, and asking students to make corrections themselves. Indirect feedback, it is claimed, forces students to be more reflective and analytical about their errors than if they simply transcribed teacher corrections (direct feedback) into the next draft of their papers. Since students are required by indirect feedback to take more responsibility for their errors, they are likely to learn more to make long-term progress in finding, correcting, and eventually avoiding errors. (Ferris 2002: 63)

La gran escasez de recursos que ofrezcan una retroalimentación indirecta a los errores que los alumnos cometen en sus redacciones, provoca que sea el profesor el que se lleve las redacciones a casa y las devuelva al alumno junto con la retroalimentación necesaria para que el aprendiz la modifique y la vuelva a entregar al profesor para una corrección final. El tiempo que requiere esta metodología reduce las oportunidades que el alumno tiene de escribir redacciones, ya que un mismo profesor enseña a varios grupos de alumnos y por lo tanto son muchas las redacciones que debe de corregir. El número de redacciones al trimestre podría estar entre las 3 o 4. Hay que tener en cuenta que el resto de destrezas (hablar, leer y escuchar) también requieren dedicación y práctica.

Como apuntan Lawley (1999) y Chacón-Beltrán (2007), la posibilidad de contar con una herramienta informática que permitiera al alumno avanzar de forma autónoma en la corrección de errores, reduce la búsqueda y la corrección de errores por parte del profesor pudiendo así dedicar más tiempo a otros aspectos como la secuenciación y planificación de ideas y la corrección de errores más complejos.

If the software enabled students to hand in assignments with no language errors then teachers' marking burdens would be reduced and more time could be spent on other aspects of the writing process.... But the ideal computer-mediated writing aid will not only help a student remove the errors from one composition on one occasion but also, and much more importantly, enable that student to avoid similar mistakes in the future. The ideal writing aid will teach. (Lawley 1999: 34)

Existen correctores gramaticales convencionales que están al alcance de la mayoría de los alumnos de los institutos de enseñanza secundaria, que sólo ofrecen una retroalimentación directa y que por lo tanto no entregan ningún tipo de explicación sobre los errores cometidos en las redacciones, sino que directamente brindan la corrección al alumno, la cual en muchos casos no es la correcta, o en ocasiones no detectan los errores. Uno de los correctores presente en la gran mayoría de ordenadores personales es el de Microsoft Office.

El principal problema que muestra este corrector para los aprendices de lenguas extranjeras, es que en su diseño no se adoptó un enfoque pedagógico que contemplara que los alumnos que están en la fase de aprendizaje del idioma, no son capaces de discriminar si las opciones de corrección que ofrece el programa son correctas o no, como haría un hablante nativo con un gran dominio de la lengua.

¹ El término *retroalimentación a través del ordenador* es una traducción del término inglés *computer-delivered feedback* usado por Hyland y Hyland (2006:1).

Al introducir en este corrector los errores *all people know* o *since is necessary continuously travel* que se encontraron en el corpus recopilado para la investigación, podemos ver que el corrector de Microsoft no detecta ningún error. Para corregir la oración *all people know* el aprendiz tendría que escribir *everybody knows*, y en el caso de la oración *since is necessary continuously travel* el alumno tendría que modificarla por *since it is continuously necessary to travel*. En estos casos se requeriría la siguiente retroalimentación.

- (1) Es un error común traducir *toda la gente* por *all people*. En inglés existe una única palabra para expresar esa idea, *everybody*, y aunque el concepto hace referencia a un grupo de gente, la palabra es singular con lo cual si el verbo que le sigue está en presente simple, se ha de tener en cuenta la forma de tercera persona del singular.
- (2) La oración *it is continuously necessary to travel* tiene los siguientes errores: omisión incorrecta del sujeto *it*, falta de la partícula *to* en el infinitivo *travel* y posición incorrecta del adverbio de frecuencia *continuously* que debería colocarse detrás del verbo *to be*.

Ofrecer este tipo de retroalimentación con los correctores gramaticales tradicionales no es posible debido a que muchos de estos correctores están basados en la tecnología *parsing* con la que se hace un análisis sintáctico del texto. De esta forma sólo se analizan errores gramaticales y no se considera el contenido semántico. En el funcionamiento de este tipo de correctores conocidos como *parser-based grammar checker*, primero se identifica cada una de las palabras del texto (función conocida como *tagging*) y luego se realiza un análisis sintáctico (función conocida como *parsing*). Por ejemplo en la expresión *I work here*, el corrector etiqueta *I* como sujeto, *work* como verbo y *here* como adverbio y analiza que la estructura es correcta. Si el usuario escribiera *she work here*, el analizador detectaría que la estructura gramatical es incorrecta y sugeriría cambiar *work* por *works*. Sin embargo errores como *all people know* no es capaz de destacarlo.

Tschichold (1999a) explica que tradicionalmente los correctores gramaticales se han diseñado con la tecnología *parsing* (análisis sintáctico):

Like all natural language processing (NLP) applications, grammar checkers handle language in a way which is very different from the way humans deal with language. In order to arrive at a (for humans) meaningful analysis of an input text, they typically submit the text to a series of independent modules which segment, tag and parse the input. The two processes of tagging (assigning exactly one part-of-speech tag to each word) and parsing (syntactic analysis) form the major hurdles in this analysis process. State-of-the-art taggers have a success rate of 95 to 98 per cent...with progress becoming more and more difficult and costly. Considering that these percentages are achieved on correct text, a higher percentage of wrong tags has to be expected when the tagger is dealing with non-native partly erroneous text. (Tschichold 1999a: 7)

Como sugiere esta autora (*ibid.*), este tipo de correctores han tenido un gran éxito en textos escritos por hablantes con un gran dominio de la lengua, sin embargo las producciones escritas de los aprendices de una lengua extranjera suelen contener muchos más errores, lo que complica en gran medida el éxito del programa.

Granger y Meunier (1994) en la evaluación que hicieron de los correctores gramaticales *Correct Grammar* y *Gramatik 5*, considerando las necesidades de hablantes de inglés como segunda lengua o como lengua extranjera, afirmaron que muchos de los correctores gramaticales estaban dirigidos a cualquier usuario, nativo o no. Sin embargo, parecía que estos habían sido diseñados pensando sólo en hablantes nativos.

El principal problema que destacan Gupta y Schulze (2009) es que el número de construcciones que el usuario podría introducir en el corrector es infinito, lo que supondría

que el programa tendría que considerar una lista infinita de estructuras. Esto hace que no se pueda acertar en un 100% en las sugerencias que se ofrecen al usuario para que corrija los errores. Debido a esto, tenemos que contar con que un estudiante de lengua extranjera va a cometer errores que el programa no va a poder corregir.

Philip Bolt llevó a cabo un estudio (Bolt 1992) para determinar los principales problemas que presentan los correctores gramaticales. Hizo un análisis de siete correctores: Correct Grammar for Windows, Right Writer for Windows, Grammatick for Windows, CorrecText, Reader, PowerEdit, LINGER. Bolt (ibid.) coincide con otros investigadores (Pennington 1992 y 1993; Jacob y Rogers 1999; Lawley 1999; Tschichold 2003) que muchos de los correctores gramaticales destinados a aprendices de inglés como lengua extranjera emiten diagnósticos erróneos o destacan errores donde no los hay.

Para Bolt (1992), uno de los principales problemas en el diseño de los correctores gramaticales es que la gramática es un fenómeno complejo.

[...] se trata de un proceso en el que las palabras se combinan y se relacionan entre sí, dando lugar a un rango de formas infinitas y por esta razón no sería realista esperar que estos programas fueran capaces de tratar de forma adecuada más de una proporción de las formas del inglés, aunque se usen diccionarios muy completos. (Bolt 1992: 49)

Correctores más actuales como *WhiteSmoke* (ver figura 1) o *Grammar Slammer* tampoco han sido diseñados para aprendices de inglés como lengua extranjera. En el caso de *WhiteSmoke*, sólo funciona a través de internet y no es muy pedagógico. En el caso de la redacción, *I would like to work in a farm. I like to work with the animals, they are beautiful. A farmer live in small villages where the people are sympatic and there aren't cars, planes, contamination and other problems that there is in big towns. But I think that if I work in a farm I will have to work very hard because is necessary. Farmers spent all their time in the farm and they are paid with few money.* *WhiteSmoke* ofrece la siguiente retroalimentación.

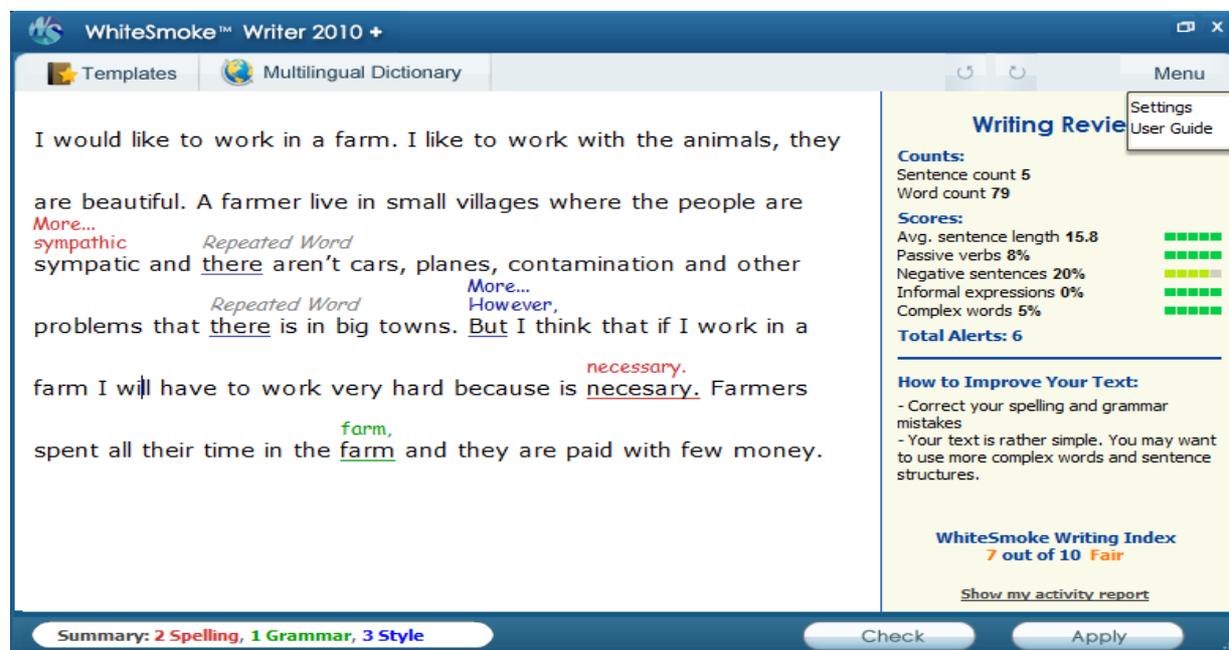


Figura 1: Corrector gramatical *WhiteSmoke*.

- (3) En el caso de *sympatic*, si hacemos click en *More ...*, *WhiteSmoke* sugiere 5 alternativas: *sympathic*, *sympatric*, *sympathica*, *sympathici* o *symplastic*. Sin embargo no sugiere *sympathetic* que sería la opción correcta.

- (4) Destaca la repetición de la palabra *there* pero no sugiere que hacer.
- (5) Si hacemos click en *But*, el programa sugiere que esta palabra debería cambiarse por *However*, *Nevertheless* o *Even so*. El mensaje que ofrece en inglés es: *a new sentence should not begin with a coordinating conjunction such as and or but. [...] coordinating conjunctions should be used to connect between items of the same class within a sentence. They should not be used to introduce a new sentence*. Sin embargo, no se explica cuál de las tres opciones elegir. De hecho, la opción *Even so* no es apropiada en este contexto.
- (6) El error ortográfico *necessary* lo corrige con *necessary*.
- (7) Finalmente, *WhiteSmoke* sugiere añadir una coma después de *farm*. Esto mejora la claridad del texto.

Sin embargo, lamentablemente, *WhiteSmoke* no detecta varios de los errores importantes: *in a farm* debería ser *on a farm*; *a farmer live* debería ser *a farmer lives*; *problems that there is* debería ser *problems that there are*; *because is* debería ser *because it is*; *Farmers spent* debería ser *Farmers spend*; *few money* debería ser *little Money*.

En el caso del corrector *Grammar Slammer*, se corrige automáticamente la palabra *necessary* por *necessary* y sugiere que *sympatic* se cambie por *synaptic* o por cualquier otra de las 30 alternativas que ofrece. También detecta el error *A farmer live in a small village* y destaca el error de concordancia entre el sujeto y el verbo. No encuentra ningún otro error.

Evidentemente, merecería la pena el esfuerzo de crear un nuevo corrector que sí fuese capaz de corregir estos errores y muchos otros más que cometen los estudiantes de inglés como lengua extranjera de un nivel A2/B1, como el autor de la redacción que acabamos de analizar.

Algunas de las ventajas que podría tener un nuevo corrector capaz de detectar los errores de estudiantes de inglés como lengua extranjera de un nivel A2/B1 son las que destacan Johnson y Sterkel (1984:36):

[...] help the writer learn principles of good writing by offering immediate, reliable, consistent, and in-depth feedback; reduce writing time by assisting the writer with editorial decisions; [...] allow students to use judgment to correct some of their errors, rather than just seeing them marked with a red pencil; [...] permit teachers to focus more on content ideas and organization in both instruction and grading than on language and mechanical errors.

El desarrollo de correctores como *e-gramm*, diseñados desde un enfoque pedagógico, permitiría que el profesor dejara de ser el único recurso capaz de ofrecer retroalimentación indirecta a los alumnos y que fuesen ellos los que de forma autónoma, decidieran cómo, cuándo y qué errores corregir en sus redacciones. Los alumnos al pegar sus textos en este corrector verán que aquellas palabras que son destacadas ofrecen una explicación para aprendices que le permitirá corregir los errores y por lo tanto mejorar su interlengua.

2. *E-gramm*

E-gramm corrige los errores de estudiantes españoles de inglés que se encuentran en un nivel A2 / B1 (clasificación en el MCER, 2001). El corrector funciona con una base de datos de errores típicos que suelen cometer los estudiantes de estos niveles y su función consiste en detectarlos en las redacciones de los alumnos para ofrecerles retroalimentación que guíe al alumno a la corrección de los mismos. Las instrucciones del corrector le indican al usuario que primero corrija los errores de ortografía usando el corrector ortográfico que incluyen los procesadores de texto como por ejemplo *Word Office* de *Microsoft* y después el usuario

puede pegar su texto en *e-gramm*. Una vez pegado, se procede a corregir primero las Secuencias Incorrectas (*Incorrect Sequences*) (ver figura 2) que según explica Lawley (2004) son combinaciones de palabras (u ocasionalmente una sola palabra) que normalmente no se usan en un inglés correcto con el significado que pretende el alumno. Por ejemplo, en la oración *the working of the actor must be very fun*, la secuencia *very fun* aparecería destacada como *Secuencia Incorrecta* y el corrector ofrecería la siguiente retroalimentación.

(8) *very fun*.

Esta secuencia no es correcta. Fue muy divertido = it was very funny (y nunca ...it was very fun...).

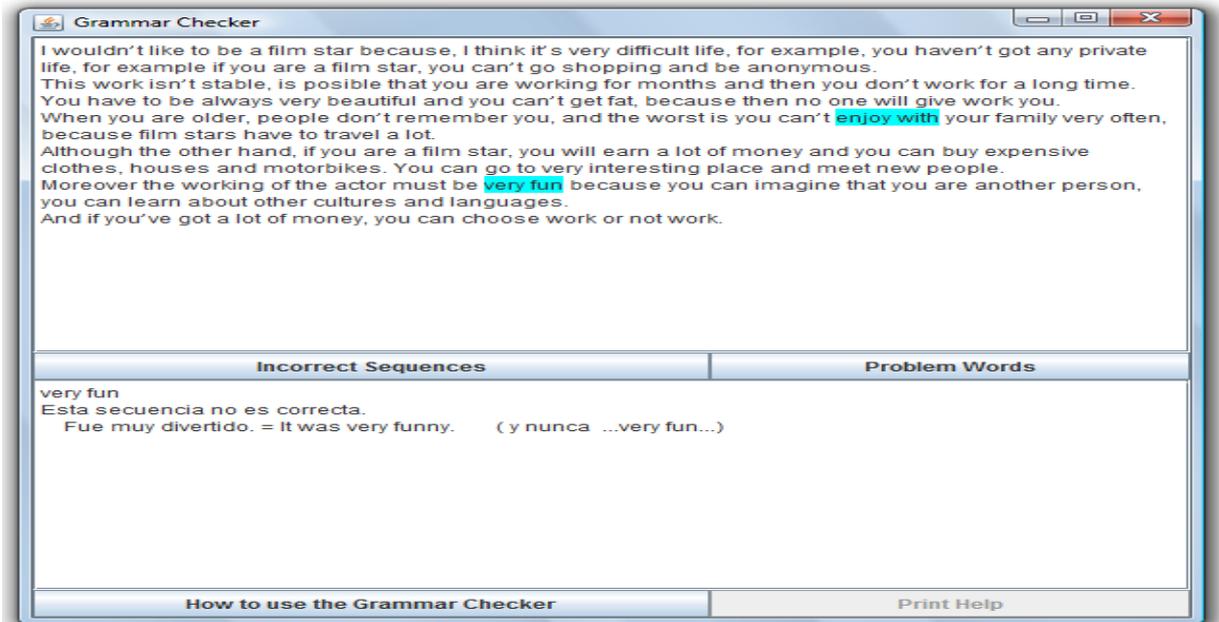


Figura 2: Corrector gramatical *e-gramm* 1.0. Uso de la herramienta *Incorrect Sequences*.

La herramienta de *Palabras Problemáticas* (*Problem Words*) (ver figura 3) destaca todas aquellas palabras consideradas como problemáticas tantas veces aparezcan en el texto. El hecho de destacarlas no quiere decir que se estén usando mal en ese contexto, sino que suelen estar asociadas con errores. Aquí podemos ver un ejemplo.

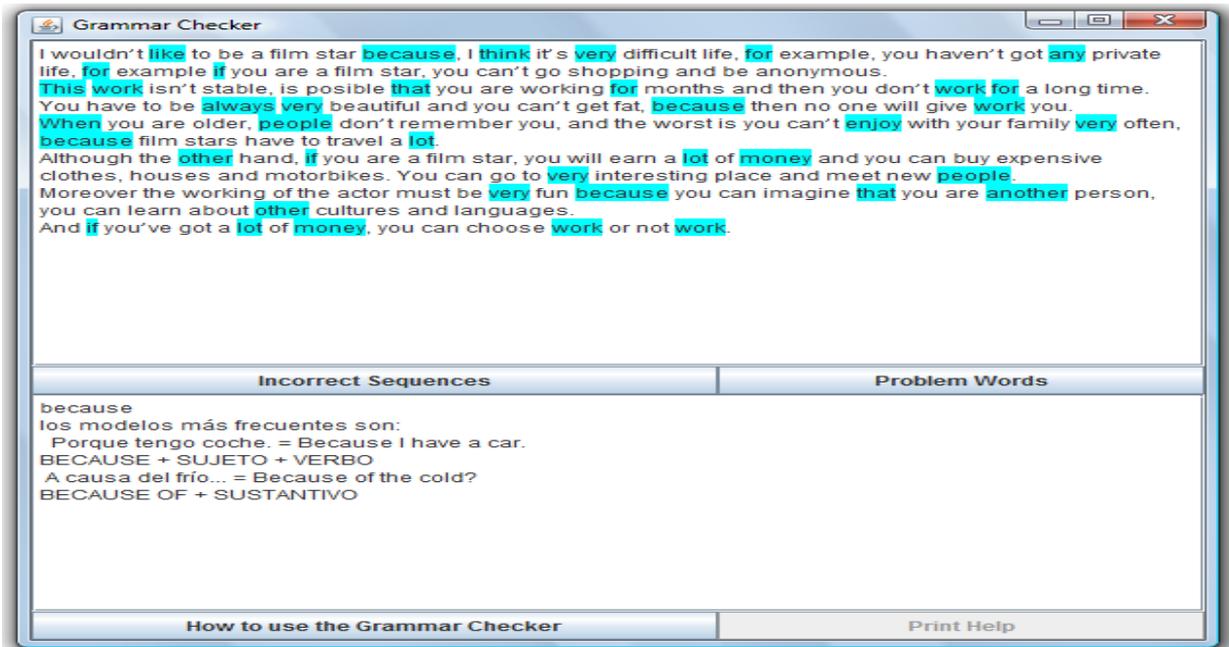


Figura 3: Corrector gramatical *e-gramm* 1.0. Uso de la herramienta *Problem Words*.

Cuando se presiona en el botón *Problem Words* aparecen subrayadas palabras asociadas con errores. Por ejemplo, *like*, *because*, *think*, *very*, etc. Si se hace clic en alguna de ellas, el programa indica la forma correcta de usar la palabra. Ejemplo.

(9) *because*.

Los modelos más frecuentes son:

porque tengo coche = because I have a car. BECAUSE + SUJETO + VERBO

a causa del frío... = because of the cold. BECAUSE OF + SUSTANTIVO

El programa no sabe si estas palabras se han usado de forma correcta o incorrecta. Se señalará cualquier uso de estas palabras, incluyendo usos correctos. Si el usuario está seguro de haberla usado correctamente no necesita detenerse a leer la información que ofrece el programa. Si tiene dudas debe leer los consejos.

Esta forma de corregir la gramática de una redacción es lo que distingue a *e-gramm* de otros correctores gramaticales convencionales. Como se ha explicado en el apartado anterior, este corrector promueve el aprendizaje autónomo y el uso de las tics en el aprendizaje de lenguas extranjeras.

Sin embargo, una de las características que más lo diferencian de otros correctores es su porcentaje de efectividad. Estudios como el que se realizaron para su evaluación y que explicamos a continuación, permiten que el porcentaje de corrección pueda llegar al 100% en el desarrollo de nuevas versiones.

3. Estudio empírico

El objetivo del estudio empírico fue el de evaluar el corrector gramatical *e-gramm* con un corpus distinto al que Lawley (2004) usó en la fase de desarrollo del programa.

En el estudio se siguieron los siguientes pasos.

1. Elaboración de un corpus de aprendices de 4.000 palabras con redacciones de alumnos de segundo de Bachillerato con un nivel A2 / B1 (clasificación en el MCER, 2001).
2. Corrección de errores por el investigador y por *e-gramm*.
3. Identificación de los errores que no corrige el corrector gramatical *e-gramm*.
4. Clasificación de los errores encontrados en el corpus.
5. Análisis de las categorías.
6. Nuevas Secuencias Incorrectas y Palabras Problemáticas para conseguir el 100% de corrección en el corpus recopilado por el investigador.

3.1. Errores

Se detectó un total de 393 errores de los cuales 63 errores eran ortográficos y 330 eran gramaticales y/o lexicosemánticos.

Entendemos por error cada una de las correcciones que el profesor destaca en el texto del alumno, ya sea de tipo ortográfico (error al deletrear la palabra), gramatical (omisión del sujeto, forma verbal incorrecta, etc.), léxico (palabra no apropiada para el contexto). Por ejemplo, en la frase *since is necessary continuously travel*, se encontraron 3 errores que destacamos en negrita.

- (10) Error nº 112. *since is **necessary** continuously travel*
- (11) Error nº 113. *since is necessary continuously **travel***
- (12) Error nº 114. *since is necessary **continuously** travel*

En el error número 112, el investigador destacó la omisión incorrecta del sujeto *it*. En el error 113 se destacó la falta de la partícula *to* en el infinitivo *travel*. En el error 114 se subrayó la posición incorrecta del adverbio de frecuencia *continuously*, que tendría que ir detrás del verbo *to be*. De manera que para que la oración *since is necessary continuously travel* fuese corregida de la siguiente manera, *since it is continuously necessary to travel*, tendríamos que corregir los tres errores anteriores con tres retroalimentaciones diferentes.

Cada uno de esos errores fue contabilizado individualmente y no como grupo. Por ejemplo, consideramos como error la palabra *spanish* por no estar escrita con mayúscula. Este error se repitió seis veces en todo el corpus, tres veces en la redacción nº 7 (errores nº 84, 85, 88) y otras tres en la redacción nº 8 (errores nº 92, 98, 108). Por ello, hay que aclarar que no se trata de 393 errores distintos, sino que muchos de ellos se repitieron a lo largo del corpus.

Se introdujo el corpus en el programa *e-gramm* para evaluar cuantos de los 330 errores no ortográficos podían ser corregidos usando este corrector gramatical. Esto se hizo usando las dos herramientas que ofrece el programa: Secuencias Incorrectas y Palabras Problemáticas. Los resultados fueron: con la herramienta de *Secuencias Incorrectas* se podrían corregir 28 errores y con la herramienta de *Palabras Problemáticas* se podrían corregir 56 errores.

El programa *e-gramm* podría corregir 84 errores de los 330. Hay que aclarar que la corrección de los errores depende del aprendiz y de su capacidad para entender las explicaciones que ofrece el corrector. Por consiguiente, estas correcciones se darían en una situación ideal con un aprendiz capaz de aplicar correctamente la retroalimentación recibida de *e-gramm*.

En porcentajes podemos decir que de los 330 errores no ortográficos, *e-gramm*, en su versión 1.0, podría corregir el 25% (84 errores). Sin embargo, quedarían sin corregir 246

errores que representan el 75% de los errores. A continuación en la figura 4 se muestra un gráfico que ilustra estos datos.

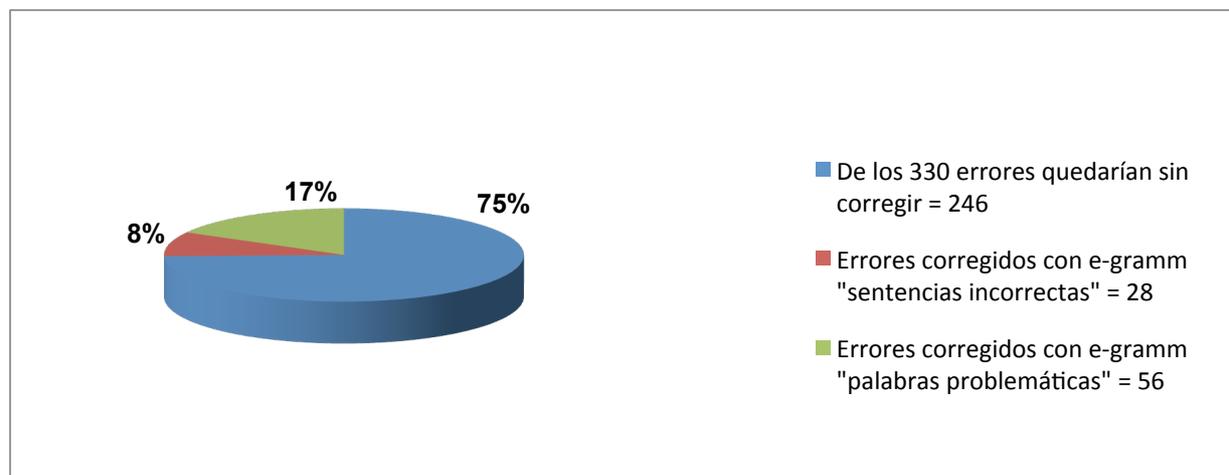


Figura 4: Porcentajes de corrección

Estos son algunos ejemplos de los 84 errores que *e-gramm* 1.0 podría corregir con las herramientas Secuencias Incorrectas y Palabras Problemáticas.

- (13) *she is living in Lorca since three years ago (she has lived in Lorca for three years)*
- (14) *she always is smiling (she is always smiling)*
- (15) *because is confident (because + sujeto + is confident)*
- (16) *The last year (last year)*
- (17) *she use to be a teacher (she used to be a teacher)*
- (18) *she likes to eat (she likes eating)*
- (19) *this things are good (these things are good)*
- (20) *all people know (everybody knows)*

3.2. Tratamiento de los errores

Quedan 246 errores (75%) en este corpus que *e-gramm* 1.0 no corrige. Cabe por tanto preguntarnos si la siguiente versión de este corrector gramatical, *e-gramm* 2.0, podría corregirlos.

La respuesta es sí. Se podría conseguir añadiendo nuevas Secuencias Incorrectas y Palabras Problemáticas al corrector.

En el caso de las Secuencias Incorrectas, se pueden añadir todas aquellas secuencias que encontramos en las redacciones de los alumnos y que no aparecen en un corpus de inglés correcto. De esta forma, cuando el corrector detecta esas secuencias, ofrece la retroalimentación necesaria para su corrección. Ejemplos.

- (21) *Always is.* Los adverbios de frecuencia suelen ir siempre detrás del verbo *to be*. Ej. *He is always tired.*
- (22) *She always try.* En la 3ª persona del singular el verbo añade *s / es / ies*. Cuando el verbo acaba en *y*, y a ésta le precede una consonante se añade *ies*. Ej. *She always tries to copy in the exams.*

- (23) *Likes go*. Los verbos que expresan gustos y preferencias como *like, love, hate*, etc van seguidos de gerundio si se generaliza o de infinitivo con *to* si se especifica. Ej. *I like doing exercise / I like to do exercise in the morning*.
- (24) *Films stars*. Los adjetivos en inglés no van en plural. En el caso *Film stars, film* funciona como un adjetivo. Otros ejemplos: *ten-year-old students; the car keys*.
- (25) *don't known*. Has escrito la forma del participio del verbo *know*. Recuerda, *know, knew, known*. Lo correcto es *don't know*.

Un tratamiento parecido es posible con 220 errores de los 246 que e-gramm no corrige ya que cuando una palabra se usa de forma incorrecta, el resultado suele ser una *secuencia incorrecta* que no aparece en un corpus de inglés correcto.

Esto nos deja con 26 errores que aparecen en secuencias que no serían incorrectas en un corpus de inglés correcto. Este tipo de errores es muy difícil de detectar por los correctores gramaticales convencionales ya que no hay errores de sintaxis sino de contenido. Ejemplos.

- (26) En la oración *Even though children whose parents are rich* el alumno debería haber escrito *Moreover, children whose parents are rich* para que la oración tuviese sentido.
- (27) En la oración *They had a fight about who was going to take the dog for a walk* el alumno debería haber escrito *They had an argument about who was going to take the dog for a walk*.
- (28) En la oración *he has got problems with her window* el alumno debería haber escrito *he has got problems with his window*.
- (29) En la oración *so that I agree with this statement* el alumno debería haber escrito *because I agree with this statement*

En estos casos, la herramienta de Palabras Problemáticas es muy útil, ya que como hemos visto en el apartado 2, el programa destacaría las palabras y daría información sobre los errores comunes que se suelen cometer. En estos casos, añadiríamos más Palabras Problemáticas a “e-gramm 2.0” para corregir los errores que encontramos en este corpus. Ejemplos.

- (30) *Even though*. No confundir con *moreover*. *Even though* significa "aunque". Ej. *They went into the sea, even though the weather was bad*. (Se metieron en el agua aunque hacía mal tiempo). *Moreover* significa "además". Ej. *The town is lovely and, moreover, their food is nice*.
- (31) *Fight*. No confundir con *argument*. *Fight* significa luchar o pelear. Ej. *They were expelled from the school because they had a fight*. (Los expulsaron del colegio por pelearse). *Argument* significa discutir. Ej. *They had an argument about who was going to take the dog for a walk*. (Discutieron sobre quién sacaría el perro a pasear)
- (32) *Her*. No confundir *her* y *his*. Como adjetivos posesivos, *her* significa su de ella y *his* su de él. *It's her house, not his* (es su casa de ella no de él).
- (33) *So that*. No confundir con *because*. *So that* se suele traducir por “de forma que, para que”. Expresa propósito. Ej. *I gave him 30€ so that he could buy a new pair of shoes*. (Le di 30 euros para que pudiera comprar unos zapatos nuevos). *Because* expresa causa, “porque”. Ej. *I gave him 30€ because it was his birthday*. (Le di 30€ porque era su cumpleaños).

De esta forma hemos mejorado la base de datos de *e-gramm* para que corrija el 100% de los errores de este pequeño corpus de 4.000 palabras. El procedimiento que hemos seguido podría repetirse con otro *corpus* de aprendices distinto que ayudaría a convertir *e-gramm* en un corrector mucho más potente ampliándolo incluso a otros niveles de enseñanza distintos de los mencionados en este artículo, A2/B1. Cada vez que se añade un nuevo error a la base de datos, *e-gramm* se convierte en una herramienta más eficaz.

4. Beneficios para los alumnos

La investigación tuvo un efecto positivo sobre los informantes. Estos se familiarizaron con los errores más comunes que solían cometer en las composiciones y que restaban puntuación en los exámenes de la PAU (Prueba de Acceso a la Universidad) por tratarse, la mayoría de ellos, de errores cometidos en el nivel A1 (clasificación en el MCER, 2001). Pudimos comprobar que estos errores más básicos que se cometieron en el primer trimestre del curso, comenzaron a desaparecer en las composiciones de finales del segundo trimestre y del tercer trimestre. Los resultados de la PAU para los 9 alumnos que se presentaron en junio fueron: 6, 5.5, 1.6, 4.7, 2.7, 8.9, 3.5, 2.6, 5.6. Es cierto que la prueba de la PAU no consistió sólo en una redacción, pero la destreza de la escritura en el examen de 2008 representó el 75% de la nota final. De las 6 preguntas que tenía el examen, en dos de ellas (la cinco y la seis) se pidió al alumno que redactase dos textos entre 25 y 50 palabras cada uno. En las preguntas tres y cuatro el alumno tuvo que completar frases.

También se pudo observar en el estudio los beneficios que el diario de errores tuvo para los alumnos. Se propuso a los estudiantes que crearan un diario de errores en el que éstos se clasificaran por categorías y donde se anotara la retroalimentación para corregirlos. Más abajo podemos ver unos ejemplos:

- (34) Categoría: omisión y/o doble sujeto. En inglés no se suele omitir el sujeto. Ej. *She speaks French because she lives in Lyon* (Habla francés porque vive en Lyon). En este caso el sujeto es *she* y no se omite como en español. Ciudadano con la repetición del sujeto en casos como *Living in England it must be very exciting* (vivir en Inglaterra debe ser muy emocionante). En este caso el error se encuentra en que la oración tiene dos sujetos *it* y *living in England*. Errores encontrados en mis redacciones: *because is very hard job; being a star it must be a, etc.*
- (35) Categoría: presencia y/o omisión incorrecta de *the*. Diferencia entre uso general y uso particular. Si se hace referencia a un grupo general no se usa *the*. Ej. *Students were given a chat about sex in the school*. (Les dieron una charla sobre sexo a los estudiantes). Si se hace referencia a un grupo particular se usa *the*. Ej. *The 16-year-old students were given a chat about sex*. (Les dieron una charla sobre sexo a los estudiantes de 16 años). Errores encontrados en mis redacciones: *the last year; the people don't known that*
- (36) Categoría: error en la preposición. Usamos *on* para los días de la semana. Dependier de se dice *depend on*. La diferencia entre *between* y *among*. Solemos usar *between* para indicar dos objetos, entre el pueblo y la ciudad (*between the town and the city*). Solemos usar *among* para grupos, entre los papeles (*among all these papers*).

Errores encontrados en mis redacciones: *in Sunday; depend of the cartoon; more famous between young people.*

Los alumnos consultarían el diario al hacer sus redacciones en casa para evitar cometer otra vez los mismos errores. De esta forma, al igual que en *e-gramm*, se fomentó el aprendizaje autónomo y la autoevaluación entre los alumnos y pudimos observar que el diario de errores ayudaba a que los estudiantes no cometieran los mismos errores de forma repetida. Promover el aprendizaje permanente entre los alumnos supone no sólo cumplir con las exigencias del currículo de la ESO y Bachillerato sino también cumplir con las exigencias de la reestructuración del sistema universitario europeo conocido como "proceso de Bolonia" (Chacón-Beltrán, 2009). Los cambios continuos que se producen en la sociedad requieren que sus ciudadanos sean capaces de adquirir nuevos conocimientos y nuevas destrezas. Una forma de hacerlo es preparando a los alumnos para que sean capaces de aprender de forma autónoma. Con herramientas como *e-gramm*, esto es posible. (Chacón-Beltrán 2009: 187)

También es de destacar que la retroalimentación que ofrece *e-gramm* podría ayudar a desarrollar la actividad de clase conocida como peer feedback (Lundstrom, K. y Baker W, 2009), en la que los alumnos, después de escribir sus redacciones, se dan retroalimentación de forma recíproca.

5. Conclusiones

En este trabajo de investigación se evaluó el corrector gramatical *e-gramm* 1.0, un prototipo que se encontraba en fase de desarrollo y que era susceptible de mejora. Para llevar a cabo esta evaluación se recopiló un corpus de aprendices de 4.000 palabras con las redacciones de alumnos de 2º de Bachillerato. El investigador detectó manualmente 330 errores (gramaticales y/o léxicosemánticos) y analizó cuántos de estos se podían corregir con la versión 1.0 de *e-gramm*.

En el análisis cuantitativo se detectó que *e-gramm* 1.0 corregía el 25% de los errores (84 errores) dejando por tanto el 75% de los errores (246 errores) sin corrección.

Ante esta situación es obvio preguntarse si la siguiente versión de *e-gramm*, versión 2.0, podría corregir este 75% de errores y cómo se podría hacer. La respuesta es que sí y además se puede conseguir que *e-gramm* detecte el 100 % de los errores de este pequeño corpus.

Para ello sólo hay que añadir nuevas Secuencias Incorrectas y Palabras Problemáticas al corrector. Como se ha explicado en el apartado 3.2 estas son las dos herramientas del programa cuyo carácter abierto y flexible permiten que se puedan implementar fácilmente.

El procedimiento que se ha seguido en este estudio ha permitido enriquecer la base de datos de *e-gramm* para que corrija más errores, llegando en nuestro caso a corregir el 100 % de los errores de este pequeño corpus de 4.000 palabras. Entre estos errores encontramos errores léxicosemánticos que son muy difíciles de detectar por un programa informático ya que éste no es capaz de entender el significado de las oraciones. Errores como *Even though, children whose parents are rich will receive loads of presents*, pasan desapercibidos para la gran mayoría de correctores gramaticales convencionales basados en la tecnología *parsing* y por lo tanto centrados en el análisis sintáctico. En la oración anterior el alumno quería decir *Moreover, children whose parents are rich will receive loads of presents*.

El desarrollo de programas como *e-gramm* supone un gran avance en el desarrollo de tecnologías informáticas que permiten desarrollar el aprendizaje autónomo en la escritura del inglés como lengua extranjera, ya que esta herramienta permite que el alumno decida cuándo, cómo y qué errores corregir.

Por esta razón, proponemos que si este procedimiento se repitiera con otro corpus de aprendices mucho mayor, se conseguiría otra versión de *e-gramm* mucho más potente, contribuyendo así a las exigencias de los currículos de la ESO y de Bachillerato y del Marco Común Europeo de Referencia para las Lenguas, que demandan tanto el uso de las TICs como el desarrollo del aprendizaje autónomo.

En el análisis cualitativo que tuvo lugar en el estudio pudimos comprobar que parte del los alumnos de bachillerato que participaron en la investigación siguen cometiendo errores muy básicos como omisión de la -s de 3ª persona del singular (*she always try to help*), error en el número (*with two Word*), presencia/omisión incorrecta de *the* (*The last year*); mal uso u omisión de *a* y *an* (*because is very hard job*). Es por esto que consideramos importante que el profesorado con alumnos de bachillerato tenga en cuenta que parte del alumnado seguirá cometiendo errores de un nivel inferior al del Bachillerato.

La escritura de redacciones puede ayudar a mejorar en este sentido y con la ayuda de programas como *e-gramm* se puede conseguir que la retroalimentación sea más individualizada y en cualquier momento del día sin tener que limitarse al profesorado. Si los alumnos son cada vez más críticos con sus redacciones podemos llegar a concluir que el uso de *e-gramm* hará que el alumno esté más preparado para actividades que requieran una actitud crítica como por ejemplo el peer feedback (retroalimentación entre compañeros).

Las conclusiones de nuestro trabajo pueden servir no sólo para mejorar las nuevas versiones de *e-gramm* sino también para futuras investigaciones en el campo del inglés como lengua extranjera. El corpus de 4.000 palabras que se creó, serviría para otras investigaciones en análisis de errores y para el diseño de materiales pedagógicos que funcionen con un corpus de errores frecuentes.

Referencias bibliográficas.

- Anderson, J.R. 1982. Acquisition of cognitive skill. *Psychological Review* 89/4: 369-406.
- Benson, P. 2001. *Teaching and researching autonomy in language learning*. London: Longman.
- Bolt, P. 1992. An Evaluation of grammar-checking programs as self-help learning aids for learners of English as a foreign language. *CALL* 5: 49-91
- Brophy, J.E. 1981. Teacher praise: a functional analysis. *Review of Educational Research* 51/1: 5-32.
- Chacón-Beltrán, R. 2007. Learner autonomy in Call: A grammar-checker for Spanish EFL students. En K. Rasulic e I. Trbojevic (eds.), *English Language and Literature Studies: Interfaces and Integrations*. 267-273. University of Belgrade: Serbia.
- Chacón-Beltrán, R. 2009. Learner Autonomy and Lifelong Learning: Technological Solutions in the European Higher Education Area. En M.L. Pérez Cañado (ed.), *English Language Teaching in the European Credit Transfer System*. 187-195. Berna, Berlin, Bruselas, Frankfurt am Main, Nueva York, Oxford, Viena: Peter Lang.
- Corder, S.P. 1981. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Decreto número 291/2007, de 14 de septiembre, por el que se establece el currículo de la Educación Secundaria Obligatoria en la Comunidad Autónoma de la Región de Murcia. (BORM 24.09.07)
- Decreto número 262/2008, de 5 de septiembre, por el que se establece el currículo del Bachillerato en la Comunidad Autónoma de la Región de Murcia. (BORM 10.09.08)
- Díaz Villa, A.M. 2005. Tipología de errores gramaticales para un corrector automático. *Procesamiento del Lenguaje Natural* 35: 409-416.
- Edge, J. 1989. *Mistakes and Correction*. London: Longman.

- Ferris, D. 2002. *Treatment of error in second language student writing*. Ann Arbor: University of Michigan Press.
- García-Heras, A. 2007. Programas informáticos de corrección gramatical en el aprendizaje de una lengua extranjera (inglés): expresión escrita. *Docencia e Investigación: revista de la Escuela Universitaria de Magisterio de Toledo* 17: 71-101.
- Granger, S. y F. Meunier. 1994. Towards a grammar checker for learners of English. In Fries U. y Tottie G. (eds), *Creating and using English language corpora*. 79-91. Amsterdam y Atlanta: Rodopi.
- Gupta, P. y M. Schulze. 2009. Human language technologies (HLT). Module 3.5. En G. Davies (ed.), *Information and communications technology for language teachers (ICT4LT)*, Slough, Thames Valley University. [Disponible en: http://www.ict4lt.org/en/en_mod3-5.htm]
- Hyland, K. y F. Hyland. 2006. *Feedback in second language writing: Contexts and issues*. New York: Cambridge University Press.
- Jacobs, G. y C. Rodgers. 1999. Treacherous Allies: Foreign Language Grammar Checkers. *CALICO Journal* 16/4: 509-531.
- Johnson, M. y K. Sterkel. 1984. Writing with the computer brings positive results. *Business Education Forum* 35: 32-36.
- Lawley, J. 1999. A new computer-mediated writing aid for Spanish students of English as a foreign language. En P. Faber, W. Gewehr, M. Jiménez y A. J. Peck (eds.), *English Language Teacher Education in Europe*. 31-43. Sonderdruck: Peter Lang.
- Lawley, J. 2004. A preliminary report on a new grammar checker to help students of English as a foreign language. *Arts and Humanities in higher Education* 3/3: 331-342.
- Lundstrom, K. y W. Baker. 2009. To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing* 18: 30-43
- Marco Común Europeo de Referencia para las lenguas: Aprendizaje, Enseñanza, Evaluación. 2002. Coeditan Secretaría General Técnica del MECD - Subdirección General de Información y Publicaciones, y Grupo ANAYA, S.A. Traducción: Common European Framework for Languages: Learning, Teaching, Assessment. Strasbourg.
- Pennington, M. 1992. Beyond off-the-shelf computer remedies for student writers: Alternatives to canned feedback. *System* 20/4: 423-437.
- Pennington, M. 1993. Computer-assisted writing on a principled basis: the case against computer-assisted text analysis for non-proficient writers. *Language and Education* 7/1: 43-59.
- Torre, S. de la. 1993. *Aprender de los errores. El tratamiento didáctico de los errores como estrategia de innovación*. Madrid: Escuela Española.
- Tschichold, C. 1999. Intelligent grammar checking for CALL. *ReCALL* 11: 5-11.
- Tschichold, C. 2003. Lexically Driven Error Detection and Correction. *CALICO Journal* 20/3: 549-559.
- Vygotsky, L.S. 1978. *Mind in Society*. Cambridge, MA: Harvard University Press.