

EVALUACIÓN AUTOMATIZADA DE NARRACIONES EN EL ÁREA DE LENGUA EXTRANJERA

AUTOMATED STORY SCORING IN FOREIGN LANGUAGE CLASS

Dr. José Luis García González

garciagjl@unican.es

Universidad de Cantabria. Facultad de Educación. Departamento de Educación

Avda. Los castros s/n. 39005 Santander.

El programa EssA fue creado en 2004 como un instrumento para ayudar a valorar narraciones en el área de lengua extranjera. El análisis de regresión múltiple, según variables léxico-gramaticales, determinó una ecuación de regresión que predice la calificación de los textos con un coeficiente de regresión del 0.79. EssA ha demostrado realizar la valoración gramatical de las narraciones con mayor objetividad y fiabilidad que los docentes. Por otra parte, proporciona feedback inmediato al estudiante sobre el uso que hace de los distintos recursos lingüísticos en el contexto de la narración, al mismo tiempo que le muestra la valoración del mismo.

Palabras clave: Tecnología educativa, evaluación automatizada, recuperación de la información, lengua extranjera, Educación primaria.

EssA software was created in 2004 as an instrument to help EFL teachers to assess narrative stories. Multiple regression analysis, using lexical-grammar variables, determined a regression equation to predict the score of each text with a regression coefficient of 0.79. The final result we found was that the EssA's analysis of the stories has proven to be more reliable and more objective compare to the scores given by the teachers. It also provides immediate feedback on those linguistics resources of the academic curriculum the student uses in their story, as well as the score.

Keywords: Technology Education, automated evaluation, information retrieval, foreign language, Primary Education.

1. Introducción.

Experimentar con métodos didácticos así como con diferentes instrumentos de evaluación fue lo que originó la creación del programa que seguidamente se va a analizar y que se denomina con el acrónimo EssA, *English short-story Analyzer* ('Analizador de historias cortas en inglés'). Está disponible, libre y gratuitamente, en la dirección de Internet: <http://essa.coconia.net>.

EssA surgió en el año 2004 como instrumento que ayudara a valorar con objetividad y fiabilidad el uso que el alumnado de L2 realizaba de los recursos lingüísticos en las composiciones escritas. Ese proyecto se vio impulsado por dos circunstancias que se observaron cuando el alumnado utilizaba el programa EssA: la motivación por el *feedback* inmediato que aquel les proporciona y por la actitud más positiva ante los resultados, independientemente de su valor, cuando era el ordenador el que les valoraba las composiciones.

La evaluación o el *feedback* es una fase clave en todo proceso de enseñanza y aprendizaje. Tradicionalmente la evaluación se realiza con controles escritos y pruebas objetivas (Fernández, 2003; Fontán, 2004). Existen otros medios que pueden complementar más adecuadamente la medición de la competencia comunicativa, como son las composiciones escritas. No obstante, cuando se trata de valorar esos relatos, el docente se enfrenta a la complejidad que conlleva evaluar ese tipo de pruebas, la subjetividad inherente a las mismas y la dedicación requerida (Savignon, 1997, p. 266; Valenti, Neri & Cucchiarelli, 2003, p. 319).

En cuanto a los tests, Labarca (1995, p. 170) afirma que no son una prueba adecuada para medir las capacidades lingüísticas de los estudiantes:

Ya que los test lingüísticos tal y como todavía existen hoy están tan alejados del mundo cultural, del uso y comprensión de la segunda lengua, y también porque las tareas propuestas en los test

lingüísticos son imprácticas o carentes de realismo, o malas imitaciones de tareas naturales habitualmente desarrolladas con otras [personas] –no aisladamente, muchos estudiantes adoptan actitudes, estrategias y atajos que tienen muy poco que ver con lo que el administrador del test asume o espera.

Todo esto requiere que se diseñen actividades basadas en el currículo y un sistema de evaluación implementado consecuentemente (Myers, 2003, pp. 3-20). Se podrían medir las cuestiones gramaticales con tests; pero la gramática no es un objetivo en sí mismo, sino su aprendizaje contextualizado. Las pruebas objetivas verifican un conocimiento aislado, que en muchos casos no es suficiente para demostrar que el estudiante es capaz de utilizarlo cuando tiene que enfrentarse a una situación comunicativa.

La presente propuesta se fundamenta en el enfoque educativo basado en competencias, de acuerdo con la Ley Orgánica 2/2006, de 3 de mayo, de Educación y su concreción en el Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las Enseñanzas Mínimas de la Educación Primaria. Profundizando más en las teorías, se hallan los enfoques constructivista y sociocultural, que fundamentan, a su vez, el currículo en Educación Primaria.

La última reforma educativa en Educación Primaria ha supuesto un nuevo enfoque que pretende dar respuesta a las demandas de la sociedad del siglo XXI. Una revisión de los principios y objetivos que marca la Ley Orgánica 2/2006, de 3 de mayo, de Educación para la Educación Primaria y cómo la han desarrollado los órdenes y decretos correspondientes en todas las comunidades autónomas de España, da una visión de la importancia que están adquiriendo las TIC.

La LOE establece en el artículo 17, (Objetivos de la Educación primaria) que enumera los objetivos de la Educación Primaria, uno en relación a la iniciación en el uso de las TIC: «Iniciarse en la utilización, para el aprendizaje, de las tecnologías de la información y la comunicación desarrollando un espíritu crítico ante los mensajes que reciben y elaboran».

El Real Decreto 1513/2006 desarrolla la LOE concretando los aspectos señalados anteriormente en la competencia digital y en el tratamiento de la información. Además de incluir esta competencia

por primera vez, otro cambio que resulta evidente es que ahora ya no se habla de conocimientos, sino de competencias, que se conciben como saberes prácticos que permiten desenvolverse en situaciones reales.

Los rasgos diferenciales de las competencias fundamentales serían los siguientes: constituye un «saber hacer» complejo y adaptativo, esto es, un saber que se aplica no de forma mecánica sino reflexiva, es susceptible de adecuarse a una diversidad de contextos y tiene un carácter integrador, abarcando conocimientos, procedimientos, emociones, valores y actitudes, que evolucionan a lo largo de la vida (Consejería de Educación de Cantabria, 2007, p. 16. Las comillas proceden del original).

El enfoque basado en competencias desplaza al enfoque teórico tradicional y la preponderancia de la memorización de contenidos hacia la habilidad para seleccionar la información, procesarla, aplicarla y generar nuevos conocimientos.

Las competencias básicas suponen los saberes, las habilidades y las actitudes básicas que todos los estudiantes tendrían que alcanzar, de acuerdo a las finalidades generales de la enseñanza obligatoria, para comprender y actuar en la sociedad actual. Además, definir los aprendizajes básicos en forma de competencias supone pasar de una identificación del aprendizaje como dominio de un contenido a la consideración de su aplicación para realizar tareas. Hoy, que se necesita menos memorizar porque la información está siempre a nuestro alcance, es necesario saber seleccionarla, procesarla y aplicarla (Marquès, 2000).

La adopción de una metodología de aula condiciona tanto la práctica, como su evaluación. Mientras que la investigación ha aportado nuevas metodologías para la enseñanza de las lenguas extranjeras cada vez más activas y comunicativas, la forma de evaluar se ha mantenido constante en el tiempo.

La disputa entre los lingüistas sobre una teoría que explique la adquisición de una segunda lengua (L2), no sólo no ha contribuido a esclarecer la cuestión sino que ha aumentado el vacío que el docente experimenta a la hora de tener que optar un método o cuando ha de decidir qué, cómo y cuándo evaluar.

1.1. La Lingüística computacional.

Con la aparición de los ordenadores, aflora la Lingüística Computacional como disciplina en la década de los cincuenta. Su denominación se origina con el término inglés *Computational Linguistics*, llevado al francés como *Linguistique Automatique* (Kock, 1974, p. 19). Ambos términos fueron adaptados, más tarde, al español como *Lingüística Computacional* y *Lingüística Automática o Automatizada*. Finalmente, se ha consolidado la denominación *Lingüística Computacional*.

La primera concepción que se tenía de la Lingüística Computacional era la de *una lingüística practicada con ayuda de los ordenadores* (Kock, 1974, p. 15). Actualmente, su objetivo es más ambicioso que el mero uso del ordenador para ayudar en los cálculos estadísticos: se persigue hacer que el ordenador sea capaz de comunicarse con los humanos. Para López (2004) «la Lingüística Computacional (Computational Linguistics) puede considerarse una disciplina de la lingüística aplicada y la Inteligencia Artificial, y tiene como objetivo la realización de aplicaciones informáticas que imiten la capacidad humana de hablar y entender».

1.2. Recuperación de la información.

Page creó en 1966 un programa (Dikli, 2006, p. 5) que llamó PEGTM, *Project Essay Grade*TM ('Proyecto para calificar ensayos'). Ese programa fue el primer paso que, posteriormente, en la década de los 80, daría lugar a la línea investigadora denominada en inglés *data mining* o *information retrieval*, 'recuperación de la información'. Manning, Raghavan y Schütze (2007) explican que el concepto recuperación de la información (RI) consiste en seleccionar documentos que sean relevantes desde el punto de vista del contenido: «encontrar material (normalmente documentos) de naturaleza desestructurada (habitualmente texto escrito) que satisface una

necesidad informativa entre muchos textos disponibles».

El concepto *information extraction* ('extracción de la información') tiene como objetivo extraer la información, que un conjunto de documentos contiene, mediante la *localización de datos determinados en el contenido de un conjunto de documentos y ofrecérselos al usuario de la manera más organizada posible* (Lavid, 2005, p. 366). Este campo de investigación ha derivado en diferentes líneas investigadoras según el objetivo perseguido: elaboración de resúmenes o de traducciones, clasificación de textos según la relevancia de la información o según el contenido, procesamiento del lenguaje natural (*natural language processing*), etc.

La clasificación de textos ha evolucionado en lo que se ha dado en llamar *Automated Essay Scoring* ('Evaluación automatizada de textos'). Shermis y Burstein (2003: XIII) lo definen como «la capacidad que el ordenador tiene para evaluar y puntuar ensayos escritos en prosa».

1.3. El currículo en Educación Primaria para el área de lengua extranjera.

Lectura-escritura y hablar-escuchar constituyen actos complementarios entre sí, aparte de la relación secuencial que existe entre ellos: escuchar, hablar, leer y escribir. El desarrollo gradual de esas cuatro habilidades responde a la necesidad de ser capaz de expresarse en otra lengua y de comprender los mensajes en la misma. También ha sido habitual tratar esas destrezas aisladamente, aunque los docentes de lenguas extranjeras cada vez las consideran más como una unidad (McCarthy, 1999).

Una persona que estudia una lengua extranjera debe conocer un vocabulario y unas reglas gramaticales para combinarlo. No obstante, López y Encabo (2001) matizan que los enunciados no tienen por qué ser gramaticalmente correctos para permitir la comunicación, si se adecuan al contexto y son comprensibles. Este es otro debate en el que no sólo los lingüistas no llegan a un acuerdo, sino que tampoco sucede entre los propios docentes. Se intenta que la persona desarrolle la habilidad para producir y entender enunciados que no son tanto gramaticales sino lo que es más

importante, apropiados al contexto en el que tienen lugar (López & Encabo, 2001, p. 58).

El Real Decreto 1513/2006, de ocho de diciembre, por el que se establecen las Enseñanzas Mínimas de Educación Primaria divide el currículo para una lengua extranjera en cuatro bloques de contenidos. EssA tiene su aplicación a la hora de valorar los bloques 2 y 3: 1-Escuchar, hablar y conversar. 2-Leer y escribir. 3-Conocimiento de la lengua. 4-Aspectos socioculturales.

1.4. Metodología de aula.

El primer método que se ha utilizado ha sido el que se denomina *clásico*. Se empleaba para el aprendizaje de las lenguas clásicas como el latín y griego. Consistía en *el aprendizaje memorístico del vocabulario, la morfología y las reglas gramaticales; el uso de la traducción directa e inversa y la preeminencia de la escritura sobre el habla* (Martín, 1997, p. 228).

La metodología está relacionada directamente con la teoría que fundamenta el quehacer en el aula, en nuestro caso, el enfoque en competencias. Esa concepción implica un cambio sustancial en la manera de enseñar, de aprender y en la enseñanza misma. El enfoque de competencias básicas [...] entendido desde una interpretación que se asienta en enfoques constructivistas y socioculturales, trae consigo una serie de cambios en los centros educativos. Dichos cambios afectan muy particularmente al proceso de enseñanza-aprendizaje y, en concreto, a la metodología. En este sentido, es necesaria una redefinición de las relaciones entre los aspectos que componen los procesos relacionados con el cómo enseñar: el currículo, el profesorado y el alumno (Consejería de Educación de Cantabria, 2007, p. 21).

La existencia de múltiples teorías y la de muchos más métodos que surgen a su amparo, complica la labor del docente a la hora de optar por una estrategia en el aula. La variedad es buena en sí, pero puede llegar a saturar cuando es excesiva y, habitualmente, excluyente entre unas alternativas y otras. Para Vez (2000, p. 321) «sabemos que la lingüística teórica es hoy un campo surcado por profundas divisiones, en el que los cultivadores de las tres perspectivas al uso –la histórica, la

generativa y la sociopragmática– se combaten con más frecuencia que colaboran. Podemos situar aquí una de las razones [...] para el actual «eclecticismo militante» que está empezando a imperar...».

Es necesaria una reflexión sobre la metodología que se trabaja en el aula, pues la manera de enseñar y de evaluar debe corresponderse. En realidad, una determina la otra. Prácticamente todos los métodos de inglés para Educación Primaria incluyen actividades comunicativas como las sugeridas. Por ello, la utilización del programa EssA es fácilmente adaptable en el aula. Únicamente, hemos de tener en cuenta la organización del aula y del Centro en cuanto a los espacios, tiempos y recursos para implementar la propuesta (Martín-Moreno, 2007, pp. 295-308).

Abordar el aprendizaje de una segunda lengua no se puede hacer desde una única dimensión. Como señala Sánchez (2009: 14), *el problema de cada método no se cifra en si es o no bueno o malo, sino más bien en qué medida toma en consideración esa complejidad*. Se trataría de buscar un equilibrio entre teoría y práctica, ya que ambos aspectos son necesarios. La reflexión que aquí hacemos sobre la metodología, simplemente busca esa complementariedad. Vez (2000, p. 194) indica que como cualquier otro campo disciplinar, la enseñanza de lenguas extranjeras no es ajena al desarrollo paradigmático de la racionalidad técnica y la racionalidad práctica. El devenir tanto de sus conocimientos teóricos como prácticos guarda relación con ambos tipos de racionalidad, si bien el peso de una y de otra no ha alcanzado, ni parece estar en condiciones de lograr, el equilibrio deseable.

El enfoque aquí propuesto persigue, principalmente, el logro comunicativo. La corrección lingüística queda relegada a un segundo plano, aunque no por ello se descuida, sino que se potencia a través del *input* continuo. Beltrán también es concluyente apoyando la línea de trabajo indicada: «Una metodología de enseñanza, basada en el enfoque comunicativo y centrado en la utilización de estrategias cognitivas y lingüísticas, incrementa el nivel de calidad en la producción de textos y por consiguiente en la comunicación del alumnado». Cualquier método comunicativo trabaja las cuatro destrezas secuencialmente (escuchar, hablar, leer y escribir), dejando para el final la expresión

escrita. EssA tiene su aplicación a la hora de valorar la última fase mediante la valoración e las composiciones escritas. Evidentemente, las restantes destrezas deber ser también valoradas, en este caso por el docente, de acuerdo con el enfoque comunicativo que se persigue y la mejora del mismo (González, 2006, pp. 199-203).

1.5. La evaluación en Educación Primaria.

La información que proporcionaban los programas era utilizada solamente como evaluación sumativa. Actualmente, su rol se considera más como evaluación formativa (Philips, 2007, p. 51; Shermis & Burstein, 2003: XV).

La evaluación, que debe ser concebida como un proceso continuo e integrado en el proceso de enseñanza-aprendizaje, ha de proporcionar información para reorientar dicho proceso, ya sea manteniendo aquellos aspectos que nos permiten conseguir buenos resultados, ya sea modificando aquellos otros que interfieran en un adecuado progreso del alumnado (Consejería de Educación de Cantabria, 2007, p. 24).

Ridgway, Mccusker y Pead (2004, p. 5) afirman que enseñar, aprender y evaluar forman parte inseparable de un mismo proceso. La evaluación debe plantearse como un medio que proporciona *feedback* sobre el proceso de enseñanza y aprendizaje (evaluación formativa) y sobre el grado de consecución de los objetivos propuestos (evaluación sumativa).

La evaluación debe considerarse en Primaria, fundamentalmente, como *feedback* en relación al progreso del alumno (Fontán, 2004, p. 82). El ideal será, por tanto, realizar de manera continuada esa valoración de los logros parciales que apuntan al producto final, con el fin de descubrir las dificultades que surgen en el proceso, atenuar o superar los obstáculos, adecuar la acción a los medios disponibles y viceversa, etc. (García, 2008, p. 1).

EssA constituye un recurso que contribuye a ese objetivo, si se emplea con regularidad en el análisis de los relatos, pues ayuda a que los estudiantes reflexionen sobre la información que en

cada apartado les proporciona el programa. Ese enfoque queda recogido en el Real Decreto que establece las Enseñanzas Mínimas de la Educación Primaria, cuando se refiere a que el alumnado debe desarrollar progresivamente la capacidad de *valorar sus propios progresos* e, incluso, *reflexionar sobre sus procesos de aprendizaje* (Consejería de Educación de Cantabria, 2007, p. 13), lo que además se relaciona directamente con la competencia *aprender a aprender*. También se evaluará la capacidad de ir valorando sus progresos, la identificación de recursos y estrategias que le ayudan a aprender mejor y el uso espontáneo de formas y estructuras sencillas y cotidianas (R.D. 1513/2006, de ocho de diciembre, p. 43095).

Los criterios de evaluación son un referente más en el proceso de evaluación, que *no deben ser traducidos automáticamente en exámenes o pruebas que evalúen básicamente la adquisición, en muchos casos memorística, de determinados contenidos* (Consejería de Educación de Cantabria, 2007, p. 19).

EssA se concibe como una herramienta que se integra en el conjunto de estrategias que se utilizan en clase. Se limita al desarrollo de la habilidad de escribir, pero en conjunción con las demás destrezas, que se han debido trabajar previamente, contribuye a la consolidación de los contenidos, y por tanto la consecución de los criterios de evaluación de esta etapa.

El Real Decreto que establece las Enseñanzas Mínimas redacta los criterios de evaluación utilizando una serie de expresiones que permiten dar una idea clara de cuáles son las expectativas y, por lo tanto, el nivel de competencia. Obsérvense, especialmente, expresiones como: *trabajado previamente, mensajes sencillos, con ayuda, siempre a partir de modelos, valorar sus progresos, estructuras sencillas...* (R.D. 1513/2006, de ocho de diciembre, p. 43095).

Las narraciones permiten contextualizar el aprendizaje de una segunda lengua y el *input* continuo, la adquisición de la misma (Ellis, 2005, pp. 7-8). Los textos, como una manifestación escrita de las historias previamente trabajadas de forma oral, desarrollan la habilidad de escribir y, su complementaria, leer. De esta manera, se están abarcando siete criterios, sobre un total de ocho.

1.6. Software.

Los siguientes programas son una muestra reducida del software que ha sido aplicado a la evaluación de textos en el campo educativo: Project Essay Grade™ (PEG), Intelligent Essay Assessor™ (IEA), Electronic Essay Rater® (E-RATER), Concept Rater™ (C-RATER), AUTOMARK, CRITERION™, MY ACCESS!™ e INTELLIMETRIC™.

La evaluación de textos es un proceso comparativo entre unas muestras de referencia, evaluadas por docentes, que «entrenan» al programa y los textos que evalúan. La exactitud del resultado se determina previamente mediante el grado de correlación que presentan las valoraciones dadas por el ordenador con las que proporcionan unos docentes evaluando los mismos relatos.

Las diferencias principales entre EssA y los programas señalados, radica en que el objetivo de aquéllos es mejorar la expresión escrita de estudiantes nativos. La valoración la realizan holísticamente, según una escala cualitativa. Por el contrario, EssA se dirige al alumnado español que aprende inglés como segunda lengua. Evalúa el uso de los recursos lingüísticos contemplados en el currículo de lengua extranjera que realizan los estudiantes en el contexto de una narración. El análisis es analítico y la valoración cuantitativa en una escala en porcentaje.

2. Método.

2.1. Objetivo.

Analizar el programa EssA como herramienta que ayude al docente en la valoración de los textos escritos por alumnado de sexto de EP.

2.2. Instrumentos.

Los instrumentos han sido los textos escritos por los estudiantes, el propio programa EssA y una plantilla de evaluación que, previamente, había sido validada por los propios docentes, como especialistas en su materia.

2.3. Muestra.

Las muestras se obtienen a lo largo de los cursos 2006-2009 de alumnado que está cursando sexto de Educación Primaria en dos centros públicos de Cantabria. Las narraciones fueron realizadas al finalizar cada curso escolar, en el mes de junio. Se les pidió que narraran una historia o un cuento de los que habían trabajado previamente en clase.

2.4. Procedimiento para la obtención de los datos.

Los estudiantes tenían 35 minutos para componer su cuento y 15 minutos más para revisar la ortografía entre iguales e individualmente. Las narraciones las redactaban directamente en la pantalla del ordenador.

Los relatos fueron valorados por EssA y, posteriormente, por cuatro docentes especialistas en lengua inglesa, que impartían esa materia en Educación Primaria. Cada docente evaluó los textos de forma individual sin tener conocimiento de la calificación de los demás.

La evaluación de los relatos se hizo según cuatro métodos de valoración. Los siguientes:

- EssA: valoraciones proporcionadas por el programa EssA.
- Gramática: las puntuaciones se obtenían mediante el uso de una plantilla que medía sólo los aspectos gramaticales.
- Expresión: calificación promedio entre la nota procedente de la plantilla gramática y un ítem extra sobre la expresión.
- Docentes: calificación de los docentes según su criterio personal.

3. Resultados.

3.1. Análisis de las muestras.

El alfa de Cronbach para las cuatro muestras es 0,937. La prueba ANOVA encontró que existían diferencias significativas en la calificación que los cuatro docentes había dado para cada texto.

Las valoraciones extremas que se han producido en el conjunto de la muestra cuando se utiliza la plantilla gramática, ascienden al 8,4% de las calificaciones, mientras que cuando los docentes utilizan su propio criterio, el porcentaje es del 10,09%.

Analizando las puntuaciones extremas (tabla 1) cuando se comparan los resultados procedentes de los métodos *gramática* y *expresión* con la calificación de EssA, se observa que *se* han producido en el 15,1% y 14,3% de las valoraciones, respectivamente. Si no se usa ninguna plantilla, los valores extremos ocurren el 19,3% de los casos.

Las diferencias entre las puntuaciones de los tres métodos utilizados pueden implicar que un estudiante pase de aprobar a suspender. El número de ocasiones en las que tal diferencia implica aprobar, asciende cuando valora EssA a 30,3%, 58,8% y 60,9%, respectivamente. Por el contrario, cuando se usan las plantillas, el 22,9% y el 11,8% están por encima del aprobado. Valorar sin usar plantillas, reduce la diferencia a favor del estudiante hasta un 4,3%.

Diferencia	EssA	<i>Gramática</i>	EssA	<i>Expresión</i>	EssA	<i>Docentes</i>
Absoluta	15,1%		14,3%		19,3%	
Por encima del aprobado	33,9%	22,2%	58,8%	11,8%	60,9%	4,3%

Tabla 1. Porcentaje de puntuaciones extremas

Se puede decir que, en el caso de que las notas entre EssA y los otros métodos discrepen, esa diferencia favorece al alumnado en un porcentaje más elevado cuando valora EssA: el 52,7% superior con respecto al método *gramática*, el 398,31% con respecto al método *expresión* y el 1316,28%, si no se usa ninguna plantilla. Esos resultados ofrecen una primera evidencia de lo que la subjetividad puede cometer en las valoraciones, atribuible a la complejidad del análisis y la

dificultad en el reconocimiento de las unidades.

Con respecto al grado de acuerdo interdocentes, cuando evalúan sin plantilla, la correlación Pearson media ha resultado ser de 0,65. Mientras que cuando se utiliza la plantilla *gramática*, el grado de acuerdo interdocentes es de 0,83.

3.2. Análisis factorial.

El análisis factorial determina dos dimensiones para explicar la interrelación entre las variables: una gramática y, la otra, expresión. El componente gramática explica el 38,36% de la varianza; y el componente expresión, el 22,38%. La suma de ambos supone el 60,75% de la varianza total de la variable dependiente.

3.3. Análisis de regresión múltiple.

Se han efectuado tres análisis de regresión múltiple mediante el método *stepwise* para determinar la influencia de cada variable, para cada procedimiento de evaluación empleado (tabla 2).

El método de valoración con la plantilla *gramática* es el que ha obtenido un índice de regresión más elevado ($r = 0,79$). Asimismo, es el que más contribuye a explicar la varianza de la variable dependiente ($R^2 = 62,6\%$). Los restantes R^2 corregidos disminuyen su valor, a pesar de que las variables independientes son las mismas y el tamaño muestral no varía. Por lo tanto, se puede afirmar que la utilización de una plantilla para valorar los contenidos gramaticales, ha repercutido considerablemente en los resultados. A medida que se elimina y se incorpora el factor *expresión* los coeficientes reflejan la subjetividad de las valoraciones.

Método evaluación	R	R ² corregida	Error típico de estimación
Gramática	0,79	0,626	1,180
Expresión	0,75	0,553	1,094
Docentes	0,71	0,492	1,130

Tabla 2. Resumen de los análisis de regresión

4. Discusión.

El método denominado *gramática* es el que ha ofrecido los resultados más elevados ($R = 0,799$), respondiendo al objetivo planteado en cuanto a un análisis objetivo y consistente de los textos; ya que en este momento, es preferible dejar en manos de los ordenadores la valoración gramatical, y en las del docente la valoración del contenido. La capacidad actual de los ordenadores para «comprender» un texto, invita a ser cautos en cuanto a incluir ese aspecto en las valoraciones.

La evaluación de los relatos ha supuesto un esfuerzo considerable para los docentes colaboradores. A pesar de su dedicación, la laboriosidad para identificar y cuantificar las unidades de análisis estuvo presente. Esa circunstancia es la que explica la diferencia entre los resultados y la que contribuye, en parte, a la subjetividad que caracteriza este tipo de pruebas. Subjetividad que aumenta todavía más cuando consideramos la expresión.

Hemos visto que la valoración realizada por el ordenador es más fiable y objetiva que la que puede alcanzar los docentes. Sólo el número de criterios y unidades de análisis que el programa considera a la hora de valorar un texto, demuestran su superioridad, ya que los docentes se verían desbordados si tuvieran que proceder con tal precisión.

La utilización de EssA en el área de inglés aportaría las siguientes ventajas: objetividad, consistencia, economía y rapidez en el feedback. El ordenador analiza objetivamente las unidades que tiene programadas y las valora según los criterios que tiene establecidos. Es consistente, porque la valoración no depende del estado de ánimo ni de otros factores que afectan al docente. En ese sentido, un texto siempre recibirá la misma valoración independientemente de cuándo, dónde y quién lo analice. Rapidez en el *feedback*, por la inmediatez del análisis con la consiguiente motivación del estudiante. Economía, por el ahorro de tiempo que supone la evaluación automatizada.

EssA es un instrumento para evaluar fácilmente integrable en la metodología de la clase de lengua

extranjera en sexto de Educación Primaria, extensible al primer ciclo de ESO, por la cercanía de los contenidos en esos niveles.

La evaluación automatizada complementa la evaluación de los docentes, proporcionando una valoración más consistente. Rudner y Gagne (2001) sugieren un equilibrio entre ambas, de manera que una compense las posibles debilidades de la otra: «Nosotros imaginamos la utilización de cualquiera de estas tecnologías como una herramienta validadora de la puntuación de cada texto por un juez humano y un ordenador. Cuando las puntuaciones difieren, el texto se marcaría para una segunda lectura».

La tecnología EAT ha iniciado su camino en el terreno de la evaluación de textos y está demostrado ser efectiva y consistente en los resultados; no obstante, su futuro, al menos a medio plazo, es ayudar al docente más que suplantarle (Dikli, 2006, p. 27; Burstein, Chodorow & Leacock, 2003, p. 1; Shermis & Burstein, 2003: XV).

5. Referencias bibliográficas.

Consejería de Educación de Cantabria (2007). *Las competencias básicas y el currículo: orientaciones generales*. Grupo de trabajo de Competencias Básicas. Cuadernos de educación de Cantabria, número 2. Edita Consejería de Educación de Cantabria.

Beltrán Nuñez, F. *Desarrollo de la Competencia Comunicativa*. México: Universidad Abierta.

Recuperado de http://www.universidadabierta.edu.mx/Biblio/B/BeltranFiliberto_CompeticiaComunicativa.htm.

Burstein, J., Chodorow, M. & Leacock, C. (2003). CriterionSM Online Essay Evaluation: An Application for Automated Evaluation of Student Essays. Proceedings of the *Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence*, Acapulco, Mexico. Recuperado de <http://www.ets.org>.

Dikli, S. (2006). An Overview of Automated Scoring of Essays. *The Journal of Technology*,

Learning, and Assessment, 5 (1), 1-35. Recuperado de <http://escholarship.bc.edu/cgi/viewcontent.cgi?article=1044&context=jtla>.

Ellis, R. (2005). Principles of Instructed Language Learning. *Asian EFL journal*, 7,3. Recuperado de http://www.asian-efl-journal.com/sept_05_re.pdf.

«Ley Orgánica 2/2006, de 3 de mayo, de Educación». *Boletín Oficial del Estado* (4 de mayo de 2006), págs. 17158-17207. Recuperado de <http://www.mepsyd.es/mecd/gabipren/documentos/A17158-17207.pdf>.

Fernández, C. & Cebreiro, B. (2003): Evaluación de la enseñanza con TIC. *Píxel-Bit. Revista de Medios y Educación*, 21, 65-72.

Fontán, M. T. (2004): Evaluar a través de Internet. *Píxel-Bit. Revista de Medios y Educación*, 24, 79-88.

García, A. (2008). Evaluación en formatos no presenciales. *Revista Iberoamericana de Educación*, 45 (1), 1-8. Recuperado de <http://www.uned.es/catedraunesco-ead/editorial/p7-1-2008.pdf>.

González, I. (2006). La evaluación y la mejora de la calidad de los procesos de comunicación. *Comunicar*, 27, 199-203.

Instituto Cervantes (2010). Centro Virtual Cervantes. *Diccionario de términos clave de ELE*. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/.

Kock, J. (1974). *Introducción a la lingüística automática en las lenguas románicas*. Madrid: Gredos.

Labarca, A. (1995). Interactive testing: a dormant giant? En Fernández-Barrientos Martín, J. & Wallhed, C. (1995). *Temas de Lingüística Aplicada*. Granada: Universidad de Granada.

Lavid, J. (2005). *Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI*. Madrid: Ediciones Cátedra.

López, L. (2004). *¿Qué es la Lingüística Computacional o PLN?* Recuperado de <http://www.aucel.com/pln/k-es.html>.

López, A. & Encabo, E. (2001). *El desarrollo de habilidades lingüísticas. Una perspectiva crítica*.

Grupo Editorial Universitario.

Manning, C., Raghavan, P. & Schütze, H. (2007). *Introduction to Information Retrieval*. Cambridge University Press. Recuperado de <http://www-csli.stanford.edu/~schuetze/information-retrieval-book.html>.

Marquès, P. (2000). *Nueva cultura, nuevas competencias para los ciudadanos. La alfabetización digital. Roles de los estudiantes hoy*. Recuperado de <http://www.peremarques.net/competen.html>.

Martín, E. (1997). *Bases para el estudio del lenguaje*. Barcelona: Octaedro Universidad.

Martín-Moreno, Q. (2007). *Organización y dirección de centros educativos innovadores. El centro educativo versátil* (pp. 295-308). Madrid: Mc Graw Hill.

Mccarthy, C. (1999). *Reading Theory as a Microcosm of the Four Skills*. Applied Linguistics Series. Recuperado de <http://iteslj.org/Articles/McCarthy-Reading.html>.

Myers, M. (2003). What Can Computers and AES Contribute to a K-12 Writing Program?, en Shermis, M. D. & Burstein, J. C. (2003). *Automated Essay Scoring. A Cross-Disciplinary Perspective* (pp. 3-20). New Jersey: Lawrence Erlbaum Associates, Publishers.

Pérez, A. (2007). *La naturaleza de las competencias básicas y sus aplicaciones pedagógicas*. Cuadernos de educación de Cantabria, número 1. Consejería de Educación de Cantabria.

Philips, S. M. (2007). *Automated Essay Scoring: A Literature Review*. Society for the Advancement of Excellence in Education (SAEE); 9-70. Recuperado de <http://www.sae.ca/pdfs/036.pdf>.

Ray, B. & Seely, C. (1998). *Fluency Through TPR Storytelling. Achieve Real Language Acquisition in School*. Command Performance Language Institute.

«Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las Enseñanzas Mínimas de la Educación Primaria». *Boletín Oficial del Estado* (8 diciembre de 2006), págs. 43053-43102. Recuperado de <http://www.mepsyd.es/educa/sistema-educativo/loe/files/educacion-primaria.pdf>.

Richards, J. (1976). The Role of Vocabulary Teaching. *TESOL Quarterly*, 10(1), 77-89.

Ridgway, J., Mccusker, S. & Pead, D. (2004). Literature Review of E-Assessment. *Futurelab Series*, 10, 1-47. Recuperado de http://www.futurelab.org.uk/resources/documents/lit_reviews/

Assessment_Review.pdf.

Rudner, L. & Gagne, P. (2001). An Overview of Three Approaches to Scoring Written Essays by Computer. *Practical Assessment, Research & Evaluation. A peer-reviewed electronic journal*, 7 (26). Recuperado de <http://pareonline.net/getvn.asp?v=7&n=26>.

Sánchez, A. (2009). *La enseñanza de idiomas en los últimos cien años. Métodos y enfoques*. Madrid: Sgel.

Savignon, S. (1991). Communicative Language Teaching: State of the Art. *TESOL Quarterly*, 25 (2), 261-277.

Savignon, S. (1997). *Communicative Competence: Theory and Classroom Practice*. Second edition. New York: McGraw-Hill.

Shermis, M. D. & Burstein, J. C. (2003). *Automated Essay Scoring. A Cross-Disciplinary Perspective*. New Jersey: Lawrence Erlbaum Associates, Publishers.

Valenti, S.; Neri, F. & Cucchiarelli, A.(2003). An Overview of Current Research on Automated Essay Grading. *Journal of Information Technology Education*, 2, 319-329. Recuperado de <http://jite.org/documents/Vol2/v2p319-330-30.pdf>.

Veiz, J. M. (2000). *Fundamentos lingüísticos en la enseñanza de lenguas extranjeras*. Barcelona: Ariel lenguas modernas.

Fecha de recepción: 2010-01-14

Fecha de evaluación: 2010-02-07

Fecha de aceptación: 2010-09-06