

La evaluación de la escritura en lengua extranjera (inglés) en el nivel superior: análisis de reportes docentes

Martínez, Julia Inés; Salinas, Julieta; Canavosio, Andrea de los Ángeles
Facultad de Lenguas, Universidad Nacional de Córdoba
Córdoba, Argentina
juliamartinez@gmail.com, juliasalinasr@gmail.com, acanavosio@gmail.com

Abstract

Several specialists agree that the assessment of foreign language writing is a complex task, which tends to generate anxiety and apprehension among teachers; as such, it is still the focus of permanent research. In this context, and due to the need to explore the assessment criteria used by the teachers of *Lengua Inglesa II*, the research project “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior” (“Criteria for Assessing Writing in a Foreign Language (English)”) was designed for 2014-2015, the main objective of which is to study the practices, attitudes and beliefs of teachers of the abovementioned course regarding the assessment of essay writing in cases of summative assessment. The purpose of this paper is to present a preliminary analysis of the reports written by the teachers, in which they explain and justify the assessment and grading of certain academic essays. For that reason, the qualitative data gathered from the reports will be analysed following the steps proposed by Hatch (2002), which imply the identification of frames, domains and topics from an inductive reading of the data obtained.

Keywords: foreign language writing, assessment, criteria

Resumen

Muchos especialistas coinciden en que la evaluación de la escritura en lengua extranjera es un campo de acción complejo, que suele generar ansiedad y aprehensión entre docentes, por lo que continúa siendo un campo de permanente investigación. En este contexto, y dada la necesidad de estudiar los criterios de evaluación utilizados por los docentes de la cátedra *Lengua Inglesa II*, se presentó el proyecto de investigación “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior” para los años 2014-2015, cuyo objetivo principal es explorar las prácticas, actitudes y creencias de los docentes de la asignatura en lo que respecta a la evaluación de la escritura de ensayos en instancias de evaluación sumativa. El propósito de este trabajo es presentar un análisis preliminar de los reportes escritos por los docentes de la cátedra, en los que se explica y justifica la evaluación y calificación de determinados ensayos académicos. Para tal fin, se analizarán los datos cualitativos reunidos a partir de los reportes siguiendo los pasos propuestos por Hatch (2002), los cuales implican la identificación de marcos, dominios y temas a partir de un recorrido de lectura inductivo de los datos obtenidos.

Palabras clave: escritura en lengua extranjera, evaluación, criterios

1. INTRODUCCIÓN

El impacto de la evaluación y las maneras de evaluar en las prácticas áulicas es una cuestión que merece ser analizada y estudiada en profundidad ya que, en muchos casos, son las características de los procesos de evaluación las que determinan las prácticas pedagógicas concretas. En lo que a la evaluación de la escritura en lengua extranjera se refiere, muchos especialistas coinciden en que se trata de un campo de acción complejo, que suele generar ansiedad y aprehensión entre docentes debido a los efectos e implicancias tanto a nivel pedagógico como ético y político; en consecuencia, la evaluación de la escritura en lengua extranjera continúa siendo un campo de permanente investigación. En este contexto, y dada la necesidad de estudiar los criterios de evaluación utilizados por los docentes de la cátedra *Lengua Inglesa II*, se presentó el proyecto de investigación “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior” para los años 2014-2015, cuyo objetivo principal es explorar las prácticas, actitudes y creencias de los docentes de *Lengua Inglesa II* en lo que respecta a la evaluación de la escritura de ensayos en instancias de evaluación sumativa. En el marco de dicho proyecto, el propósito de este trabajo es presentar un análisis preliminar de los reportes escritos por los docentes de la cátedra, en los que se explica y justifica la evaluación y calificación de determinados ensayos académicos. Para tal fin, se analizarán los datos cualitativos reunidos a partir de los reportes siguiendo los pasos propuestos por Hatch (2002), los cuales implican la identificación de marcos, dominios y temas a partir de un recorrido de lectura inductivo de los datos obtenidos.

2. MARCO TEÓRICO

Los reportes docentes utilizados en este trabajo de investigación se elaboraron luego de la primera instancia de evaluación parcial de *Lengua Inglesa II*. Siguiendo a Goodman & Swann (2003) y Hyland (2003), este tipo de evaluación de la escritura puede considerarse como una *evaluación sumativa* porque se califica formalmente y generalmente se lleva a cabo al final del proceso de enseñanza (en este caso, al finalizar un cuatrimestre) con el fin de determinar si se han logrado los objetivos planteados para el curso o etapa del curso. Por otra parte, este tipo de evaluación puede identificarse como una *evaluación directa* de la escritura, debido a que la escritura se evalúa a través una producción textual individual y no a través de ejercicios de componentes lingüísticos que hacen a la escritura (Williams, 2005, p. 120). Asimismo, la *evaluación directa* de muestras concretas de escritura suele considerarse como una de las maneras más efectivas de lograr una apreciación de la habilidad escrituraria de un estudiante (Hamp-Lyons, 2001; Williams, 2005), especialmente por tratarse de una instancia real de escritura en la que existe una interacción entre escritores y evaluadores; en otras palabras, se trata de un tipo de evaluación en el que el componente humano es central (Hamp-Lyons, 2001, p. 120-121). Finalmente, vale la pena mencionar la distinción que se realiza habitualmente entre *enfoque holístico* y *enfoque analítico* en lo que a abordajes de evaluación de un texto se refiere. La *evaluación holística* aborda al texto como un todo e implica la asignación de una nota o evaluación final general. Esta calificación se basa en la premisa de que el constructo de la escritura es una entidad única que puede capturarse mediante una escala única que integra las cualidades inherentes de la escritura y que la calidad sólo puede ser reconocida por lectores expertos

(y no mediante instrumentos objetivos). Entre los problemas típicos de este tipo de calificación se destacan la aplicación de diferentes criterios por parte de los evaluadores y las diferencias en el contexto en el que se llevan a cabo las evaluaciones, que suelen derivar en resultados disímiles (Hamp-Lyons & Kroll, 1996; Hamp-Lyons, 2003). El *enfoque analítico* supone la identificación de diferentes aspectos del texto con el fin de establecer las fortalezas y debilidades de cada uno de los aspectos, para luego determinar una evaluación final. Este enfoque implica la calificación de rasgos múltiples y considera el constructo de la escritura como algo complejo y multifacético. Una de las ventajas de esta forma de evaluar los textos es que le permite a los evaluados analizar las fortalezas y debilidades en sus propios textos, y de este modo, orientar estrategias de revisión y mejora de los propios productos y procesos de escritura (Deane et.al., 2008; Hamp-Lyons, 2003; Williams, 2005). Hamp-Lyons (2003) afirma que, a la hora de tratar los problemas y las necesidades específicos de la escritura en una segunda lengua, este tipo de calificación resulta más conveniente. Además, afirma que sirve como herramienta para investigar lo que sucede dentro del proceso de evaluación de escritura y que es una herramienta esencial para los docentes de escritores en lengua extranjera, ya que les brinda información valiosa que facilitará la toma de decisiones.

3. ANTECEDENTES

Si bien se han llevado a cabo numerosos estudios sobre evaluación, mencionaremos aquí sólo algunos cuyos resultados se relacionan con el objetivo principal de nuestro trabajo: analizar las actitudes y creencias de los docentes evaluadores de escritura. Por ejemplo, Edgington (2005) utilizó el protocolo de pensamiento en voz alta para analizar más de 500 respuestas docentes a composiciones de alumnos. Este investigador clasificó las respuestas en 13 categorías y destacó que la mayoría de los docentes establecen una amplia interacción con el alumno al leer su texto y que tienden a evaluar, clarificar y hacer preguntas al leerlo. Además, Edgington (2005) y Baker (2014) demostraron que, en general, una preocupación de los docentes al momento de evaluar y calificar la escritura en lengua extranjera es la de promover el aprendizaje, lo que hace que los procesos de enseñanza, retroalimentación y evaluación se encuentren permanentemente interrelacionados y se confundan los objetivos de cada una de estas etapas en muchos casos. Por otro lado, Hall & Sheyholislami (2013), Mei (2010) y Wolfe (2003) observaron diferencias importantes entre evaluadores no sólo en lo que respecta a los modos de calificar, sino también a las formas en las que conciben y abordan la evaluación de la escritura. Hall & Sheyholislami (2013) utilizaron la teoría de la valoración para analizar la variación entre evaluadores y sus criterios de evaluación al calificar exámenes de escritura. Los resultados de esta investigación sugieren que las decisiones de los evaluadores se vieron influenciadas por rasgos textuales que no estaban relacionados con la tarea, por interpretaciones personales que cada evaluador hace sobre qué es “escribir bien” y por sus propios valores. De manera similar, Mei (2010) estudió el proceso de evaluación de la escritura al usar una escala analítica. En este estudio se usó el protocolo de pensamiento en voz alta, además de analizar comentarios docentes en el margen de la hoja y hacer entrevistas a los evaluadores. Los resultados demostraron que, a pesar de que había algunas coincidencias en la manera en que calificaban, se destacaban principalmente las diferencias en la manera en que evaluaban el

contenido de los ensayos y en las estrategias que utilizaron para resolver dudas respecto de la nota a asignar.

4. METODOLOGÍA

4.1 Sujetos e instrumentos

En esta investigación participaron todos los docentes de la cátedra *Lengua Inglesa II* que están a cargo de la evaluación de ensayos en instancias de evaluación (profesores asistentes, adjuntos y titulares): siete docentes en total. Los instrumentos utilizados fueron los reportes escritos por los docentes, explicando y justificando la evaluación y calificación de determinados ensayos (un ensayo reprobado y un ensayo aprobado, seleccionados al azar) luego de la primera evaluación parcial de 2014. La consigna del ejercicio de escritura de ensayo fue la siguiente: *“The local College has been holding debate sessions about employment opportunities and discrimination at the workplace. As part of an awareness campaign, you have been asked to submit an essay about the similarities and differences between skilled and unskilled labor in terms of job opportunities, working conditions or gender bias policies. Write a well-developed academic essay of about 350-words”*.

4.2 Procedimiento

Cada docente seleccionó al azar 2 de los ensayos evaluados con calificaciones diferentes (un ensayo aprobado y un ensayo desaprobado). Luego, cada docente completó un reporte escrito explicando y justificando la evaluación y calificación del ensayo. Este tipo de instrumento de recolección de datos comparte algunas características con los reportes narrativos en diarios o bitácoras; sin embargo, en este caso, no se trató de un reporte extendido en el tiempo, sino de un reporte focalizado en tareas de evaluación específicas (Hatch, 2002).

4.3 Análisis de datos

Cada docente de la cátedra debía completar un reporte en el cual justificara y explicara la calificación de los ensayos corregidos. Posteriormente, se analizaron los resultados, se contabilizaron y clasificaron las marcas, correcciones y comentarios, para luego compararlos y contrastarlos a fin de determinar los aspectos en los que existe discordancia entre la evaluación de los docentes.

Los datos cualitativos correspondientes a los reportes se analizaron de manera inductiva, sin categorías o tipologías establecidas *a priori* y siguiendo los pasos propuestos por Hatch (2002), los cuales implican la identificación de marcos, dominios y temas a partir de un recorrido de lectura inductivo de los datos obtenidos.

5. RESULTADOS

En una primera instancia, los docentes debían identificar las marcas o comentarios realizados en los ensayos seleccionados, contabilizando los casos de precisión, contenido, y organización, categorías de importancia a la hora de corregir ensayos académicos. Además de contabilizar los comentarios, los docentes debían proporcionar un ejemplo que ilustrara dicho comentario o marca. Asimismo, los docentes debían clasificar las categorías según el orden de importancia que le atribuyeran. De un total de 14 reportes, 286 anotaciones se referían a precisión, contenido, y organización. Del mismo modo, el 71% de los reportes (diez casos) indicaron que el aspecto más importante a la hora de corregir ensayos académicos era la organización, seguido por el contenido y luego la precisión. Un 29% (cuatro casos) señaló otra jerarquía, comenzando por el contenido para continuar con la precisión y finalmente la organización.

Otra de las preguntas incluidas en el reporte apuntaba a que los docentes establecieran, en términos generales, sus expectativas con relación a la actitud de los estudiantes con respecto a la retroalimentación recibida. En reiteradas oportunidades los docentes indicaron que esperaban que los estudiantes controlaran su ensayo haciendo hincapié en la retroalimentación recibida y analizando los comentarios, de modo que pudieran tomar esta instancia como una oportunidad de aprendizaje y así mejorar su habilidad escrituraria. Varios docentes también advirtieron la posibilidad de que los estudiantes se contactaran con el docente para aclarar dudas e interpretar la retroalimentación recibida.

Por otro lado, los docentes debían manifestar lo que suponían que los estudiantes realmente hacen con la retroalimentación recibida. Varios docentes indicaron que los estudiantes se enfocan únicamente en la calificación y que, por lo tanto, no analizan la retroalimentación recibida. En otros casos, algunos docentes manifestaron que los estudiantes leen los comentarios pero no se acercan al docente para aclarar dudas o solucionar los problemas encontrados, o incluso se olvidan de la retroalimentación en el futuro cercano. Según algunos docentes, hay estudiantes que se enfocan en los errores pero no los incorporan para así mejorar la habilidad de escritura. Algunos docentes indicaron que en ciertos casos, los estudiantes responden a la retroalimentación en función de lo que los docentes esperan que hagan. Por último, señalaron que algunos estudiantes nunca se encuentran con sus ensayos, ya que no los retiran después de los parciales o finales.

Luego del análisis de las marcas y comentarios de los docentes con relación a todos los ensayos, se clasificaron los ensayos entre aquellos aprobados (50%) y los reprobados (50%), para luego completar las tablas correspondientes.

En primer lugar, se analizaron siete (7) ensayos aprobados. Los docentes debían contabilizar las marcas sobre precisión, contenido, y organización, además de proporcionar un ejemplo en cada caso. En cuanto a la precisión, los docentes computaron 79 marcas, en su mayoría relacionadas con sintaxis o a elementos ausentes, como por ejemplo: *“Some of the aspects which worth being mentioned”*. Con respecto al contenido, se contabilizaron 11 marcas, particularmente relacionadas con ideas imprecisas, así como también generalizaciones, tales como *“Although nowadays finding a job is hard for anyone, if you are too specialized in a field it may be hard to find a job, especially if you wish to work near your home. For example, not all cities need ten new specialists in treating pneumonia every year”*. Finalmente, con relación a la organización, se contaron 7 marcas, entre las

que se pueden mencionar ejemplos como *“To begin with, a skilled worker, someone who has spent several years at university will share the workplace with someone who has never gone to a college”* (oración tópica inadecuada).

Se indagó a los docentes sobre los aspectos importantes a la hora de corregir un ensayo académico, pero sus respuestas fueron las mismas que se recabaron para las preguntas sobre todos los ensayos en general. Lo mismo sucede con respecto a las expectativas que tienen los docentes con relación a la respuesta que dan los estudiantes a la retroalimentación recibida, salvo que algunos docentes agregaron que los estudiantes comprenden cómo mejorar sus ensayos y que consultan el diccionario y reflexionan sobre la importancia del desarrollo en el proceso escriturario.

Del mismo modo, se analizaron 7 (siete) reportes sobre ensayos reprobados. Las preguntas fueron las mismas que para los ensayos aprobados. En cuanto a las marcas de precisión, se contabilizaron 126 marcas, entre las que se destacó la sintaxis, en casos como: *“that implies be able to work flexible hours in a peaceful enviroment”*. En cuanto al contenido, se resaltaron 49 marcas, tales como aquellas relacionadas con ensayos fuera de foco y con problemas de lengua, como por ejemplo: *“To begin with, recruitment fairs are similar to the vast majority of jobs interviews. They are done following certain steps like massive interview, an individual one and then, a role-play where you are supposed to to your ‘future’ job”*. Finalmente, 14 marcas hacían referencia a la organización, como por ejemplo: *“Nowadays, technology is applied in almost every job and, as a result, the gap between skilled and unskilled job opportunities and working conditions grows larger by the day”* (tesis del ensayo inadecuada).

No se presentaron cambios con respecto a lo que los docentes consideran de importancia a la hora de corregir un ensayo académico, pero sí agregaron comentarios con respecto a lo que esperan que los estudiantes hagan con la retroalimentación. Algunos docentes añadieron que el estudiante debería releer la consigna cuidadosamente para así poder comprender por qué algunas ideas están fuera de foco, así como también reconocer los problemas de organización. Otros docentes sugirieron que los estudiantes deberían reflexionar sobre su desempeño para, de este modo, revisar los aspectos a mejorar.

Entre las fortalezas detectadas por los docentes, particularmente para los ensayos aprobados, se pueden mencionar la organización y el hecho de que dichos ensayos se enfocan en el tema asignado, utilizando vocabulario específico. Asimismo, también advierten la presencia de ideas interesantes y específicas para el tema. En el caso de los ensayos reprobados, la mayoría de los docentes concuerdan en considerar la organización del ensayo como el aspecto más fuerte. En algunos casos, si bien el ensayo no explora los dos aspectos del tema a desarrollar, de todas maneras la organización es apropiada. Inclusive uno de los docentes consideró la organización como la única fortaleza en uno de los ensayos corregidos. Otra de las fortalezas apunta a la expresión de ideas y el uso de expresiones y/o vocabulario específico, lo cual contribuye a mejorar el ensayo.

Con respecto a las debilidades que los docentes detectaron en los ensayos aprobados, se mencionaron problemas relacionados con la precisión en el uso de la lengua, así como también errores elementales que, si bien no dificultaron la comprensión de las ideas presentadas, resultan relevantes para el nivel del curso. Otra de las debilidades hace alusión a la falta de ideas críticas y sólidas. También se hace referencia a la falta de aspectos formales tales como el título, los márgenes o el número de palabras al final del ensayo, aspectos a cumplir no sólo en instancias de evaluación

formal, sino también en las entregas de ensayos durante el cursado de la materia. En cuanto a los ensayos reprobados, se destacan entre las debilidades los problemas relacionados con la falta de oraciones tópicas, así como también problemas de precisión en el uso de la lengua. Varios docentes reportaron que los estudiantes habían cometido errores relacionados con sintaxis, tiempos verbales, vocabulario, puntuación, etc. En algunos casos, dichos errores no entorpecieron la comprensión del texto, aunque en otros se presentaron como errores elementales.

En cuanto a contenido y organización, algunos docentes mencionaron problemas relacionados con el uso incorrecto de conectores, los cuales deberían ser dominados por los estudiantes en este nivel, así como también instancias en las que uno de los puntos a contrastar no fue mencionado o en las que el alumno tuvo dificultades en expresar la oración tópica de cada párrafo. Finalmente, en un caso se analizó la solidez de las ideas, lo cual se correspondía con el uso de estructuras simples en la producción escrita de los estudiantes.

Por último, la mayoría de los docentes informó que las debilidades de los ensayos reprobados se relacionaban con organización y contenido. Varios ensayos no presentaron un análisis profundo del tema y no expresaron las ideas de manera clara. Asimismo, se hizo hincapié en la irrelevancia de algunas ideas y en el hecho de que algunos estudiantes no desarrollaron el tema que debían abordar; por lo tanto, el ensayo estaba fuera de foco y no daba respuesta a la pregunta del examen.

6. CONCLUSIONES

Luego de analizar los reportes docentes se puede concluir que, si bien la mayoría de los ellos (71%, 10 casos) sugiere que los docentes siguen el mismo criterio en cuanto a la jerarquía de los aspectos a tener en cuenta a la hora de corregir textos académicos, un 29% (4 casos) dejó en evidencia que algunos docentes jerarquizan estos aspectos de otra manera. A pesar de que en las reuniones de cátedra y antes de cada evaluación se fijan y reiteran estos criterios, no caben dudas de que será necesario tomar otras medidas a nivel cátedra para garantizar la igualdad de criterios a la hora de corregir textos académicos.

Además, es importante destacar que, tanto en los ensayos aprobados como en los desaprobados, la mayoría de los comentarios/marcas encontrados se relacionan con la precisión (uso de la lengua). Si bien en la mayoría de los reportes se indicó que éste es el aspecto menos importante a tener en cuenta, es probable que el contexto en el que se llevó a cabo la investigación (formación de futuros profesionales de una lengua extranjera) sea la razón por la que se advierte una importante presencia de este tipo de comentarios/marcas.

Cuando los docentes debieron indicar qué consideraban como fortalezas de los ensayos corregidos, mencionaron aspectos tales como la organización del texto, la inclusión de ideas interesantes, el enfoque adecuado y el uso de expresiones o vocabulario específico. Por otro lado, al indicar las debilidades de las composiciones, hicieron referencia a la falta de precisión en el uso de la lengua, a la ausencia o escasez de ideas críticas, a la ausencia de oraciones tópicas y a problemas con aspectos formales.

Respecto de las expectativas de los docentes en cuanto a la reacción de los estudiantes frente al *feedback*, en términos generales, los docentes esperan que los estudiantes revisen sus ensayos enfocándose en la retroalimentación recibida y analizando los comentarios, de modo de poder tomar

esta instancia como una instancia de aprendizaje. Varios docentes también mencionaron la posibilidad de que los estudiantes se contacten con el docente para aclarar dudas e interpretar la retroalimentación recibida.

Vale aclarar que cuando los docentes debieron explicitar qué es lo que creen que los estudiantes efectivamente hacen con el *feedback* recibido, indicaron, entre otras cosas, que los estudiantes sólo se enfocan en la nota y no analizan el *feedback* recibido o que leen los comentarios, pero no se acercan al docente para aclarar dudas o solucionar los problemas encontrados. También comentaron que algunos estudiantes se enfocan en los errores, pero no los incorporan para así mejorar la habilidad de escritura, y que un grupo de estudiantes hace lo que los docentes esperan que hagan. Por último, los docentes mencionaron que algunos estudiantes nunca se encuentran con sus ensayos, ya que no los retiran después de los exámenes parciales o finales.

Referencias

- Baker, N. (2014). "Get it off my stack": Teacher's tools for grading papers. *Assessing Writing*, 19, 36-50.
- Deane, P; Odendhal, N.; Quinlan, T.; Fowles, M.; Welsh, C. & Bivens-Tatum, J. (2008). *Cognitive models of writing: Writing proficiency as a complex integrated skill*. Princeton, NJ: ETS.
- Edgington, A. (2005). Understanding teachers reading and response through a protocol analysis study. *Journal of Writing Assessment*, 2 (2), 125-148.
- Goodman, S. & Swann, J. (2003). Planning the assessment of student writing. En Coffin, C; Curry, M.J.; Goodman, S.; Hewings, A.; Lillis, T.M. & Swann, J. (Eds), *Teaching academic writing. A toolkit for higher education* (pp. 73-100). USA: Routledge.
- Hall, C. & Sheyholislami, J. (2013). Using appraisal theory to identify rater values: An examination of rater comments on ESL test essays. *Journal of Writing Assessment*, 6 (1), 1-15.
- Hamp-Lyons, L. (2003). Writing teachers as assessors of writing. En Kroll, B. (Ed.), *Exploring the dynamics of second language writing*. USA: Cambridge University Press.
- Hamp- Lyons, L. (2001). Fourth generation writing assessment. En Silva, T. & Matsuda, P. (Eds.), *On second language writing*. New Jersey: Lawrence Erlbaum Associates.
- Hamp-Lyons, L. & Kroll, B. (1996). Issues in ESL writing assessment: An overview. *College ESL*, 6 (1), 52-72.
- Hatch, J. A. (2002). *Doing qualitative research in education settings*. New York: State University of New York Press.
- Hyland, K. (2003). *Second language writing*. New York: Cambridge University Press.
- Mei, W. (2010). Investigating raters' use of analytic descriptors in assessing writing. *Reflections on English Language Teaching*, 9 (2), 69-104.
- Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. Chicago: McGraw-Hill.
- Wolfe, E. W. (2003). Uncovering rater's cognitive processing and focus using think-aloud protocols. *Journal of Writing Assessment*, 2 (1), 37-56.