



## **Criterios uniformados de evaluación para la corrección de párrafos académicos: Una propuesta de investigación**

**Spataro, Claudia y Bersano, Natalia**

Fecha de recepción: 30/11/2018

Fecha de aprobación: 11/12/2018

Eje: Currículum, perspectivas didácticas y evaluación

### **Resumen**

El lugar central que ocupa la evaluación en relación con la carrera de los alumnos justifica un análisis profundo de los criterios de evaluación que garantice objetividad y consistencia por parte de los evaluadores a la hora de calificar. Este aspecto constituyó el objeto de estudio del proyecto de investigación intercátedra *Hacia la elaboración de criterios uniformados de evaluación que aseguren fiabilidad entre los calificadores de las pruebas escritas en las cátedras de Lengua Inglesa I, Práctica Gramatical del Inglés, Gramática Inglesa I, Práctica de la Pronunciación del Inglés, Lengua Castellana I, y Lectocomprensión en Lengua Extranjera IV (Inglés)* avalado por Secyt y llevado a cabo en Facultad de Lenguas, Universidad Nacional de Córdoba en el periodo 2016-2017. Uno de los subgrupos de trabajo fue conformado por profesores de la cátedra Lengua Inglesa I quienes trabajamos con la evaluación de escritura de párrafos académicos. Lengua Inglesa I es una asignatura de primer año de las carreras de Profesorado de Lengua Inglesa, Licenciatura en Lengua Inglesa y Traductorado Público Nacional de Inglés de la Facultad de Lenguas, U.N.C., que tiene como objetivo lograr que los alumnos desarrollen las cuatro macrohabilidades de la lengua (lectura, escucha, escritura y habla) a fin de alcanzar un nivel intermedio en su conocimiento de inglés. En esta ponencia presentaremos la metodología de recolección y análisis de datos que nos permitió, considerando distintos antecedentes (Romano, Morra & Dalla Costa, 2014; Lee, 2014; Tang, Chiachieh y Liu, Yeu-Ting, 2018) y tipos de retroalimentación (Hyland & Hyland, 2006), diseñar dos posibles constructos para homogeneizar criterios uniformados de calificación de los párrafos académicos de nuestros alumnos de Lengua Inglesa I.



## Introducción

Numerosos estudios de investigación han argumentado extensamente acerca de la fiabilidad de las calificaciones y de la probada influencia negativa que ejerce la subjetividad en la percepción y actitud de los que corrigen (Bachman y Cohen, 1990; Gass, 1994; Alderson, Clapham & Wall, 1995); es por ello que, más allá de las discusiones acerca de las concepciones y de las formas que puedan considerarse más pertinentes o apropiadas para llevar a cabo la acción de evaluar por parte del docente, esa acción tiene una importancia fundamental ya que en ella se produce el encuentro entre los criterios sostenidos por la institución educativa y por el que “enseña”, por un lado, y lo que le es posible mostrar al alumno como adquisición durante un periodo preestablecido, por el otro.

No podemos ignorar el hecho de que es, precisamente, como resultado del proceso de evaluación, que el alumno es promovido a un nivel siguiente en la sucesión en que han sido ordenadas esas adquisiciones, o bien sufre el impedimento para continuar y, en consecuencia, debe volver sobre la etapa anterior para revisar, corregir, practicar aquellas competencias y/o conocimientos que aún no han sido adquiridos para superar la presente etapa y ser promovidos a la siguiente. Es relativamente simple concluir que, al menos parcialmente, el hecho de permanecer, detenerse, avanzar en la carrera al ritmo previsto o a un ritmo menor depende del juicio que el profesor elabore, de acuerdo con determinados criterios, sobre lo que demuestra el alumno en la situación diseñada a los efectos de la evaluación. Este aspecto constituyó el objeto de estudio del proyecto de investigación intercátedra *Hacia la elaboración de criterios uniformados de evaluación que aseguren fiabilidad entre los calificadores de las pruebas escritas en las cátedras de Lengua Inglesa I, Práctica Gramatical del Inglés, Gramática Inglesa I, Práctica de la Pronunciación del Inglés, Lengua Castellana I, y Lectocomprensión en Lengua Extranjera IV (Inglés)* avalado por Secyt y llevado a cabo en Facultad de Lenguas, Universidad Nacional de Córdoba en el periodo 2016-2017.

Uno de los subgrupos de trabajo fue conformado por profesores de la cátedra Lengua Inglesa I quienes trabajamos con la evaluación de escritura de párrafos académicos. Lengua Inglesa I es una asignatura de primer año de las carreras de



Profesorado de Lengua Inglesa, Licenciatura en Lengua Inglesa y Traductorado Público Nacional de Inglés de la Facultad de Lenguas, U.N.C., que tiene como objetivo lograr que los alumnos desarrollen las cuatro macrohabilidades de la lengua (lectura, escucha, escritura y habla) a fin de alcanzar un nivel intermedio en su conocimiento de inglés.

### **Encuadre teórico**

La evaluación de la escritura en lengua extranjera ha sido un campo de permanente investigación dada su variedad según los contextos en los que se lleva a cabo. Existen distintas formas de evaluar la producción escrita de los alumnos de una lengua extranjera. Entre las formas más efectivas de corrección de la escritura de los alumnos de inglés como una segunda lengua (ESL) o lengua extranjera (ILE) por parte del docente, Harmer (2007) menciona cinco tipos: 1. corrección selectiva, en donde el docente-evaluador solo señala los errores o secciones más relevantes; 2. uso de grillas de evaluación con puntaje específico para las categorías delimitadas (gramática, vocabulario, coherencia), 3. uso de símbolos de corrección; 4. reformulación correcta de la palabra, frase u oración con error; y 5. referencia a diccionarios monolingües o libros de gramática para que los alumnos consulten específicamente un error y descubran cómo corregirlo. En la práctica, todas estas formas se utilizan por docentes de ESL y ILE según el contexto, la instancia y el propósito de la corrección y todas presentan distintas fortalezas y debilidades.

La conocida distinción de Hendrickson (1978) entre *retroalimentación directa* e *indirecta* también es ampliamente utilizada en la literatura del área (Hyland & Hyland, 2006). Al utilizar retroalimentación directa, el evaluador proporciona la forma o estructura correcta (Hendrickson, 1978). De esta forma, el alumno puede corroborar cuál fue su error y cómo se resuelve de manera inmediata. Diversos estudios han demostrado que este tipo de corrección permite que el alumno resuelva los errores de manera exitosa (Bitchener, Young y Cameron, 2005). En el caso de la retroalimentación indirecta, el docente-evaluador señala el tipo de error pero no lo corrige (Ferris, 2006). El uso de símbolos de corrección, por ejemplo, es instrumental para implementar este tipo de retroalimentación en donde se señala el error en el



texto para que el alumno, en una segunda etapa del proceso de escritura, lo identifique y corrija. Ferris y Roberts (2001) indican que los expertos en el área favorecen el uso de retroalimentación indirecta para los errores léxico-gramaticales puesto que permiten al alumno mejorar su precisión en cuanto al uso de la lengua a largo plazo. Tanto la retroalimentación directa como la indirecta pueden utilizarse para la *evaluación sumativa*, la cual es objeto del presente trabajo ya que califica formalmente y se lleva a cabo, generalmente, al final del proceso de enseñanza con el fin de determinar si se han logrado los objetivos planteados para todo el curso (Romano, Morra & Dalla Costa, 2014).

Finalmente, una distinción frecuente en cuanto a los abordajes de evaluación de un texto es la de *enfoque holístico y analítico*. Las escalas *holísticas* presentan bandas que consideran al texto como un todo y conlleva la asignación de una nota o evaluación final general, con descriptores más o menos detallados de los distintos aspectos. Romano, Morra & Dalla Costa sostienen que:

Esta calificación se basa en la premisa de que el constructo de la escritura es una entidad única que puede capturarse mediante una escala que integra las cualidades inherentes de la escritura y que la calidad solo puede ser reconocida por lectores expertos” (2014:3).

Entre los problemas más comunes 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.

Por el contrario, las escalas *analíticas* consideran distintas categorías o campos específicos que permiten determinar las fortalezas y debilidades de distintos aspectos del texto por separado 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”. (Romano, Morra y Dalla Costa, 2014:3). Hamp-Lyons (2003) menciona que este tipo de enfoque es más efectivo para centrarse en los problemas y necesidades específicas de los escritores en una segunda lengua. Una de las ventajas de esta forma de evaluar los textos es que les permite a los estudiantes evaluados analizar las fortalezas y debilidades de sus



propios textos, y de esta manera, orientar estrategias de revisión y mejora de los propios productos y procesos de escritura. El autor también sostiene que este tipo de escala sirve como herramienta para investigar lo que sucede en el 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.

Si bien se han llevado a cabo numerosos estudios sobre evaluación, aquellos que resultan de especial interés para esta investigación son los que han considerado la relativa efectividad de diferentes criterios de calificación (Barkaoui, 2007; Crusan, 2015; Cushing Weigle, 2002; Eckes, 2008; White, 1998) y los que han analizado los distintos procesos que llevan a cabo los docentes lectores al momento de evaluar y calificar los textos producidos por los estudiantes (Kim, Schatschneider, Wanzek, Gatlin y Al Otaiba, 2017; Kuiken y Vedder, 2014). Estos últimos han demostrado que, en general, una preocupación de los docentes al momento de evaluar y calificar la escritura en lengua extranjera es la de fomentar 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 (Baker, 2014).

En un volumen de una de las publicaciones especializadas en evaluación de escritura, *Assessing Writing*, un estudio etnográfico llevado a cabo por Baker (2014) pone de manifiesto que los objetivos de los docentes al momento de evaluar están orientados tanto hacia el producto textual (y la calidad del mismo) como hacia el proceso de escritura (y su desarrollo). Es por ello que los docentes que participaron en este estudio reportaron, por ejemplo, la necesidad de justificar su calificación, especialmente cuando se trata de una baja calificación, y de anticipar críticas o desacuerdos por parte de los estudiantes. En otra publicación de la misma revista, Lee (2014) recalca la necesidad de contar con más investigaciones situadas en contextos de escritura en lengua extranjera (específicamente, en inglés como lengua extranjera), que analicen prácticas concretas de evaluación y retroalimentación de la escritura. Autores como Eguiluz Pacheco y De Vega Santos (2009), en un estudio sobre los criterios para la evaluación de la producción escrita, señalaron que las



pruebas subjetivas pueden presentar un cierto grado de imprecisión en su corrección. Para asegurar que se cumplan igualmente los requisitos de validez y fiabilidad, afirmaron que es necesario establecer criterios unívocos de valoración que permitan la objetivación de los resultados.

Teniendo en cuenta estos antecedentes, los cuales dan cuenta de la complejidad inherente a la evaluación de la escritura, así como también la importancia de la fiabilidad de cualquier instrumento de medición de la evaluación escrita, el subgrupo de trabajo conformado por profesores de la cátedra Lengua Inglesa I quienes trabajamos con la evaluación de escritura de párrafos académicos planteamos las siguientes preguntas de investigación:

¿Qué sustento tienen las inquietudes planteadas por algunos alumnos de la Facultad de Lenguas con respecto a una aparente divergencia en relación con a los criterios de corrección y calificación y, en consecuencia, entre las puntuaciones asignadas en los exámenes escritos por los distintos evaluadores?;

¿Cuál es el grado de convergencia entre las puntuaciones asignadas por los distintos evaluadores en las pruebas escritas de las cátedras objeto de estudio en este proyecto?;

¿En caso de ausencia significativa de consistencia entre los distintos profesores-calificadores en cuanto a la puntuación otorgada a una misma prueba escrita, es posible determinar las causas más comunes que la provocan?;

¿Es posible elaborar criterios claros, uniformados y precisos de calificación que ayuden, por una parte, a los profesores evaluadores a lograr coincidencia en la puntuación y/o calificación de las pruebas escritas y, por otra, a los alumnos a confiar en la calificación obtenida?

Teniendo en cuenta el marco teórico, los antecedentes y las preguntas de investigación, el subgrupo de trabajo del proyecto de investigación intercátedra conformado por profesores de la cátedra *Lengua Inglesa I* se focalizó en la evaluación sumativa de los párrafos académicos utilizando tanto retroalimentación directa como indirecta y en el diseño de posibles constructos analíticos para homogeneizar criterios uniformados de calificación, tal como lo detallan los objetivos del trabajo.



## **Objetivo general**

Abordar la elaboración de criterios de evaluación uniformados con el objetivo de lograr fiabilidad interevaluadora en las Cátedras de Lengua Inglesa I, Práctica Gramatical del Inglés, Gramática Inglesa I, Práctica de la Pronunciación del Inglés, Lengua Castellana I y Lectocomprensión en Lengua Extranjera IV (Inglés) en la Facultad de Lenguas de la Universidad Nacional de Córdoba (UNC)

## **Objetivos específicos**

1. Describir y analizar los objetivos, contenidos programáticos, metodología de enseñanza, metodología y criterios de evaluación, tal cual se explicitan en el programa vigente correspondiente a cada una de las asignaturas objeto de análisis en este proyecto.
2. Verificar el grado de convergencia entre las puntuaciones asignadas por los distintos profesores-calificadores a cada una de las pruebas escritas.
3. Analizar, interpretar y/o describir las posibles causas que conduzcan a la ausencia de criterios unificados de corrección, según el tipo de actividad de que se trate y el tipo de error detectado.
4. Elaborar criterios de corrección uniformados referenciales de suficiencia de la competencia lingüística esperada en forma de escalas de estimación (de niveles de habilidad), que reflejen los requerimientos de suficiencia en términos de ausencia de errores tipificados y/o presencia de errores tipificados cuantificables.
5. Verificar la eficacia y utilidad de dichos criterios uniformados de suficiencia de la competencia lingüística esperada, en términos de uniformidad lograda en las calificaciones, independientemente de la identidad del profesor-calificador.

## **Metodología**

### *Sujetos*

Lengua Inglesa I está dividida en dos cátedras, la cátedra A cuenta con 6 comisiones de la mañana y la cátedra B con dos de la tarde. Aproximadamente 500 alumnos cursan la asignatura por año y la asignatura tiene una carga horaria de 10



horas semanales. Las cátedras cuentan con un número elevado de estudiantes y como el cuerpo docente a cargo suele fluctuar dependiendo de variables tales como licencia por salud, estudios, embarazos, entre otros, se consideró necesario acordar ciertos criterios de corrección que aseguraran que las correcciones se logaran de la manera más efectiva y objetiva posible. Por este motivo, 5 de los profesores de la asignatura (María José Morchio, Claudia Spataro, Ana Granata, Agustín Massa y Natalia Bersano) formamos parte del subgrupo 1 del proyecto. La Profesora Titular de la cátedra B, Natalia Dalla Costa, actuó como evaluador externo.

### **Recolección y análisis de datos**

En un primer momento, y tal como sucede habitualmente, todos los profesores de las Cátedras de Lengua Inglesa I procedieron a diseñar en conjunto el examen final (“la prueba”) correspondiente al turno de examen febrero de 2017. Dicho examen fue elaborado de acuerdo con los contenidos disciplinares y formato de evaluación correspondientes al ciclo lectivo 2016. Una vez que el examen fue administrado y realizado por los alumnos en la mesa de examen, el director del proyecto seleccionó aleatoriamente 14 exámenes de alumnos “Regulares”. A los fines de preservar la versión original de la muestra, se procedió a fotocopiar la sección de composición de párrafos académicos de cada uno de los exámenes seleccionados antes de que comenzara el proceso de corrección de los mismos. Se ocultó el nombre del alumno a fines de preservar su identidad.

El subgrupo de trabajo se abocó a realizar la descripción y el análisis de los objetivos, contenidos disciplinares, metodología de trabajo, formas y criterios de evaluación de la asignatura objeto de estudio en cada caso. Seguidamente, se asignó a los docentes miembros del subgrupo de trabajo a corregir 7 párrafos de los exámenes que fueron fotocopiados y preservados oportunamente.

Para la recolección de datos cualitativos se organizaron grupos de discusión estructurados siguiendo la dinámica de una conversación entre el grupo de docentes participantes. En dichas instancias, el director del proyecto fue registrando los intercambios y, posteriormente, comparando los puntajes obtenidos por cada



docente y señalando las discrepancias dadas en cada corrección así como también las similitudes.

Se realizaron 3 sesiones (mayo 2017, octubre 2017 y febrero 2018) en las que participaron todos los integrantes del subgrupo 1 junto con el director y la codirectora. Para la primera y segunda sesión se utilizaron los mismos párrafos mientras que para la última sesión se corrigieron párrafos nuevos (los 7 restantes) para, de este modo, poder poner a prueba la grilla final de evaluación elaborada por los integrantes del subgrupo del proyecto.

### **Conclusión**

El concepto de fiabilidad es aplicable a cualquier instrumento y, por lo tanto, es también aplicable al instrumento que emplea el docente para evaluar. De este modo, la fiabilidad es una cualidad esencial que debe estar presente en todos los exámenes de carácter académico-científico, y que deciden, como en el caso que nos convoca, la promoción al curso siguiente.

La descripción de los antecedentes y metodología del presente proyecto permite direccionar a aquellos investigadores interesados en el área a evaluar sus criterios de calificación de producciones escritas a fin de poder responder importantes interrogantes: ¿cuál es el grado de convergencia entre las puntuaciones asignadas por los distintos evaluadores en las pruebas escritas de las cátedras objeto de estudio en este proyecto?; ¿en caso de ausencia significativa de consistencia entre los distintos profesores-calificadores en cuanto a la puntuación otorgada a una misma prueba escrita, es posible determinar las causas más comunes que la provoca? Las respuestas a estas preguntas podrían, de algún modo, guiar la forma y método de corrección de un equipo de trabajo con el fin de aunar criterios y lograr un alto grado de fiabilidad en la evaluación.

### **Referencias bibliográficas**

- Alderson, J. C., C. Clapham & D. Wall (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. F. & A. D. Cohen (1998). *Interfaces Between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press.



- Baker, N. (2014). Get it off my stack: Teacher's tools for grading papers. *Assessing Writing*, 19, 36-50.
- Barkaoui, K. (2007). Rating scale impact on EFL essay marking: A mixed-method study. *Assessing Writing*, 12, 86–107.
- Bitchener, J. S.; Young, S. y Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*. 14(3), pp.215-241.
- Crusan, D. (ed). (2015). The use of rubrics to assess writing: Issues and challenges. *Assessing Writing*, 26, 1-82.
- Cushing Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Eckes, T. (2008). Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing*, 25, 155–185.
- Eguiluz Pacheco, J. y De Vega Santos, C. (2009). Criterios para la evaluación de la producción escrita. *Didáctica del español como lengua extranjera*, 9.
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. En: HYLAND K. and F. HYLAND, eds. *Feedback in second language writing. Contexts and issues*. Cambridge: Cambridge University Press, pp.81-104.
- Ferris, D. R. y Roberts, B. (2001). Error feedback in L2 writing classes: how explicit does it need to be? *Journal of Second Language Writing*, 10(3), pp.161-184.
- Gass, S. (1994). The reliability of grammaticality judgements. En E. Tarone, S. Gass and A. Cohen (eds.). *Research Methodology in Second Language Acquisition*. 303-322. Hillsdale, N. J: Lawrence Earlbaum Associates.
- 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.
- Harmer, J. 2007. *The Practice of English Language Teaching* (4th edition). Harlow: Longman.
- Hendrickson, J. (1978). Error correction in foreign language teaching: recent theory, research and practice. *Modern Language Journal*. 62(8), pp.387-398.
- Hyland, K. & Hyland, F. (2006). *Feedback in Second Language Writing. Context and Issues*. Cambridge: CUP



- Kim, Y., Schatschneider, C., Wanzek, J., Gatlin, B. y Al Otaiba, S. (2017). Writing evaluation: rater and task effects on the reliability of writing scores for children in Grades 3 and 4. *Reading and Writing, 30*(6), 1287-1310.
- Kuiken, F., & Vedder, I. (2014). Rating written performance: What do raters do and why? *Language Testing, 31*, 329–348.
- Lee I. Feedback in writing: Issues and challenges. (2014). *Assessing Writing, 19*, 1-5.
- Romano, M, Morra, A & Dalla Costa, N. (2014). La evaluación de la escritura en lengua extranjera (inglés) en el nivel superior: una propuesta de investigación. En Actas del Congreso Nacional Cátedra UNESCO Subsede Rosario "La lectura y la escritura en las sociedades del siglo XXI". Disponible en <http://rephip.unr.edu.ar/handle/2133/4751>
- Tang, Chiachieh y Liu, Yeu-Ting. (2018). Effects of indirect coded corrective feedback with and without short affective teacher comments on L2 writing performance, learner uptake and motivation. *Assessing Writing, 35*, 26-40.
- White, E. M. (1998). *Teaching and assessing writing: Recent advances in understanding, evaluating and improving student performance*. USA: Calendar Islands Publisher.