



Evaluación de la escritura académica en ILE con un instrumento analítico

Salinas, Julieta; Cad, Ana Cecilia y Canavosio, Andrea de los Ángeles

Fecha de recepción: 30/11/2018

Fecha de aprobación: 11/12/2018

Eje: Currículum, perspectivas didácticas y evaluación

Resumen

El rol de la evaluación así como las maneras a través de las cuales este proceso puede materializarse se presentan como una problemática a abordar que se convierte en objeto de estudio de diversas investigaciones. Esto se debe a su importancia tanto por los datos que proporciona como también por su impacto en las prácticas pedagógicas y en los procesos de aprendizaje de los estudiantes. El caso de la evaluación de la escritura en lengua extranjera en particular se presenta como un campo de permanente investigación por la variedad de contextos en los que se lleva a cabo.

En nuestro contexto educativo, la materia Lengua Inglesa II de las carreras de Profesorado, Traductorado y Licenciatura en Inglés de la Facultad de Lenguas (UNC), la evaluación de la escritura es clave en los procesos de enseñanza y de aprendizaje y se lleva a cabo de diversas maneras y en diferentes instancias de evaluación tanto formativa como sumativa. Es por ello que, en el marco del proyecto de investigación titulado “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior: aplicación de instrumentos de evaluación”, se propuso diseñar y analizar la aplicación de dos instrumentos de evaluación de ensayos académicos para lograr garantizar la confiabilidad del sistema de calificación así como su utilidad. En este caso, presentaremos un análisis descriptivo de la implementación de un instrumento de naturaleza analítica y los resultados obtenidos.

Introducción

El proceso de evaluación en el nivel superior está revestido de gran importancia dada su naturaleza y su sentido dentro del contexto universitario. Por lo tanto, las discusiones sobre la forma en la que la evaluación es diseñada y su puesta en



práctica son comunes dentro del ámbito. Se entiende que al graduarse, el alumno ha superado situaciones de evaluación que acreditan el logro del conocimiento práctico y teórico para el ejercicio de su profesión en el ámbito laboral o académico pertinente a su área de acción. Por este motivo resulta necesario adoptar a la evaluación como objeto de estudio a fin de ser capaces de perfeccionar el instrumento evaluador y lograr mayor confiabilidad en su aplicación. La evaluación de la escritura en lengua extranjera concretamente se constituye en foco dentro del marco de esta investigación debido a su variedad y complejidad.

Este trabajo de investigación se desarrolla dentro de la cátedra de Lengua Inglesa II de las carreras del Profesorado, Traductorado y Licenciatura en inglés de la Facultad de Lenguas, UNC. El proceso de escritura es una parte integral de esta materia y se lleva a cabo tanto en instancias de evaluación formativa (parciales y tareas de escritura) como en instancias de evaluación sumativa (finales). Considerando la importancia de la escritura en nuestra materia y el número de docentes que corrigen dichos escritos, se torna fundamental legitimar el uso de los instrumentos evaluativos a través del uso y aplicación de criterios comunes a todos los docentes de la cátedra para así lograr confiabilidad y validez dados por la uniformidad de los criterios a seguir a la hora de calificar al alumnado. Teniendo lo antedicho en consideración, en nuestro proyecto de investigación, “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior: aplicación de instrumentos de evaluación”, se propuso diseñar y analizar la aplicación de dos instrumentos de evaluación de ensayos académicos para lograr garantizar la confiabilidad del sistema de calificación así como su utilidad. En este caso, presentaremos un análisis descriptivo de la aplicación de un instrumento de naturaleza analítica y los resultados obtenidos.

Evaluación de la escritura

La evaluación de composiciones académicas es un componente esencial en muchas materias de la Facultad de Lenguas ya que este tipo de respuesta de carácter abierto se considera una buena fuente para poder inferir el nivel de proficiencia de la habilidad escrituraria del alumnado. Los resultados de estos exámenes pueden tener serias consecuencias para la vida de los estudiantes ya que pueden afectar sus



decisiones académicas y profesionales. Por lo tanto esto constituye una razón de peso para lograr entender el comportamiento del evaluador y poder así disminuir el nivel de subjetividad al colocar notas. Es importante considerar primero qué entendemos por evaluación. En palabras de Gimeno: Evaluar hace referencia a cualquier proceso por medio del que algunas o varias características de un alumno, de un grupo de estudiantes, de un ambiente educativo, de objetos educativos, de materiales, de profesores, de programas, etc. reciben la atención del que evalúa, se analizan y se valoran sus características y condiciones en función de unos criterios o puntos de referencia para emitir un juicio relevante para la educación (1992 citado en Fernández March, 2008:4).

La evaluación educativa es definida por otros autores como ‘cualquier método que se use para averiguar qué conocimiento poseen los estudiantes en un momento dado’ (Collins y O’Brien, 2003:29). De manera similar, Crooks (2001) se refiere a la evaluación como ‘cualquier proceso que nos brinde información sobre los logros y el progreso de los estudiantes’.

Dentro de nuestro contexto educativo, la evaluación de la escritura se realiza durante parciales y finales. En ambos casos estamos presentes frente a que Williams (2005) denomina la *evaluación directa* de la escritura ya que se evalúa el proceso escriturario a través una producción textual individual y no a través de ejercicios de componentes lingüísticos que hacen uso de la escritura (ejercicios basados en usos gramaticales, de vocabulario, etc.); estas actividades son definidas como *evaluación indirecta* de la escritura (Williams, 2005:120). La *evaluación directa* da muestras concretas de que la escritura suele ser considerada como una de las maneras más efectivas de lograr una apreciación de la habilidad escrituraria de un estudiante (Hamp-Lyons, 2001; Williams, 2005)

Otro aspecto a considerar es el propósito de la evaluación. Se pueden distinguir, entre otros, dos tipos de evaluación: *sumativa* (evaluación *del* aprendizaje) y *formativa* (evaluación *para* el aprendizaje). La evaluación sumativa se hace al final de un ciclo de enseñanza y se utiliza para estimar cuánto han aprendido los alumnos o si se han logrado los objetivos planteados para el curso. Se realiza generalmente para determinar si el alumno ha alcanzado el nivel necesario para continuar al siguiente nivel y suele hacerse a través de exámenes, pruebas, proyectos o tareas.



Su resultado se expresa generalmente en una nota (Bloom, Hastings y Madaus, 1971; Black y Wiliam, 2009; Crooks, 2001; Shepard, 2005).

Por el otro lado, la evaluación formativa se realiza a lo largo del curso y su objetivo es primordialmente obtener retroalimentación sobre los procesos de enseñanza y aprendizaje (E/A), identificando sus fortalezas y debilidades (Collins y O'Brien, 2003; O'Malley y Pierce, 1996). Es un proceso continuo durante el cual se recolecta información sobre el desempeño de los alumnos a través de diferentes instrumentos de evaluación para intervenir en base a los resultados. Al hacerse durante el proceso de E/A, su objetivo no es solo medir el grado de proficiencia de los estudiantes sino visualizar su nivel de avance, la efectividad del curso, identificar problemas y sugerir soluciones.

Escala analítica

Las escalas de medición pueden categorizarse teniendo en cuenta cómo nos permiten obtener resultados. Por un lado en las escalas holísticas los evaluadores consideran la producción escrita como un todo, mientras que en las escalas analíticas invitan a los evaluadores a tomar las funciones del lenguaje, tales como la gramática, lengua y estructura por separado (Alderson, Clapham and Wall, 1995)

Las escalas analíticas incorporan criterios evaluativos definidos por separado. En el caso de los instrumentos para evaluar la escritura académica, este tipo de escala suele incluir un cierto número de elementos tales como organización, contenido, cohesión, uso de lengua, resolución de la tarea (Weigle, 2002). Por esta característica, la escala analítica presenta mayor grado de confiabilidad. Cada uno de estos componentes integra un criterio de evaluación por lo que cada uno recibe una nota. Algunos elementos pueden ser integrados dentro de un mismo descriptor de criterio; por ejemplo: vocabulario y gramática pueden conformar el criterio uso del lenguaje. Según Cohen (1994) esto constituye una ventaja del uso de la escala analítica ya que al contar con componentes individuales es fácil entrenar diferentes evaluadores para que logren entender su uso. Otra ventaja de escala analítica es que le permite a los evaluadores identificar tanto las fortalezas como las debilidades de la producción de los alumnos (Carr, 2000). Knock (2007) postula que las escalas de evaluación analítica que presentan un detallado nivel de descriptores pueden ser



de gran utilidad en la evaluación del proceso escriturario ya que les brinda a los evaluadores una base explícita desde la cual asignar notas.

Dentro del marco del proyecto de investigación “Criterios de evaluación de la escritura en lengua extranjera (inglés) en el nivel superior: aplicación de instrumentos de evaluación”, presentamos los resultados de la aplicación de una escala analítica en nuestro contexto educativo, la materia Lengua Inglesa II, de las carreras de Profesorado, Traductorado y Licenciatura de la Facultad de Lenguas, Universidad Nacional de Córdoba.

Diseño y piloteo de la escala

En las últimas décadas, la validez de la evaluación de la escritura académica ha sido abordada desde diferentes ángulos. Sin embargo, Milanovic y otros (1996) subrayan la necesidad de prestar más atención al comportamiento del evaluador ya que sus correcciones suelen ser la principal causa de errores en las notas asignadas a los alumnos (Mei, 2010). Por lo tanto, Milanovic y otros (1996) abogan por lograr un mejor entendimiento del valor asignado a un trabajo, la toma de decisiones por parte del evaluador para asignar un puntaje y la naturaleza idiosincrática de las apreciaciones de los evaluadores. En este trabajo se hablará de las estrategias empleadas por un grupo de docentes de la Facultad de Lenguas, UNC, de la cátedra Lengua Inglesa II con el fin de mostrar las estrategias empleadas para minimizar la subjetividad del criterio de evaluación y de colocar notas.

En una primera instancia del proyecto de investigación se diseñaron instrumentos de evaluación para ser aplicados en dos momentos diferentes. En primer lugar, durante el año 2016, se llevó a cabo una sesión de entrenamiento para abordar la implementación de una escala de naturaleza holística que fue aplicada a cinco ensayos seleccionados por un evaluador externo. Cada docente de la cátedra debió evaluar los ensayos implementando dicha escala. Luego se llevó a cabo un análisis correlacional de los resultados. En una segunda instancia, durante el año 2017, se repitió el mismo procedimiento pero esta vez aplicando una escala analítica, para que de esta manera los resultados obtenidos en ambas instancias puedan ser comparables en términos de consistencia entre evaluadores.



Como punto de partida para la implementación de la escala analítica, la escala base seleccionada para la prueba piloto fue la escala analítica de Tribble (1996 en Kern 2000). Esta escala base, junto con otras, fue socializada entre las diferentes docentes de la cátedra en la cual se decide adoptar la escala propuesta por Tribble (1996). Su elección se basa en los resultados de una reunión docente en la cual se analizaron distintas escalas, las docentes realizaron diferentes comentarios sobre las distintas posibilidades, y se concluyó que ese modelo de escala podría llegar a ser el más útil para nuestro contexto educativo. Uno de los aspectos que fomentó la elección de esta escala fue que contaba con descriptores más detallados. Esto es de particular importancia ya que se pretendía adoptar este instrumento de evaluación en instancias de evaluación formativa también y, por lo tanto, una descripción más detallada favorecería la identificación de áreas problemáticas, lo cual permite orientar de manera más certera la práctica correspondiente de cada alumno. El nivel de detalle al mismo tiempo no deja espacio para la libre interpretación de los descriptores lo que permite generar un buen grado de confiabilidad.

Durante la instancia de socialización, se realizaron diferentes cambios en la escala modelo. En una primera instancia, se decidió agregar una columna en donde las docentes pudieran visualizar la nota que le correspondería a un alumno que haya obtenido un puntaje determinado para facilitar la asignación del puntaje del examen. Entre las docentes también surgió la necesidad de revisar el puntaje propuesto por la escala original ya que si un ensayo desarrollaba el tema de forma adecuada pero no organizaba su trabajo de forma apropiada lograba alcanzar una buena nota sin que su trabajo reflejara realmente un buen empleo de la segunda lengua. Segundo, se decidió incluir los descriptores de gramática y vocabulario como aspectos diferenciados. Tercero, en términos de organización del ensayo, se colocó un descriptor para tener en cuenta los conectores utilizados dentro de un párrafo y entre los párrafos.

Otras de las modificaciones propuestas durante el proceso de elaboración de las versiones preliminares y hasta llegar a la versión actual tuvieron que ver con cambios en el fraseo de los descriptores para adecuarlo al contexto específico en el que se utiliza, por ejemplo, en el criterio de contenido se cambió el término *preciso* por la palabra *apropiado*. En algunos casos, el criterio consensuado por los



docentes llevó a generar nuevos descriptores dentro de cada criterio, mientras que en algunos casos algunos descriptores fueron fraseados nuevamente o eliminados ya que no resultaban útiles para el propósito de evaluación de la cátedra. Por ejemplo, dentro del criterio de contenido, no hay nada que haga referencia a la información redundante o no relacionada con el tema de pregunta del ensayo. Dado nuestro contexto, varias docentes señalaron que la escala analítica debe tener un descriptor específico para conclusiones ya que los alumnos experimentan problemas en esta sección específica de la composición. Este descriptor se agregó dentro del criterio de organización. Al mismo tiempo manifestaron que, en el caso de haber ensayos que presenten falta de secciones como la conclusión o el título, no hay dónde consignarlo en la escala.

Además, se agregaron los puntajes que se utilizan normalmente para calificar los ensayos en la cátedra (tanto las calificaciones como los puntajes discretos). Fue necesario adaptar esta escala para que pueda ser aplicada en la evaluación de ensayos en lengua inglesa de carreras universitarias de profesorado, traductorado y licenciatura de inglés. Se realizaron cambios a nivel de criterios, nota y descriptores. En relación con el puntaje, en la escala modelo se le asigna 30 puntos a la sección de organización, lo que para algunos docentes resultaba excesivo, por lo que se reasignó el puntaje y se distribuyó más equitativamente entre criterio y contenido. Luego de la instancia de piloteo, los comentarios de los docentes indicaron la practicidad de la escala ya que su nivel de detalle contribuyó a su fácil aplicación, algo bueno que no siempre sucede con escalas analíticas.

Conclusiones

En este trabajo se realiza un resumen de los pasos dados por un grupo de docentes de Lengua Inglesa II a fin de optimizar un instrumento de evaluación de escritura de carácter analítico. Este instrumento tiene como uno de sus objetivos primordiales garantizar un mayor grado de confiabilidad entre los diferentes docentes evaluadores. Los comentarios realizados por los docentes de Lengua Inglesa II fueron los datos que se tuvieron en cuenta para la selección de la escala analítica, para determinar la utilidad de la misma al evaluar una instancia de examen parcial, y poder determinar así la validez, utilidad y confiabilidad de dicha escala. Los



comentarios previos a su implementación tuvieron como principal objetivo hacer de esta escala una herramienta práctica en la que los descriptores guiaran rápidamente la nota que se correspondiera con el desempeño del alumno. Luego, en la instancia posterior a la implementación de la escala, los docentes de la cátedra se manifestaron de forma positiva en relación al uso de la escala analítica. Entre las ventajas mencionadas se encuentran su practicidad y fácil uso. Estas escalas, por su nivel de detalle, resultan ser de gran utilidad tanto con fines evaluativos como diagnósticos. Al mismo tiempo, se advirtió que a diferentes ensayos se les puede otorgar la misma nota pero por diferentes razones.

El análisis detallado de la escala propuesta por Tribble y los cambios que se introdujeron con el fin de adaptarla a nuestro contexto de uso pone de manifiesto la estrecha relación entre el énfasis que reciben ciertos elementos a la hora de ser presentados y practicados en clase y su peso en establecer la nota dentro de cada criterio. El proceso evaluativo no puede estar divorciado de la práctica áulica.

Referencias bibliográficas

- Alderson, C J, Clapham, C and Wall, D (1995) *Language Test Construction and Evaluation*, Cambridge: Cambridge University Press.
- Black, P., y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bloom, B.S., J.T. Hastings and G.F. Madaus. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Co.
- Carr, N. (2000). A comparison of the effects of analytic and holistic rating scale types in the context of composition texts. *Issues in Applied Linguistics*, 11, 207-241.
- Cohen, A. (1994). *Assessing language ability in the classroom*. Boston, MA: Heinle & Heinle. The descriptive essay rubric. Retrieved June 2, 2012, from www.gnaedu.org. The descriptive writing rubric. Retrieved May 29, 2012, from msjohnsononline.weebly.com/uploads/.../1c_rubric_decriptive_writin.Pdf
- Collins, J., y O'Brien, N. (2003). *The Greenwood dictionary of education*. Westport, CT: Greenwood Press.
- Crooks, T. (2001). The validity of formative assessment. Paper presented to *The British Educational Research Association Annual Conference, University of Leeds*, 13-15 September, 2001. Recuperado de: <http://www.leeds.ac.uk/educol/documents/00001862.htm>.
- Fernández March, A. (2008). Evaluación de los aprendizajes en la universidad: nuevos enfoques. Recuperado el 20 de julio de 2017, de: <http://www.unizar.es/ice/rec-info/materiales-acciones.html>
- Hamp-Lyons, L. (1991). Scoring procedures for ESL contexts. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex.



- Knoch, U. (2007). Do empirically developed rating scales function differently to conventional rating scales for academic writing? *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, 5, 1-36.
- Kern, R. (2000). *Literacy and language teaching*. Oxford: Oxford University Press.
- Mei, W. (2010). Investigating raters' use of analytic descriptors in assessing writing. *Reflections on English Language Teaching*, 9 (2), 69-104
- O'Malley, J y Pierce, L. (1996). *Authentic Assessment for English Language Learners: Practical Approaches for Instructions*. Boston: Addison-Wesley Publishing Company.
- Shepard, L.A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 66-70.
- Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. Chicago: McGraw-Hill.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.