

## Capítulo 6

### Evaluación de la escritura en lengua extranjera (inglés): Análisis de criterios docentes

María Elisa Romano

[romano.mariaelisa@gmail.com](mailto:romano.mariaelisa@gmail.com)

Julia Inés Martínez

[juliamartinez@gmail.com](mailto:juliamartinez@gmail.com)

Facultad de Lenguas, Universidad Nacional de Córdoba

Córdoba, Argentina

#### Resumen

La evaluación de la escritura en lengua extranjera constituye 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. Las investigaciones sobre esta temática recalcan la multiplicidad de factores que intervienen en el proceso de evaluación y calificación de la escritura en lengua extranjera, entre los que se destaca la importancia de la conceptualización que tienen los docentes de la habilidad escrituraria en la lengua extranjera.

Los estudios sobre procesos que llevan a cabo los docentes lectores al momento de evaluar y calificar los textos producidos por los estudiantes han demostrado, en general, que una preocupación de los docentes al momento de evaluar y calificar 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, se han observado diferencias entre evaluadores en lo que respecta no sólo a los modos de calificar, sino también a las formas en las que conciben y abordan la evaluación. Teniendo en cuenta estos resultados, los cuales dan cuenta de la complejidad inherente a la evaluación de la escritura, el objetivo de este trabajo es indagar sobre los criterios utilizados por docentes de grado para evaluar ensayos académicos en lengua extranjera (inglés) y la relación entre dichos criterios y los aspectos textuales que ponderan en el proceso de calificación. Para lograr este objetivo, se analizarán reportes docentes sobre ensayos aprobados y desaprobados en una misma instancia de evaluación. Dicho análisis implicará la comparación y el contraste entre los reportes de los docentes, así como también la relación entre los criterios reportados y las calificaciones establecidas en cada caso.

#### Introducción

Debido al rol fundamental que tienen los procesos y las prácticas de evaluación en el ámbito académico, su impacto en las prácticas áulicas es una problemática susceptible de ser analizada y explorada en profundidad, ya que, en muchos casos, son las características de los procesos de evaluación las que determinan las prácticas pedagógicas concretas. Según Crusan (2010b), “lo que no aparece en las evaluaciones tiende a desaparecer de las aulas con el tiempo” (247). La importancia de la evaluación es una realidad que docentes e investigadores no podemos desconocer y sobre la cual debemos reflexionar de manera sistemática.

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 fue 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 indagar sobre los criterios utilizados por docentes de grado para evaluar ensayos académicos en lengua extranjera (inglés) y la relación entre dichos criterios y los aspectos textuales que ponderan en el proceso de calificación. Para lograr este objetivo, se analizarán reportes docentes sobre ensayos aprobados y desaprobados en una misma instancia de evaluación sumativa. Dicho análisis implicará la comparación y el contraste entre los reportes de los distintos docentes, así como también la relación entre los criterios reportados y las calificaciones establecidas en cada caso.

#### Marco teórico y antecedentes

La evaluación de la escritura objeto de la presente investigación (evaluación de ensayos en exámenes parciales y finales) puede considerarse, siguiendo a Goodman & Swann (2003) y Hyland (2003), como una *evaluación sumativa*, dado que se 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. Este tipo de evaluación es identificada, además, como una *evaluación directa* de la escritura; es decir, se evalúa la escritura a través una producción textual individual y no a través de ejercicios de componentes lingüísticos que hacen a la escritura (como, por ejemplo, ejercicios de vocabulario, gramática, etc.), lo cual se define como *evaluación indirecta* de la escritura (Williams, 2005: 120). La *evaluación directa* de muestras concretas de 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), 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: 120-121).

Dos términos de uso frecuente en el área de evaluación (no sólo de evaluación en escritura) son confiabilidad (*reliability*) y validez (*validity*). El primero de estos conceptos hace referencia a la medida en la cual una determinada respuesta (un texto, en el caso de la escritura) podrá ser evaluada siempre de la misma manera y alcanzando el mismo resultado. Este punto es muy importante y especialmente controversial en el caso de la evaluación de la escritura, ya que implica que, ante un mismo ensayo o párrafo, dos evaluadores diferentes deberían arribar al mismo resultado. Una tarea de escritura puede considerarse confiable si mide lo que intenta medir de manera consistente, tanto en lo que respecta a los estudiantes (el mismo estudiante en diferentes ocasiones) como en lo que respecta a los evaluadores (la misma tarea a la vista de distintos evaluadores) (Hyland, 2003: 215).

El segundo concepto implica, a grandes rasgos, que la prueba debe evaluar lo que se pretende evaluar, y no otras habilidades y/o contenidos (Williams, 2005: 121). Hamp-Lyons (1990) identifica cuatro tipos de validez: validez aparente (*face validity*), validez de contenido (*content validity*), validez de criterio (*criterion validity*) y validez de constructo (*construct validity*), pero sostiene que la validez de constructo, que implica que la evaluación refleja la realidad psicológica del comportamiento en el área que se evalúa, es un tipo de validez que engloba a todas las demás. Más recientemente, la autora (Hamp-Lyons, 2003) ha propuesto una nueva visión de la validez que es más integradora y que incluye la validez de contenido, el control de la calidad, la validez de constructo, la utilidad de la evaluación, la capacidad de la evaluación de tener significado y relevancia en más de un contexto (generalización de validez) y la validez consecucional (*consequential validity*). Messick (1996) destaca la importancia de que las evaluaciones tengan “representación de constructo” (*construct representation*) y “relevancia de constructo” (*construct relevance*). La primera característica se refiere al hecho de que la evaluación debería ser tan similar a la habilidad real como sea posible, mientras que la segunda se refiere al hecho de que la evaluación debe evaluar sólo las habilidades que son parte del constructo (no debería haber dificultades en la evaluación que no sean parte de la habilidad/destreza misma).

En lo que respecta a abordajes de evaluación de un texto, una distinción básica es aquella entre el *enfoque holístico* y el *enfoque analítico*. La *evaluación holística* aborda al texto como un todo y conlleva 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 sola escala 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* implica la identificación de distintos aspectos del texto, de modo de determinar las fortalezas y debilidades de cada uno de los aspectos, para luego establecer 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 les permite a los evaluados analizar las fortalezas y debilidades en sus propios textos, y de esta manera, 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. Asimismo, sostiene 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. Sin embargo, algunos investigadores afirman que la evaluación analítica detallada a través de una lista de criterios requiere más tiempo, esfuerzo y entrenamiento, sin evidenciar mayor efectividad que la evaluación holística en la calificación de textos escritos (Rezaei & Lovorn, 2010; Robb Singer & Lemahieu, 2011; Freedman, 1981; Hunter, Jones & Randhawa, 1996).

Si bien se han llevado a cabo numerosos estudios sobre evaluación, aquellos que resultan de especial interés para esta investigación son los que han analizado la relativa efectividad de diferentes modos de calificación (Cushing Weigle, 2002; White, 1998; Wolcott & Legg, 1998) y los que han explorado los procesos que llevan a cabo docentes lectores al momento de evaluar y calificar los textos producidos por los estudiantes. 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 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 (Baker, 2014). El 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.

Por otro lado, se han observado diferencias importantes entre evaluadores en lo que respecta no sólo a los modos de calificar, sino también a las formas en las que conciben y abordan la evaluación de la escritura (Ferris, 2014; Hall & Sheyholislami, 2013). Entre los factores sobresalientes que suelen determinar las diferencias entre evaluadores se encuentran aspectos relacionados con la personalidad (Alaei, Ahamadi & Zadeh, 2014; Brown, 1991; Carrell, 1995) y con factores contextuales, tales como el tiempo disponible y las herramientas que se utilizan para evaluar (Baker, 2014). El estudio realizado por Edgington (2005), por ejemplo, da cuenta de la influencia que ejercen esos y otros factores (tales como la relación docente-alumno y las experiencias áulicas) en las estrategias que utilizan los docentes para leer y evaluar los textos de los estudiantes.

La experiencia en docencia y evaluación también determina diferencias en los resultados y el proceso de calificación de ensayos (Wolfe, 2003). Otros estudios han demostrado, sin embargo, que estos factores pueden ser controlados en gran medida a través de un entrenamiento sistemático de los evaluadores (Crusan, 2010a; Cushing Weigle, 2002; Fahim & Bijani, 2011; Panou, 2013; Trace, Janssen & Meier, 2015) y una permanente revisión de los instrumentos que se utilizan para evaluar, ya sean holísticos o analíticos, ya que las categorías que se definen y establecen en escalas o guías de evaluación pueden ir transformándose a partir de las prácticas individuales (Mei, 2010).

La evaluación de la escritura es parte integral de la asignatura *Lengua Inglesa II* de las carreras de Profesorado, Traductorado y Licenciatura de la Facultad de Lenguas, UNC, y se lleva a cabo de diversas maneras y en diferentes instancias. Durante el ciclo

lectivo, los docentes corrigen ensayos escritos por los alumnos como parte de las tareas para realizar fuera del aula. Si bien estas evaluaciones son informales, en el sentido de que no llevan una nota ni determinan la aprobación del curso, sirven a los alumnos como instancias de práctica previas a las evaluaciones formales. En cuanto a este último tipo de evaluación, el alumno regular de *Lengua Inglesa II* debe aprobar tres evaluaciones parciales, dos de las cuales incluyen la escritura de un ensayo. Por otra parte, el examen final escrito de la materia consta de una sección de escritura de ensayo eliminatoria y una sección de escritura adicional para alumnos libres que también es eliminatoria. Como podemos observar, la evaluación de la escritura se constituye en un elemento casi fundamental de nuestra asignatura y, como tal, requiere un alto índice de confiabilidad y validez.

Teniendo en cuenta los estudios previos mencionados y el contexto educativo particular en el que nos encontramos, nos planteamos el objetivo de 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 (exámenes parciales y finales). En este trabajo en particular, nos proponemos indagar sobre los criterios utilizados por docentes de grado para evaluar ensayos académicos en lengua extranjera (inglés) y la relación entre dichos criterios y los aspectos textuales que ponderan en el proceso de calificación.

## Metodología

### Sujetos y materiales

Participaron en esta investigación todos los docentes de la cátedra *Lengua Inglesa II* que están a cargo de la evaluación de ensayos en instancias de evaluación sumativa (profesores asistentes, adjuntos y titulares): siete docentes en total. Además de los textos evaluados, se recolectaron reportes escritos por los docentes (ver Apéndice), explicando y justificando la evaluación y calificación de determinados ensayos (un ensayo aprobado y un ensayo desaprobado) 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”*.<sup>4</sup>

---

<sup>4</sup> La universidad local ha llevado a cabo sesiones de debate sobre oportunidades de empleo y discriminación en el lugar de trabajo. Como parte de una campaña de concientización, se le pide que escriba un ensayo sobre las similitudes y diferencias entre trabajo calificado y no calificado con relación a las oportunidades de empleo, las condiciones de trabajo o las políticas sobre discriminación de género. Escriba un ensayo académico bien desarrollado de aproximadamente 350 palabras.

## Procedimiento

Cada docente seleccionó al azar 2 (dos) de los ensayos evaluados con calificaciones diferentes (un ensayo reprobado y un ensayo aprobado) luego de la primera evaluación parcial y completó un reporte escrito explicando y justificando la evaluación y calificación del ensayo. Previamente, se realizó una reunión entre los docentes a fin de familiarizarlos con las características y los requerimientos de este tipo de reporte. El reporte tiene una extensión de entre una o dos carillas y describe en detalle las razones por las cuales el docente considera que la calificación propuesta es la adecuada. 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 trataría de un reporte extendido en el tiempo, sino de un reporte focalizado en una tarea de evaluación específica (Hatch, 2002).

## Análisis de datos

Cada docente de la cátedra que participó de la investigación debió completar un reporte en el cual justificara y explicara la calificación de los ensayos corregidos. Posteriormente, se analizaron los resultados y se contabilizaron y clasificaron las marcas, correcciones y comentarios, para luego compararlos y contrastarlos con el fin de determinar cuáles son los aspectos acerca de los cuales existe acuerdo y desacuerdo entre los docentes evaluadores.

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.

## Resultados

En total se analizaron 14 reportes escritos por los docentes de la cátedra. En los reportes que los docentes tenían que completar debían identificar las marcas y los comentarios realizados en los ensayos seleccionados, contabilizando los casos de precisión, contenido, y organización, todas categorías de importancia a la hora de corregir ensayos académicos. Además de contabilizar los comentarios y marcas, los docentes debían proporcionar un ejemplo que los ilustrara y, finalmente, clasificar las categorías según el orden de importancia que le atribuyeran.

En lo que respecta a los aspectos textuales que ponderan los docentes al momento de evaluar, la mayoría de los reportes (10 reportes, 71%) ubican en primer lugar la organización de los ensayos, es decir, la forma en la que se estructura el texto según las convenciones del género. En el caso de los ensayos que fueron evaluados en esta tarea (de comparación y contraste), se estudiaron tres maneras de organizar estos textos, generalmente conocidas como *block method*, *point by point method* e *item by item method*. En segundo lugar, estos docentes ubicaron el contenido del texto. Este aspecto hace referencia a las ideas principales y secundarias del texto, la relación lógica entre ellas y su pertinencia en relación con la consigna. En tercer y último lugar de esta lista

de prioridades se ubica la precisión lingüística, es decir, el uso correcto de la lengua, y la complejidad y adecuación de las estructuras gramaticales y del léxico.

Un porcentaje menor de reportes presenta una escala de prioridades diferente: para algunos docentes (4 reportes, 29%), lo más importante al momento de evaluar un texto es el contenido del mismo, seguido de la precisión lingüística o el uso de la lengua y, finalmente, la organización textual:

Total de reportes	Total de comentarios/ marcas (precisión + contenido + organización)	Aspectos importantes a tener en cuenta a la hora de corregir un texto académico
14	286: -205 (72%) sobre precisión -60 (21%) sobre contenido -21 (7%) sobre organización	10 reportes (71%): organización, contenido, precisión 4 reportes (29%): contenido, precisión, organización

Tabla 1

El análisis de los reportes y los ensayos corregidos por cada docente nos indica que la ponderación de aspectos del texto no varía entre textos aprobados y desaprobados para un mismo docente. La jerarquía que orienta la evaluación parece mantenerse estable, independientemente de la calidad del texto que se evalúa.

La cantidad de comentarios, sin embargo, varía entre los textos aprobados y desaprobados, presentando estos últimos una mayor cantidad de comentarios en todas las categorías:

	Ensayos aprobados	Ensayos desaprobados
<b>Precisión</b>	79	126
<b>Contenido</b>	11	49
<b>Organización</b>	7	14

Tabla 2

Con respecto a las fortalezas y debilidades de los textos evaluados, entre las fortalezas más frecuentemente mencionadas por los docentes en relación con los ensayos aprobados se encuentra la organización del texto; por ejemplo: *“Sigue el pattern requerido (...), tiene una Thesis Statement clara y bien definida y todos los párrafos del cuerpo del ensayo tienen una buena Topic Sentence”* (D1). Algunos docentes mencionan también fortalezas relacionadas con el contenido (ideas, enfoque del tema, etc.), tales como *“El ensayo comienza con un hook histórico muy bueno, que se relaciona muy bien con la Thesis Statement. Las ideas del autor son muy interesantes”* (D2). El aspecto menos mencionado es la precisión lingüística, que, de todos modos, es mencionada por algunos docentes, quienes resaltan, especialmente, el uso de vocabulario específico como un aspecto destacable. Las debilidades mencionadas en los casos en los que los

textos superan el puntaje mínimo para aprobar apuntan, en general, a falencias en la precisión lingüística: *“El ensayo tiene algunos errores gramaticales”* (D6). También se mencionan algunos aspectos relacionados con el contenido y, finalmente, el incumplimiento de aspectos formales tales como márgenes, extensión, etc.

Fortalezas de los ensayos aprobados	Debilidades de los ensayos aprobados
-organización	-problemas relacionados con la precisión (incluidos errores elementales para el nivel)
-enfoque en el tema asignado (contenido)	-falta de ideas críticas y sólidas (contenido)
-ideas interesantes y específicas al tema (contenido)	-incumplimiento de los aspectos formales (título, márgenes, cantidad de palabras, etc.)
-uso apropiado de vocabulario específico al tema (precisión)	-----

Tabla 3

En cuanto a los ensayos desaprobados, los reportes indican que las fortalezas observadas en estos textos tienen que ver, por lo general, con la organización textual; por ejemplo, *“Está bien organizado”* (D7). En algunos casos, se mencionan la expresión de ideas y el uso de vocabulario específico como aspectos positivos del ensayo. Por ejemplo: *“El ensayo aborda el tema asignado”* (D5) y *“El estudiante utiliza algunas expresiones específicas relacionadas con el tema”* (D2). Las debilidades mencionadas también apuntan a la organización del texto, como ilustran los siguientes ejemplos: *“Algunas ideas no están organizadas lógicamente”* (D6); *“Los párrafos del cuerpo del ensayo no tienen Topic Sentences apropiadas y la conclusión no es un párrafo sino una oración”* (D1). Sin embargo, también incluyen referencias frecuentes a problemas relacionados con el contenido de los textos: *“El ensayo no tiene una Thesis Statement bien enfocada; por lo tanto, la mayoría de las ideas incluidas en el ensayo no se relacionan con el tema asignado”* (D3), o *“El estudiante no respondió la pregunta”* (D4). La organización del texto aparece, por lo tanto, como la fortaleza mencionada con más frecuencia, tanto en ensayos aprobados como en ensayos desaprobados. En lo que respecta a las debilidades, aquí se observa una diferencia entre los ensayos que superan el puntaje mínimo y aquellos que no, ya que en el caso de los primeros, las debilidades mencionadas con más frecuencia tienen que ver con la precisión lingüística o el uso del idioma, mientras que en el caso de los segundos, la categoría que más se menciona es la organización, seguida de cuestiones relacionadas con el contenido (relevancia de las ideas, enfoque adecuado del tema dado, desarrollo apropiado del tema dado, etc.). Este resultado nos presenta nuevamente una preocupación de la mayoría de los evaluadores por la organización del texto, reforzando la escala de prioridades reportada en la pregunta B e, indirectamente, reflejando coherencia en la aplicación de dicha escala, ya que, al tratarse del aspecto considerado como primordial, es de esperar que aparezca como una debilidad en aquellos ensayos que se consideran reprobados.

<b>Fortalezas de los ensayos desaprobados</b>	<b>Debilidades de los ensayos desaprobados</b>
-organización	-falta de oraciones tópicas (organización)
-expresión de ideas (precisión)	-problemas de precisión en el uso de la lengua (incluidos errores elementales para el nivel)
-uso de expresiones o vocabulario específico (precisión)	-omisión de uno de los temas que debía tratarse (contenido)
-----	-falta de un análisis profundo del tema (contenido)
-----	-irrelevancia de algunas ideas (contenido)
-----	-no desarrollo del tema dado (contenido)

**Tabla 4**

En todos los reportes analizados, las fortalezas y debilidades expresadas por los docentes coinciden con la jerarquía de los aspectos considerados importantes a la hora de evaluar un texto académico; es decir, al referirse a los aspectos positivos y/o negativos de los textos evaluados, aquellos docentes que dicen priorizar la organización mencionan en primer lugar fortalezas o debilidades relacionadas con esa categoría, refiriéndose luego a aspectos de contenido o uso del idioma. De igual manera, aquellos docentes que reportaron priorizar el contenido por sobre los demás aspectos, también hacen referencia a este aspecto cuando se refieren a las fortalezas y debilidades de los textos evaluados. Esta tendencia parece reflejar una coherencia interna entre una escala de prioridades de aspectos a considerar y la mirada que se hace del texto. Esto, sin embargo, debe ser analizado junto con los resultados de otras fuentes de datos, tales como entrevistas y ensayos corregidos por estos docentes, con el fin de comprobar esta correspondencia entre criterios y prácticas.

## Conclusiones

El análisis de los datos obtenidos a partir de la exploración nos presenta un panorama bastante consistente en lo que respecta a los criterios de evaluación aplicados por docentes en la cátedra objeto de estudio.

Si bien en la mayoría de los reportes se indicó que la precisión lingüística es el aspecto menos importante a tener en cuenta, la presencia de una mayor cantidad de comentarios de este tipo puede estar relacionada tanto con el contexto educativo en el que se lleva a cabo la evaluación (formación de futuros profesionales de una lengua extranjera) como con la naturaleza de la asignatura *Lengua Inglesa II*. Si tenemos estos aspectos en cuenta, podemos observar que la precisión gramatical, el uso de vocabulario adecuado y pertinente y la estructura oracional de cierta complejidad constituyen no sólo rasgos de la escritura sino también objetos de estudio en sí mismos, por lo que los docentes, al notar un error o imprecisión, están fuertemente orientados a marcarlo (de manera que el estudiante no internalice esos errores en el uso del lenguaje). Esta cantidad de comentarios sobre precisión lingüística refleja también una forma de evaluar y una

toma de decisiones por parte de los docentes; se podría acordar responder solamente a un número mínimo de errores de ese tipo (observando, por ejemplo, los errores más serios o recurrentes), y hacer más visible la ponderación de los aspectos textuales a través de la inclusión más frecuente de comentarios explícitos sobre contenido y organización.

La existencia de una coherencia entre criterios y calificación sugiere que los criterios que aplican los docentes al momento de evaluar no parecen variar según el texto; en general, las correcciones presentes en los ensayos reflejan la ponderación expresada en los reportes, lo que nos indica, a su vez, que los criterios establecidos son de orden interno y que no parecen variar a partir de factores externos, tales como la calidad del texto.

Las mayores diferencias entre evaluadores surgen con relación a la categoría organización: se deduce de los reportes y los comentarios en ensayos que no todas las docentes contemplan los mismos aspectos. Es también una categoría que no es ponderada de igual manera: en la mayoría de los reportes aparece como uno de los primeros aspectos que se tiene en cuenta, mientras que en algunos se ubica en último lugar. Una hipótesis que planteamos en relación con este punto es que el concepto de organización textual no es concebido de la misma manera por todas las docentes, apareciendo de manera difusa cuestiones relacionadas con la macro-estructura textual y las características del género, y aspectos de coherencia y cohesión interna del texto. A pesar de que en las reuniones de cátedra y antes de cada evaluación se fijan y reiteran los criterios a seguir, es evidente que será necesario apelar a otros recursos a nivel cátedra para garantizar la igualdad de criterios a la hora de corregir textos académicos.

Finalmente, la presencia de más comentarios y/o marcas en ensayos desaprobados puede deberse a la necesidad de justificar el aplazo, pero también a una tradición o práctica general en la evaluación de la escritura que plantea la necesidad de escribir sobre el error o problemas en el texto, más que sobre las fortalezas o los aspectos positivos del mismo.

## Referencias bibliográficas

- Alaei, M.; Ahamadi, M. & Zadeh, N. (2014). The impact of rater's personality traits on holistic and analytic scores: Does genre make any difference too? *Procedia - Social and Behavioral Sciences*, 98 1240 – 1248.
- Baker, N. (2014). "Get it off my stack": Teacher's tools for grading papers. *Assessing Writing*, 19, 36-50.
- Brown, I. D. (1991). Do English and ESL learners faculties rate writing samples differently? *TESOL Quarterly*, 25(4), 587-603.
- Carrell, P. (1995). The effect of writers' personalities and raters' personalities on the holistic evaluation of writing. *Assessing Writing*, 2 (2), 153-190.
- Crusan, D. (2010a). *Assessment in the second language writing classroom*. Michigan: Ann Arbor.
- Crusan, D. (2010b). Assess thyself least others assess thee. En Silva, T. & Matsuda P. (Eds.), *Practicing theory in second language writing* (pp. 245-262). West Lafayette, Indiana: Parlor Press.

Cushing Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

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.

Fahim, M. & Bijani, H. (2011). The effect of rater training on rates' severity and bias in second language writing assessment. *Iranian Journal of Language Testing*, 1(1), 1-16.

Ferris, D. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, 19, 6-23.

Freedman, S. W. (1981). Influences on evaluators of expository essays: Beyond the text. *Research in the Teaching of English*, 13 (3), 245-255.

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. (1990). Second language writing: assessment issues. En Kroll, B. (Ed.) *Second language writing. Research insights for the classroom*. 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. (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. & 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.

Hunter, D. M., Jones, R., & Randhawa, B.S. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *The Canadian Journal of Program Evaluation*, 11 (2), 61-85.

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.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.

Rezaei, A. & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15, 18-39.

Robb Singer, N. & Lemahieu, P. (2011). The effect of scoring order on the independence of holistic and analytic scores. *Journal of Writing Assessment*, 4 (1), 1-13.

Trace, J., Janssen, G. & Meier, V. (2015). Measuring the impact of rater negotiation in writing performance assessment. *Language Testing*, 25, 465-493.

White, E. M. (1998). *Teaching and assessing writing: Recent advances in understanding, evaluating and improving student performance*. USA: Calendar Islands Publisher.

Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. Chicago: McGraw-Hill.

Wolcott, W. & Legg, S. (1998). *An overview of writing assessment: Theory, research and practice*. Urbana, IL: National Council of Teachers of English.

Wolfe, E. W. (2003). Uncovering rater's cognitive processing and focus using think-aloud protocols. *Journal of Writing Assessment*, 2 (1), 37-56.

## Apéndice

Fecha: \_\_\_\_\_

REPORTE SOBRE EL ENSAYO N° \_\_\_\_\_

EVALUADOR: \_\_\_\_\_

PUNTAJE FINAL: \_\_\_\_\_

A. Identifique los comentarios/marcas que hizo en el ensayo y responda las siguientes preguntas:

1. ¿Cuántos comentarios/marcas se relacionan con la precisión (uso de la lengua)? \_\_\_\_\_

Escriba un ejemplo de este tipo de comentario: \_\_\_\_\_

2. ¿Cuántos comentarios/marcas se relacionan con el contenido (ideas, desarrollo de ideas, enfoque de la información)?

Escriba un ejemplo de este tipo de comentario: \_\_\_\_\_

3. ¿Cuántos comentarios/marcas se relacionan con la organización (*thesis statement*, estructura de párrafos, estructura general del texto)?

Escriba un ejemplo de este tipo de comentario: \_\_\_\_\_

B. De los tres aspectos mencionados anteriormente (precisión, contenido, organización), ¿cuál es, en su opinión, el aspecto más importante a tener en cuenta a la hora de evaluar un texto académico? Enumere los tres aspectos en orden de importancia (1 para el más importante y 3 para el menos importante):

1. \_\_\_\_\_ 2. \_\_\_\_\_ 3. \_\_\_\_\_

C. Describa de manera breve (no más de cuatro líneas) cuáles considera que son las fortalezas del ensayo que está analizando:

---

---

---

---

D. Describa de manera breve (no más de cuatro líneas) cuáles considera que son las debilidades del ensayo que está analizando:

---

---

---

---

E. ¿Qué espera que los estudiantes hagan con la retroalimentación brindada sobre este ensayo?

---

---

---

---

F. ¿Qué cree que los estudiantes hacen realmente con los ensayos que usted corrige?

---

---

---

---