

La evaluación del aprendizaje: La importancia de unificar criterios de corrección y calificación en el área de gramática inglesa

Autores: Fabián Negrelli, Cecilia Ferreras, Natalia Dalla Costa y María José Morchio.

Institución: Facultad de Lenguas - Universidad Nacional de Córdoba - Argentina

Palabras clave: evaluación - aprendizaje - criterios - calificación - gramática inglesa

En este trabajo presentaremos los resultados de un proyecto de investigación bianual 2008-2009 y que fuera avalado por SECyT UNC. En dicho trabajo abordamos la elaboración de criterios de evaluación estandarizados con el objetivo de evitar la disparidad de criterios existentes y asegurar fiabilidad en las calificaciones de los exámenes de la asignatura *Gramática Inglesa I*, dictada en el segundo año de las carreras de Profesorado, Traductorado y Licenciatura en Inglés en la Facultad de Lenguas de la UNC. Los sujetos estudiados fueron, por un lado, un grupo de 60 alumnos regulares que cursaron la asignatura durante el ciclo lectivo 2008 y, por otro, todos los profesores a cargo del dictado, desarrollo y evaluación de los contenidos teórico-prácticos de la asignatura. A tales efectos, se recolectó la muestra, que consistía en 60 exámenes finales escritos elegidos en forma aleatoria, los cuales fueron corregidos individualmente por cada uno de los docentes. De esta manera, se pudo verificar que en algunos casos no existía uniformidad de criterios de evaluación y, por ende, la calificación dada a cada uno de los alumnos, variaba según el sujeto-calificador. Seguidamente, se procedió a analizar, describir e interpretar los distintos tipos de errores detectados en la muestra a los fines de ordenarlos jerárquicamente según su relevancia para elaborar, de este modo, bandas consensuadas de suficiencia con respecto a la habilidad gramatical requerida por parte de los alumnos para acceder a la promoción de la asignatura. Los resultados de este trabajo de investigación condujeron a implicancias relevantes para mejorar la calidad del proceso de calificación de las evaluaciones de los alumnos, asentando dichas calificaciones sobre un constructo informado de fiabilidad.

Introducción

Todo proceso de evaluación académica adquiere sentido en tanto facilite la identificación de áreas, procesos, decisiones, etc. que necesitan modificarse para mejorar la calidad de los resultados obtenidos. Ello será posible en la medida en que se avance, a partir de indicadores de rendimiento globales, como en el caso de la enseñanza universitaria, hacia el conocimiento más específico y en detalle del encadenamiento de situaciones e instancias del proceso enseñanza-aprendizaje que producen esos resultados (Ramsden, 1998; Apodaca & Lobato, 1997; Villar Angulo y Alegre de la Rosa, 2004).

Evaluar un aprendizaje es, pues, una acción encaminada a estimar, apreciar o juzgar el valor o mérito que tiene el cambio en el conocimiento, capacidades o actitudes de los estudiantes. Cuando se aplica la evaluación a la enseñanza universitaria se amplía el campo de ideas, términos y significados relacionados y derivados de la evaluación. Así también se habla de. (a) *medir*, como la asignación de un número a un objeto, según una regla aceptable, o (b) de *codificar*, como la atribución de un valor a una actuación en una prueba. La medición en la enseñanza es la comparación de una tarea de aprendizaje (tipos de percepciones, comprensiones, conocimientos declarativos y procedimentales, o capacidades de respuesta que un estudiante debe poseer para tener éxito en un aprendizaje) con su respectiva unidad (tal como las puntuaciones de una prueba educativa), con el fin de averiguar cuántas veces la segunda (unidad) está contenida en la primera (experiencia). El concepto *medición* se refiere a un amplio rango de tareas de aprendizaje (destrezas y competencias específicas) que tienen distintas valoraciones para los profesores, incluso de una misma asignatura o cátedra. En consecuencia, la medición requiere un análisis sistemático (ampliación de conceptos de validez y fiabilidad) y una reflexión crítica acerca del rasgo, habilidad o tarea que está midiendo el ítem de una prueba. Desde esta perspectiva, una discusión a los postulados previos a la evaluación del aprendizaje tiene como objeto analizar: (i) Los principios éticos de funcionamiento de la evaluación, entre otros, la *igualdad moral* de los estudiantes evaluados en una materia, la *autonomía moral* de los agentes (profesores y estudiantes) implicados en una asignatura, la *imparcialidad* en la representación de los intereses de los estudiantes, la reciprocidad en el trato dispensado a los estudiantes en las revisiones de los exámenes (House, 1980), y (ii) la presentación lógica de las cuestiones de un test, prueba o examen.

Es en este sentido que creímos que el estudio de algunos aspectos relacionados con los procesos de evaluación en el área de gramática inglesa en la Facultad de Lenguas de la Universidad Nacional de Córdoba (UNC) contribuiría a tomar

decisiones para superar la problemática de la falta de uniformidad de criterios de evaluación.

La fiabilidad y la validez son dos cualidades esenciales que deben tenerse todos los exámenes de carácter científico y que deciden, como en el caso que nos convoca, la promoción al curso siguiente. Por definición, podemos decir que una prueba es fiable, cuando es estable, equivalente o muestra consistencia interna. Una prueba alcanza un elevado coeficiente de fiabilidad si los errores de medida quedan reducidos al mínimo. Las pruebas se consideran fiables cuando, midan lo que midan, proporcionan puntuaciones comparables cuando se repite su aplicación, o se compara con otra equivalente. La fiabilidad debe entenderse como el término que describe la consistencia que existe entre las medidas, la ausencia de error. Se dice que un test es fiable cuando mide con la misma precisión, da los mismos resultados, en sucesivas aplicaciones realizadas en situaciones similares, independientemente del sujeto calificador. Podemos decir que a más fiabilidad, más estables y consistentes son los resultados de las pruebas entre una aplicación y otra. En cuanto a la validez de un examen, podemos decir que ésta está garantizada cuando una prueba está concebida, elaborada y aplicada de tal manera que mida lo que se propone medir. En este contexto, el presente trabajo de investigación ha intentado encauzar los intereses e inquietudes - en cuanto a lograr fiabilidad y validez en las evaluaciones que suministren a sus alumnos - de un grupo de profesores que desarrollan su actividad docente en el área de gramática inglesa de las carreras de Profesorado, Traductorado y Licenciatura en Inglés de la Facultad de Lenguas de la UNC; por otro lado, también han formado parte de este proyecto docentes que integraron el equipo de gestión de nuestra Facultad (durante el periodo junio 2005- junio 2008) - tal es el caso de la Vicedecana, de la Secretaria Académica (Asesora de este proyecto) y de la Pro-secretaria Académica - quienes desde su profundo conocimiento de la realidad institucional y desde su formación académica y experiencia en el campo de la investigación educativa, principalmente en el área de evaluación, han realizado un invaluable aporte a este proyecto. La inquietud ha estado, entonces, orientada a desarrollar actividades de investigación que condujeran a establecer criterios de evaluación y calificación unificados que ayudaran, por una parte, a los profesores evaluadores a lograr coincidencia interpretativa en la calificación y, por otra, a los alumnos a creer en la fiabilidad y validez de la calificación obtenida. En este punto, debemos remarcar que los criterios de evaluación constituyen el referente fundamental para determinar el grado de consecución de los objetivos generales de la asignatura que hayan

alcanzado los alumnos. Son indicadores para la evaluación de los aprendizajes de los alumnos en los diferentes niveles de concreción curricular. Dichos criterios deben ser considerados puntos de referencia que hacen posible la calificación de lo que nos proponemos evaluar, en otras palabras, deben ser referentes de valor argumentados que nos ayuden a conocer en qué medida un sujeto alcanza el dominio de cada área.

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, 1990; Gass, 1994; Alderson, Clapham & Wall, 1995; Bachman & Palmer, 1996; Bachman & Cohen, 1998; Bachman, 2002; Hughes, 2003; entre otros); 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 central, 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. Como resultado, 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, así, debe volver sobre la etapa anterior para revisar, corregir, aquellas competencias y/o conocimientos que aún no han sido adquiridos y que forman parte de la habilidad gramatical necesaria -entendida ésta como “la combinación de conocimiento gramatical y competencia estratégica” (Purpura, 2004:86) 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 lugar central que ocupan los momentos de evaluación en relación con la carrera académica de los alumnos justificaba un análisis profundo de la situación y un estudio descriptivo que permitiera garantizar objetividad por parte de los evaluadores a la hora de evaluar y/o calificar lo que denominaremos genéricamente la “prueba” (remitiéndonos más a una categoría descriptiva que a algo conceptualmente definido), y que, al mismo tiempo, garantizara imparcialidad en la representación de los intereses de los alumnos.

Objetivos de la investigación

El objetivo general de nuestro proyecto fue abordar la elaboración de criterios de evaluación estandarizados con el fin de evitar la disparidad de criterios existentes y asegurar validez y fiabilidad en la calificación de los exámenes del área gramática inglesa. Entre los objetivos específicos que se plantearon al comenzar el desarrollo de este proyecto, podemos citar los siguientes:

- (i) Describir y analizar los objetivos de las carreras de Profesorado, Traductorado y Licenciatura en Inglés de la Facultad de Lenguas de la UNC y el perfil del egresado en cada una de las carreras mencionadas, según consta en el plan de Estudios vigente;
- (ii) Describir y analizar los objetivos, contenidos programáticos mínimos, metodología del proceso enseñanza-aprendizaje y formas de evaluación de la asignatura Gramática Inglesa I, tal cual se explicitan en el programa vigente;
- (iii) Establecer el nivel de habilidad gramatical de la lengua inglesa que se espera de los alumnos al finalizar el cursado de la asignatura *Gramática Inglesa I*;
- (iv) Verificar la presencia o ausencia de uniformidad de criterios de evaluación entre los profesores que califican la habilidad gramatical adquirida por los alumnos en la asignatura *Gramática Inglesa I*;
- (v) Analizar, describir, interpretar y clasificar los errores detectados en la muestra para luego ordenarlos según su relevancia a fin de lograr el nivel de habilidad gramatical requerida para acceder a la promoción de la asignatura;
- (vi) Verificar la eficacia y utilidad del instrumento en términos de uniformidad lograda en las calificaciones independientemente de la identidad del profesor-calificador.

Materiales y métodos

El grupo de sujetos estudiados en este proyecto estuvo conformado, por un lado, por 60 alumnos inscriptos en la asignatura *Gramática Inglesa I* – materia común en las carreras de Profesorado, Traductorado o Licenciatura en Inglés de la Facultad de Lenguas de la UNC. Dichos alumnos habían regularizado la asignatura en el ciclo lectivo 2007. Por otro lado se estudió a un grupo de profesores (miembros de este proyecto) integrantes de las cátedras de Práctica Gramatical del Inglés y/o Gramática Inglesa I.

Los siguientes materiales han sido estudiados y/o analizados hasta el momento para la consecución de este proyecto:

1. Plan de Estudios vigente para las carreras de Profesorado, Traductorado y Licenciatura en Inglés de la Facultad de Lenguas de la UNC;
2. Programa 2007 de la asignatura *Gramática Inglesa I*;
3. Exámenes Finales Regulares escritos de la asignatura *Gramática Inglesa I* correspondientes al turno marzo 2008;
4. Taxonomías descriptivas y explicativas de errores.

Métodos de análisis

Los miembros del equipo que dictan la asignatura *Gramática Inglesa I* procedieron a diseñar y administrar el examen (“la prueba”) en un todo de acuerdo con los contenidos disciplinares y formato de evaluación impartidos durante el ciclo lectivo 2007, a los fines de recolectar la muestra. Dicha muestra consistió en tres grupos de 20 exámenes cada uno elegidos en forma aleatoria, correspondientes al turno de Exámenes Finales Regulares de marzo de 2008. En dicha oportunidad, una vez que los alumnos habían resuelto el examen y a los fines de preservar la versión original que contenía las respuestas de cada uno de los alumnos, (es decir, previo al proceso de corrección que realizara cualquiera de los docentes que integraba la mesa examinadora), se procedió a fotocopiar cada uno de los 60 exámenes seleccionados.

Algunos miembros del equipo de investigación (miembros externos a la cátedra de Gramática Inglesa I se avocaron en sucesivas reuniones a realizar la descripción y análisis de los objetivos y perfil del egresado de las carreras de Profesorado, Traductorado y Licenciatura en inglés (Plan de Estudio vigente), para continuar, luego, con la descripción y el análisis de los objetivos, contenidos disciplinares, metodología de trabajo, formas y criterios de evaluación de la asignatura *Gramática Inglesa I*, según el programa vigente de la asignatura. En su informe, llegaron a la conclusión que tanto los objetivos, como así también la metodología de enseñanza y de evaluación planteados en el Programa vigente de la asignatura respondían a las pautas y descriptores establecidos en el Plan de Estudios vigente para las carreras de Profesorado, Traductorado y Licenciatura en Inglés. Sin embargo, puntualizaron el hecho de que los criterios de evaluación no estaban lo suficientemente descriptos o establecidos para asegurar fiabilidad en la corrección de los exámenes independientemente del docente que corrigiera esa prueba, por lo cual recomendaron revisar este punto del programa, explicitando de manera suficientemente pormenorizada y teniendo en cuenta los resultados de la presente investigación, para los ciclos lectivos venideros.

Por otra parte, los alumnos que formaron parte de este proyecto fueron invitados a responder un cuestionario en el que se hacía especial hincapié en su

experiencia personal en el cursado de la asignatura Gramática Inglesa I. En las respuestas de los alumnos a dicho cuestionario se notó un acuerdo general en que a lo largo de su trayectoria como estudiantes en la Facultad de Lenguas habían experimentado a menudo en las distintas cátedras – Gramática Inglesa no era una excepción- la falta de uniformidad en cuanto a los criterios de corrección de sus evaluaciones (valoración de errores, respeto por las ideas, distintas exigencias en cuanto a la profundidad en el desarrollo teórico de un tema, etc.) con respecto a las de otros compañeros, dependiendo de quiénes habían sido los docentes que habían corregido los exámenes. En otro orden de cosas, los alumnos valoraron muy positivamente el hecho de que este proyecto de investigación se llevara a cabo, ya que hubo absoluta coincidencia entre ellos en la importancia que debe otorgársele a la validez y a la fiabilidad en cualquier examen. Un aspecto positivo que es menester señalar en esta oportunidad es que, en general, los alumnos coincidieron en que el examen final de la asignatura Gramática Inglesa I que ellos habían rendido había valorado sus conocimientos disciplinares de la asignatura, la habilidad para aplicar las adquisiciones logradas, y la adecuada exposición de las mismas. Asimismo, valoraron positivamente que los ejercicios incluidos en el examen reflejaban el tipo de ejercicios incluidos en las evaluaciones parciales impartidas durante el transcurso del ciclo lectivo. Es deseable destacar que las opiniones vertidas por los alumnos – ya sea en forma oral en los distintos encuentros como en forma escrita a través del cuestionario mencionado – contribuyeron a realizar importantes cambios con respecto a la metodología y criterios de evaluación en el programa de estudios del ciclo lectivo 2009 como así también a otorgar la promoción de la asignatura sin examen final, lo cual contribuyó a realizar una evaluación más pormenorizada y más profunda del proceso de enseñanza-aprendizaje de los contenidos disciplinares de la asignatura y a aumentar el nivel de presentismo de los alumnos a las clases teórico-prácticas.

Primera etapa de corrección de los exámenes

Seguidamente, cuatro docentes (el Profesor Titular, la Profesora Asistente y la Profesora adscripta de la cátedra Gramática Inglesa I, como así también la Prof. Adjunta de la cátedra Práctica Gramatical del inglés) procedieron a corregir el primer grupo de exámenes que fueran fotocopiados y preservados oportunamente para este fin. Así, los cuatro docentes corrigieron en forma individual -sin ningún tipo de influencia o ingerencia de los otros tres docentes y sin establecer ni acordar ningún tipo de criterios de corrección (excepto los establecidos en el Programa de

la asignatura) - los 20 exámenes que formaban parte del primer grupo de la muestra recolectada. Posteriormente, se realizó una reunión con los miembros del equipo de investigación a los fines de verificar la presencia o ausencia de uniformidad de criterios de evaluación entre los distintos sujetos evaluadores.

La siguiente tabla ilustra los resultados obtenidos en esta primera etapa:

Sujeto	Examinado r 1		Examinado r 2		Examinado r 3		Examinado r 4	
	%	Calif.	%	Calif.	%	Calif.	%	Calif.
1	72	7	67	6	63	6	69	7
2	69	7	67	6	73	7	63	6
3	65	6	70	7	63	6	69	7
4	72	7	73	7	69	7	74	7
5	71	7	73	7	69	7	76	8
6	70	7	62	5	63	6	63	6
7	58	4	57	4	58	4	66	6
8	57	4	59	5	58	4	69	7
9	79	8	79	8	70	7	82	9
10	79	8	80	8	79	8	85	9
11	64	6	65	6	60	5	67	6
12	71	7	77	8	75	7	79	8
13	69	7	76	8	73	7	72	7
14	66	6	63	6	66	6	79	8
15	65	6	57	4	57	4	67	6
16	57	4	66	6	66	6	65	6
17	66	6	72	7	73	7	72	7
18	78	8	78	8	80	8	80	8
19	58	4	58	4	60	5	65	6
20	45	3	51	3	50	3	57	4

Como se desprende del análisis de los resultados obtenidos en esta primera etapa, todos los miembros del equipo coincidimos en que:

(1) en sólo dos casos (10%) (sujetos nº 4 y 18) hubo coincidencia con respecto a la calificación final obtenida; por lo tanto, eran fundadas las quejas y/o percepciones de los alumnos en cuanto a la ausencia de criterios unificados de corrección en la asignatura *Gramática Inglesa I*.

Veamos la representación del porcentaje obtenido:



Figura 1. Resultados obtenidos en la primera etapa de corrección de los exámenes

(2) Existía una necesidad imperiosa de mejorar y compartir las prácticas evaluativas de la cátedra, sobre todo en el caso del evaluador N° 4 (donde se nota mayor disparidad en cuanto a las notas o porcentajes otorgados en cada caso con respecto a los otros tres evaluadores), quien se había incorporado a la cátedra recientemente y, por ende, carecía de la experiencia que poseían los otros tres evaluadores, quienes se venían desempeñando en el área de gramática inglesa durante varios años.

(3) Los docentes de la cátedra debían trabajar en pos de desarrollar métodos o procedimientos de corrección fiables que garantizaran a los alumnos resultados justos.

(4) Era necesario describir los distintos tipos de errores cometidos por los alumnos y acordar la importancia de dichos errores en la calificación final con el objeto de elaborar una normativa de aplicación que lograra la máxima estandarización con respecto a los criterios de evaluación que debían seguir todos los miembros de la cátedra.

(5) Todos los esfuerzos debían dirigirse hacia el desarrollo de una evaluación "criterial", que debía consistir en la fijación de criterios externos bien formulados y concretos para proceder a evaluar un aprendizaje, tomando como punto de referencia un criterio marcado y/o las fases en que éste se haya podido desglosar, al margen del nivel personal del grupo.

Algunas especificaciones en cuanto a la "prueba":

El examen analizado fue un examen escrito y puede definirse como una *prueba de rendimiento* (Consejo de Europa, 2001), es decir que dicho examen se consideraba como el resultado de una evaluación sumativa (la que tiene lugar al final de un

determinado proceso didáctico; constata los resultados de ese proceso y sirve de base para adoptar decisiones de certificación, promoción o repetición, y de selección) + una evaluación del aprovechamiento. Esta prueba de rendimiento (o examen final) tenía como finalidad determinar los conocimientos lingüísticos de cada alumno y las posibilidades que se observaban en este estudiante para desenvolverse en el mundo real, aplicando sus conocimientos, aplicando sus conocimientos previos y lo que había aprendido a lo largo del cursado de la asignatura *Gramática Inglesa I*. Era, por lo tanto, una prueba que pretendía evaluar el aprovechamiento que el alumno había realizado de los contenidos disciplinares y estratégicos desarrollados durante el curso a los fines de evaluar si el estudiante había alcanzado los objetivos específicos fijados en el programa del curso para poder certificar su promoción al curso siguiente.

Por otra parte, queremos destacar que el examen objeto de esta investigación incluyó tareas de corrección objetiva (ejercicios de llenado de blancos, de selección de la opción correcta, etc.) y tareas de corrección subjetiva (ejercicios de producción y de justificación teórica), en los que el alumno debía demostrar tanto sus conocimientos lingüísticos como su capacidad de pensamiento crítico y utilización de estrategias de aprendizaje para resolver una situación dada exitosamente. En todos los casos el estudiante debía utilizar la lengua como objeto de estudio.

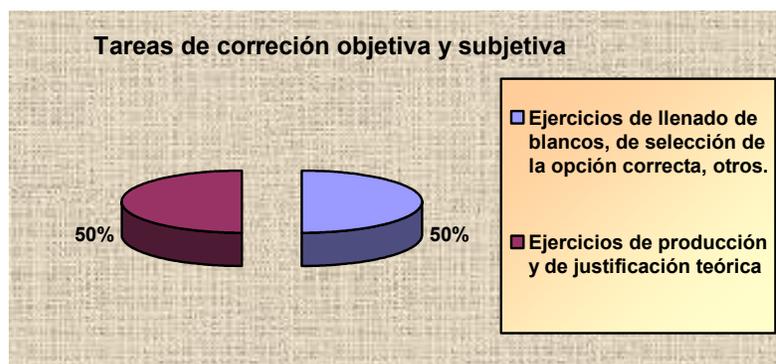


Figura 2. Distribución de tareas de corrección en el examen objeto de la investigación

En cuanto al contenido de la prueba, se prestó especial atención al hecho de que ésta reflejara una muestra significativa de lo que se había trabajado en clase a lo largo del curso, de acuerdo con el cronograma y programa de estudios vigentes.

Para otorgar el puntaje obtenido y la nota correspondiente a cada alumno/a se utilizó el baremo oficial vigente en la Facultad de Lenguas de la UNC .

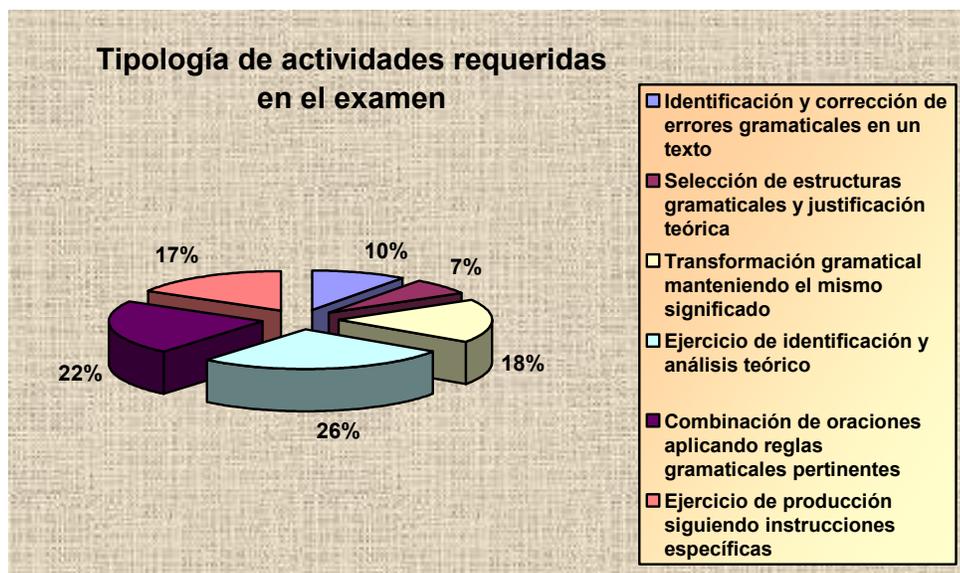


Figura 3. Contenido de la “prueba” o examen y puntaje asignado a cada actividad.

Segunda etapa de corrección de los exámenes

Teniendo en cuenta los resultados obtenidos en la primera etapa del proyecto y luego de haber llevado a cabo distintos talleres, en los cuales se discutieron, se analizaron y se trataron de interpretar las posibles causas en cuanto a la disparidad en los porcentajes y en las notas brindadas en cada caso por cada uno de los 4 evaluadores, se diseñaron diferentes taxonomías descriptivas de los errores más frecuentes observados en los exámenes y se llegó a un consenso con respecto al grado de incidencia de cada tipo de error en la nota o porcentaje obtenido según su relevancia para acceder a la promoción de la asignatura. Seguidamente, los sujetos evaluadores procedieron a corregir el segundo grupo de exámenes. La siguiente tabla ilustra los resultados obtenidos en esta segunda etapa:

Sujeto	Examinador 1		Examinador 2		Examinador 3		Examinador 4	
	%	Calif.	%	Calif.	%	Calif.	%	Calif.
1	74	7	77	8	70	7	77	8
2	74	7	73	7	66	6	74	7
3	74	7	78	8	81	9	78	8
4	85	9	86	9	84	9	77	8
5	71	7	70	7	70	7	77	8
6	68	6	76	7	64	6	74	7
7	84	9	85	9	83	9	84	8
8	84	9	83	9	86	9	79	8

9	66	6	68	6	68	6	67	6
10	69	7	68	6	68	6	70	7
11	77	8	73	7	80	8	77	8
12	77	8	73	7	78	8	69	7
13	67	6	76	7	78	8	74	7
14	80	8	78	8	82	9	80	8
15	79	8	74	7	80	8	77	8
16	84	9	86	9	85	9	82	9
17	80	8	80	8	78	8	80	8
18	79	8	78	8	77	8	78	8
19	85	9	82	9	77	8	85	9
20	80	8	80	8	77	8	81	9

Como puede observarse, si bien en esta segunda etapa existe un mayor número de alumnos que obtuvieron la misma calificación, independientemente del docente-calificador, el porcentaje en que no hubo coincidencia (80% = 16 casos), demostró que era necesario continuar trabajando para pulir el instrumento de evaluación.

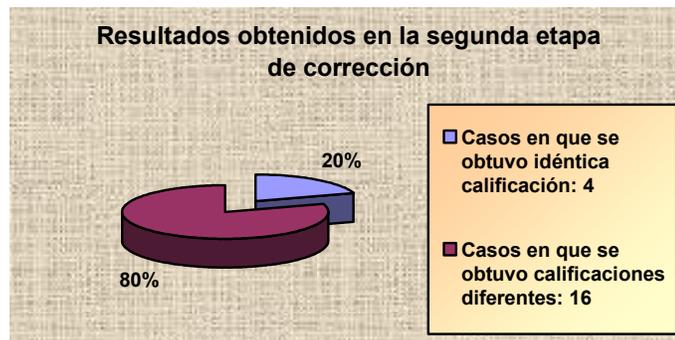


Figura 4. Porcentajes de resultados obtenidos del segundo grupo de exámenes, segunda etapa del proyecto.

Es por ello que sobre la base de estos resultados, la próxima tarea que desarrolló este equipo de investigación fue llevar a cabo el diseño y elaboración de bandas o descriptores referenciales de suficiencia con respecto a la habilidad gramatical adquirida en forma de escalas de estimación ordinales que contuvieran requerimientos plenos de habilidad gramatical (léase ausencia de errores) o requerimientos parciales de habilidad gramatical (con mención a número de errores tipificados) o falta de habilidad gramatical. Para ello, se acordó realizar una calificación pormenorizada de cada ejercicio, es decir llevar

a cabo una calificación que puntuara sobre cada uno de los puntos y dentro de cada punto sobre las diversas características consideradas esenciales en el tema propuesto.

Tercera etapa de corrección de los exámenes

Seguidamente, los sujetos evaluadores procedieron a corregir el segundo grupo de exámenes. La siguiente tabla ilustra los resultados obtenidos en esta segunda etapa:

Sujeto	Examinador 1		Examinador 2		Examinador 3		Examinador 4	
	%	Calif.	%	Calif.	%	Calif.	%	Calif.
1	83	9	85	9	75	7	85	9
2	74	7	75	7	73	7	69	7
3	80	8	83	9	78	8	80	8
4	72	7	72	7	69	7	74	7
5	74	7	75	7	76	7	69	7
6	80	8	80	8	75	7	79	8
7	82	9	81	9	81	9	82	9
8	74	7	73	7	69	7	76	7
9	77	8	73	7	83	9	79	8
10	80	8	80	8	80	8	78	8
11	75	7	73	7	74	7	79	8
12	74	7	75	7	63	6	68	6
13	68	6	67	6	68	6	63	6
14	60	5	59	5	61	5	60	5
15	68	6	51	4	64	6	72	7
16	67	6	66	6	68	6	67	6
17	73	7	71	7	74	7	70	7
18	75	7	76	7	81	9	73	7
19	76	8	74	7	78	8	68	6
20	77	8	78	8	77	8	77	8

El grupo de investigadores analizó los resultados en términos de grado de uniformidad en la calificación que fuera asignada a cada uno de los veinte exámenes por cada profesor-calificador durante la tercera etapa. En esta

oportunidad, se puede observar un mayor grado de coincidencia en comparación con las dos etapas anteriores (10% = 2 casos en la primera etapa, 20% = 4 casos en la segunda etapa; 60 % = 12 casos en la tercera etapa) en cuanto a los porcentajes y calificaciones finales otorgadas a cada alumno en cada caso, lo cual refleja un mayor consenso en lo que concierne a la interpretación de los errores cometidos por los alumnos y, por ende, un gran paso hacia la unificación de criterios. Sin embargo, en algunos casos aún se observa cierta disparidad – en algunos casos con una marcada diferencia en el porcentaje asignado -, lo cual incide indefectiblemente en la variación de la nota final.

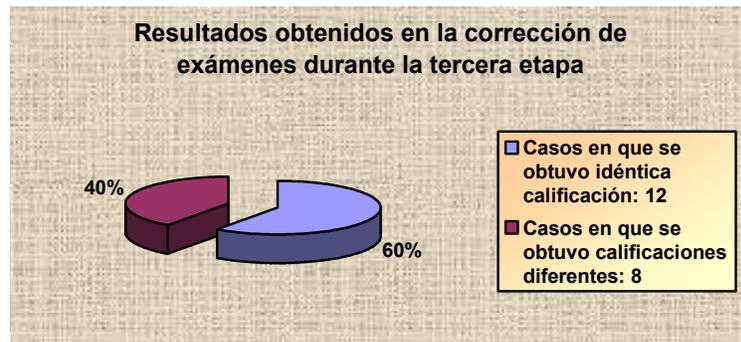


Figura 5. Porcentajes de resultados obtenidos durante la tercera etapa del proyecto.

Conclusiones

Se espera que los resultados del presente proyecto conduzcan a implicaciones relevantes para el mejoramiento de la calidad del proceso de evaluación de la asignatura *Gramática Inglesa I* en la Facultad de Lenguas de la UNC, constituyendo una base de mayor solidez para perfilar líneas de acción en el trabajo docente. Asimismo, se espera que los resultados de esta investigación resulten de utilidad para mejorar la práctica educativa en otras asignaturas a cargo de los docentes integrantes de este proyecto. En términos más específicos, se pretende aportar criterios unificados para calificar a los alumnos, ya que una calificación sólo es fiable si se asienta sobre un constructo informado de validación (Bailey, 1998; McNamara, 2000; Celce-Murcia, 2001; Purpura, 2004) . En este sentido, este proyecto resulta significativo al dar una respuesta a las constantes quejas planteadas por los alumnos respecto a la disparidad en la forma de calificar de los docentes de la Facultad de Lenguas. En otras palabras, desde la perspectiva que trasciende al equipo mismo de investigación, alcanzar los objetivos propuestos constituiría un invaluable aporte para reflexionar sobre la propia práctica docente a distintos niveles (institución, departamentos, cátedras) a los fines de conscientizar a los docentes de la necesidad e importancia de mejorar la calidad del sistema de evaluación en nuestra Facultad y llevar a cabo acciones que conduzcan

a tal fin. Los últimos resultados analizados demuestran que aún quedan ciertos aspectos o criterios para unificar y rediscutir, por lo cual se seguirá trabajando en pos de lograr la unificación deseada con los miembros de la cátedra de Gramática Inglesa I.

Referencias bibliográficas

APODACA, P. & C. LOBATO (eds.). *Calidad en la Universidad: Orientación y Evaluación*. Barcelona: Leartes. 1995.

ALDERSON, J. C., C. CLAPHAM & D. WALL. *Language Test Construction and Evaluation*. Cambridge: C.U.P. 1995.

BACHMAN, L. F. *Fundamental Considerations in Language Testing*. Oxford: O.U.P. 1990.

BACHMAN, L. F. "Alternative interpretations of alternative assessments: some validity issues in educational performance assessments". *Educational Measurement: Issues and Practice*, 21 (3), 5-19. 2002.

BACHMAN, L. F. & A. D. COHEN. *Interfaces Between Second Language Acquisition and Language Testing Research*. Cambridge: C.U.P. 1998.

BACHMAN, L. F. & A. S. PALMER. *Language Testing in Practice*. Oxford: O.U.P. 1996.

BAILEY, K. M. *Learning about Language Assessment: Delemmas, Decisions, and Directions*. Boston, MA: Heinle and Heinle. 1998.

CELCE-MURCIA, M. (editor). *Teaching English as a Second or Foreign Language*. 3rd ed. New York: Heinle & Heinle. 1999

COHEN (eds.) *Research Methodology in Second Language Acquisition*. Hillsdale, N.J: Lawrence Erlbaum Associates.

CONSEJO de EUROPA. Marco Común Europeo de referencia para las lenguas: aprendizaje, enseñanza, y evaluación. Ministerio de Educación, Cultura y Deporte, Subdirección General de Cooperación Internacional y Grupo Anaya, 2002, Cap. 9. 2001.

GASS, S. "The reliability of grammaticality judgements." In E. Tarone, S. Gass and A. Cohen (eds.). *Research Methodology in Second Language Acquisition*. (pp. 303-322). Hillsdale, N.J: Lawrence Earlbaum Associates. 1994.

HOUSE, E. R. *Evaluating with validity*. Beverly Hills: Sage . 1980.

HUGHES, A. *Testing for Language Teachers*. (2nd edition). Cambridge: C.U.P. 2003.

McNAMARA, T. *Language Testing*. Oxford: O.U.P. 2000.

PURPURA, J.E. *Assessing Grammar*. Cambridge: C.U.P. 2004

RAMSDEN, P. *Learning to Teach in Higher Education*. London: Routledge. 1998.

VILLAR ANGULO, L. M. & ALEGRE de la ROSA, O. M. *Manual para la excelencia en la enseñanza superior*. España: McGraw-Hill. 2004.