CIEE: RETOS E INNOVACIÓN DE SU EXAMEN DE CLASIFICACIÓN ON–LINE

ANTONIO RODRÍGUEZ GARCÍA
CIEE

Desde el año 2005, año de su creación, CIEE, en su centro de Sevilla, lleva aplicando su examen de clasificación on-line a fin de lograr la creación de grupos homogéneos en sus diversos programas (*Advanced Liberal Arts, Business and Society*, *Language and Society* y *Teaching and Development* e *Internacional Business and Cultural Program*). Los mencionados programas exigen requisitos académicos y lingüísticos diferentes (desde 0 semestres de español hasta 6 semestres de español). Así, el examen on-line nos sirve para colocar a nuestros estudiantes en los diferentes cursos intensivos de lengua y clases de contenido que CIEE ofrece durante los diferentes semestres. Asimismo, este examen es usado como piedra angular en las diferentes recomendaciones académicas hechas a los alumnos para la amplia gama de clases impartidas a lo largo de dichos semestres.

El diseño y creación de este tipo de examen on-line presenta algunos retos susceptibles de ser analizados minuciosamente, a saber, el contenido de los items gramaticales y léxicos de las secciones de gramática y vocabulario, la elección de textos apropiados para la sección de comprensión auditiva para los diferentes programas, el perfeccionamiento de la aplicación informática en orden a generar información lingüística individual instantánea, etc.

El primer reto que se plantea es la inevitable heterogeneidad de niveles (nuestro espectro de estudiantes cubre desde el nivel de principiantes hasta el de superior), así como enfoques metodológicos que nuestros estudiantes traen de sus diferentes universidades estadounidenses. Por ello, resulta incuestionable plantearse la pregunta de qué contenidos gramaticales y léxicos, así como el número de ellos que se deben incluir en los diferentes items y por qué. La justificación de su inclusión es un aspecto clave a la hora de discernir aquellos que discriminan niveles de competencia en español y ayudan a encontrar la homogeneidad que se busca. Al respecto de nuestros items gramaticales, se pueden afirmar que pueden que resulten novedosos en su concepción pero ello es algo que se analizará en este trabajo posteriormente.

Pero antes de entrar en el análisis de dichos ítems, sería conveniente estudiar la dinámica de nuestro examen desde que el estudiante recibe el correo electrónico con las instrucciones para acceder al examen, la estructura de nuestro OLPT (*O*n-line *L*anguage *P*lacement *T*est), secciones de éste, sistema de puntuación, criterios de corrección, proceso de clasificación de los estudiantes, etc. a fin de familiarizar al lector con todos los aspectos colaterales del OLPT.

El proceso es como sigue: el estudiante recibe un e-mail en inglés mientras está en EEUU para cumplimentar el OLPT (tiene un plazo de 30 días, reducido en nuestra última edición a 15 días). La razón de ello se fundamenta, como se ha mencionado anteriormente, en el amplio perfil de nuestros estudiantes, lo cual explica que dicho e-mail esté redactado en inglés a fin de que todos los estudiantes puedan leer y entender las instrucciones claramente. CIEE considera que la evaluación del nivel de español de nuestros estudiantes empieza una vez que el estudiante entra en el examen propiamente dicho y comienza a completarlo. Dicho e-mail contiene el *link*, *password* y *login* para acceder al examen, así como instrucciones básicas sobre el formato del examen y el propósito de éste. Una vez que el estudiante accede al examen, en él puede encontrar por orden secuencial un mensaje de bienvenida, instrucciones técnicas (enlace para bajar Windows Media Player en caso de que el estudiante carezca de él, información detallada sobre el objetivo del examen, el sistema de puntuación de cada sección, criterios de corrección, código de honestidad, académica, etc.).

El OLPT consta de 5 secciones:

- Sección gramatical (común)
- Sección léxica (común)
- Comprensión auditiva (común)
- Comprensión escrita (texto según programa y nivel)
- Expresión escrita (tema según programa y nivel)
- Test oral individual de 10 minutos a la llegada del estudiante a Sevilla.

El valor de cada sección es el siguiente:

- Gramática (30 puntos)
- Vocabulario (10 puntos)
- Comprensión auditiva (15 puntos)
- Comprensión escrita (15 puntos)
- Expresión escrita (30 puntos)

El contenido de los ítems se puede desglosar del siguiente modo:

La sección gramatical contiene 30 ítems gramaticales de opción múltiple (corrección automática) con 3 distractores por ítem. La sección léxica contiene 10 items léxicos de opción múltiple con 5 falsos amigos y 5 colocaciones léxicas (corrección automática) con 3 distractores por cada ítem individual. La sección de comprensión auditiva consta de un vídeo de producción propia, cuya duración es de unos 7 minutos con 15 preguntas de opción múltiple y tres distractores por cada ítem individual.

Capítulo aparte merecen las secciones de producción escrita y oral. La sección de comprensión escrita, que consta de un texto de longitud media, extraído, editado de revistas o periódicos españoles con 3 preguntas abiertas, es una sección donde se intenta evaluar la habilidad de cada estudiante para extraer e inferir ideas principales de un texto instando al estudiante a usar sus propias palabras y penalizando la acción de cortar y copiar. A diferencia de las anteriores secciones, cada texto se adecua al supuesto nivel de competencia de los diferentes programas, el cual, como se mencionó anteriormente, depende de un número mínimo de semestres de español cursados por el estudiante. La sección de expresión escrita gira en torno a la redacción de una composición que oscila entre las 150 y 300 palabras, según cada programa. En ella, los estudiantes deben elegir un tema entre tres que se ofrecen y cuyo contenido varía según cada programa. Se evalúa la capacidad de los estudiantes para expresarse en un español correcto y fluido. Finalmente, se encuentra la sección de expresión oral, en la cual se evalúa la capacidad de expresión y comprensión oral mediante una entrevista individual de 10 minutos. Dicha entrevista sigue el modelo y protocolo del OPI (Oral Proficiency Interview).

La razón por la cual en las secciones de opción múltiple no existe discriminación de programas en lo referente al contenido de los ítems, comunes para todos los estudiantes, se basa en dos criterios. El primero, como se mencionó previamente, se sustenta en la idea de que los diferentes enfoques metodológicos empleados en las diferentes universidades no nos sirven como pauta a seguir para establecer niveles. Hemos comprobado que los mismos semestres de español cursados por un estudiante de una universidad no implican similitud de competencia en comparación con los mismos semestres estudiados por otro estudiante de una universidad distinta. Todo ello sin entrar en variables personales, por otra parte lógicas, sobre el grado de competencia en español de cada estudiante. Sin embargo, sí hemos podido constatar que los estudiantes de una universidad, digamos X, tienen, como regla general, un nivel superior a los estudiantes de la universidad, digamos Y, aun teniendo los mismos semestres de español cursados. Incluso, entre los estudiantes de la misma universidad con igualdad de semestres estudiados de español existen grandes variaciones. En segundo lugar, creemos que los aspectos gramaticales y léxicos incluidos en la sección de gramática y vocabulario son los suficientemente representativos de todos los niveles de español como para discriminar niveles. Todo ello se puede hacer extensible a la sección de comprensión auditiva, cuyas cuestiones están diseñadas en torno a un grado de dificultad creciente para que puedan distinguir niveles de competencia. De todos modos, en CIEE pensamos que las restantes secciones de nuestro examen, secciones de producción (comprensión escrita, expresión escrita y examen oral), nos ayudan a completar con un alto grado de fiabilidad el nivel de dominio de español de cada estudiante.

En el título de esta comunicación se menciona un sustantivo como *innovación*, el cual queda justificado por la inclusión de determinadas novedades con respecto a otros exámenes de clasificación existentes. Entre las innovaciones que presenta nuestro examen on-.line se puede encontrar una que merece, a nuestro juicio, especial consideración, a saber, la mezcla de contenidos gramaticales diferentes en un mismo ítem. Asimismo, podemos encontrar la presentación de las diferentes opciones (respuesta correcta + 3 distractores) mediante un menú desplegable en cada hueco. Con la mezcla de cuestiones gramaticales diferentes se intenta ofrecer muestras de lengua que sean reales y naturales, las cuales pueden ser evaluadas individualmente dado que la aplicación informática garantiza que se pueda medir qué hueco gramatical es correcto y cuál es incorrecto, separando así lo que el estudiante sabe o no sabe en un mismo ítem gramatical. Pensamos que la forma tradicional de un aspecto gramatical por ítem es antinatural ya que crea un contexto mental que automatiza al estudiante y convierte la sección en algo monótono y carente de fiabilidad. Veáse una muestra debajo:



Se intenta evitar que el estudiante mecanice sus respuestas en la sección gramatical intuyendo lo que será incluido en los diferentes ítems. Por citar un ejemplo aleatorio, el estudiante entra en el examen y sabe que está contestando los ítems de *ser y estar*, luego posteriormente intuye que vienen los de *por y para*, a continuación los de los pronombres, etc. Si el examen debe ser fiel ejemplo de lo que es una muestra de lengua natural y fluida, parece lógico preguntarse la razón por la que el contenido de un ítem siempre debe girar sobre un mismo aspecto gramatical. En nuestra opinión, tanto en la lengua escrita como hablada se emplean usos mezclados de diferentes cuestiones gramaticales con total naturalidad y esa muestra de lengua real y cotidiana es la que nuestro examen pretende reflejar en sus ítems gramaticales... La contextualización de todos los items gramaticales (intentamos reproducir ejemplos de la realidad lingüística cotidiana del español actual en la forma de diálogos) nos parece de obligado cumplimiento como factor clave a la hora del logro de una reproducción fidedigna de lo que es la lengua española en uso. Dicho de otra forma, se intenta que nuestro examen cumpla con unos de los preceptos básicos: el estudiante, fuera del examen, tiene ese mismo nivel de competencia en la comunicación cotidiana.

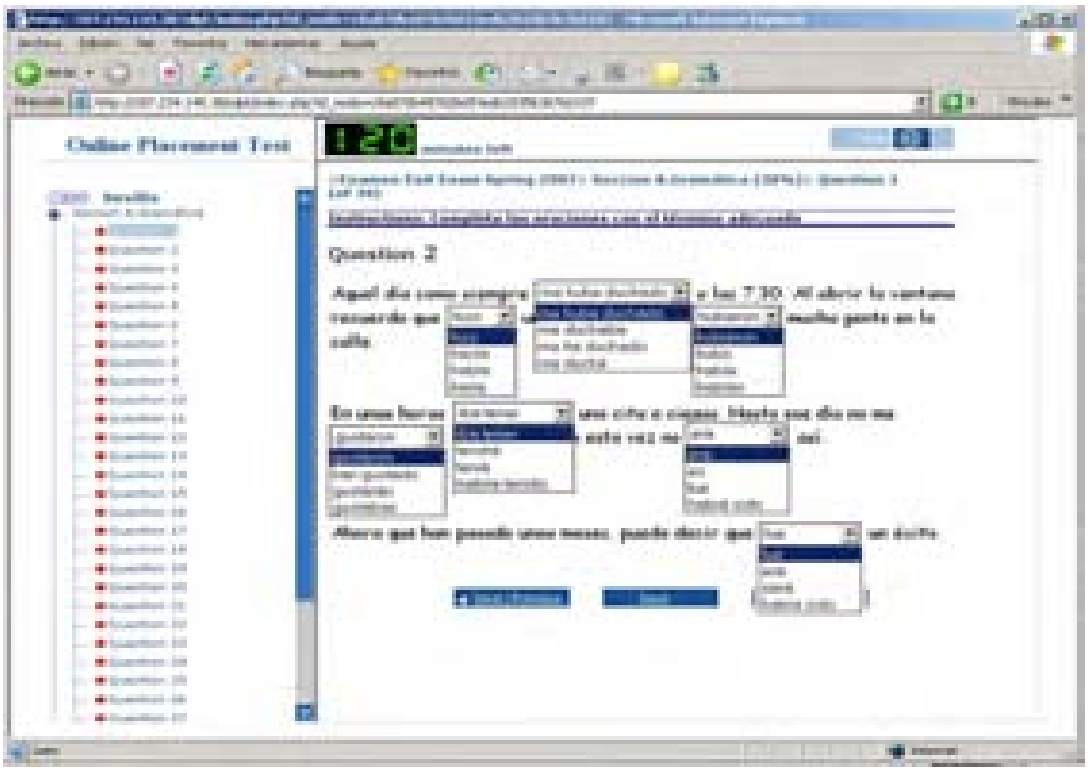
Asimismo, en CIEE Sevilla pensamos que la combinación de aspectos gramaticales en un mismo ítem gramatical ayuda a solventar uno de los problemas que presenta la elaboración de un examen de clasificación, a saber, acotar la lengua para elegir los contenidos pertinentes para dicho examen. La combinación de aspectos gramaticales reduce el número de ítems pero no de cuestiones gramaticales, algo que en nuestra opinión ahorra las engorrosas secciones gramaticales de 50 ó 60 items, que en muchos caso no hacen sino aburrir y cansar al estudiante antes incluso de terminar la primera sección del examen. Esta última cuestión suscita una importante cuestión para un examen de estas características: ¿Cuántos usos de *ser* y *estar*, por ejemplo, se deberían incluir en un examen de clasificación que pretende cubrir el espectro de estudiantes de todos los niveles? ¿Se debería incluir un alto porcentaje de esos usos si lo que se pretende evaluar en la sección gramatical es el conocimiento o dominio de los diferentes aspectos gramaticales? ¿Se debería incluir un aspecto gramatical más de una vez para cerciorarnos de que el estudiante lo domina y no es fruto del azar? La acotación de contenidos gramaticales, aspecto farragoso en el diseño de un examen de esta índole, no queda resuelta con nuestra aportación pero sí se logra que los contenidos gramaticales estén mezclados en un mismo ítem de forma más natural, así como incrementar la variedad de aspectos gramaticales incluidos en la sección.

De modo similar, la aplicación de un menú desplegable es extensible a ítems de más de un hueco pero con el mismo aspecto gramatical, a saber, los pretéritos o los conectores de discurso, por nombrar sólo algunos de ellos. La inclusión de una pequeña conversación con 4 ó 5 huecos referidos a los pretéritos nos

parece una forma menos forzada de evaluar el supuesto dominio del estudiante de esa parcela gramatical que la tradicional forma consistente en un ítem por uso y en muchos casos, totalmente descontextualizado. Como se mencionó anteriormente, los resultados individuales de cada ítem gramatical se pueden extraer perfectamente, por cuanto la aplicación informática de nuestro examen permite la corrección automática individualizada dentro del mismo ítem de la respuesta correcta y la incorrecta. Esta cuestión es de incalculable valor lingüístico a la hora de aislar los usos del mismo aspecto gramatical que realmente sabe el estudiante y lo que no, evitando que un estudiante pueda resolver acertadamente un ítem gramatical de más de un hueco sabiendo la respuesta correcta de uno de los huecos. Con la aportación de nuestro menú desplegable logramos un triple propósito: ofrecer una muestra de lengua con más de un hueco por ítem y la combinación de contenidos gramaticales dotando a la sección de mayor veracidad y naturalidad, incrementar el número de aspectos gramaticales de la sección gramatical y, finalmente, evitar la respuesta acertada de un ítem de más de un hueco simplemente sabiendo uno, como es el caso de otros exámenes más tradicionales. Véase una muestra de lo que sería un ítem gramatical sobre preposiciones sin usar el menú desplegable y otro ejemplo, extraído del examen de un estudiante, sobre los pretéritos con el mencionado menú desplegable:

(1) -Juan, el otro día Teresa me pidió_____ el número de teléfono de mi mujer. Se va a casar_____ Pedro. No me lo puedo creer.
-Carlos, piensa que ella está muy enamorada_____ él y siempre está pensando _____ su futura boda.

a. por / NA / con / de
b. NA / a / con / sobre
c. NA/ con / de / en
d. Por / con / de / en



Todo lo expuesto anteriormente se puede hacer extensible a la sección de vocabulario. Las secciones de vocabulario de muchos exámenes de vocabulario juegan en ocasiones el papel de comparsa por cuanto existe la absurda tendencia de considerarla una sección menor, con ítems que parecen elegidos de modo arbitrario, lo cual no puede quedar más lejos de la realidad. En el examen de CIEE se han buscado ítems que realmente ayuden a obtener una visión global del conocimiento léxico del estudiante y cuyo dominio sirva para lograr una expresión clara, correcta y que no interrumpa la secuencia lógica de ideas. En este sentido, hemos optado por los falsos amigos y colocaciones léxicas específicas. Dichos ítems abarcan desde el nivel inicial hasta el superior. La inserción de otros ítems (sinónimos, antónimos, expresiones idiomáticas, etc.) quedan evaluados en la sección de redacción, en la cual se incluye una escala analítica con varios descriptores y entre ellos, el de vocabulario. Una vez más, intentamos que la difícil tarea de

acotar el contenido de una sección sea lo más representativa y amplia posible a fin de obtener la muestra más fiable de la habilidad y dominio de la lengua española de cada estudiante.

Otra de las innovaciones que nuestro examen presenta y que, a buen seguro, muchos de los profesores agradecen es la suma automática de los totales de cada sección. La engorrosa labor de sumar manualmente los totales de cada sección, los posibles errores en dicha suma, deslices a la hora de corregir un ítem, etc., quedan relegados al olvido con la suma automática de cada sección y del total de puntos del examen. Al ser las tres primeras secciones de nuestro examen de opción múltiple, dichas secciones quedan sumadas automáticamente al completarlas el estudiante. Así, los profesores encargados de evaluar las secciones de corrección subjetiva –comprensión escrita y redacción– entran en cada examen asignado, previamente han recibido una contraseña individual, corrigen la sección de cada estudiante y marcan la puntuación de cada sección en una escala situada a la derecha de la pantalla. Cada escala consta de varios descriptores acorde a la puntuación consensuada por los profesores en un taller de corrección que se celebra justo antes de cada edición de nuestro examen, por lo que el profesor sólo tiene que marcar en dicha escala la puntuación que considere oportuna y ésta se añade directamente al total de la sección y como consecuencia, al total del examen. De modo, que una vez que el profesor ha terminado la corrección del examen, el sistema ya dispone, en cuestión de segundos, de la puntuación total de todos los exámenes corregidos.

Concerniente a la obtención de los totales de cada sección, así como a la puntuación total de cada estudiante, la aplicación empleada para nuestro examen genera datos exportables fácilmente a una hoja de Excel, las cuales ofrecen el nombre y apellidos de cada estudiante, universidad, puntuación en cada sección, puntuación total y respuesta individual a cada ítem de opción múltiple, profesor encargado de la corrección de cada test, etc. Ello supone un ahorro de tiempo innegable por cuanto un usuario medio de Excel puede en cuestión de minutos ordenar dicha hoja de Excel, clasificar y distribuir a los estudiantes en los diferentes grupos para las clases que organiza CIEE. No olvidemos que la génesis de este examen es clasificar, según niveles, a los estudiantes para las diferentes clases de lengua y de contenido que ofrecemos en los diferentes cursos intensivos distribuidos a lo largo de las tres primeras semanas de su estancia en Sevilla.

Asimismo, nuestra oferta de clases durante el semestre para los demás cursos ofrecidos tiene como base principal los resultados obtenidos en dicho test. La información que genera la aplicación se completa con el tiempo empleado por cada estudiante en cumplimentar el examen (el estudiante dispone de 120 minutos para completarlo por lo que un reloj situado en la parte superior del examen le indica el tiempo empleado y el restante), y la fecha de finalización, aspecto éste ultimo de importancia, ya que en programas como el nuestro la cuestión administrativa no debe ser dejada al margen y es por ello, que ante la masificación de determinadas clases y frente a la imposibilidad material y física de contar con clases con un número tan elevado de estudiantes, el criterio de establecer prioridad, criterio que aparece claramente reflejado en las instrucciones del examen, para matricularse en clases según la fecha de finalización del examen parecer funcionar como un elemento disuasorio de gran eficacia.

La generación de datos automáticos sobre la respuesta individual de cada estudiante es de suma utililidad a la hora de realizar análisis detallados de carácter estadístico sobre el examen. Dicho análisis contempla la obtención de la media, mediana, moda, rango, índice de dificultad, índice de discriminación, etc. En resumen, herramientas estadísticas empleadas en el análisis sobre la fiabilidad y validez de nuestro examen. A este respecto, la obtención de datos por parte de la aplicación del índice de dificultad, así como el de discriminación de forma casi automática supone la constante actualización de los diferentes ítems que componen las secciones de respuesta múltiple, cuestión esta última de capital importancia para sustentar la mencionada fiabilidad de un examen.

En CIEE pensamos que evaluar no debe ser el único fin de cualquier examen sino el punto de partida para empezar a enseñar y corregir errores de los estudiantes. Esta característica, una vez más, es la base sobre la que la aplicación compone, mediante un complejo sistema, un informe individual detallado de cada estudiante. Dicho informe generado en cuestión de minutos, contiene información sobre los contenidos gramaticales de los diferentes items de cada sección, ofreciendo una evaluación ítem por ítem de en qué aspectos gramaticales muestra cada estudiante mayor fortaleza y, por el contrario, en cuáles una mayor debilidad. Se ofrecen porcentajes de aciertos por categorías gramaticales, –subjuntivo, preoposiciones, pretéritos, *por* y *para*, *ser* y *estar*, etc.–, lo cual supone una información valiosísima para los diferentes profesores de cada clase. En la imagen de debajo se puede ver el mencionado informe lingüístico de carácter individual que la aplicación genera y que es enviado a los profesores.



La tradicional información de nombre, apellido, totales por sección y puntuación total es completada con datos estadísticos y lingüísticos sobre cada estudiante, información de incalculable valor para el profesor a la hora de enfatizar aspectos gramaticales específicos, metodología a emplear, fortalezas, debilidades. Creemos que esta herramienta lingüística supone un paso adelante en pos de obtener información detallada, rápida y fiable sobre las destrezas gramaticales y léxicas de cada estudiante sin tener que esperar a que pasen varios días antes de que el profesor conozca el nivel de competencia de sus estudiantes. Sin embargo, dicho informe queda completado por otro de carácter más tradicional (los profesores encargados de corregir completan con comentarios de carácter lingüístico una hoja de Excel en las secciones de corrección objetiva) una vez que han corregido dichas secciones, comprensión escrita y expresión escrita. De este modo, el profesor encargado de impartir la clase, desde el primer día, tiene toda la posible información actualizada de contenido lingüístico sobre sus estudiantes. No descartamos, de hecho ya estamos trabajando en ello, la creación de un informe automático que contenga comentarios sobre las mencionadas secciones de corrección subjetiva que complete la automatización de la aplicación en todas las secciones del examen. En definitiva, se intenta obtener la mayor cantidad de información de carácter lingüístico y estadístico sobre los estudiantes, así como de nuestro examen, sin olvidar el inmenso ahorro de tiempo que supone la generación de datos como los mencionados en esta comunicación. Todo ello con el claro propósito de mejorar nuestras clases y dotar a nuestro examen de la mayor fiabilidad, validez y viabilidad posible.

Otra de las innovaciones que presenta nuestro examen es la inclusión de un vídeo de producción propia. Dicho vídeo rompe con la dinámica de extraer una muestra de lengua en forma de copia visual de programas de televisión en los cuales la temática y registro se aleja, en muchos casos, de los que se supone es un ejemplo actual y natural de la lengua española. Nuestro vídeo muestra una conversación a la que nuestros estudiantes pueden estar expuestos a lo largo de su estancia en Sevilla y España. Contempla los tres registros, tres macrofunciones: descripción, narración y argumentación, secuencia de imágenes que apoyen el contenido del video y sobre todo, la presencia en todo momento de los interlocutores en pantalla, algo que resulta fundamental y natural para los hablantes no nativos a efectos de poder visualizar los labios de la persona que habla. De modo similar, considerando los efectos positivos que ha tenido hasta ahora nuestro examen on-line sobre la estructura de nuestros cursos, pensamos que la evaluación de competencia lingüística de nuestros estudiantes deber ser completada mediante la aplicación del denominado examen de salida. Este examen pretende evaluar el nivel de progreso de nuestros estudiantes una vez haya terminado el curso, independientemente de los exámenes finales que hayan hecho en sus diferentes clase de lengua o de contenido. En dicho examen aplicaremos las mismas herramientas informáticas que se aplican en el examen de colocación.

CIEE a través de su examen on-line intenta ofrecer el mejor examen posible cumpliendo los tres parámetros previamente mencionados, a saber, fiabilidad, validez y viabilidad. Sin embargo, ello no implica que pensemos que la consecución del examen perfecto es uno de los desafíos más difíciles de lograr y es un reto al que nos enfrentamos edición tras edición de nuestro test. Nuestro grado de satisfacción es moderado ya que creemos que existen muchos interrogantes en el campo del *testing*

moderno que no han sido resueltos o, en su defecto, enmascarados bajo axiomas universales, algunos de ellos cuestionables. Sirva esta breve introducción para dar paso a algunas de las preguntas que en la creación y aplicación de nuestro examen nos hemos encontrado, las cuales suponen debates apasionantes. Son temas que, pensamos, nos mueven a seguir investigando y, por qué no, realizar todos los cambios pertinentes en nuestro examen si concluimos que estos son necesarios para mejorarlo y actualizarlo.

Sirva como muestra de nuestra constante búsqueda del mejor examen posible algunas de las siguientes preguntas. Estamos seguros de que el lector encontrará algunas interesantes como punto de partida para reflexionar, o al menos, cuestionar algunos de los axiomas incorporados al campo del testing:

- ¿Qué se entiende por un examen bueno/válido?
- ¿Cuál es la diferencia entre un examen bueno y uno prestigioso?
- ¿En qué secciones produce realmente el estudiante?
- ¿Por qué la producción se limita, en muchos casos, a la sección de expresión escrita y el test oral?
- ¿Se deben incluir variantes dialectales de Latioamérica? (ustedes, suburbios, etc.)
- ¿No se convierten las pruebas de comprensión lectora en una prueba de sinónimos?
- ¿A qué llamamos una muestra real de lengua?
- ¿Qué instrumentos se adaptan mejor al concepto de competencia lingüística?
- ¿Se deben eliminar las secciones de opción múltiple?
- ¿Cuántos ítems habría que incluir en una sección para que ésta fuera fiel reflejo del nivel del estudiante?
- ¿Qué conclusiones extraemos de señalar una opción? (grandes contradicciones entre el ítem respondido y la producción en la redacción)

A modo de breve conclusión, la creación y aplicación de nuestro examen nos ha motivado a intentar investigar, innovar y aplicar, en la medida de lo posible, todos los aspectos en este trabajo mencionados. Pensamos que estamos en el punto de partida de un estudio detallado en el campo del *testing* que sólo nos puede conducir a mejorar nuestro examen on-line en futuras ediciones. Aparte de los retos y aspectos innovadores del examen on-line, CIEE no pretende hacer de su examen un fetiche y convertirlo en objeto de culto por cuanto más allá de sus diferencias con respecto a otros; el examen de colocación on-line de CIEE en un proyecto de investigación que intenta servirse de las ventajas que la informática e Internet ofrecen en la actualidad. Se intenta agilizar y dinamizar la evaluación de cientos de estudiantes, lograr la mayor fiabilidad posible e intentar servir de referente para el progreso de nuestros estudiantes a lo largo de su estancia en Sevilla.

Bibliografía

ALDERSON, C. (2005): *Assessing Listening*, Cambridge: Cambridge University Press.

ALDERSON, C. et al. (1998): *Exámenes de Idiomas*, Madrid: Cambridge University Press.

BACHMAN, L. (1990): *Fundamental Considerations in Language Testing*, Oxford: OUP.

BACHMAN, L. (2004): *Statistical Analyses for Language Assessment*, Cambridge: Cambridge University Press.

BROWN, J. (2005): *Testing in Language Programs*, New York: McGraw Hill.

BUCHMAN, L. y A. PALMER (2005): *Language Testing in Practice*, Oxford: OUP.

BUCK, G. (2006): *Assessing Listening*, Cambridge: Cambridge University Press.

FLUTCHER, G. y F. DAVIDSON (2007): *Language Testing and Assesment*, Oxon: Routledge.

MAHUKE, K. y C. DUFFY (1996): *The Heinneman ELT TOEFL Preparation Course*, Oxford: Mc Millan Heineman.

READ, J. (2005): *Assessing Listening*, Cambridge: Cambridge University Press.