

ADAPTACIÓN DEL EXAMEN DE INGLÉS DE LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD A UN ENTORNO INFORMÁTICO: ESTUDIO SOBRE LAS TIPOLOGÍAS DE PREGUNTA

JESÚS GARCÍA LABORDA
ANA GIMENO SANZ
Universidad Politécnica de Valencia

RESUMEN. En el trabajo que aquí presentamos se investigan las posibilidades de adaptar el examen de inglés de las Pruebas de Acceso a la Universidad a un entorno informático que permita la evaluación de conocimientos tanto escritos como orales. Para llevar a cabo dicha adaptación de contenidos se analizarán diversas plataformas existentes de exámenes asistidos por ordenador (Computer Assisted Language Testing - CALT), para posteriormente realizar una propuesta de modalidad de examen que se adecue a los condicionantes específicos de la prueba de inglés de las PAU, teniendo en cuenta las últimas innovaciones tecnológicas de reconocimiento de voz. Se demostrará que, aunque algunas de las modalidades de pregunta que se utilizan en la actualidad en dicho examen son adecuadas para medir los conocimientos de una lengua extranjera, un entorno informático ofrece un abanico de posibilidades que el formato impreso no puede ofrecer. Se habrán de encontrar y extrapolar de otros tipos de examen de idiomas de carácter más general un nuevo diseño de ítems que se pueda incorporar a una aplicación informática. Esto, por lo tanto, hace necesario establecer una taxonomía jerárquica de la tipología de ítems que sean metodológicamente adecuados para medir los conocimientos de los estudiantes que aspiran a estudiar en la universidad española. La ponencia concluirá con el esbozo del enfoque para emprender esta tarea y una propuesta inicial para ser aplicada en CALT.

PALABRAS CLAVE. Exámenes asistidos por ordenador (CALT), diseño y tipología de ítems, Pruebas de Acceso a la Universidad, examen de inglés

ABSTRACT. Although English language was included in Spanish university entry exams at a later stage with regard to when the state exams were implemented nationally, the English exam has remained practically unchanged for many years. Since it was introduced, the English language exam has been primarily based on reading comprehension and, to a lesser extent, written production. The current type of questions used in the Valencian Community to measure secondary students' level of knowledge relies on short answers to questions, multiple choice, true/false, vocabulary recognition and writing a short text based on personal opinion. This presentation will focus on the adaptation of the university entry English exam to an online environment based on the findings and research in Computer Assisted Language Testing. The authors will analyse different exercise typologies to suit both the requirements of the university entry exam and that of the actual online environment. It will conclude with a description of the hierarchical taxonomy of question types and the necessary item design that would be desirable in order to efficiently evaluate the language skills of future university students.

KEYWORDS: Computer Assisted Language Testing (CALT), item design, university entry examinations, English exam

1. INTRODUCCIÓN

Aunque la asignatura de inglés se incorporó como materia en las Pruebas de Acceso a la Universidad (PAU)¹ relativamente tarde respecto a la instauración de dichas pruebas en el sistema educativo español, son ya muchos los años que se han venido utilizando tipologías de examen que han variado en muy poco. En lo que respecta a la Comunidad Valenciana, el examen de inglés ha sufrido ligeras modificaciones al tener que adaptarse éste a las distintas regulaciones de educación secundaria vigentes en cada momento (LGE, LOPEG, LOGSE, LOCE y LOE). La prueba de inglés, no obstante, sigue basándose en la lectura de un texto y en la contestación a una serie de preguntas de comprensión por medio de cuestiones de respuesta corta, de verdadero o falso, de opción múltiple, reconocimiento de vocabulario y la producción de un texto de opinión.

Una simple mirada al tipo de pruebas que se realizan en toda España da muy breves variaciones sobre los tipos de ítems y tareas utilizados en nuestro país. García Laborda (2006a) analizó precisamente estas tareas en 10 comunidades autónomas españolas obteniendo un total de 14 ítems diferentes que solamente incluyen 3 destrezas: lectura, la que predomina en todas las comunidades analizadas, escritura y, de una manera casi anecdótica, comprensión oral (Cataluña). Precisamente, García Laborda (2006a) y otros estudios relacionados con la diferencia en el diagnóstico de lenguas extranjeras han puesto de manifiesto las diferencias significativas que aparecen en las puntuaciones y las consiguientes consecuencias en torno a la capacidad lingüística de estudiantes de idiomas debido a los distintos tipos de ítems que se evalúan (Christie, J. y Delaruelle 1997; Brindley y Slatyer 2002). Estas diferencias aparecen incluso en exámenes adaptativos, tal y como señalan en sus estudios Wang, 2001 y Yi *et al.*, 2001. Por tanto, dada la independencia educativa y las inmensas diferencias que hay entre los exámenes que integran las PAU [y el proceso de desarrollo de estrategias que conlleva (Van Gelderen y Oostdam 2005)] en las distintas comunidades autónomas españolas, parece imposible que todos los alumnos españoles se enfrenten a estas pruebas en condiciones similares, lo que hace pensar que estas desigualdades afectan a unos alumnos más que a otros.

Al parecer, para las autoridades educativas, así como para autores como Fulchery Reiter Xing (2003), en lo que se refiere a exámenes de producción oral, o Hambleton (2004), en lo que se refiere a exámenes adaptativos de selección múltiple, la cuestión es simplemente buscar criterios de igualdad y una selección del banco de tareas adecuadas. Más aún, para Hambleton (2004) tiene más importancia el banco de ítems utilizado en un examen que el diseño de las tareas incluidas en el mismo (también Veldkamp y Van der Linden 2000). Sin embargo, ninguno de estos estudios es concluyente sino que, más bien, sugieren que esto es así. Como reacción a una propuesta criterial como la del ministerio de educación español (y, en su defecto, las consejerías de educación) existen autores que reclaman una Selectividad de calidad y más equitativa basada en acciones comunes para todo el territorio español (García Laborda 2005).

El sistema de examen para alcanzar esa supuesta igualdad podría encontrarse en el uso de plataformas de exámenes asistidos por ordenador e Internet, basado en exámenes C-test. Sin embargo, para algunos autores como Davies (2003) es necesaria la inclusión de tareas tradicionales junto con aquellas más modernas (como los tests adaptativos) para obtener una mejor evaluación. Para otros autores, es muy importante establecer un control de dificultad en el diseño de ítems en exámenes orales (Iwashita *et al.* 2001) con el fin de equiparar la validez de la prueba oral y el resto de las secciones en exámenes combinados orales y escritos lo que, a su vez, facilita el establecimiento de una rúbrica adecuada y equivalente para todas las partes del examen (Folse 2006).

En otros foros (García Laborda 2005; 2006b) también se ha presentado la posibilidad de replantearse la Selectividad e incluir pruebas no realizadas hasta la actualidad, como tareas de producción oral, a través de plataformas informáticas en red que, además, permiten realizar pruebas adaptativas en algunas secciones del examen con un resultado mucho más informativo que el que se aplica en la actualidad (Hambleton y Xing 2006) y permiten analizar destrezas nunca examinadas anteriormente. Por tanto, en el trabajo que aquí presentamos se investigan las posibilidades de adaptar el examen de inglés de las PAU a un entorno informático que permita la evaluación de conocimientos tanto escritos como orales y, por tanto, vaya más allá del examen adaptativo ya casi tradicional. Asimismo, se hace con el convencimiento de que los alumnos son capaces de dar el paso del papel y lápiz al teclado (Mead y Drasgow 1993; Chapelle *et al.* 2003; Choi *et al.* 2003) o aceptar los recursos audiovisuales de la pantalla de su ordenador como interlocutores válidos para un examen oral de idiomas (Kenyon y Malabonga 2001).

2. ADAPTACIÓN DE TAREAS E ÍTEMS TRADICIONALES A ENTORNOS INFORMÁTICOS

Para llevar a cabo dicha adaptación de contenidos se analizarán las diversas pruebas existentes de exámenes asistidos por ordenador (Computer Assisted Language Testing - CALT), para posteriormente realizar una propuesta de modalidad de examen que se adecue a los condicionantes específicos de la prueba de inglés de las PAU, teniendo en cuenta las últimas innovaciones tecnológicas de reconocimiento de voz. Se demostrará que, aunque algunas de las modalidades de pregunta que se utilizan en la actualidad en dicho examen son adecuadas para medir los conocimientos de una lengua extranjera, un entorno informático ofrece un abanico de posibilidades que el formato impreso no puede ofrecer. Se habrán de encontrar y extrapolar de otros tipos de examen de idiomas de carácter más general nuevas tipologías de ítem que se puedan incorporar a una aplicación informática. Esto, por lo tanto, hace necesario establecer una taxonomía jerárquica de la tipología de ítems que sean metodológicamente adecuados para medir los conocimientos de los estudiantes que aspiran a estudiar en la universidad española. Este trabajo concluirá con el esbozo del enfoque para emprender esta tarea y una propuesta inicial para ser aplicada en CALT.

2.1. Las Pruebas de Acceso a la Universidad en España

Como se ha dicho más arriba, tras analizar las pruebas de inglés en 10 comunidades autónomas, García Laborda (2006a) elabora una lista de ítems incluidos en dichos exámenes. Se ha podido constatar que en el resto de los distritos universitarios no se ofrecen grandes diferencias con respecto a dicha lista. En general se puede establecer una lista de tareas basada en las destrezas utilizadas (Tabla 1). Como se puede observar, esta lista muestra que el peso otorgado a cada destreza no es equitativo, centrándose fundamentalmente en la lectura y su comprensión.

Lista de tareas	
Comprensión lectora	<ul style="list-style-type: none"> • Justificar el título del texto • Redactar respuestas a preguntas abiertas relacionadas con el texto • Escribir definiciones de palabras • Contestar preguntas de Verdadero / Falso y justificar respuesta • Resumir el texto • Relacionar información extraída del texto • Proporcionar sinónimos / antónimos • Contestar preguntas de selección múltiple • Rellenar huecos con información extraída del texto
Conocimientos Gramaticales	<ul style="list-style-type: none"> • Rescribir texto (sin variar significado) • Realizar modificaciones gramaticales (varios)
Producción oral	<ul style="list-style-type: none"> • Identificar fonéticamente las palabras
Producción escrita	<ul style="list-style-type: none"> • Redactar un texto de opinión propia
Comprensión oral	<ul style="list-style-type: none"> • Relacionar información • Contestar preguntas de selección múltiple

Tabla 1. Lista de tareas basada en las destrezas utilizadas en las PAU.

La cuestión fundamental es cómo y qué preguntas específicas de las tradicionalmente utilizadas sería recomendable transferir a un formato electrónico para su inclusión en una plataforma de evaluación de idiomas. De ser así, habría también que determinar qué tipos de interfaces se deberían diseñar para ajustar los cambios de estas pruebas a su nuevo entorno. Aunque bien es cierto que una gran parte de este trabajo ya ha sido realizado para otras

plataformas de evaluación como TOEFL o IELTS o, incluso, PLUS o BULATS, las características de las PAU y sus normas reguladoras hace imprescindible diseñar un nuevo sistema que se ajuste adecuadamente a los fines del examen de lengua extranjera de éstas.

2.2. Cambiando los parámetros de taxonomía

Es evidente que la presentación de ítems debe ser modificada cuando se trata de adaptarlos a interfaces informáticas. Además, en la actualidad, igual que sucede en cualquier plataforma de *e-learning*, todas pueden incorporar elementos de un repertorio multimedia que favorezca las respuestas de los alumnos. De acuerdo con esto, cabría determinar tres tipos de interfaces fundamentales:

1. *Interfaces de selección múltiple.* Son aquellas en las que el estudiante debe elegir la opción (u opciones, según se desee) que se considere correcta o más adecuada. Este tipo de interfaz puede utilizarse casi con cualquier tipo de pregunta de comprensión ya sea lectora u oral. Como podemos observar en el ejemplo de abajo, una pregunta puede tener una serie de respuestas asignadas de entre las que el alumno habrá que elegir la más adecuada.

Pregunta original abierta	Adaptación a cerrada
Where does Peter live?	Where does Peter live? a) Rome b) Luxemburg c) Southampton

Tabla 2. *Interfaz de selección múltiple.*

2. *Interfaces diseñadas para respuestas libres o alfanuméricas.* En principio, durante la década de los años 90, se utilizaron como cadenas predeterminadas que no permitían variaciones de los patrones preestablecidos. Generalmente se presentaban en cuadrados en los que la herramienta informática detectaba las variaciones indicando su incorrección. En la actualidad estos ítems se corrigen de manera automática gracias al uso de corpus lingüísticos como los usados para las correcciones gramaticales en procesadores de texto (p. ej. TOEFL) o por correctores que reciben archivos de texto. En la actualidad estas interfaces han ido adquiriendo una considerable importancia debido a la mejora en los sistemas de corrección automática, por la flexibilidad de las respuestas y por su sencillez de almacenaje y distribución.

3. *Interfaces diseñadas para preguntas orales.* Tienden a tener una presencia cada vez mayor. Lo más importante es que permitan la optimización de los elementos multimedia. En la actualidad, permiten incluir recursos, tareas e ítems audiovisuales centrados en la interfaz con las que los estudiantes actúan directa o indirectamente. En lo referente a los botones e iconos suelen ser ligeramente más complicados que los anteriores pero tanto el aspecto ergonómico como la claridad de uso obligan a una gran sencillez.

3. ADAPTANDO TAXONOMÍA Y TIPOS DE INTERFAZ

Así pues, descritos los tipos de ítem y las tres clases de interfaz genéricas, conviene adaptar los elementos descritos en la taxonomía básica presentada en el punto 2.1 a las interfaces del punto 2.2. Esta clasificación es fundamental para poder llegar al diseño de las plantillas informáticas dentro de las que se enmarcará cada tipo de prueba. Algunos de estos ítems pueden ser variables por lo que hemos tratado de marcar su mejor ubicación. Por ejemplo, en el caso de “Where does Peter live?” podríamos encuadrar el tipo de pregunta tanto en la interfaz de tipo 1 como en la de tipo 2 aunque, teniendo en cuenta las

características de las PAU (B1 en el Marco Común Europeo de Referencia para las Lenguas), probablemente sería más adecuada una interfaz de tipo 2 que en una de tipo 1. A continuación se indica a qué tipo de interfaz corresponden las tareas reseñadas en el apartado anterior.

Lista de tareas		Tipo de interfaz		
		Tipo 1	Tipo 2	Tipo 3
Comprensión lectora	Justificar el título del texto		X	
	Contestar preguntas abiertas		X	
	Escribir definiciones de palabras	X	X	
	Contestar preguntas de Verdadero / Falso	X		
	Resumir el texto		X	
	Relacionar información extraída del texto	X	X	
	Proporcionar sinónimos / antónimos	X	X	
	Selección múltiple	X		
	Rellenar huecos con información extraída del texto		X	
Realizar modificaciones gramaticales (varios)		X	X	
Identificar fonéticamente una palabra		X		
Redactar un texto			X	
Comprensión oral	Relacionar información	X	X	X
	Contestar preguntas de selección múltiple	X		

Tabla 3. Adaptación de los ítems utilizados en la actualidad en las PAU a las distintas interfaces de examen.

Dado que nos referimos en todo momento a un entorno asistido por ordenador a través de una plataforma de exámenes, podemos añadir a los ítems anteriores preguntas de respuesta oral como las que se muestran a continuación en la Tabla 4:

Lista de tareas		Tipo de interfaz		
		Tipo 1	Tipo 2	Tipo 3
Expresión oral	Respuestas orales a preguntas o elementos audiovisuales			X
	Descripción oral de objetos y situaciones			X

Tabla 4. Adaptación de ítems propuestos para pruebas de Selectividad a interfaces de examen.

No cabe duda de que esta adaptación conlleva el encuadrar los ítems en el tipo de plantilla informática que mejor se adapte a sus objetivos y forma de expresar la respuesta.

4. DISEÑO DE PLANTILLAS INFORMÁTICAS

Debido a que cada vez son mayores las posibilidades de explotación didáctica que ofrecen las tecnologías de la información y de las comunicaciones, los exámenes asistidos por ordenador pueden ser tan, e incluso más ricos pedagógicamente hablando, que los tradicionales sobre papel. El formato electrónico permite utilizar –y reutilizar, si así se desea– una gran cantidad de recursos multimedia que sirvan de apoyo para crear una gran variedad de modelos de examen destinados a evaluar eficazmente el nivel de conocimiento de una lengua extranjera. Desde secuencias de audio y vídeo para generar actividades de comprensión oral, hasta grabaciones de la propia producción oral de los estudiantes, pasando

por la escritura de texto libre como respuesta a preguntas abiertas. Hemos de distinguir, por lo tanto, tres tipologías de ejercicios bien diferenciados:

- a) Aquellos que permiten la *corrección automática* por parte del sistema (ejercicios de opción múltiple, relleno de huecos, relacionar ítems, etc.).
- b) Aquellos que requieren la *intervención humana* para su corrección (resúmenes, textos de opinión, preguntas abiertas).
- c) Aquellos cuya corrección puede realizar un programa informático comparando con *modelos predefinidos* (como por ejemplo cuando se evalúa la pronunciación de un estudiante mediante un programa de reconocimiento de voz por medio de representaciones gráficas).

Tal y como se ha demostrado en el *Proyecto INGENIO*¹, la programación de ejercicios basada en plantillas informáticas puede ser una vía muy útil, a la vez que rápida y eficiente, para generar bancos de ítems para evaluar una lengua extranjera (Gimeno 2002). En estos momentos existe un amplio abanico de posibilidades para generar ejercicios que supongan un importante avance respecto a los ejercicios que caracterizaban las primeras etapas de la enseñanza asistida por ordenador, definidos comúnmente con el apelativo “drill and kill”. Efectivamente, la tecnología ha avanzado exponencialmente desde los años 80 y los medios con los que contamos son infinitamente más fiables para la evaluación del conocimiento humano.

Desde la perspectiva del funcionamiento de un ejercicio, podemos distinguir una decena de tipologías de ejercicios que abarcarían una gran parte del espectro de actividades que se requieren para evaluar una lengua extranjera, y que se adecuarían a las necesidades de las PAU y, por consiguiente, a las programaciones de las asignaturas de lengua extranjera de 2º de Bachillerato.

5. RECOMENDACIONES PARA EL DISEÑO DE UN EXAMEN DE LENGUA EXTRANJERA ASISTIDO POR ORDENADOR

A continuación relacionaremos las tipologías de ejercicio según el tipo de corrección con la tipología de ejercicios enumerada en la Tabla 1.

Destreza	Tipos de ejercicio	Tipo de corrección
Comprensión lectora	Título justificado	Corrección automática
	Preguntas abiertas	Intervención humana
	Definiciones de palabras	Intervención humana
	Verdadero / Falso	Corrección automática
	Resumen del texto	Intervención humana
	Relacionar	Corrección automática
	Sinónimos	Corrección automática
	Elección múltiple	Corrección automática
Conocimientos Gramaticales	Rellenar huecos	Corrección automática
	Modificaciones Gramaticales (varios)	Corrección automática
Producción oral	Identificación fonética	Modelo predefinido
Producción escrita	Redacción	Intervención humana
Comprensión oral	Relacionar	Corrección automática
	Elección múltiple	Corrección automática

Tabla 5. *Tipologías de ejercicio según el tipo de corrección.*

Podemos, por lo tanto, comprobar que la gran mayoría de las tipologías de ejercicios que se viene utilizando en las PAU en la actualidad permite la corrección automática y sólo una tercera parte requeriría la intervención humana, siendo un solo caso en el que habría que

evaluar la articulación de palabras a través de un sistema electrónico de análisis de registros. Esta realidad nos hace constatar que no sería irreal crear un sistema en el que aleatoriamente se generaran exámenes a partir de bancos de ítems que, de forma proporcional y equilibrada, midiera los conocimientos de una lengua extranjera en las cuatro destrezas tradicionales y que naturalmente contara con un sistema informatizado para la corrección humana de ejercicios por parte de un experto en la materia. El sistema aquí propuesto reduciría sustancialmente el tiempo requerido para la corrección masiva de exámenes y, a la larga, abarataría también los costes de producción.

6. CONCLUSIONES

Como conclusión, quisiéramos apuntar simplemente que los exámenes de lengua extranjera de las PAU habrían de renovarse sustancialmente e incorporar todas las ventajas que suponen las nuevas tecnologías para medir de una manera más justa, equitativa y objetiva el nivel de conocimientos que de una lengua extranjera tienen los estudiantes que aspiran a formarse en una universidad española, máxime cuando se exige desde la propia Comisión Europea que haya la máxima convergencia posible entre unos estados y otros.

NOTAS

1. El *Proyecto INGENIO* fue llevado a cabo por el Grupo de Investigación CAMILLE de la Universidad Politécnica de Valencia

BIBLIOGRAFÍA

- Brindley, G., y H. Slatyer (2002). "Exploring task difficulty in ESL listening assessment". *Language Testing*. Special Issue: Interpretations, intended uses and designs in task-based language assessment, 19 (4): 369-394.
- Chapelle, C. A., J. Jamieson, y V. Hegelheimer (2003). "Validation of a web-based ESL test". *Language Testing*, 20(4) : 409-439.
- Christie, J., y S. Delaruelle (1997) *Assessment and moderation*. Sydney: National Centre for English Language Teaching and Research, Macquarie University.
- Davies, A. (2003) "Three heresies of language testing research". *Language Testing*, 20 (4): 355-368.
- Fulcher, G., y R.M. Reiter (2003) "Task difficulty in speaking tests". *Language Testing*, 20 (3): 321-344.
- Folse, K. S. (2006) *The Art of teaching Speaking*. Michigan: Michigan U. P.
- García Laborda, J. (2005) "Un análisis cualitativo de la Selectividad abierto a la esperanza" en Herrera Soler, H. y J. García Laborda, J. (Eds.) *Estudios y Criterios para una Selectividad de Calidad en el examen de inglés*. Valencia : Servicio de Publicaciones de la Universidad Politécnica de Valencia. 27-40.
- García Laborda, J. (2006a) "Analizando críticamente la Selectividad de inglés ¿Todos los estudiantes españoles tienen las mismas posibilidades?" *Tesol Spain*, 30.3: 9-12.
- García Laborda, J. (2006b) "¿Qué pueden aportar las nuevas tecnologías al examen de selectividad de inglés? Un análisis de fortalezas y oportunidades". *Revista de CC Educación*, 206: 151-166.

- Gimeno, A. (2002) *Call Software Design and Implementation: the template approach*. Valencia: Servicio de Publicaciones de la Universidad Politécnica de Valencia.
- Gimeno, A. (2005) "New challenges in developing an online CALL authoring shell, content manager and courseware: the INGENIO model". *The EUROCALL Review*, 7, 2-11. Accesible en Internet el 1 de diciembre de 2006 en <http://www.eurocall-languages.org/news/newsletter/7/index.html>
- Hambleton, R. y D. Xing (2004) *Computer-based test designs with optimal and non-optimal tests for making pass-fail decisions*. Research Report, University of Massachusetts, Amherst, MA.
- Iwashita, N., T. McNamara y C. Elder (2001). "Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design". *Language Learning*, 51 (3): 401-36.
- Kenyon, D. y V. Malabonga (2001) "Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments". *Language Learning and Technology*, 5 (2): 60-83.
- Mead, A. D. y F. Drasgow (1993) "Equivalence of computerized and paper-and-pencil cognitive ability tests: a meta-analysis". *Psychological Bulletin*, 114: 449-58.
- Veldkamp, B. P., y W. J. van der Linden (2000) "Designing item pools for computerized adaptive testing. In W. J. van der Linden & C. A. Glas (Eds.), *Computer adaptive testing: Theory and practice*. Boston: Kluwer Academic.
- Van Gelderen, A., y R. Oostdam (2005) "Effects of fluency training on the application of linguistic operations in writing". *L1-Educational Studies in Language and Literature*, 5(2): 215-240.
- Wang, T., y M. J. Kolen (2001) "Evaluating comparability in computerized adaptive testing: Issues, criteria and an example". *Journal of Educational Measurement*, 38(1): 19-49.
- Xing, D., y R. K. Hambleton (2004). "Impact of test design, item quality, and item bank size on the psychometric properties of computer-based credentialing examinations". *Educational and Psychological Measurement*, 64(1): 5-21.
- Yi, Q., T. Wang y J. Ban (2001). "Effects of scale transformation and test-termination rule on the precision of ability estimation in computerized adaptive testing". *Journal of Educational Measurement*, 38(3): 267-292.