

APLICACIONES DE LAS TECNOLOGÍAS EMERGENTES AL LANGUAGE TESTING

Sara Ferrari – Universidad Católica de Brescia, Italia

Las actuales tendencias para integrar el uso de las tecnologías de la información y de las comunicaciones en la enseñanza de las lenguas extranjeras ha llevado algunas realidades didácticas a diseñar las pruebas de evaluación lingüística en el ordenador. Las variables situacionales relacionadas con este tema han sido investigadas sobre todo por la lengua inglesa pero poco ha sido publicado sobre las pruebas informatizadas de español. En este artículo se relacionan los aciertos obtenidos gracias a una investigación realizada para comprobar la validez de la utilización del ordenador como soporte en la prueba de español para estudiantes universitarios italianos. A continuación se describen tres software para la evaluación de las destrezas en la lengua española y se proponen investigaciones para optimizar la aplicación de las tecnologías al language testing de español.

Nowadays most of the Linguistic Centres use new technology in order to teach foreign languages and often to value language competences. Since the '60 there have been a lot of researches about this topic, especially in English, but just a few of them are about Spanish Computer Assisted Language Tests. This article describes an investigation about the validity of using a CALL to test Spanish competencies in Italian university students (beginners). It follows an analysis of three software programmed to test language abilities in Spanish. It also proposes the increment of new investigations in order to optimise the application of new technology for Spanish language tests.

La creciente incorporación de las tecnologías informáticas en las disciplinas humanísticas ha brindado a los docentes la oportunidad de tener un acceso rápido, fiable y eficaz a un elevado número de datos reales y de herramientas útiles para la didáctica de una forma que hubiera resultado impensable hace tan sólo unas décadas. La literatura que trata estos temas afirma que, en línea general, los discentes tienen una actitud positiva hacia el ordenador en el aula pero todavía poco ha sido escrito sobre la reacción de los examinados que completan una prueba de lengua en el ordenador.

Los primeros intentos de informatizar las pruebas lingüísticas remontan a los años '60 pero su efectiva difusión tiene lugar sólo a principios de los años '80 con la llegada del *personal computer*. En la década sucesiva la mutua conexión de los ordenadores gracias a la *World Wide Web* integra, con su riqueza de enlaces, las potencialidades del ordenador que ya se ha convertido en una herramienta para todo el mundo. Aunque al principio la ventaja de una sumministrazione centralizada de la prueba quedaba reducida por una escasa interactividad, pronto se desarrollaron los programas implementando las posibilidades de interacción del examinando con la misma prueba.

En 1995 el invento de *JavaScript* facilitó la inserción, en las pruebas lingüísticas, de las respuestas breves también, de los ejercicios de reorden y de asociación. Un ulterior progreso se logró con la aparición de los tests

adaptativos utilizados incluso por Internet. La reciente realización de software programados con la *Item Response Theory* (teoría estadística sobre la cual se basan tales pruebas) de hecho ha permitido la suministración de ejercicios calibrados al nivel del examinando, reduciendo la tensión emotiva y estimando con mayor precisión la competencia lingüística.

El amplio consenso que han recibido las tecnologías emergentes también en la evaluación lingüística no debe crear la ilusión de que se pueda simplemente pasar una prueba preparada en papel, al ordenador. Al contrario, se hace necesario verificar la efectiva equivalencia de los dos formatos (tradicional, con papel y lápiz, y en el ordenador) e integrar en la didáctica clases en las aulas multimedia que contemplen también unas ejercitaciones al ordenador con pruebas similares.

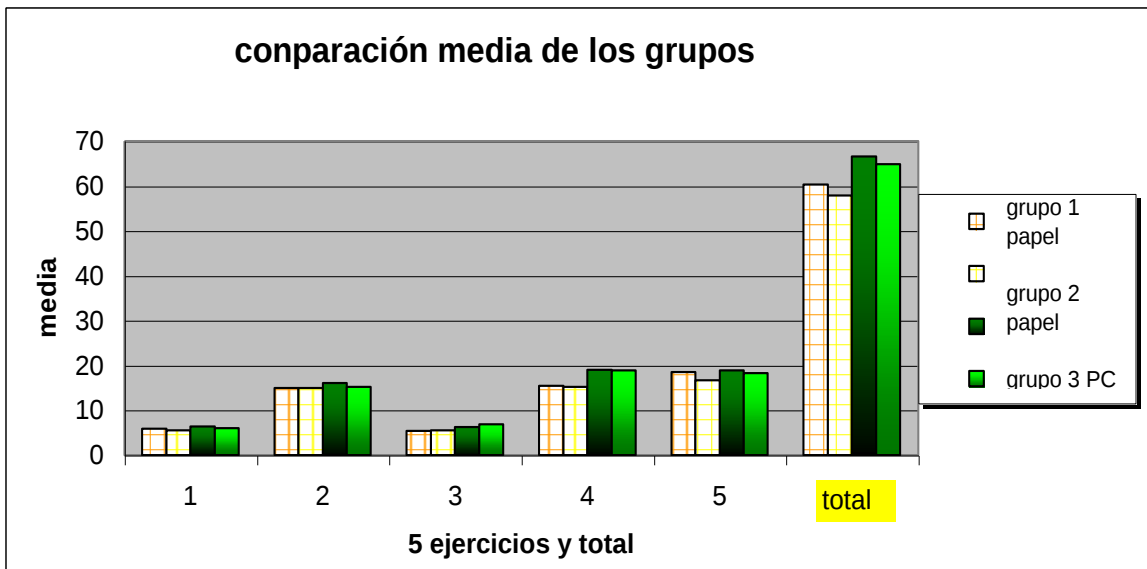
En efecto se ha demostrado que la familiarización con el medio informático y con el software elegido reduce la ansiedad que una situación de evaluación suele suscitar al encontrarse frente a instrumentos inusualesⁱ. En cuanto se haya comprobado la equivalencia es oportuno analizar las ventajas y los límites de una posible suministración de la prueba al ordenador para evaluar luego su efectiva utilidad de empleo. Sin embargo, después de haber llevado a cabo todas esas laboriosas operaciones, no hay que tener miedo en descartar la posibilidad de someter a los estudiantes a la prueba al ordenador. De hecho, en una didáctica cada vez más centrada en el contexto en el que vive el estudiante es necesario tener en cuenta que, aunque si por una sola variable situacional no es oportuna la utilización del ordenador, es mejor volver al tradicional formato papel y lápiz.

Por lo que he podido constatar, análogas investigaciones, relativas a la lengua inglesa, han sido realizadas por varios centros universitarios pero no han sido extendidas a la lengua española. Durante los tres años de doctorado de investigación en Lingüística Aplicada en la Università Cattolica de Milán, he intentado subvenir a esta carencia indagando sobre la efectiva equivalencia de una prueba de español suministrada con ordenador y una tradicional con papel y lápiz.

1 VALIDEZ DE UNA PRUEBA DE ESPAÑOL CON EL ORDENADOR

La investigación para verificar la validez de una prueba de español propuesta con el ordenador fue llevada a cabo con 104 estudiantes principiantes de facultades distintas de las de idiomas. Se trataba entonces de alumnos de unos 20 años, no excesivamente motivados en el aprendizaje del español pero sí en pasar el examen final. Después de algunos meses de clases tradicionales y de laboratorio multimedia, sometí una pequeña muestra de alumnos (11) a la pruebaⁱⁱ realizada en el ordenador y otra muestra (otros 11 estudiantes) a la misma prueba con papel y lápiz. Los resultados de este *pretest*ⁱⁱⁱ me ayudaron a optimizar no solo los ítems sino también la estructura de la prueba y su modalidad de suministración. Por lo tanto, poco después sometí la población real de los estudiantes que se presentaron al examen dividiéndolos en cuatro grupos de 26 alumnos. La prueba contaba con distintos tipos de ejercicios: el primero de comprensión oral y relleno, el segundo de comprensión escrita con verdadero o falso, el tercero de colocación de las preposiciones con elección múltiple, el cuarto de conocimiento de los verbos con relleno y el último de transformación del género y del número con producción escrita. Para completar la prueba los estudiantes tenían una hora de tiempo pero la mayoría terminó antes de los 50 minutos. Además, de momento que los que estaban delante del ordenador no podían revisar sus

respuestas después de haber terminado cada ejercicio, entregaban la prueba unos 10 minutos antes de los compañeros con papel y lápiz. Éstos, teniendo la oportunidad de corregirse cometieron más errores. De hecho, si bien la investigación quería demostrar que el ordenador no es un obstáculo para las pruebas de español, los resultados de esta experimentación mostraron que los estudiantes con papel y lápiz habían cometido más errores de los que usaban el ordenador. El gráfico representa la media de los 4 grupos evidenciando la diferencia entre ellas.



Otra diferencia significativa de las dos modalidades, desde el punto de vista del examinando, es la velocidad con la que él recibe la puntuación de la prueba. Mientras los alumnos que completaron la prueba con papel y lápiz tuvieron que esperar algunos días para después volver a la universidad y leer en los tablones sus resultados, sus colegas con el ordenador obtuvieron inmediatamente la puntuación de cada ejercicio con algunos consejos prácticos sobre cómo estudiar el argumento tratado. Veamos un ejemplo de lo que podía aparecer en la pantalla al terminar el ejercicio 4:

4. Ejercicio de relleno:

7 respuestas exactas sobre 14 (50%), has totalizado 15 puntos

Si tu puntuación es > de 19: tu resultado es bueno, ¡enhorabuena!

En caso contrario: repasa los verbos al presente del indicativo y al infinito,

practica la comprensión escrita y profundiza el léxico. Lee y escribe más.

¡Cuidado con los falsos amigos!

La misma ventaja debida a la rapidez de la corrección automática (ajena a las interpretaciones subjetivas) la encontré yo que tuve que dedicar una media hora para corregir cada prueba escrita en papel sin ocuparme de la puntuación de los exámenes al ordenador, ya listos en el archivo del software.

Además aunque en aquel entonces la versión del programa^{iv} no estaba dotada de herramientas estadísticas, ahora sí las tiene. Entre las demás opciones el software desarrolla el ítem análisis^v que después del examen tuve que hacer con Excel. El análisis de cada respuesta fue muy provechoso dado que proporcionó muchísimos datos útiles para la verificación de la actividad didáctica propuesta hasta entonces, estimulando al mismo tiempo el ajuste de la programación de las clases siguientes.

Sin embargo, además de las evidentes ventajas, la utilización del ordenador durante los exámenes presenta algunas desventajas. La principal está a cargo del docente que tiene que familiarizar bien con el software y proponer ejercicios que la máquina sea capaz de elaborar: especialmente respuestas cerradas. Cada ordenador del aula multimedia, que tiene sitio solo para un número limitado de estudiantes frente a los centenares que pueden presentarse en algunos exámenes, tiene que funcionar perfectamente y permitir al estudiante el acceso sólo al software que presenta la prueba y no a otras herramientas como diccionarios instalados en el ordenador o internet.

De todas maneras nos parece que el balance de la experimentación sea positivo dado que confirma la hipótesis inicial según la cual, en determinadas situaciones, el ordenador no es de obstáculo a la resolución de una prueba de español sino que brinda notables ventajas no solo en la corrección sino también en la posibilidad de enriquecer los ejercicios con imágenes y sonidos, cosa que era difícil proponer en la modalidad tradicional con papel y lápiz. Además, si consideramos los tests adaptativos, cuya duración es inferior aunque crezca la aproximación al real nivel de competencia comunicativa del discente, nos encontramos frente a otra constatación de las potencialidades que el elaborador puede poner al servicio del docente en fase de evaluación.

2 ALGUNOS SOFTWARE PARA LAS PRUEBAS DE ESPAÑOL INFORMATIZADAS

Visto que, como acabamos de demostrar, en determinadas situaciones el auxilio del ordenador para la resolución de una prueba de español, resulta ventajosa, vamos a describir algunas realidades donde se ha adoptado esa aplicación. Además del software que he utilizado en mi investigación: la *Banca Dati Multimediale* del Centro Linguistico Interfacoltà de la universidad Ca' Foscari de Venecia (que sigue siendo implementada por muchos profesores), en Italia se puede acceder a algunas pruebas de español suportadas por diversos programas. Entre los demás, analizo *L-Test*, *AlfaCert* y *DIALANG*, realizados con fines distintos pero fruto de años de experimentaciones.

2.1 L-TEST

L-Test es un instrumento de evaluación del conocimiento de la lengua española.^{vi} Consiste en la presentación de 120 preguntas de elección múltiple (a las cuales el usuario tiene que contestar dentro de 20-30 segundos), propuestas según los resultados obtenidos progresivamente, elegidas entre 1.000 distintas y ordenadas según un grado de dificultad creciente. El examen comprende dos sesiones de trabajo de media hora cada una, la primera dedicada enteramente a la comprensión oral y la segunda a la comprensión escrita y a la gramática. El programa evalúa el nivel de conocimiento de la lengua, según los parámetros de competencia lingüística establecidos por el Consejo de Europa. El método adaptativo elegido por *L-Test* difiere un poco del modelo *IRT* clásico. Realizado con fines comerciales, *L-Test* es particularmente útil en fase de selección inicial de un grupo de estudiantes en

los distintos niveles de conocimiento de la lengua. El software no resulta entonces adecuado para la evaluación final de un curso.

2.2 ALFACERT

Reconocido en toda Italia por los proyectos de integración de los aspectos tecnológicos con los pedagógicos en la didáctica de las lenguas extranjeras, el Centro Lingüístico de los Polos Científico-Didácticos de la Romagna ha realizado, entre otros, un software capaz de soportar pruebas de inglés, francés, español y alemán.

Las pruebas del proyecto *AlfaCert* Español B1 y Español B2 son paralelos al Certificado Inicial de Español (C.I.E.) y al Diploma Básico de Español (D.B.E.) acreditados por el Ministerio de Educación, Cultura y Deporte de España, y que corresponden a los niveles B1 (Umbral) y B2 (Avanzado) establecidos por el Consejo de Europa en el documento Marco de Referencia Europeo para la Enseñanza, Aprendizaje y Evaluación de Lenguas.

La primera acredita la competencia lingüística suficiente para desenvolverse en situaciones que requieren un nivel de comunicación elemental mientras que la segunda se refiere a situaciones de la vida cotidiana, en circunstancias normales de comunicación que no requieren uso especializado de la lengua.

Las dos pruebas están divididas en tres bloques que valoran distintas habilidades:

1. Comprensión de lectura y expresión escrita.
2. Contenidos gramaticales, léxicos y semánticos.
3. Comprensión auditiva y expresión oral.

La duración de cada bloque, que va de 40 minutos a dos horas, la división de las partes y la puntuación de la prueba han sido diseñadas con mucho cuidado para brindar al candidato las condiciones más favorables en la resolución y en la interpretación de la prueba.

2.3 DIALANG

DIALANG es el sistema de evaluación lingüística más importante basado en el Marco de Referencia Europeo. Ha sido desarrollado por más de 20 importantes instituciones con el soporte de la Comisión Europea y contempla 14 lenguas europeas. Con *DIALANG*, accesible gratuitamente por Internet^{vii} desde hace pocos meses, el usuario puede conocer sus puntos fuertes y débiles en una lengua extranjera y averiguar su nivel de conocimientos en la misma. En concreto, el programa ofrece pruebas de diferentes habilidades lingüísticas, cuidadosamente diseñadas y validadas, junto a una extensa gama de comentarios y consejos especializados sobre cómo mejorar los conocimientos.

El programa propone también actividades de auto-evaluación siguiendo las actuales teorías pedagógicas que afirman que ser capaz de valorar los propios conocimientos lingüísticos es una parte importante del aprendizaje de una lengua.

DIALANG no expide certificados pero brinda al usuario la posibilidad de averiguar qué nivel tiene de manera que pueda decidir cómo desarrollar mejor su dominio de la lengua. A continuación de la sección dedicada a la auto-evaluación, el test suministra pruebas adecuadas al nivel del usuario en la habilidad lingüística que él elige. Esto reduce las posibilidades de que le sean presentados ejercicios que sean demasiado fáciles o demasiado difíciles.

Otro recurso importante que no se encuentra en la mayoría de los demás sistemas de *language testing* es el feedback proporcionado. En concreto,

después de completar la prueba, el sistema ofrece al usuario una amplia gama de comentarios y consejos. Ante todo, una revisión de los resultados:

- el nivel en la escala del Consejo de Europa y una indicación de lo que significa;
- el resultado de la prueba de nivel y lo que implica con relación a la amplitud de los propios conocimientos de palabras;
- comentarios sobre la auto-evaluación;
- una revisión de las respuestas.

A continuación el programa da una serie de consejos individualizados:

- para la mayor parte de las habilidades, descripciones detalladas de lo que las personas del nivel del usuario pueden normalmente hacer en una lengua que están intentando dominar;
- descripciones de lo que estudiantes de otros niveles pueden hacer;
- estrategias específicas para mejorar los conocimientos lingüísticos;
- las razones por las que puede haber diferencias entre esa auto-evaluación y los resultados de la prueba para que el usuario sea capaz de interpretar mejor sus resultados.

Además de este útilísimo programa de *testing*, fruto de años de experiencias y experimentos con miles de estudiantes europeos, la web de *DIALANG* contiene enlaces a recursos para los estudiantes de lengua.

3 INVESTIGACIONES ACTUALES

La creciente demanda de certificaciones de conocimiento del español ha llevado a varias realidades a promover investigaciones para la aplicación de las tecnologías informáticas como herramientas útiles a tal fin. Sin embargo, al momento son todavía pocas las instituciones que utilizan el ordenador en fase de *testing* de la lengua española y muchas veces no comparten resultados y programas.

A nivel italiano, vista la exigencia de establecer un sistema de certificación que tenga una tipología estructural única para todos los centros lingüísticos, la Asociación Italiana de los Centros Lingüísticos Universitarios^{viii} ha redactado un proyecto para la certificación de las competencias lingüísticas en las universidades italianas (CERCLU) que contempla pruebas informatizadas.

Además de una certificación común quedan todavía por aclarar muchas dudas relativas a la utilización de las nuevas tecnologías para la evaluación lingüística. Entre otras habría que investigar sobre la efectiva autenticidad y seguridad de las pruebas informatizadas; al mismo tiempo habría que aprovechar mejor del análisis estadístico de los resultados de las pruebas actualizando los modelos psicométricos relativos a la evaluación lingüística.

Para concluir recuerdo que a la base de estas investigaciones queda la voluntad institucional y docente de ponerse al día para ofrecer con profesionalidad cursos para adquirir, integrándolas, dos destrezas entre las más requeridas: la competencia informática y el conocimiento de las lenguas extranjeras.

BIBLIOGRAFÍA

- Gagliardi, C. 1997. *Metodologia della ricerca in linguistica inglese*. Pescara: Libreria Universitaria Editrice.
- Alderson, C. 2000. "Technology in testing: the present and the future". *System* 28: 25-34.
- Bachman, L. F., y Palmer A. S. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
- Dean Brown, J. 1997. "Computers in Language Testing: present research and some future directions". *Language Learning and Teaching* 1[Documento de Internet disponible en <http://polyglot.cal.msu.edu/llt/vol1num1/brown/default.html>].
- Fernández Álvarez, M. 2002. "Las nuevas tecnologías en el área de testing: Tests Informatizados y Tests Adaptativos Informatizados (TAIs)". *Nuevas Tecnologías y Nuevos Retos Educativos, Actas de Conference and Media Days*, Granada, 16-21 octubre 2002. [Documento de Internet disponible en www.ugr.es/~icem2002/Ponencias/FernandezAlvarez.PDF].
- Ferrari, S. 2001. *Validità di test per la lingua spagnola su supporto informatico*. Tesis doctoral. Milán: Università Cattolica del Sacro Cuore.
- Guerrero, M. y Del Vecchio, A. 1996. "Handbook of Spanish Language Proficiency Tests", Albuquerque: EAC West, New Mexico Highlands University. [Documento de Internet disponible en <http://www.ncbe.gwu.edu/miscpubs/eacwest/slptests.htm>].
- McNamara, T. 2000. *Language Testing*, Oxford: Oxford University Press.
- Porcelli, G., y Dolci, R. 1999. *Multimedialità e insegnamenti linguistici*, Turín: UTET Libreria.
- Russell, M. y Haney, W. 1997. "Testing writing on computers: an experiment comparing student performance, on tests conducted via computer and via paper-and-pencil". *Education Policy Analysis Archives* 3.
- San Vicente, F. 2003. *AlfaCert. Alfabetizzazione e Certificazione Linguistica. L'autovalutazione Linguistica nel Quadro Comune Europeo di Riferimento*. Bologna: CLUEB.
- Schiavinato, G. y Ongaro, V. 1999. "Software per il testing adattivo multimediale". Venecia: Il Progetto Itals.
- Taylor, C., Jammieson, J., Eignor, D. y Kirsch I. 1998. "The relationship between ordenadorfamiliarity and performance on ordenadorbased TOEFL test tasks". *Research Reports* 77-89.

ENLACES:

Applied Linguistics Research Group, <http://www.surrey.ac.uk/ALRG/index.html>
Association of Language Testers in Europe, <http://www.alte.org/>
CLI, <http://www.cli.unive.it>
CLIRO, <http://www.cliro.unibo.it>
DIALANG, <http://www.dialang.org>
ILTA, <http://www.surrey.ac.uk/ELI/ilta.html>
Language Testing, <http://www.ling.lancs.ac.uk/staff/visitors/kenji/test.htm>
Assessment & Evaluation on the Internet, http://www.cua.edu/www/eric_ae/intass.html

ⁱ Para averiguar si existe una relación entre la familiaridad con el ordenador y los resultados de la prueba, los colaboradores del *TOEFL* (Taylor C., Jammieson J., Eignor D., Kirsch I., 1998) llevaron a cabo una investigación *ad hoc* confrontando la *performance* en inglés de alumnos con distintos niveles de competencia informática. Los datos relevados de esta investigación llevaron a la conclusión evidente de que no hay diferencias significativas en los resultados obtenidos en

la prueba completada con ordenador según el conocimiento informático si todos los examinandos han aprendido antes cómo utilizar ese software.

ⁱⁱ Para realizar la prueba de español seguí, entre otros, los útiles consejos de Bachman (1996), mientras para la modalidad de la experimentación sobre los dos grupos de estudiantes universitarios utilicé el esquema propuesto por Gagliardi (1997).

ⁱⁱⁱ En las experimentaciones con personas hay que someter siempre a la prueba una muestra significativa de individuos antes de realizar la investigación sobre la población real. Después de haber optimizado el proyecto de la misma prueba gracias a los resultados de ese *pretest*, tenemos la garantía de una mejor sumministrazione de la prueba con la población objetivo.

^{iv} El software *Banca Dati Multimediale*, utilizado para la prueba de español de mi investigación, fue amablemente ofrecido por el Centro Lingüístico Inter facultad de la Universidad Ca' Foscari de Venecia. Se trata de un programa, realizado por Gino Schiavinato, que propone tests adaptativos en varias lenguas pero que se puede utilizar también con pruebas no adaptativas.

^v El ítem análisis trata de proporcionar datos sobre los resultados de cada ítem ayudando así al profesor a eliminar los ítems demasiado fáciles o difíciles y a mejorar los ambiguos.

^{vi} Editado por TEAM TEACHING, *L-Test* está constituido por un CD-rom y un manual de uso.

^{vii} <http://www.dialang.org>

^{viii} Associazione Italiana dei Centri Linguistici Universitari: <http://www.aiclu.it>