

Ministerio
de Educación, Cultura
y Deporte

MOSAICO

REVISTA PARA LA PROMOCIÓN Y APOYO A LA ENSEÑANZA DEL ESPAÑOL

35/2017



Directora de la publicación:

Guadalupe Melgosa Fernández
(Consejera de Educación en Bélgica,
Países Bajos y Luxemburgo)

Coordinador:

José Ángel Piña Sánchez
(Asesor Técnico Docente)

Revisión y edición:

José Ángel Piña Sánchez
Clara Migúelez Sotillo

Diseño y maquetación:

Verónica Pérez Pérez

Colaboran en esta Publicación como autores de los artículos y recursos presentados:

Lorena Albert Ferrando
Blanca Berjano Rodríguez
Carmen Daza Márquez
Nathalie Mus
July De Wilde
Begoña García Migura
Yeray González Plasencia
Vicenta González
Ángel Hernando Domingo
Helena Legaz Torregosa
Paula Lorente
Matilde Martínez Sallés
Susana Martínez
Enrique Morales Lara
Antonio Ramos Álvarez

Catálogo de publicaciones del Ministerio mecd.gob.es

Catálogo general de publicaciones oficiales publicacionesoficiales.boe.es

El fichero digital de este boletín puede descargarse desde la página web de la Consejería de Educación:

www.educacion.gob.es/belgica/publicaciones-materiales/publicaciones.html

Las ilustraciones son de dominio público, están libres de derechos de autor o constan de los permisos adecuados para su reproducción.

Se permite la copia total o parcial de este boletín siempre y cuando:

- se cite la procedencia;
- no se proceda a cobro o contraprestación de ningún tipo;
- se informe a la Consejería de Educación en Bélgica, Países Bajos y Luxemburgo de la incorporación, ofreciendo los datos que permitan la vigilancia del cumplimiento de lo expuesto en el punto anterior.

La Consejería no se responsabiliza de las opiniones aparecidas en los artículos de esta revista.

[Imágenes de la portada y contraportada: barrio "Les Marolles" en Bruselas]



MINISTERIO DE EDUCACIÓN,
CULTURA Y DEPORTE
Subsecretaría
Subdirección General de Cooperación Internacional
Edita:

© SECRETARÍA GENERAL TÉCNICA
Subdirección General de Documentación y Publicaciones

Edición: Diciembre de 2017

NIPO: 030-14-283-8 (electrónico)

EMBAJADA DE ESPAÑA
CONSEJERÍA DE EDUCACIÓN EN BÉLGICA,
PAÍSES BAJOS Y LUXEMBURGO

Bd. Bischoffsheim 39, B.15-16
1000 Bruselas
TEL: (+32) (0)22232033
FAX: (+32) (0)22232117

www.mecd.gob.es/belgica

consejeriabelgica.be@mecd.es

asesoriabelgica.be@mecd.es

INDICE

PRÓLOGO	4
Guadalupe Melgosa Fernández Consejera de Educación en Bélgica, Países Bajos y Luxemburgo	
ARTÍCULOS	
<i>“La autonomía del aprendizaje y cómo fomentarla en las destrezas orales”</i> Carmen Daza Márquez y Nathalie Mus	6
<i>“Formatos de evaluación en comprensión auditiva”</i> July De Wilde	16
<i>“Cómo fomentar la reflexión intercultural en la enseñanza de ELE: el concurso audiovisual Spanish in”</i> Yeray Ramos Plasencia y Antonio Ramos Álvarez	28
<i>“Camino de ida y vuelta en formación continua de ELE: de la teoría a la práctica pasando por la refle”</i> Vicenta González y Paula Lorente	42
<i>“La práctica reflexiva: propuestas metodológicas en ELE Lovaina”</i> Begoña García Migura y Helena Legaz Torregosa	55
<i>“¿Tareas finales para evaluar las competencias?”</i> Matilde Martínez Sallés	67
PROPUESTAS DIDÁCTICAS	
<i>“Aprendizaje Integrado de Contenido y Lenguas Extranjeras en contexto de inmersión: la leyenda “El be”</i> Lorena Albert Ferrando	74
<i>“Los tipos de “se”: propuestas para las clases de ELE”</i> Blanca Berjano Rodríguez	81
<i>“TelELEdiario: producción de un Telediario en 4’ en clase de ELE”</i> Ángel Hernando Domingo	91
<i>“El uso de la ropa en el aula de ELE”</i> Susana Martínez	95
PARA TENER EN CUENTA	
<i>“El español y los signos de interrogación y exclamación, del dólar y de la arroba”</i> Enrique Morales Lara	98
INFORMACIONES	100

Formatos de evaluación en comprensión auditiva

Este trabajo analiza la correlación positiva entre dos factores a la hora de evaluar la comprensión audiovisual: el formato de la evaluación (tipo test o preguntas abiertas) y los resultados.

Se comparan dos pruebas de un mismo ejercicio realizado en clase de ELE (N=66): una primera de preguntas abiertas (N=17) y una segunda que incluye preguntas de selección múltiple e ítems dicotómicos (N=49). En este último caso se analiza además la correlación entre el número de visionados y los resultados obtenidos. Tras revisar la literatura se presentan dos hipótesis: (1) los estudiantes sometidos a la prueba tipo preguntas abiertas puntúan menos que los otros; (2) el número de visionados influye en los resultados. En contra de la primera hipótesis planteada, los participantes en la prueba tipo test no obtienen mejores resultados. La segunda hipótesis sí se comprueba aunque el turno adicional favorece en primer lugar a los estudiantes menos fuertes.

El presente estudio proporciona datos empíricos para el análisis de las posibles incidencias que tienen el formato de prueba y el número de repeticiones en los resultados de pruebas de comprensión auditiva en ELE, un tema de investigación poco explorado, en particular en comparación con el inglés como lengua extranjera.

July De Wilde / Universidad de Gante, Departamento de Traducción, Interpretación y Comunicación, sección de español / july.dewilde@ugent.be

1. Introducción

En este trabajo se analiza la correlación positiva entre dos factores a la hora de evaluar la comprensión audiovisual: el formato de la evaluación (tipo test o preguntas abiertas) y los resultados obtenidos de la misma. Por tanto, se comparan diversos formatos utilizados para evaluar la comprensión audiovisual (CA¹ de ahora en adelante) del español como lengua extranjera (en lo que sigue ELE).

Para ello comparamos dos pruebas de un mismo ejercicio de CA: una primera de preguntas abiertas y una segunda que incluye preguntas de selección múltiple e ítems dicotómicos. Para esta segunda, la prueba tipo test, también analizamos la correlación positiva entre el número de visionados y los resultados obtenidos.

En la primera parte situamos en qué contexto se enmarca la investigación: ¿cuál es el contexto curricular y el perfil de los participantes? Después, y tras revisar la literatura sobre el tema, formulamos las hipótesis a la base de la presente investigación. En la última parte detallamos el diseño de la prueba antes de presentar los resultados y la discusión.

1. En lo que sigo hablamos únicamente de la *comprensión oral sin colaboración* (Buck 2001: 98), es decir de una situación de audición en que el oyente no interviene en la construcción del discurso. Se distingue por lo general este tipo de audición de aquella otra llamada *con colaboración* donde la interacción es central. Esta última suele evaluarse junto con la expresión oral.

2. Metodología: Contexto, grupos de análisis y procedimientos

Este estudio ha sido llevado a cabo en el marco de una asignatura de CA impartida a estudiantes de ELE de nacionalidad belga cuya lengua materna (L1) es el neerlandés. Los estudiantes cursan la carrera de Lingüística Aplicada en el Departamento de Traducción, Interpretación y Comunicación de la Universidad de Gante (Bélgica). Los 66 estudiantes que participaron en la investigación se especializan en dos lenguas. Han elegido el español, junto a una segunda LE, siendo ésta francés (23 estudiantes), inglés (28 estudiantes) o alemán (7 estudiantes). Cuando tuvo lugar la prueba (noviembre de 2012), el perfil medio de los estudiantes era el de estudiantes de segundo curso, de edades comprendidas entre los 19 y los 20 años y que llevaban estudiando español un poco más de un año completo, a un ritmo intensivo de 8 horas por semana.

En cuanto a la recopilación de datos, los 66 participantes visionaron un mismo fragmento sobre el Museo de la Trufa en Navarra y realizaron las mismas actividades de preaudición (véase el apéndice 1 para una breve descripción de estas actividades). Después fueron sometidos a evaluaciones con formato diferente. El objetivo era ver si dicho formato de evaluación ejercía una influencia en los resultados de sus tareas. De discusiones previas con los estudiantes resultaba que la mayoría (56 de los 66) prefiere el formato de selección múltiple al de las preguntas abiertas ya que perciben dicho formato como marcadamente “más fácil”.

17 participantes realizaron una *prueba A* de preguntas abiertas, tales como las dadas en el apéndice 2. Otros 49 realizaron una *prueba B* agrupando preguntas de selección múltiple e ítems dicotómicos, tales como las mostradas en el apéndice 3. Los resultados de la prueba B han sido comparados con los de la prueba A. Aparte del formato, también se verificó el posible efecto que puede tener el número de visionados. Así, de los estudiantes sometidos a la prueba B, 31

visionaron el fragmento dos y 18 estudiantes tres veces. Este último grupo no consultó el formulario de preguntas sino hasta finalizado el primer turno.

En lo que sigue hacemos un repaso de la literatura existente sobre el tema que nos ocupa, y a continuación explicitamos las dos preguntas que constituyen la base de la presente investigación. Después comentaremos la metodología y presentaremos los resultados. En la conclusión volvemos críticamente sobre los resultados.

3. Marco teórico

Considerada durante años como una destreza pasiva, la comprensión audiovisual se considera hoy como un elemento clave en la adquisición de una lengua extranjera (Krashen 2002): implica actividades mentales tan complejas que el que escucha es tan activo como el que produce el mensaje (Vandergrift 2007). La comprensión oral es el resultado de la interacción entre un número de fuentes de información que incluyen “el input acústico, el conocimiento lingüístico, el conocimiento del mundo, el contexto de comunicación y el uso de los oyentes de cualquier información a su disposición que les parezca relevante para ayudarles a interpretar lo que el hablante está diciendo” (Pastor Villalba 2009: 2).

“Considerada durante años como una destreza pasiva, la comprensión audiovisual se considera hoy como un elemento clave en la adquisición de una lengua extranjera (Krashen 2002): implica actividades mentales tan complejas que el que escucha es tan activo como el que produce el mensaje”

Es decir, en la CA intervienen factores de índole muy diferente: cognitiva, lingüística, propia del aprendizaje de la lengua extranjera, sociocultural, psicológica, material (las grabaciones), lo que hace difícil la enseñanza, pero también, y sobre todo, la evaluación de esta destreza comunicativa. La enseñanza y la evaluación de la CA se justifican sin embargo en tanto que forman parte esencial de la competencia comunicativa (Consejo de Europa 2002: 60), por “su potencial efecto positivo en las prácticas de clase” (Alderson y Bachman, en Buck 2001: xi) o por el desempeño de los aprendientes en futuras situaciones comunicativas fuera del aula.

A pesar de su evidente importancia, la CA sigue investigándose esencialmente en relación con especulaciones teóricas, lo que se debe en parte a su carácter indirecto. En efecto, tratándose de la destreza menos explícita, resulta difícil medirla; intervienen muchas otras subdestrezas que constituyen variables difícilmente controlables en un estudio empírico. Si resulta factible controlar variables relacionadas con el tipo de texto (velocidad del habla, entonación, acento, rasgos lingüísticos como el vocabulario o la sintaxis), o con la *situación externa* (ambiente, ruido exterior), no es igualmente fácil controlar variables relacionadas con la cognición del aprendiente (su conocimiento previo, su memoria) o con su *metacognición* (la capacidad del aprendiente de evaluar el nivel de dificultad del texto audiovisual y de adaptarse a él²). La interconexión de todas estas subdestrezas también explica por qué es difícil evaluar un ejercicio de CA, la cual siempre será indirecta: al estudiante se le pide

“La interconexión de todas estas subdestrezas también explica por qué es difícil evaluar un ejercicio de CA, la cual siempre será indirecta: al estudiante se le pide que realice otras actividades que permitan hacer patente su comprensión. Así, en un ejercicio de selección múltiple interviene también, en gran medida, la habilidad lectora del aprendiente.

Por otra parte, cuando se trata de preguntas abiertas, al estudiante se le pide que demuestre sus habilidades escritas. Es imposible, pues, aislar las subestrategias que miden la comprensión de un texto en sentido estricto”.

que realice otras actividades que permitan hacer patente su comprensión. Así, en un ejercicio de selección múltiple interviene también, en gran medida, la habilidad lectora del aprendiente. Por otra parte, cuando se trata de preguntas abiertas, al estudiante se le pide que demuestre sus habilidades escritas. Es imposible, pues, aislar las subestrategias que miden la comprensión de un texto en sentido estricto (Buck 2001: 100). En términos de Pastor Villalobos (2009: 7) la relación entre “la actuación y las habilidades que queremos evaluar tiende a ser débil e incierta”.

No sólo las *formas* de medir, sino también el *formato* de las evaluaciones se presta a debate. Las preguntas abiertas permiten ir más allá de la superficie del texto y fomentar un aprendizaje más profundo (Herbert & Burt 2002). Redactar los ítems no suele presentar demasiadas dificultades pero corregirlos es costoso. No sólo el tiempo dedicado, sino también la puntuación puede hacer difícil el uso de las preguntas abiertas, en tanto que implica una puntuación subjetiva³ y es necesario desarrollar criterios lo más objetivos posibles (Brown & Hudson 2002).

En cuanto a las preguntas de selección múltiple, sus ventajas son bien conocidas: son altamente fiables y su estructura cerrada permite una puntuación objetiva (Bailey 1998). También sus desventajas han sido documentadas: escribir ítems eficaces no es tarea fácil y además, no toda la información se presta a interpretación unívoca (véase Burton 2005). Los detractores también objetan que se evalúa la capacidad de reconocer más que de conocer y que, además, no hay manera de excluir el azar (Currie & Chiramanee 2010). Sería además un instrumento de evaluación superficial en tanto que no permite reconstruir el origen del error: resulta difícil determinar si una respuesta errónea se debe a la comprensión deficiente de algún texto audiovisual o, en cambio, guarda relación con una interpretación errónea de las posibles respuestas presentadas en el formulario (Brown & Hudson 2002). Otros refutan que es un formato poco auténtico: en situaciones comunicativas auténticas no se nos presentan nunca alternativas de las cuales escogemos la mejor opción.

2. Para una visión general de todas las variables implicadas, véase Buck (2001).

3. La evaluación es objetiva si no se requiere el juicio de un evaluador, como en un ejercicio de selección múltiple, en el que una respuesta correcta equivale, por ejemplo, a un punto. La evaluación subjetiva, en cambio, es una valoración que realiza un examinador.

Los ítems dicotómicos de tipo verdadero/falso presentan en parte las mismas ventajas e inconvenientes. Al igual que en el caso de las preguntas de selección múltiple, el carácter cerrado de los ítems dicotómicos permite una puntuación rápida y objetiva. A diferencia de las preguntas de selección múltiple, su redacción no suele ser tan difícil. Pero así como con las preguntas de selección múltiple, resulta difícil rastrear el origen exacto del error. Para disminuir el efecto del azar, se suele pedir al estudiante que justifique su respuesta, sobre todo cuando las frases son falsas.

De lo anterior se puede concluir que los distintos formatos tienen a la vez ventajas y desventajas. Se han investigado poco los posibles efectos de formas alternativas de medir la CA, incluso si se acepta generalmente que los estudiantes procesan de manera diferente según estén confrontados con un formato de evaluación u otro (Hansen y Jensen 1994). Brindley y Slatyer (2001) concluyen que el grado de dificultad de diversas tareas evaluativas de CA se ve influenciado por el formato de la evaluación. Teng (1998) ha comprobado que el formato de la evaluación, pero también el tipo de texto influyen considerablemente en la puntuación.

Cheng (2004), tras comparar tres formatos de evaluación –(1) preguntas abiertas, (2) preguntas de selección múltiple y (3) preguntas de completación de selección múltiple⁴– concluye que el formato interviene en la puntuación: las mejores notas se obtienen con las preguntas de completación de selección múltiple y las peores en el caso de las preguntas abiertas. Estas conclusiones concuerdan con las de Berne (1993, en Brindley y Slatyer 2002), que muestra mejores resultados para preguntas de selección múltiple y peores resultados en el caso de preguntas abiertas o algún tipo de completación parcial.

También el factor psicológico interviene: muchos aprendientes se sienten más cómodos con preguntas (semi)cerradas porque reconocen la respuesta correcta (o su ausencia en los ítems dicotómicos falsos) en una de las opciones impresas en el formulario (Traub y MacRury 1990). Esta impresión también ha sido comprobada en el aula: antes de conocer el formato de la prueba, 56 de los participantes de la presente investigación habían manifestado su preferencia por el formato cerrado tipo test (cf. prueba B).

En cuanto al número de repeticiones del texto audiovisual, existe cierto consenso sobre la correlación

positiva entre el número de turnos y la actuación de los estudiantes. Sakai (2009) demuestra que tras repetir un fragmento, se facilita considerablemente la comprensión. Aunque muchos investigadores coinciden en que se debe repetir el fragmento, no existe consenso sobre el número de repeticiones necesarias. También la cuestión de si leer las preguntas de antemano lleva a mejores resultados o no sigue siendo objeto de debate. Buck (1990) encontró que aunque los examinados afirman tener una mejor comprensión al conocer las preguntas de antemano, los resultados puntuados no lo reflejan. Se trata entonces más bien de un efecto positivo psicológico. Estos datos no coinciden con los de Sherman (1997), quien afirma una mejora más fuerte en la actuación cuando se mostraban las preguntas a los examinados después del primer turno, y antes del segundo.

Los estudios mencionados anteriormente se ocupan esencialmente de la CA en la clase de inglés como LE. El presente trabajo analiza si los resultados también se aplican al ejercicio de CA en clase de ELE, verificando posibles efectos que puede tener el formato de la evaluación en el rendimiento del aprendiente de ELE. Estas son las dos hipótesis predictivas que guían la presente investigación:

1. Los estudiantes que participan en la prueba A (preguntas abiertas) obtienen un puntaje menos elevado que los que participan en la prueba B (selección múltiple + verdadero / falso);
2. El número de repeticiones tiene una incidencia en la actuación. Los participantes en la prueba B que visionaron el fragmento tres veces y consultaron las preguntas después del primer turno obtienen mejores resultados que los que visionaron dos veces y consultaron el formulario antes del primer visionado.

“Los estudios mencionados anteriormente se ocupan esencialmente de la CA en la clase de inglés como LE. El presente trabajo analiza si los resultados también se aplican al ejercicio de CA en clase de ELE, verificando posibles efectos que puede tener el formato de la evaluación en el rendimiento del aprendiente de ELE”.

4. En el original “open-ended questions”, “traditional multiple choice” y “multiple choice cloze” respectivamente (Cheng 2004).

4. Diseño y condiciones situacionales de las pruebas⁵

Al determinar el texto audiovisual que se empleará para la evaluación, es importante plantearse primero qué se quiere evaluar y para qué fin. La prueba realizada para esta investigación se hizo en el marco de una asignatura de CA de 24 horas lectivas presenciales. Los progresos de los estudiantes se evalúan de modo formativo mediante dos pruebas de evaluación continua y la preparación de un portafolio. El timing de las pruebas formativas es importante: la primera se realiza a las cuatro semanas de clase, la segunda durante la penúltima semana de la asignatura (después de diez semanas de clase). Las tareas incluidas en el portafolio tienen una serie de objetivos precisos: (i) el registro del proceso de aprendizaje mediante documentos realizados de forma autónoma, (ii) incrementar la reflexión mediante la conciencia metalingüística del aprendiente e (iii) incrementar la autoevaluación fomentando su responsabilidad en el proceso de aprendizaje.

Al realizar la prueba los aprendientes sabían que se trataba de un ejercicio no puntuado y se les impuso trabajar de forma individual. A los estudiantes se les comunicó que simularíamos las condiciones de trabajo de una prueba formal que tendría lugar la semana después. De esta manera hemos querido reducir posibles efectos psicológicos, como son la ansiedad y el estrés, de los que se ha afirmado que son más pronunciados para la CA que para otras actividades en clase de lengua (Elkhafaifi 2005).

Como input lingüístico elegimos un fragmento de una duración de un minuto y 22 segundos, sacado del programa semanal Panorama y emitido en la televisión española (TVE) en 2010. El texto trata del Museo de la Trufa en Navarra que en 2010 acababa de abrir sus puertas. La elección de este texto audiovisual respondió a dos condicionantes.

Primero, y en vistas del perfil de los participantes, optamos por un tema relativamente desconocido para no favorecer ni perjudicar determinados aprendientes que tendrían menos o más preconocimientos sobre el tema. Como bien señala Buck (2001), es difícil controlar los conocimientos previos en situaciones de prueba, pero una estrategia puede ser la elección de un tema del que se considera que es desconocido por la (casi) mayoría de los estudiantes. De esta forma

se evitan discrepancias marcadas dentro del grupo. Para compensar en lo posible dificultades debidas a esta falta de conocimiento previo diseñamos dos actividades de preaudición que aseguraran la suficiente familiarización con el tema (véase también el apéndice 1). Estas actividades las realizamos de forma idéntica en todos los grupos. Incluyen un apoyo pictográfico que activa el léxico y genera la emisión de hipótesis. Al desarrollar dichas actividades partimos de la idea de la rentabilidad de las imágenes: la proyección de imágenes sirve de patrón para interpretar conceptos y estructuras conceptuales abstractas y vinculan la información a aspectos emocionales o afectivos que otorgan viveza e intensidad. En nuestro caso, tratándose de “la trufa”, un vocablo que se presta a ambigüedad (¿se trata del chocolate o del hongo?), y de un campo semántico poco familiarizado, consideramos necesario incluir un apoyo visual.

El segundo factor que motivó la elección del texto audiovisual guarda relación con sus características gramaticales (velocidad de habla, acento no marcado de los locutores, pocos falsos comienzos, muchas repeticiones), discursivas (clara separación de los puntos esenciales y los detalles facilitando la comprensión de la estructura), pragmáticas (ausencia de inferencias contextualizadas, de lenguaje figurado) y sociolingüísticas (registro neutro, pocas expresiones idiomáticas).

Tratándose de formatos muy desiguales, administramos el tiempo de la prueba de modo diferente. Los estudiantes sometidos a la prueba A recibieron tres minutos para leer las preguntas y anticipar posibles respuestas. También las instrucciones orales incluían información en este sentido, señalando a los participantes de antemano el tiempo previsto para y entre cada uno de los tres visionados y motivándolos a que anticiparan en lo posible respuestas plausibles. Se les prohibió tomar nota durante el primer visionado, pero terminando este disponían de tres minutos donde sí podían tomar apuntes o formular ya partes de las respuestas. A partir del segundo visionado tenían la posibilidad de tomar apuntes. Entre el segundo y tercer visionado disponían de tres minutos. Terminado el último turno, se les concedió un periodo de cinco minutos para finalizar las respuestas y pasarlas todas a limpio.

5. La discusión de las características de la prueba está basada en parte en el marco de Buck (2001: 107 y 116-153) y Pastor Villalba (2009) para la terminología castellana.

Los participantes están familiarizados con este tipo de procedimiento ya que durante el curso se practican con frecuencia ejercicios con este mismo formato. La prueba A agrupa preguntas “locales”⁶ lo suficiente espaciadas para dejar a los estudiantes el tiempo necesario. El orden de los datos solicitados en la ficha respeta la cronología de la secuencia de vídeo. En los casos ambiguos, añadimos entre paréntesis el número de elementos necesarios para que la respuesta sea correcta. Considerando que se trata de una evaluación subjetiva, se pretendía reducir el riesgo de parcialidad o subjetividad decidiendo de antemano, y para cada pregunta separada, qué respuestas eran válidas y qué elementos se consideraban necesarios o suficientes⁷. Incluimos una segunda medida: corregimos primero las pruebas A antes de corregir las preguntas cerradas de la prueba B.

En la prueba B se incluyeron preguntas locales orientadas a evaluar, entre otras cosas, la comprensión gramatical. Por ejemplo, para interpretar correctamente la pregunta A1 (véase el apéndice 3) es necesario que el estudiante (re)conozca la perífrasis acabar de + infinitivo al principio del texto audiovisual. En cambio, para A2 resulta esencial la discriminación de formas acústicas muy cercanas. También se incluyeron preguntas orientadas a evaluar la comprensión de material cifrado (véanse por ejemplo A5, B3 y B6). Además de eso, se incluyeron preguntas orientadas a evaluar la capacidad de inferir información correcta no mencionada de manera explícita (B1, B5 y B7). Para cada pregunta de selección múltiple se ofrecían dos distractores aparte de la clave. Respetamos en lo posible las sugerencias deducibles por el tipo de respuestas ofrecidas: la posición de la respuesta clave frente a los distractores se determinó al azar, todos los distractores eran plausibles sintáctica y semánticamente, no se incluyeron fórmulas de tipo “todas las anteriores” o “ninguna de estas”. En las preguntas de selección múltiple, se pidió al estudiante que marcara la respuesta “más adecuada”, que no la correcta. Se fijó y se comunicó de antemano el tiempo previsto para leer las preguntas (cuatro minutos), entre cada turno (tres minutos) y después del último turno (dos minutos). A los estudiantes que visionaron el fragmento tres veces se les mostró el formulario solo después del primer turno.

5. Discusión de los resultados

Tras la corrección de las pruebas puntuadas sobre un total de 20, hemos clasificado los resultados en cinco categorías: la primera agrupa los estudiantes con un puntaje inferior a 10, la segunda comprende los resultados entre 10 y 11.9; la tercera reúne los puntajes comprendidos entre 12 y 13.9; la cuarta incluye los puntajes entre 14 y 15.9 y a la última se asignaron los puntajes superiores a 16. En cuanto a la primera, la posible diferencia entre los participantes a la prueba A versus B, no hemos comprobado diferencias sustanciales. A diferencia de lo esperado, los resultados de ambos formatos son relativamente homogéneos (tabla 1).

	Prueba A		Prueba B	
	Total	%	Total	%
Categoría 1: - 10	3	17,65	7	14,29
Categoría 2: 10-11,9	5	29,41	15	30,61
Categoría 3: 12-13,9	3	17,65	11	22,45
Categoría 4: 14-15,9	4	23,53	10	20,41
Categoría 5: 16 y +	2	11,76	6	12,24

Tabla 1 – Comparación prueba A / B

6. Utilizamos el término siguiendo a Shohamy & Inbar (1991) quienes distinguen entre (a) las preguntas globales que requieren que los examinados sinteticen partes de la información y saquen conclusiones; (b) las preguntas locales que requieren que los examinados identifiquen partes de la información o comprendan segmentos de palabras individuales o bloques de información restringidos y (c) preguntas triviales que piden, por parte del examinado, que comprenda detalles precisos pero irrelevantes para el tema general del texto audiovisual (en Buck 2001: 136-137).

7. No seguimos las pautas que ofrece Buck para optimizar la corrección de preguntas abiertas y que consiste en considerar, en lo posible, las preguntas abiertas como ítems casi dicotómicos por medio de una primera evaluación radical considerando la pregunta o bien como correcta o bien como incorrecta. Tras esta primera evaluación se hace una segunda, en donde se analizan detenidamente las respuestas ambiguas que no son asimilables a los otros dos grupos (Buck 2001: 141).

Dos elementos en particular pueden aclarar estos resultados, que van en contra de lo comprobado en estudios anteriores. Un primer elemento concierne al grado de familiarización de los estudiantes con el formato de las preguntas abiertas. En el marco de esta asignatura, los estudiantes están familiarizados ante todo con las preguntas abiertas y mucho menos con las preguntas tipo selección múltiple e ítems dicotómicos. Como han señalado Powers y Rock (1999), los estudiantes necesitan un tiempo de adaptación para familiarizarse con el formato típico de las pruebas tipo test. Aparte de la preparación, juegan también las expectativas sobre el tipo de evaluación (selección múltiple o preguntas abiertas): estas determinan en gran medida la forma en que se prepara el estudiante (Entwistle 1991). Por otra parte, puede haber influenciado la motivación, una variable que no incluimos en la presente investigación. Que pueda ser esencial la motivación, lo demuestra Baleghizadeh (2011) quien encontró una correlación estadísticamente significativa entre la relación intrínseca y el rendimiento en actividades de CA. Estudios de seguimiento deberían incluir también este tipo de información con el fin de incrementar información relevante sobre los participantes.

Un segundo elemento guarda relación con la estructura de las pruebas, y en particular con la cronología de las preguntas. En la prueba A, el orden de los datos solicitados en la ficha respeta rigurosamente la cronología de la secuencia de vídeo. En cambio, en la prueba B la información requerida para alguna pregunta de la sección A (selección múltiple) se solapa con la de la sección B (ítems dicotómicos). Tampoco descartamos que el sistema de corrección establecido de antemano favoreciera a los participantes de la prueba A. Aunque no permitimos que los estudiantes respondieran en su lengua nativa —una práctica propuesta para evitar que se evalúe demasiado la producción escrita en CA (Buck, 2001: 100)—, sí se ha adoptado un sistema muy flexible de corrección.

En cuanto a la segunda hipótesis, que verifica la posible incidencia del número de secuencias de vídeo en los resultados, concluimos que existe cierta correlación positiva.

	Prueba B dos visionados		Prueba B tres visionados	
	Total	%	Total	%
Categoría 1: - 10	7	22,58	0	0
Categoría 2: 10-11,9	7	22,58	8	44,4
Categoría 3: 12-13,9	4	12,90	3	38,9
Categoría 4: 14-15,9	7	22,58	7	16,7
Categoría 5: 16 y +	6	19,35	0	0

Tabla 2 – Comparación prueba B dos visionados / prueba B tres visionados

Como se puede observar en la tabla 2, las diferencias más pronunciadas se sitúan en las categoría 1, 2, 3 y 5. Ningún estudiante reprueba tras tres turnos. En cambio, el número relativo de estudiantes comprendidos en las categorías 2 y 3, es más elevado. De estas cifras se deduce que el turno adicional beneficia sobre todo a los estudiantes menos sobresalientes. Otro dato llamativo: ningún estudiante obtiene resultados sobresalientes (categoría 5), lo cual parece sugerir que el turno adicional no conlleva mejor rendimiento para los estudiantes con mejores rendimientos. Es más, se podría concluir —por paradójico que parezca— que el turno adicional perjudica a este tipo de estudiante.

6. Conclusión

El primer objetivo era hacer hincapié en la urgencia de contribuir empíricamente al debate en torno a los formatos para evaluar la CA en el aula de ELE. Hemos señalado al respecto algunas dificultades de este tipo de investigación: la CA combina muchas subdestrezas interconectadas de índole muy diferente; además, sólo se le puede evaluar de manera indirecta.

El segundo objetivo era comparar, sobre bases empíricas, diversos formatos de evaluar la CA en la clase de ELE, comparando dos pruebas: una primera que incluye respuestas de selección múltiple e ítems dicotómicos y una segunda compuesta por preguntas abiertas. Ambos tipos tienen tanto ventajas como desventajas, que hemos señalado en su debido lugar. En los estudios anteriores, enfocados en el aprendizaje del inglés como LE, se ha subrayado la importancia del formato en CA. En la literatura especializada existe cierto consenso para decir que las preguntas tipo test se asocian con mejores resultados que las preguntas abiertas. Aparte del formato se ha verificado también la interrelación entre el número de secuencias y los resultados, sobre la cual también existe cierto consenso, aunque menos pronunciado, en la literatura.

Esta literatura ha sido el punto de partida para nuestras dos hipótesis: (1) que los estudiantes sometidos a la prueba tipo preguntas abiertas puntúan menos en comparación con los otros y que (2) cuantos más visionados mejores resultados. Los resultados no confirman la primera hipótesis. Concluimos que los estudiantes frente a preguntas abiertas no puntúan menos que los participantes a la prueba tipo test, incluso si avanzamos factores que deberían incluirse en investigaciones de seguimiento con el fin de afinar el análisis y complementar los datos. En cuanto a la segunda hipótesis, el número de repeticiones resultó determinante, sobre todo para los estudiante menos fuertes.

Referencias bibliográficas

- BACHMAN, L. y PALMER, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- BAILEY, K. (1998). *Learning about language assessment: Dilemmas, decisions and directions*. New York: Heinle & Heinle.
- BALEGHIZADEH, S. (2011): "The relationship among listening performance, Metacognitive strategy use and motivation from a self-determination theory perspective". En *Theory and Practice in Language Studies*, nº. 1(1), 61-67.
- BORDÓN, T. (2001): "La evaluación de la comprensión auditiva en aprendices de E/LE". En *Carabela*, nº. 49, 77-101.
- (2004): "La evaluación de la expresión oral y de la comprensión auditiva". En *Vademécum para la formación de profesores. Enseñar español como segunda lengua (L2) / lengua extranjera (LE)*, SGEL, 983-1001.
- BRINDLEY, G. y SLATYER, H. (2002): "Exploring task difficulty in ESL listening assessment". En *Language Testing*, nº. 19 (4), 369-394.
- BROWN, J. y HUDSON, T. (1998): "The alternatives in Language Assessment". En *Tesol Quarterly*, nº. 32(4), 653-675.
- (2002). *Criterion-referenced language testing*. Cambridge, UK: Cambridge University Press.
- BUCK, G. (1991): "The testing of listening comprehension: An introspective study." En *Language Testing*, nº. 8, 78-91.
- (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- CHENG, H. (2004): "A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in English". En *Foreign Language Annals*, nº. 37(4), 545-553.
- CONSEJO DE EUROPA. (2000). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: Anaya (En línea: http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf [Consulta: 03/10./2015]).
- CURRIE, M. y CHIRAMANEE, T. (2010): "The effect of multiple-choice item format on the measurement of knowledge of language structure". En *Language Testing*, nº. 27(4), 471-491.
- ELKHAFIFI, H. (2005): "Listening comprehension and anxiety in the Arabic language classroom". En *The Modern Language Journal*, nº. 89(2), 206-220.
- ENTWISTLE, N. (1991): "Approaches to learning and perceptions of the learning environment. Introduction to the special issue". En *Higher Education*, nº. 22, 201-204.
- HANSEN, C. y JENSEN, C. (1994): "Evaluating lecture comprehension". En *Flowerdew, J. (ed). Academic listening*. Cambridge University Press, 241-68.
- HERBERT, D. y BURT, J. (2002): "The effects of different review opportunities on schematization of knowledge". En *Learning and Instruction*, nº. 13, 73-92.
- KRASHEN, S. (2002). *Explorations in language acquisition and use: The Taipei lectures*. Taipei: Crane.
- PASTOR VILLALBA, C. (2009): "La evaluación de la comprensión oral en el aula de ELE". En *MarcoELE. Revista de didáctica español lengua extranjera*, nº. 9. En línea: http://marcoele.com/descargas/9/pastor_comprensionoral.pdf [Consulta: 04.11.2014].
- POWERS, D. y ROCK, D. (1999): "Effects of coaching on SAT I. Reasoning test scores". En *Journal of Educational Measurement*, nº. 36, 93-118.
- SAKAI, H. (2009): "Effect of repetition of exposure and proficiency level in L2 listening tests". En *TESOL Quarterly*, nº. 43(2), 360-372.

SHERMAN, J. (1997): "The effect of question preview in listening comprehension tests". En *Language Testing*, n°. 14(2), 185-213.

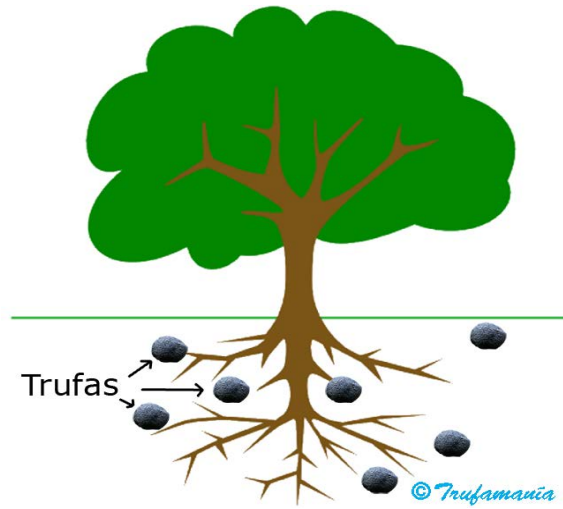
TRAUB, R.E., y MACRURY, K. (1990): "Multiple choice vs. free response in the testing of scholastic achievement". En *Test und Trends 8: Jahrbuch der Pädagogischen Diagnostik*, n°. 8, Weinheim und Base: Beltz Verlag, 128-159.

VANDERGRIFT, L. (2007): "Recent development in second and foreign language listening comprehension research". En *Language Teaching*, n°. 40, 191-210.

APÉNDICES

1. Breve descripción de las actividades de preaudición

.....



Actividades de preaudición

1. Al estudiante se le pidió que observara estas dos imágenes y que comunicara oralmente en qué pensaba. [Anticipación del léxico, emisión de hipótesis]

2. Al estudiante se le presentaron algunas secuencias aisladas (sin sonido, sin movimiento) del fragmento. Se invitó al estudiante a que formulara hipótesis sobre el contenido, que intentara inferir información. Para facilitar esta actividad, se presentó al estudiante la siguiente lista de preguntas:

- ¿Cuál es el tema?
- ¿Quién habla (sexo, edad)?
- ¿Cuántos testimonios hay?
- ¿Dónde están?
- Aparte de las personas, ¿qué ves (objetos, imágenes, etc.)?
- ¿Qué palabras (o estructuras) crees que escucharás en este fragmento?

2. Prueba A (preguntas abiertas)

.....

[Nota: no se ha respetado aquí el espaciado original]

El Museo de la Trufa - TVE, Panorama / Navarra Natural

¿Qué se menciona sobre el Museo de la Trufa?

A la trufa, ¿cómo la denominan en este reportaje?

En opinión de Pedro Jesús Oroquieta, ¿cuáles son los dos objetivos del museo? En dos niveles diferentes:

Según Koldo Roderó, ¿cuáles son los “valores” de la trufa?

¿Cuántos visitantes se esperan en el museo?

En palabras de Antonio Barbarín, ¿qué se quiere combinar con este museo?

¿Cuánto ha costado el edificio?

3. Prueba B (selección múltiple e ítems dicotómicos)

El Museo de la Trufa - TVE, Panorama / Navarra Natural

A – Contesta a las preguntas. Selecciona la opción más adecuada (A, B o C).

A1. El Museo de la Trufa...

- A. abrirá sus puertas próximamente.
- B. abrió sus puertas hace poco tiempo.
- C. abrió sus puertas hace un mes.

A2. A la trufa la denominan...

- A. el diamante negro de la cina (= especie de planta).
- B. el diamante negro de la cocina.
- C. el diamante negro de la colina (= especie de monte).

A3. El Museo de la Trufa se ha construido...

- A. entre las sierras de Lokiz y Metauten.
- B. en Metauten, la capital de seis valles.
- C. al abrigo de la Sierra de Lokiz.

A4. En opinión de Pedro Jesús Oroquieta ...

- A. este museo es el primero a nivel nacional que tiene como objetivo ser el centro del cultivo y promoción de la trufa.
- B. el museo tiene dos objetivos: en un nivel nacional, ser el centro del cultivo y promoción de la trufa y, en un nivel regional, promover el desarrollo del valle.
- C. el museo tiene dos objetivos: ser el centro de referencia y asegurar la promoción de la trufa.

A5. En esta región la trufa...

- A. se cultiva desde hace 50 años.
- B. se cultiva desde hace 5 años.
- C. se cultiva desde hace 15 años.

A6. Según Koldo Rodero...

- A. la trufa reúne muchos aromas que la gastronomía todavía desconoce.
- B. en la trufa se condensan todos los aromas del bosque.
- C. la trufa, como actor principal o complemento, puede participar en la promoción de la gastronomía local.

A7. En palabras de Antonio Barbarín, con este museo se quiere...

- A. promover tanto el desarrollo agrario como el turismo.
- B. promover tanto el desarrollo necesario como el turismo.
- C. promover el desarrollo agrario de la región: primero para el valle y luego para Tierra Estella.

B – Lee las frases y decide si son verdaderas (V) o falsas (F). Si son correctas, justifica tu respuesta. Si son falsas, corrígela:

[Nota: no se ha respetado aquí el espaciado original]

1. En España hay otros dos museos dedicados a la trufa.
2. En el museo la información se presenta por medio de paneles y vídeos.
3. Se esperan 5 000 visitantes al año.
4. La trufa crece de forma natural en la zona.
5. El museo es la primera iniciativa de un plan de desarrollo rural en esta región.
6. El edificio ha costado 7 700 euros.
7. El museo es una iniciativa completamente privada.