

CASAÑ NÚÑEZ, JUAN CARLOS

UNIVERSIDAD POLITÉCNICA DE VALENCIA

## UN MARCO TEÓRICO SOBRE EL USO DE PREGUNTAS DE COMPRESIÓN AUDIOVISUAL INTEGRADAS EN EL VÍDEO COMO SUBTÍTULOS: UN ESTUDIO MIXTO

### BIODATA

Máster Universitario en Enseñanza del Español como Lengua Extranjera por la Universidad Internacional Menéndez Pelayo-Instituto Cervantes, Certificado de Aptitud Pedagógica y licenciatura en Filología Hispánica por la Universidad de Valencia y licenciatura en Filología Inglesa por la Universidad Jaume I (Castellón). Doctorando en la Universidad Politécnica de Valencia.

Ha desarrollado su labor profesional en las áreas de la enseñanza del español como lengua extranjera y de la educación en España, Estados Unidos, Francia, Polonia y Portugal. Ha publicado ediciones didactizadas de guiones cinematográficos, artículos y secuencias didácticas. Ha colaborado en la edición de manuales electrónicos para estudiantes de 2º de Bachillerato y de un corpus sociolingüístico del español. Ha sido locutor para manuales de español como lengua extranjera.

### RESUMEN

Escuchar, leer y escribir simultáneamente es extremadamente difícil, especialmente en una lengua extranjera (Underwood, 1989). Una actividad de comprensión audiovisual puede serlo todavía más. Por una parte, existe un conflicto de atención visual entre ver la imagen y completar una actividad escrita. Por otra parte, prestar atención a la imagen, al audio, y a la tarea simultáneamente puede llevar a la saturación de la memoria de trabajo (Vandergrift y Goh, 2012).

Este trabajo forma parte de una investigación mayor que explora el uso de preguntas de comprensión audiovisual integradas en la imagen como subtítulos y sincronizadas con los fragmentos relevantes. Esta técnica docente novedosa podría aportar varios beneficios frente a visionados con las preguntas en papel, entre ellos, la reducción del conflicto de atención visual y las necesidades cognitivas de la actividad.

En este estudio mixto de tipo convergente (Creswell y Plano, 2011; Creswell, 2012, 2013), seis expertos en didáctica del español como lengua extranjera, lingüística española y/o en el uso de la tecnología completaron un cuestionario con ítems cuantitativos y cualitativos. El cuestionario tenía como objetivo averiguar la opinión de los expertos sobre un marco teórico que describe el uso de preguntas subtítuladas. Los resultados cuantitativos y cualitativos sugieren que, globalmente, los expertos están de acuerdo con la mayor parte de los planteamientos; los resultados también revelan puntos susceptibles de mejora. La consideración de las valoraciones y comentarios de los expertos ha llevado a la reformulación de la propuesta inicial.

**PALABRAS CLAVE:** comprensión auditiva, comprensión audiovisual, métodos mixtos, subtitulación, vídeo

#### A FRAMEWORK FOR USING AUDIOVISUAL COMPREHENSION QUESTIONS EMBEDDED IN THE VIDEO AS SUBTITLES: A MIXED METHODS STUDY<sup>1</sup>

Listening, reading and writing simultaneously is extremely difficult, especially in a foreign language (Underwood, 1989). An audiovisual comprehension activity can be even harder. On the one hand, there is a conflict of visual attention between viewing the video and completing a written activity. On the other hand, paying attention to the image, the audio, and the task simultaneously can lead to working memory overload (Vandergrift & Goh, 2012).

This paper is part of a larger research study that explores the use of audiovisual comprehension questions embedded in the image as subtitles and synchronized with the relevant fragments. This innovative teaching technique could provide various benefits over viewings with questions on paper, among them, the reduction of the conflict of visual attention and the cognitive requirements of the activity.

In this convergent mixed methods study (Creswell & Plano, 2011; Creswell, 2012, 2013), six experts in teaching Spanish as a foreign language, Spanish linguistics and/or the use of technology completed a questionnaire designed to elicit both quantitative and qualitative data. The purpose of the survey form was to find out what experts thought about a theoretical framework describing the use of subtitled questions. Quantitative and qualitative results suggest that, overall, the experts agree with most of the framework; results also reveal areas for potential improvement. The consideration of the comments led to the reformulation of the initial proposal.

**KEY WORDS:** listening, audiovisual comprehension, mixed methods, subtitling, video

---

<sup>1</sup> English versions of the introduction, the study design, the framework and the conclusion can be found in the appendix.

## 1. INTRODUCCIÓN

Escuchar, leer y escribir simultáneamente es extremadamente difícil, especialmente en una lengua extranjera (Underwood, 1989). Escuchar, ver, leer y escribir puede serlo todavía más. Por una parte, existe un conflicto de atención visual entre observar la imagen y completar una actividad escrita de forma simultánea. Este conflicto contribuye a explicar el grado de (des)atención a la imagen en investigaciones que contabilizan el tiempo que los estudiantes miran en dirección a la pantalla en pruebas de comprensión con vídeo. Ockey (2007) no proporciona la media de tiempo que los participantes observan el vídeo, pero a partir de los datos se puede calcular que fue el 44,9 %, Wagner (2007) obtiene el 69 % y Wagner (2010) el 47,9%. Si la comprensión auditiva es un proceso de interpretación de información auditiva y visual, como proponen especialistas como Riley (1979), Martín Peris (1991), Rubin (1995a) y Lynch (2012), estos grados de (des)atención pueden repercutir negativamente tanto en la comprensión del vídeo como en el desarrollo de esta actividad comunicativa. Por otra parte, la memoria de trabajo tiene una capacidad reducidísima (Sweller, Ayres y Kalyuga, 2011) y esto debe tenerse presente en la preparación de las actividades simultáneas al visionado. Como señalan Vandergrift y Goh (2012), prestar atención a la imagen, al audio, y a la tarea simultáneamente puede llevar a la saturación de la memoria de trabajo.

Este trabajo forma parte de una investigación más amplia que explora el uso de preguntas de comprensión audiovisual integradas en la imagen como subtítulos y sincronizadas con los fragmentos de vídeo relevantes para responderlas (ver figura 1). Se considera que este método docente podría aportar varios beneficios en la fase de visionado frente a una explotación en la que el vídeo se proyecta en

una pantalla común y los materiales de la actividad se presentan en un folio. Entre ellos, podría reducir el conflicto de atención entre ver el vídeo y realizar la tarea, al aproximar espacialmente las preguntas y los fragmentos relevantes, y podría disminuir las necesidades cognitivas de la actividad al no tener que prestar atención a varias preguntas de forma simultánea. Se trata de una técnica sobre la que no se ha encontrado referencias ni en la didáctica de la comprensión auditiva/audiovisual en el aula de lenguas extranjeras (en adelante LE) ni en la didáctica del vídeo en el aula de LE, a excepción de las del propio autor (Casañ, 2007, 2012). En este artículo se presenta un marco teórico sobre este procedimiento y se describe uno de los métodos contemplados para su validación, refutación o reformulación: la valoración por parte de un grupo de expertos en didáctica del español como lengua extranjera (en adelante ELE), lingüística española y/o en el uso de la tecnología.





Fig. 1. Ejemplos de preguntas de comprensión subtituladas. Los fotogramas pertenecen al largometraje *Ópera prima* de Fernando Trueba.

## 2. ANTECEDENTES

### 2.1. EL CONCEPTO DE COMPRENSIÓN AUDITIVA

Frente a la expresión oral y escrita, que tienen productos observables, la comprensión auditiva ocurre de manera interna, invisible a los ojos del observador. Como señala Lynch (2009), incluso cuando el oyente responde con un gesto o una interjección que sugieren que ha comprendido, no se puede tener la seguridad de que haya sido así. Debido a este carácter invisible, es difícil analizar su naturaleza y llegar a una definición o descripción definitiva. El objetivo de este apartado es recoger propuestas que puedan aplicarse en el aula de lenguas extranjeras y que den cuenta de la doble dimensión (auditiva y visual) de la comprensión de grabaciones audiovisuales.

El *Diccionario de la lengua española* (en adelante *DRAE*) y el *Diccionario de Uso del Español* (en adelante *DUE*) ofrecen propuestas similares: el *DRAE* define el término *escuchar* como “prestar atención a lo que se oye” (RAE, 2014: 937) y el *DUE* como “atender para oír cierta cosa” (Moliner, 2007: 1235). Estas definiciones ponen el énfasis en que escuchar es un proceso consciente, activo, que requiere de la atención del oyente. De este modo se opone a “oír”, que es un acto pasivo, automático.

Glenn (1989) analiza cincuenta definiciones de *listening* propuestas entre 1925 y 1988 y extrae siete conceptos sobre los que existe cierto consenso entre los especialistas: *perception*, *attention*, *interpretation*, *remembering*, *response*, *spoken sounds* y *visual cues*. Rubin (1995a: 7) propone una definición que recoge el carácter activo de la comprensión auditiva y varios de los conceptos obtenidos por Glenn (1989): “an active process in which listeners select and interpret information which comes from auditory and visual cues in order to define what is going on and what the speakers are trying to express”.

En términos similares se pronuncia Martín Peris: “la comprensión de mensajes orales requiere prestar atención a elementos lingüísticos, paralingüísticos y extralingüísticos” (1991: 2); Lynch (2012: 76): “effective listening / viewing requires the simultaneous interpretation of visual (gestural and facial) information, the speaker’s words and ambient sounds”, y Rost y Wilson (2013: 86):

In live face-to-face communication as in multimedia text processing, the listener must attend to multiple layers of input, including the verbal system (written, auditory and articulatory verbal codes), the non-verbal system (visual and symbolic content), and the physical context (including people and objects referred to).

La conciencia de esta dimensión visual lleva a algunos autores a sugerir nuevas maneras de referirse a la comprensión auditiva. Riley (1979) titula un artículo «Viewing comprehension: "L'oeil écoute"», Allan (1984) utiliza también el término "Viewing comprehension" y Harris (2003) nombra otro artículo "Listening with Your Eyes". Lynch (2012: 76) explica que utiliza el término *listening* por la convención existente en la bibliografía y sugiere que en su lugar sería más apropiado hablar de "a modern skill of listening / viewing".

El *Marco Común Europeo de Referencia* (en adelante *MCER*) redefine y amplía la concepción tradicional de cuatro destrezas (hablar, escribir, leer y escuchar) y distingue entre expresión oral y escrita, interacción oral y escrita, mediación oral y escrita, y comprensión auditiva y audiovisual. En la comprensión auditiva, "el usuario de la lengua como oyente recibe y procesa una información de entrada (*input*) en forma de enunciado emitida por uno o más hablantes" (Consejo de Europa, 2002: 68); y en la comprensión audiovisual, "el usuario recibe simultáneamente una información de entrada (*input*) auditiva y visual" (Consejo de Europa, 2002: 73). El *MCER* presenta ejemplos de actividades de comprensión auditiva (68) y audiovisual (73), y proporciona varias escalas ilustrativas para la comprensión auditiva (69-70) y una para la comprensión audiovisual (73). Como señala Casañ (2009a), algunos de los ejemplos de actividades de comprensión auditiva se contradicen con las definiciones del propio *MCER*. Así, "escuchar medios de comunicación (radio, televisión, grabaciones, cine)" (68) se incluye dentro de la comprensión auditiva aunque el oyente recibe *input* tanto por el canal oral como el visual en al menos dos casos (televisión y cine). Algo similar ocurre con "escuchar conferencias y presentaciones en público" (68). Se presenta como ejemplo de comprensión auditiva, pero sería más adecuado situarlo dentro de la comprensión audiovisual. Asimismo, podría deducirse que se ha

identificado comprensión audiovisual con comprensión de materiales en vídeo porque la única escala propuesta para la comprensión audiovisual es "ver televisión y cine" (73).

En este artículo se concibe la comprensión auditiva como un proceso de interpretación de información auditiva y visual, como proponen especialistas como Riley (1979), Martín Peris (1991), Rubin (1995a) y Lynch (2012). También se aboga por una nueva terminología que recoja el carácter multimodal de la mayor parte de las situaciones de comprensión auditiva, como sugieren autores como Riley (1979), Harris (2003) y Lynch (2012). En relación con la comprensión de grabaciones audiovisuales, la utilización de los términos *comprensión auditiva* o *listening* es especialmente desafortunada. Por un lado, es innegable que intervienen tanto la vista como el oído. Por otro lado, no es posible prescindir de la dimensión visual. Como expone Chion (1993), lo que define actualmente al vídeo es la imagen y no el audio, esto es, una grabación audiovisual sin audio sigue siendo una grabación audiovisual, pero lo contrario la convertiría en una grabación sonora.

En este trabajo se sugiere emplear *comprensión auditiva* para referirse a la comprensión de información a través del canal oral (una conversación telefónica, escuchar la radio o hablar a través de una puerta de madera), utilizar *comprensión audiovisual* para describir la comprensión de información simultáneamente por el canal oral y el visual (interacción cara a cara, sea de forma presencial o telemática, o comprensión de material en vídeo), y acuñar el término *comprensión auditiva / audiovisual* para hablar de forma genérica y para manifestar que lo visual forma parte de la mayor parte de las situaciones de comprensión auditiva. En lo que resta de artículo se sigue esta propuesta, aunque somos conscientes de que el término *comprensión auditiva* predomina en la bibliografía.

## 2.2. LA DIMENSIÓN VISUAL ES PARTE DEL MENSAJE

Brown y Yule (1983) señalan que “the visual environment gives an enormously important extra dimension of information” (85). En el mismo sentido se pronuncia Lynch (2010: 80): “the power of the visual component of communication should not be underestimated”. Riley (1979: 84) enfatiza esta relevancia de forma más rotunda: “such factors as proxemics, kinesics and deictics are all part of the message. They are not just a sort of gloss on the verbal component”.

¿Qué valores tiene la imagen? Para Brown y Yule (1983) la dimensión visual proporciona información sobre los interlocutores, el contexto, y la relación entre los hablantes:

Not only does it permit the listener to see what the participants look like, whether they are young or old, rich or poor, indoors or out of doors, together with all the details of the features of the physical context, it also permits him to see the physical relationships between the participants, how close they stand or sit to each other, whether or not they touch each other, whether they lean towards each other in a friendly fashion as they speak, or whether they turn away from each other.

En un análisis sobre el rol de lo visual en la comprensión auditiva / audiovisual, Riley (1979) identifica y describe seis funciones comunicativas de lo visual:

- a) *Deictic*: referencias a elementos presentes en la situación comunicativa.
- b) *Interactional*: señales para expresar el cambio de turno entre los hablantes.
- c) *Modal*: la manera de comprometerse con lo que se dice (por ejemplo, decir algo con cara de enfadado o sonriendo).
- d) *Indexical*: proporciona información sobre uno mismo (por ejemplo, edad, tamaño o estado emocional)
- e) *Linguistic*: conductas no verbales que tienen un significado preciso.
- f) *Situational*: el impacto o relevancia del contexto espacio-temporal en la conducta comunicativa.

Como explica Riley (1979), estas funciones no son exclusivas del canal visual. Por ejemplo, el hablante puede indicar que se dirige a cierta persona mirándola a los ojos y/o con un vocativo.

Whittaker (1995) identifica tres posibles maneras en las que la información visual puede facilitar la comunicación: “by providing (1) cognitive cues to enhance shared understanding, (2) social cues for providing information about interpersonal communication and affect, or (3) conversation process cues to facilitate speaker change” (523). Según el autor, la información sobre el grado de comprensión y el turno de palabra se puede comunicar eficazmente de manera verbal, pero no ocurre lo mismo con la información de carácter afectivo: “the removal of the visual channel changes the outcome of tasks that require access to affect” (524). El autor sugiere una explicación: el carácter involuntario de gran parte de la información afectiva. Así, “although speech can signal affect, speakers omit the full range of *affective* cues, when using audio only communication” (524). Whittaker también indica otra función de la dimensión visual que suele pasar desapercibida: las personas utilizan la información visual para determinar la disponibilidad de otra persona para iniciar una conversación.

¿Cuál es la “relación” entre lo visual y lo auditivo? La comprensión auditiva/audiovisual es un proceso de adición y combinación de información visual y auditiva. Es adición en tanto que la fusión entre lo visual y auditivo no es total, y tanto la parte visual como la auditiva pueden subsistir de forma independiente. Es combinación en tanto que “el sonido hace ver la imagen de un modo diferente a lo que esta muestra sin él, la imagen, por su parte, hace oír el sonido de modo distinto a como éste resonaría en la oscuridad” (Chion, 1993: 31). Este autor acuña el concepto de valor añadido para describir este fenómeno de enriquecimiento mutuo entre imagen y sonido. Como consecuencia de este valor añadido, en ejercicios de

descripción de secuencias audiovisuales se observa la aparición de fenómenos de “ilusión retrospectiva”, como el recuerdo de sonidos que no existieron o el olvido de otros que sí se produjeron (Chion, 1993).

En el artículo “Hearing Lips and Seeing Voices”, McGurk y MacDonald (1976) demuestran la influencia de la visión en la percepción del habla. Estos autores replican un estudio previo en el que se había grabado un primer plano de una persona pronunciando repetidamente las sílabas [ga] y [ba]. Cuando se presentaba solo el audio, [ga] o [ba], o había correspondencia entre el audio y el movimiento de los labios, los informantes identificaban correctamente la sílaba [ga] o [ba]. Sin embargo, cuando el audio era [ba] y los labios de la persona pronunciaban [ga], la mayor parte de los informantes decían que escuchaban [da]; y cuando el audio era [ga] y los labios pronunciaban [da], la mayor parte de los informantes decían que escuchaban [bagba] o [gaba]. Las ilusiones son tan fuertes que no desaparecen aunque se conozca su existencia: “by merely closing the eyes, a previously heard [da] becomes [ba] only to revert to [da] when the eyes are open again” (747). Un ejemplo de lo contrario, es decir, de la influencia del sonido sobre la imagen, es el que se produce en las películas cuando se sugiere con la música que el héroe está corriendo un grave peligro (ejemplo de Rodríguez, 1998).

### 2.3. EL LUGAR DEL VÍDEO EN LA DIDÁCTICA DE LA COMPRESIÓN AUDITIVA / AUDIOVISUAL

El uso de materiales en vídeo debería predominar sobre las grabaciones de audio para desarrollar la comprensión auditiva/audiovisual en el aula de lenguas extranjeras (Mendelsohn, 1994; Rubin, 1995b; Ur, 1999; Lynch, 2009). En primer lugar, es

coherente con una concepción de la comprensión auditiva/audiovisual como proceso de interpretación de información auditiva y visual, como defienden autores como Martín Peris (1991), Rubin (1995a) o Lynch (2012). En segundo lugar, permite observar la realidad de la mayor parte de nuestras interacciones orales: conversamos con personas (y no solo con voces) y en un contexto determinado. En tercer lugar, el vídeo tiene un efecto positivo en la motivación (Ur, 1984/1994, 1999; Flowerdew y Miller, 2005; Casañ, 2007; Vandergrift y Goh 2012). En una encuesta de análisis de necesidades, Casañ (2007) pregunta a cuarenta y dos estudiantes de ELE de la Universidad de Wrocław (Polonia) qué tipo de actividades prefieren realizar, y obtiene que la opción más valorada entre las propuestas es “trabajar a partir de fragmentos de películas, cortos o series españolas” y que realizar “actividades de comprensión oral extraídas de libros de texto” es muy poco seleccionada. En cuarto lugar, pero no por ello menos importante, hay argumentos a favor del aprendizaje multimodal.

El aprendizaje multimedia “occurs when people build mental representations from words (such as spoken text or printed text) and pictures (such as illustrations, photos, animation, or video)” (Mayer, 2006a: 2). Su justificación se escenifica en el principio multimedia de la teoría cognitiva del aprendizaje multimedia (en adelante TCAM) desarrollada por Mayer: bajo ciertas condiciones, los estudiantes aprenden mejor a partir de texto e imagen que solo texto (Mayer, 2001, 2006a; Fletcher y Tobias, 2006). Se sostiene que lo multimedia “takes advantage of the full capacity of humans for processing information. When we present material only in the verbal mode, we are ignoring the potential contribution of our capacity to also process material in the visual mode” (Mayer, 2006a: 4). La TCAM no ha sido desarrollada teniendo en cuenta las especificidades del aprendizaje de lenguas extranjeras, pero, según

Plass y Jones (2006), numerosos estudios indican que el principio multimedia se extiende al aprendizaje de lenguas extranjeras.

#### 2.4. INVESTIGACIONES SOBRE EL IMPACTO DE LA IMAGEN, LA VALORACIÓN DE LA IMAGEN Y LA CONDUCTA VISUAL

En los últimos años ha aumentado el número de investigaciones en torno al impacto de la imagen del vídeo en la comprensión de LE, sobre todo, en el área de la evaluación de la comprensión auditiva/audiovisual. Estos estudios suelen utilizar un diseño cuasiexperimental en el que un grupo responde a un test de comprensión a partir de una grabación audiovisual y el otro grupo contesta al mismo test a partir de una versión audio de la grabación audiovisual.

Hay estudios que no encuentran diferencias entre la presentación de las grabaciones en versión audio o en vídeo. Gruba (1993) diseña un test de *listening* a partir de un texto audiovisual (una conferencia sobre terrorismo internacional preparada *ad hoc*) y lo administra con vídeo a cuarenta y tres estudiantes de inglés como segunda lengua y con una versión solo audio a cuarenta y seis. Aunque la media de los participantes que completan el test a partir del vídeo es algo superior (7,16 frente a 6,39), esta diferencia no es estadísticamente significativa. Coniam (2001) lleva a cabo un estudio similar: cuarenta y siete informantes realizan un test de *listening* con un texto audiovisual (un debate auténtico) y cincuenta y siete con la versión solo audio del texto. El autor tampoco encuentra diferencias significativas en los resultados, pero en este caso el grupo que completa el test a partir de la versión solo audio obtiene una media ligeramente superior (21,76 frente a 19,85). En un cuestionario posterior a la prueba, Coniam pregunta a los participantes qué medio prefieren, y halla que "test takers who had taken the test by

one medium felt they might have scored better had they been exposed to the other medium" (10). Cabe señalar que el modo de visualización es distinto en las dos investigaciones: en el estudio de Gruba se utiliza una pantalla compartida, y en el de Coniam cada participante dispone de un monitor para minimizar el impacto de tener que levantar y agachar la cabeza para observar el vídeo y responder a las preguntas.

Otros estudios revelan que la imagen puede tener efectos negativos en la comprensión. Suvorov (2008) investiga la influencia de los *context visuals*. Este tipo de información visual proporciona información sobre "the context for the verbal exchanges, such as the participants, the setting, and text type" (10), y se distingue de los *content visuals*, que están relacionados con "the content of the verbal interaction and may include still photos, pictures, drawings, diagrams, etc." (10). Para ello diseña un test de *listening* con seis textos creados *ad hoc*: una conferencia y un diálogo en formato de grabación de audio, otra conferencia y otro diálogo en formato de grabación de audio acompañado de una fotografía, y una tercera conferencia y un tercer diálogo en formato vídeo. El autor no encuentra diferencias significativas entre las puntuaciones obtenidas en los textos orales y en los textos orales acompañados de una fotografía, sin embargo, las puntuaciones en las pruebas con textos audiovisuales son significativamente inferiores a las de las otras dos modalidades. Surovov administra un cuestionario después del test para conocer las preferencias de los participantes y encuentra que casi la mitad (catorce de treinta y cuatro) preferiría un test a partir de una grabación sonora, doce a partir de un vídeo, y siete a partir de una grabación sonora acompañada de una fotografía. El autor añade que no hay correspondencia entre las preferencias y las puntuaciones obtenidas.



Por último, hay autores que encuentran que la imagen tiene una repercusión positiva en las pruebas de comprensión. Sueyoshi y Hardison (2005) investigan la importancia de los gestos faciales y corporales en la comprensión de una conferencia sobre cerámica preparada por ellos. En el experimento participan cuarenta y dos estudiantes de inglés como segunda lengua de nivel intermedio-bajo y nivel avanzado que realizan el mismo test de comprensión. Algunos efectúan la prueba a partir de un visionado de la conferencia que muestra los gestos faciales y corporales del hablante, otros a partir de un visionado que solo permite observar los gestos faciales, y el resto a partir de una audición sin apoyo visual. Los estudiantes obtienen mejores resultados con las versiones audiovisuales que con la versión solo audio, y las diferencias son estadísticamente significativas. Los informantes de nivel avanzado obtienen los mejores resultados en la versión que solo muestra los gestos faciales, mientras que los participantes con un nivel intermedio-bajo consiguen las puntuaciones más altas en la versión que muestra los gestos corporales y faciales. En un cuestionario posterior a la prueba, los estudiantes señalan que haber observado los gestos corporales y faciales del hablante les ha facilitado la comprensión.

Algunos autores se han preocupado por investigar cuál es el comportamiento visual de estudiantes de LE cuando efectúan una actividad y/o una prueba de forma simultánea a la emisión del vídeo. Alderson, Clapham y Wall (1995) diseñan un test de comprensión con vídeo y lo pilotan obteniendo resultados que parecen aceptables. Sin embargo, en sucesivas administraciones del test les llama la atención que muchos examinandos no miran el vídeo, en su lugar, leen la hoja de preguntas y responden a partir de lo que escuchan. Esto lleva a los autores a conjeturar que la imagen puede ser redundante y a realizar dos estudios: "the first compared a group's performance on the video listening test with its

performance on the same test using only the soundtrack. (...) The second asked the same students which version of the test they preferred" (219). No encuentran diferencias en los resultados del test de comprensión pero los informantes muestran su preferencia por la versión sonora. El motivo es que consideran que la presencia de la imagen dificulta la ya de por sí compleja tarea de escuchar, leer las preguntas y escribir las respuestas. En consecuencia, Alderson *et al.* se inclinan por prescindir del vídeo.

Ockey (2007) investiga el comportamiento visual en dos situaciones distintas: audición acompañada de imágenes estáticas y visionado con vídeo. Para ello, diseña un test de *listening* con imágenes estáticas y otro con vídeo a partir de fragmentos diferentes de un mismo texto audiovisual (una lección sobre política norteamericana). Los test se administran por ordenador y se graba a seis voluntarios mientras los efectúan. Los resultados sugieren comportamientos visuales diferentes. En el caso de las imágenes estáticas, todos los informantes las observan poco tiempo (entre el 2,2 % y el 16,7 % del tiempo). En el caso del vídeo, la mayor parte de los participantes presta cierta atención a la imagen y lo hace de manera muy dispar (entre el 2,5 % y el 73,6 % del tiempo). Aunque Ockey no indica la media de tiempo que los participantes observan el vídeo, a partir de los datos se puede calcular que fue el 44,9 %.

Para investigar el comportamiento visual en un test de comprensión de una LE, Wagner (2007) coloca una cámara de vídeo al lado del monitor donde se reproduce el vídeo. Después de la prueba, mide con un cronómetro el tiempo que cada estudiante mantiene la mirada en la pantalla. Halla que los 36 participantes miran hacia el vídeo una media del 69 % del tiempo de emisión y que la conducta visual se mantiene relativamente constante durante todo el visionado. A partir de estos resultados, Wagner sugiere que los informantes perciben las imágenes como útiles para responder a las

preguntas. También comenta la variabilidad en la conducta visual: 14 de los 36 miran hacia el vídeo al menos un 90 % y 10 de los 36 menos del 50 %. El autor reconoce algunas limitaciones de la investigación, entre ellas, que mirar en dirección a la pantalla no implica necesariamente prestar atención a las imágenes del vídeo, y la posible influencia de la cámara de vídeo en los participantes. Wagner (2010) realiza un experimento similar y obtiene un tiempo de observación inferior: los 56 participantes miran el vídeo una media del 47,9 % del tiempo de reproducción. También encuentra una correlación negativa débil pero estadísticamente significativa entre el tiempo de observación del vídeo y la puntuación en el test. Wagner sugiere que podría deberse a que los participantes con un nivel de comprensión más bajo se ven obligados a ver el vídeo más tiempo para compensar sus dificultades para comprender la información lingüística. Las respuestas a un cuestionario posterior al test muestran que los examinados tienen una actitud positiva hacia el uso de textos audiovisuales en las pruebas de comprensión.

Casañ (2012) propone una tarea para reflexionar a los asistentes a un seminario sobre la didáctica de la comprensión auditiva/audiovisual en el aula de ELE: grabarles en vídeo mientras realizan una actividad de comprensión a partir de un fragmento de un vídeo auténtico en español (un cortometraje de animación). Los participantes son hablantes nativos de español y no nativos con un nivel C2, es decir, en la mayor parte de los casos se trata de comprensión de la L1. El autor no contabiliza el tiempo que cada persona mira en dirección al vídeo pero hay momentos en los que casi nadie lo hace. En la reflexión posterior a la tarea, los asistentes reconocen la dificultad de completar la tarea y ver simultáneamente el vídeo. Con esta tarea de reflexión Casañ pretende demostrar que una práctica válida con grabaciones de audio puede no serlo con grabaciones audiovisuales, y que hay un conflicto de atención entre ver la imagen y completar la tarea de forma simultánea. El problema

principal de la actividad radica en que requiere mucha lectura y escritura.

A luz de la revisión bibliográfica, las investigaciones sobre el impacto de la imagen no son concluyentes y aportan resultados aparentemente contradictorios. Algunas investigaciones no hallan diferencias significativas entre efectuar un mismo test con una grabación audiovisual y hacerlo con una versión solo audio del mismo vídeo (Gruba, 1993; Coniam, 2001), otras encuentran que se obtiene mejores resultados con la versión solo audio (Suvorov, 2008) y las hay que revelan mejores puntuaciones medias con la versión audiovisual (Sueyoshi y Hardison, 2005). En relación con la valoración de la imagen por parte de los examinados, es globalmente positiva en algunos estudios (Sueyoshi y Hardison, 2005; Wagner, 2010) y negativa en otros (Alderson *et al.* 1995). En cuanto a la conducta visual de los participantes, hay variaciones entre distintas investigaciones y también dentro de algunas de ellas. No obstante, la media del tiempo de visualización de los participantes sugiere menos atención al vídeo de la que sería deseable: en el estudio de Ockey (2007), según nuestros cálculos, la media de observación es el 44,9 % del tiempo, Wagner (2007) obtiene el 69 % y Wagner (2010) el 47,9%. Además, como se ha señalado, mirar en dirección a la pantalla no implica necesariamente estar concentrado en ella. Si la comprensión auditiva/audiovisual comprende la interpretación de información auditiva y visual, un grado de desatención tan elevado de la información visual (55,1 %, 31 % y 52,1 % respectivamente) puede influir negativamente tanto en la comprensión del vídeo como en el desarrollo de esta actividad comunicativa.

Estas divergencias pueden deberse a multitud de factores, entre ellos, la focalización de las preguntas (en información auditiva y/o audiovisual), la complejidad de la tarea, la manera como se ha

efectuado el visionado (visionado en una pantalla común, individualmente en un monitor de ordenador, o individualmente en una pantalla que integra el vídeo y las actividades), el origen del texto (auténtico o grabado por los investigadores), el tipo de texto, la calidad del *input* oral/visual (de la cinta de vídeo a la alta definición), la perceptibilidad del *input* visual (tamaño de la pantalla, ubicación de la pantalla y los participantes, distancia visual adecuada), mayor o menor solidaridad entre lo visual y lo oral en los textos audiovisuales, la preconcepción de los informantes sobre la necesidad de ver la imagen para responder a unas preguntas que normalmente se centran en información auditiva, el estrés generado por el test, la influencia de la cámara de vídeo en la conducta visual, el nivel de lengua de los participantes, etc.

## 2.5. LAS ACTIVIDADES DE COMPRESIÓN EN LA FASE DE VISIONADO EN EL AULA

Numerosos especialistas en didáctica de la comprensión auditiva / audiovisual sugieren la explotación de los textos orales en tres fases (por ejemplo, Underwood, 1989; Giovannini, Martín Peris, Rodríguez y Simón, 1996; Rost, 2002; Flowerdew y Miller, 2005; Richards y Burns, 2012). El previsionado tiene el objetivo de activar, enriquecer y/o crear conocimiento que facilite la comprensión del texto. En esta fase también se explica cómo va a efectuarse la audición/visionado (rol del estudiante, propósito, actividades, tipo de respuesta, etc.) porque estos factores condicionan el modo en que el oyente desarrolla los procesos de comprensión y su actuación. En el posvisionado se comprueba si se ha entendido aquello que interesaba comprender, se reflexiona sobre la experiencia de comprensión auditiva/audiovisual, se resuelve las dudas que hayan podido surgir, y también se puede plantear actividades centradas en otras actividades de la lengua o en la forma

lingüística a partir del texto. Según Richards y Burns (2012: 100), la fase de audición / visionado tiene los siguientes objetivos:

It should guide the listener through a listening text and help the listener understand it and respond to it. It should also involve tasks where listeners do something with the information they take from the text, such as responding, labeling, sequencing, selecting, or filling out a form.

Underwood (1989) enfatiza la necesidad de que las actividades de la fase de audición no compliquen la propia audición, dado que escuchar el texto oral, leer las preguntas y escribir las respuestas simultáneamente es extremadamente difícil, especialmente en una lengua extranjera y cuando no se escribe exactamente lo que se oye. Escuchar, leer, escribir y ver la imagen del vídeo puede serlo todavía más si la tarea no tiene en consideración esta complejidad.

En primer lugar, hay un conflicto de atención visual entre observar la imagen del vídeo y leer/escribir. Este conflicto es el responsable principal del grado de desatención a la imagen del vídeo referido arriba. Su intensidad será mayor cuanto mayor sea la cantidad de lectura y escritura que conlleve la tarea de forma simultánea a la emisión del vídeo, y cuanto mayor sea la distancia entre la imagen del vídeo y los materiales de la actividad porque se necesita más tiempo para desplazar la vista de un estímulo a otro. Por consiguiente, se debería limitar la cantidad de lectura y escritura de las actividades simultáneas al visionado, y si los medios lo permiten, aproximar el vídeo y los materiales de la actividad o integrarlos dentro de una misma pantalla. También puede ser efectiva la técnica del visionado con pausas sugerida por autores como Tomalin (1986), Lonergan (1984), Stoller (1992), Stempleski y Tomalin (2001). Casañ (2012) distingue entre las pausas que posibilitan una producción oral inmediata, y aquellas que facilitan una respuesta escrita que se comenta después del visionado.

En segundo lugar, aunque estrechamente relacionado con el primero, prestar atención a la imagen, al audio, y a la tarea simultáneamente puede llevar a la sobrecarga de la memoria de trabajo (Vandergrift y Goh, 2012). La memoria de trabajo se define como “a limited capacity system allowing the temporary storage and manipulation of information necessary for such complex tasks as comprehension, learning and reasoning” (Baddeley, 2000: 418). Para Miller (1956) el límite está en siete elementos, más o menos dos. Este número no se refiere a bits de información sino a *chunks*, grupos de información relacionados que funcionan independientemente del número de bits que los compongan. Cowan (2001), después de revisar la concepción de *chunk* y distintos estudios sobre límites de la memoria de trabajo, concluye que el límite en adultos es de tres a cinco. Como señalan Sweller *et al.* (2011), desde un punto de vista pedagógico, lo fundamental no es el número exacto, sino que la capacidad es reducidísima. Esta limitación debe tenerse en cuenta en la preparación de las actividades simultáneas al visionado. Aquellas actividades que requieran un consumo elevado de memoria son más apropiadas para la fase de posvisionado. Field (2008) explica que los oyentes inexpertos tienen los procesos de descodificación poco automatizados y tienen que dedicar buena parte de la memoria de trabajo a la descodificación. Por lo tanto, son estos estudiantes los más susceptibles de ver su memoria de trabajo saturada.

En resumen, para limitar el conflicto de atención entre ver el vídeo y completar la tarea simultáneamente, y reducir la posibilidad de que las actividades consuman recursos vitales para procesar el *input* audiovisual, en lugar de guiar al estudiante a lo largo del texto, se ha establecido varias recomendaciones. En primer lugar, las actividades escritas deberían implicar poca lectura y escritura, y deberían requerir mantener pocos elementos activos en la memoria de trabajo. En segundo lugar, se sugiere la aproximación espacial del

vídeo y los materiales o su integración dentro de una misma pantalla. Y en último lugar, se aconseja la utilización de la técnica del visionado con pausas.

## 2.6. EL ESTUDIO ACTUAL

Como se ha señalado en la introducción, no se ha encontrado referencias sobre la utilización de preguntas de comprensión audiovisual subtituladas y sincronizadas en la bibliografía, salvo las del propio autor (Casañ, 2007, 2012). A partir de la revisión de la literatura sobre la didáctica de la comprensión auditiva / audiovisual en LE y la didáctica del vídeo en LE y de la propia experiencia con este técnica, se ha elaborado un primer marco teórico que describe el procedimiento y sus posibles beneficios y limitaciones.

Este estudio mixto tiene el propósito conocer la opinión de un grupo expertos en didáctica del ELE, lingüística española y/o en el uso de la tecnología sobre el marco teórico provisional. Con este objetivo se ha diseñado un cuestionario con ítems cuantitativos que miden el grado de acuerdo con distintos aspectos del marco teórico, y con ítems cualitativos que recogen comentarios sobre estos y otros elementos. Se recopilan datos cuantitativos y cualitativos para combinar los resultados y de este modo conseguir una visión más completa de las valoraciones que la que podría obtenerse a partir de un solo tipo de datos. Se utiliza un *convergent design* (Creswell y Plano, 2011; Creswell, 2012, 2013): un tipo de diseño en el que los datos cuantitativos y cualitativos se recogen durante la misma fase de la investigación, se analizan de forma separada, y se interpretan de forma conjunta. Se pretende dar respuesta a las siguientes preguntas:

a) ¿En qué medida los expertos consultados están de acuerdo con el marco teórico?

b) ¿Qué aspectos del marco teórico se debería modificar, según los expertos consultados?

### 3. MÉTODO

#### 3.1. EL INSTRUMENTO

Debido a que los expertos que iban a ser contactos residían en países distintos a los del investigador, era conveniente que el instrumento pudiera obtener los datos en un solo encuentro presencial y, si fuera necesario, utilizarse a distancia. Al mismo tiempo el instrumento debía ser capaz de recoger las valoraciones de los expertos sobre el marco. Se decidió diseñar un cuestionario con preguntas abiertas y cerradas. Para su elaboración se ha intentado seguir las recomendaciones generales de Cohen, Manion y Morrison (2011) y Dörnyei (2007) sobre el diseño cuestionarios.

El cuestionario presenta cuatro partes: título, texto introductorio, exposición del marco teórico acompañada de los ítems y un apartado en el que se puede añadir información adicional.

El título "Preguntas de comprensión audiovisual integradas como subtítulos y sincronizadas con los fragmentos correspondientes en actividades de comprensión en el aula de LE" intenta ser lo más informativo posible y permite activar el conocimiento previo de los informantes.

Debajo del título aparece un texto introductorio en el que se explica el propósito del cuestionario, se dan las instrucciones, se pregunta si

se prefiere revelar la participación o responder de forma anónima, y se agradece la participación. El texto introductorio es el siguiente:

En las páginas siguientes se describe un primer desarrollo de un marco teórico sobre el uso de preguntas de comprensión audiovisual integradas como subtítulos y sincronizadas con los fragmentos correspondientes en actividades de comprensión en el aula de LE.

Este marco se considera en construcción y pendiente de validación. Uno de los procedimientos elegidos para validar, refutar o reformular esta propuesta metodológica consiste en conocer las valoraciones de expertos en didáctica del ELE, lingüística española y/o en el uso de la tecnología.

Le pido que responda a las preguntas que aparecen en recuadros debajo de algunos apartados.

En función de sus preferencias, su nombre puede ser revelado o permanecer en el anonimato. Subraye la opción que prefiera:

Revelar participación.

Permanecer en el anonimato.

Le agradezco su colaboración, le tendré informado sobre los resultados de esta investigación.

Los ítems del cuestionario<sup>2</sup> se han dividido en cuatro áreas: introducción (ítems 1 y 1b), descripción del procedimiento (ítems 2-10 y 2b), posibles beneficios (ítems 11-17, 11b-17b), posibles limitaciones e interrogantes (ítems 18-22, 18b, 19b, 21b y 22b), y preguntas generales (ítems 23 y 24). Se ha optado por una división por temáticas porque cambiar de tema supone un esfuerzo para el encuestado (Murillo, 2004). Se ha utilizado tres tipos de ítems: una

<sup>2</sup> El marco teórico y los ítems se presentan en el apartado de resultados para facilitar su lectura conjunta.

pregunta dicotómica excluyente (ítem 1), escalas Likert de cinco puntos (ítems 2, 11-19, 21-24), e ítems abiertos (ítem 1b, 2b, 3-10, 11b-19b, 20, 21b, 22b).

Las escalas Likert se han usado ampliamente para conocer el grado de acuerdo con distintos planteamientos del marco teórico. No se han empleado en aquellos casos en los se trataba de convenciones establecidas en la subtitulación del español (ítems 3-9) o la respuesta se consideraba obvia (ítem 20). En esos casos, se ha estimado más conveniente pedir comentarios. Se ha intentado que las opciones de las escalas Likert estuvieran equilibradas en términos positivos y negativos y que pudiera identificarse fácilmente un centro. También se ha prescindido de términos que expresan acuerdo o desacuerdo total (siempre, nunca, etc.) porque la mayoría de las personas tiende a evitarlos para no parecer extremistas (Cohen *et al.*, 2011). Los términos seleccionados para las escalas Likert son los siguientes: muy en desacuerdo, en desacuerdo, ni de acuerdo ni en desacuerdo, de acuerdo y muy de acuerdo.

Los ítems de respuesta abierta pueden aportar mucha más riqueza que los ítems cerrados gracias a la libertad de expresión, y pueden ayudar a identificar aspectos que no se había considerado previamente (Dörnyei, 2007), aunque plantean la dificultad del tratamiento de los datos (Cohen *et al.*, 2011). Se han utilizado abundantemente porque el objetivo del cuestionario, conocer las valoraciones de un grupo de expertos sobre el marco teórico en construcción, los hacían especialmente adecuados. Se han empleado para pedir referencias bibliográficas (ítem 1b) y comentarios (ítems 2b, 3-10, 11b-19b, 20, 21b, 22b).

El último apartado tiene el propósito de permitir que los expertos puedan hacer cualquier observación adicional relacionada con el procedimiento descrito o las preguntas planteadas.

### 3.2. LOS EXPERTOS

Para determinar el número de expertos necesarios se ha tomado como referencia el número recomendado en los juicios de expertos. No existe acuerdo sobre la cantidad: algunos autores sugieren un rango entre dos y veinte (Gable y Wolf, 1993, citado en McGartland *et al.*, 2003) y otros proponen un mínimo de cinco (Escobar-Pérez y Cuervo-Martínez, 2008). Este estudio ha contado con la participación de seis especialistas en didáctica del ELE, lingüística española y/o en el uso de la tecnología de cinco instituciones diferentes. Todos salvo uno han manifestado que no tenían inconveniente en revelar su identidad. En interés de la transparencia, se nombran a continuación en orden alfabético:

Dr. Rafael Arnold, director del Instituto de Romanística de la Universidad de Rostock.

Dra. Ana María Gimeno, profesora titular del Departamento de Lingüística Aplicada de la Universidad Politécnica de Valencia.

Dra. María Virginia González, directora del Máster de Estudios Hispánicos Avanzados: Aplicaciones e Investigación de la Universidad de Valencia.

Dr. Francisco Javier Satorre, profesor titular del Departamento de Filología Española de la Universidad de Valencia.

Dr. Agustín Yagüe, profesor titular de español como lengua extranjera de la Escuela Oficial de Idiomas de Barcelona.

El experto que ha preferido mantenerse en el anonimato es Doctor en Ingeniería Informática y profesor universitario.

### 3.3. ADMINISTRACIÓN DEL CUESTIONARIO

Dörnyei (2007) explica que el procedimiento de administración del cuestionario juega un papel significativo en la calidad de las respuestas. Asimismo, sugiere varias estrategias para asegurar la cooperación de los informantes que hemos intentado seguir. En primer lugar, se envió un correo electrónico a los expertos donde se explicaba el propósito y la naturaleza del cuestionario, y se preguntaba si tendrían disponibilidad e interés en participar. De las siete personas contactadas seis aceptaron colaborar. A continuación, se intentó concertar un encuentro para administrar el cuestionario de forma presencial. Fue posible reunirse personalmente con todos los especialistas en sus centros de trabajo excepto uno. A este experto se le envió una copia digital que devolvió cumplimentada por correo electrónico. Antes de entregar el cuestionario a los expertos, se saludó, se agradeció la participación y se mostraron dos vídeos. El primero consistía en un fragmento de una grabación realizada por Casañ (Casañ, 2012) en la que resultaba patente el conflicto de atención visual entre completar una actividad de comprensión y ver un vídeo. En la segunda grabación se ejemplificaba el procedimiento descrito en el marco teórico. El fragmento audiovisual seleccionado fue la primera secuencia del largometraje *Ópera Prima* de Fernando Trueba. Entre las razones para su elección figuran que reproduce un intercambio comunicativo completo, que no asume ningún conocimiento previo, que es breve (dos minutos aproximadamente), y que se había utilizado en trabajos previos (Casañ, 2009a, 2009b). Para ilustrar la técnica docente se subtitularon cuatro preguntas con el programa Adobe Premiere Pro CS5.5 para Windows: dos de comprensión de información específica, una de comprensión de ideas generales y una de interpretación. El experto que completó el cuestionario a distancia recibió una explicación acompañada de un fotograma de

cada uno de los vídeos. En el momento de entregar la copia del cuestionario, se recordó su propósito y se subrayó la importancia de las respuestas. Mientras rellenaban el cuestionario, el administrador permaneció en silencio, releyó el cuestionario y estuvo a plena disposición de los informantes. Cuando terminaron de responder, el administrador volvió a agradecerles su participación.

### 3.4. ANÁLISIS DE DATOS

Los datos cuantitativos (ítems 1, 2, 11-19, 21-24) se han introducido en la versión 21 del programa SPSS para Windows, y se ha calculado la frecuencia, el porcentaje y la moda. No se ha averiguado la media por tratarse de datos nominales y ordinales. Para la presentación de los datos cuantitativos se ha optado por la utilización de tablas. Según Dörnyei (2007: 289), las tablas pueden aportar una descripción más precisa y rica que las figuras aunque son menos "digestible".

En relación con el tratamiento de los datos cualitativos (ítems 1b, 2b, 3-10, 11b-19b, 20, 21b y 22b), Robson (1993: 370) afirma que "there is no clear and accepted set of conventions for analysis corresponding to those observed with quantitative data". Álvarez-Gayou (2003) señala que este análisis debe ser sistemático y seguir un orden. Para el tratamiento de los datos cualitativos se ha seguido las cuatro fases propuestas por Fernández (2006: 3-4): (1) obtener la información, (2) capturar, transcribir y ordenar la información, (3) codificar la información y (4) integrar la información. En el apartado de resultados se presenta de forma sintética aquellos datos que se ha considerado más relevantes para la investigación.

## 4. RESULTADOS

Con el objetivo de facilitar la lectura integrada del marco teórico, el ítem y los resultados correspondientes, se presenta estos elementos de forma sucesiva.

### I. INTRODUCCIÓN

Este marco teórico provisional se ha elaborado a partir de la revisión de bibliografía y de la propia experiencia docente con el procedimiento. Cabe destacar que a excepción de los comentarios del propio autor (Casañ, 2007, 2012), no se ha encontrado referencias sobre el procedimiento ni en la didáctica de la comprensión auditiva/audiovisual ni en la didáctica del vídeo en el aula de LE.

**TABLA 1.** Ítem 1. ¿Conoce estudios sobre la integración de preguntas de comprensión como subtítulos y sincronizadas con los fragmentos a los que se refieren en la didáctica de la comprensión auditiva / audiovisual o en la didáctica del vídeo en el aula de LE?

	Frecuencia	Porcentaje	Moda
Válidos			
Sí	1	16.7	
No	5	83.3	X
Total	6	100.0	

ÍTEM 1b. Si ha respondido sí, por favor, apunte la(s) referencia(s)

Un experto indica las revistas *Computer Assisted Language Learning (CALL)* y *Computer Assisted Language Instruction Consortium (CALICO)*. Posteriormente hubo oportunidad de volver a hablar con esta persona y matizó que sí conocía estudios sobre subtitulación pero no sobre la subtitulación de las preguntas.

Para confirmar si había investigaciones relacionados con las preguntas subtituladas en estas revistas, se leyeron los títulos de los

artículos que contenían los términos *listening, video, viewing, comprehension, synchronized / asynchronous* y *question(s)*, y los resúmenes de los artículos cuyos títulos sugerían que podrían contener alguna referencia. Se encontraron artículos sobre la subtitulación de la voz del vídeo, de palabras clave y de *advanced organisers*, pero no sobre la subtitulación de las preguntas de comprensión audiovisual.

### II. DESCRIPCIÓN DEL PROCEDIMIENTO

Los enunciados de las preguntas de comprensión audiovisual aparecen en pantalla segundos antes de que empiece el fragmento al que se refieren, permanecen visibles durante la duración del fragmento y desaparecen cuando termina. De acuerdo con Buck (2001), en casi todas las situaciones en las que se escucha fuera del aula, los oyentes saben por qué están escuchando y la información que necesitan, y los estudiantes valoran positivamente poder leer las preguntas antes de la audición. Este último aspecto lo confirma un estudio anterior de Sherman (1997) sobre el efecto de la lectura previa de las preguntas en tests de comprensión auditiva. En las respuestas de los examinados a un cuestionario se aprecia "the high psychological value of question preview for test-takers and, conversely, the negative impressions apparently created by conditions of uncertainty" (206). Presentar cada pregunta poco antes y durante el fragmento al que se refieren, permite que los estudiantes tengan presente el objetivo en cada fragmento, y respetar la preferencia de los estudiantes. Si las preguntas fueran visibles mucho antes y/o después, y durante el fragmento con el que se relacionan, no dejarían de contribuir a evitar el conflicto de atención entre ver el vídeo y leer la pregunta en un folio, pero verían disminuido su efecto de facilitar la comprensión a través de la focalización en los fragmentos más relevantes en cada momento. Si las preguntas no aparecieran durante el fragmento al que se refieren, podrían perjudicar la comprensión porque podrían inducir al estudiante a buscar la respuesta en el fragmento equivocado. La temporalización propuesta se supone la más adecuada para facilitar la concentración y poder guiar al estudiante a lo largo del texto, y estos son, según Underwood (1989), los objetivos de las actividades de comprensión. Para evitar posibles confusiones, se considera necesario informar a los estudiantes del criterio seguido en la colocación de las preguntas.



**TABLA 2.** Ítem 2. ¿En qué medida está de acuerdo con que la temporalización propuesta es la más adecuada para facilitar la concentración y guiar al estudiante a lo largo del vídeo?

	Frecuencia	Porcentaje	Modas
Válidos De acuerdo	3	50.0	X
Muy de acuerdo	3	50.0	X
Total	6	100.0	

ÍTEM 2b. *Comentarios de los expertos*

Se enfatiza la necesidad de que las preguntas estén sincronizadas porque de otro modo producirían “ruido” e “interferencias”, y se recomienda que el enunciado aparezca en un momento propicio para la lectura, es decir, en un instante en el que no ocurra nada importante. De lo contrario “habría un conflicto entre ver/escuchar frente a leer (y sería muy probablemente razón para frustrar a los estudiantes)”. También se matiza la afirmación de Buck (2001) de que en casi todas las situaciones en las que se escucha fuera del aula, los oyentes sepan por qué están escuchando y la información que necesitan, dado que el abanico de situaciones fuera del aula es muy amplio: “en un aeropuerto o en una frutería puede haber un proceso de anticipación/previsión, pero no siempre en la conversación con un amigo/un conocido o con la audición de informaciones en los medios de comunicación”.

Ciertos aspectos de la configuración de los enunciados de las preguntas se basan en convenciones establecidas en el uso de subtítulos (según Díaz, 2003):

a) Los subtítulos suelen estar compuestos por un máximo de dos líneas para limitar su impacto en la imagen. En el subtítulado intralingüístico para sordos a veces

pueden ser tres y hasta cuatro. Los enunciados de las preguntas serán preferentemente monolineales o bilineales.

ÍTEM 3. *Comentarios de los expertos*

Se valora positivamente esta convención (“totalmente de acuerdo”, “muy de acuerdo”), se aboga por la brevedad (“el límite debiera ser dos líneas”, “las preguntas han de ser cortas”) y se precisa que estas consideraciones son distintas en otros idiomas, por ejemplo, los subtítulos en japonés pueden aparecer en horizontal o en vertical. Uno de los expertos se pregunta cuál puede ser el papel de descodificación lectora en la comprensión auditiva / audiovisual y advierte que “son muchos los estudiantes que apuntan que han entendido los textos auditivos / audiovisuales, pero que no han entendido las preguntas escritas, sobre todo con una gran base léxica”.

b) Los subtítulos aparecen en la parte inferior de la pantalla, salvo cuando haya información visual imprescindible en esa área. Para evitar que los subtítulos dañen la imagen estética del texto audiovisual, algunas distribuidoras encajan la película entre dos franjas negras, una por encima y otra por debajo. En la franja inferior se proyectan los subtítulos. En la edición se intentará situar los enunciados de las preguntas debajo de la fotografía para no invadirla.

ÍTEM 4. *Comentarios de los expertos*

Esta pauta se considera adecuada (“me parece correcto”, “de acuerdo”), especialmente la proyección del subtítulo en una franja negra.

c) Generalmente los subtítulos constan de veintiocho a cuarenta espacios y caracteres. Un número superior de caracteres suele implicar una reducción del tamaño de letra y una pérdida de legibilidad.

#### ÍTEM 5. *Comentarios de los expertos*

Se estima conveniente seguir esta indicación (“totalmente de acuerdo”, “sí”).

d) El tipo de letra debe ser claramente legible. Se suele utilizar Helvética, Arial o Times New Roman.

#### ÍTEM 6. *Comentarios de los expertos*

Se evalúa positivamente esta consideración (“debe ser claramente legible”, “totalmente de acuerdo”). Algunos expertos sugieren la utilización de una fuente sin remates por varias razones. En primer lugar, son más “amables” que las fuentes con remates y esto es “relevante para un aprendiente en un estado de tensión”. En segundo lugar, se utilizan habitualmente con textos breves. Y en último lugar, “al parecer soportan también mejor el pixelado”.

e) El color ha de ser fácilmente legible. El blanco y el amarillo son los más utilizados.

#### ÍTEM 7. *Comentarios de los expertos*

Cuando el subtítulo se proyecta sobre una franja negra todos los expertos consideran que tanto el blanco como el amarillo son

adecuados, salvo uno que aboga por el uso del blanco porque “es más estable que el amarillo (un mismo amarillo puede adquirir tonos diferentes en función del sistema de visionado: proyector, resolución de pantalla si el material audiovisual se proyecta desde un ordenador, etc.)”. Si el subtítulo invade la imagen, depende del vídeo (“en una película con mucha nieve el blanco no es buena idea”).

f) A partir de estudios sobre la velocidad de lectura del espectador medio, se recomienda que el subtítulo de una línea, que suele tener siete u ocho palabras, se mantenga en pantalla durante aproximadamente cuatro segundos, y el de dos líneas, que suele tener entre catorce y dieciséis palabras, se mantenga unos seis segundos. El estudiante de LE necesitará más o menos tiempo en función de su competencia de lectura en la LE, pero esta convención sirve de orientación para determinar el tiempo que precisa el estudiante para comprender cada pregunta, y que a su vez, se utiliza para establecer el tiempo que cada pregunta debe aparecer antes del fragmento al que se refiere. En el caso de enunciados superiores a dos líneas se utiliza una equivalencia temporal a partir de la “regla de los seis segundos” para subtítulos de dos líneas.

#### ÍTEM 8. *Comentarios de los expertos*

La mayor parte de los expertos considera adecuada esta temporalización. Uno de los especialistas recomienda “preguntas breves” que no excedan las dos líneas. Otro profesor advierte que “los estudiantes no siempre prestan atención al *input* visual”, y que esto puede reducir el tiempo disponible para leer el subtítulo e incluso suponer que pase inadvertido.

g) Se debe segmentar los subtítulos del modo más lógico posible para facilitar la lectura.

### ÍTEM 9. *Comentarios de los expertos*

Se valora positivamente esta convención (“totalmente de acuerdo”, “muy de acuerdo”).

#### Preguntas en pantalla y preguntas en papel

La aparición de las preguntas en pantalla es perfectamente compatible con su entrega en papel. Las preguntas de comprensión auditiva tienen poca significación por sí solas, y solo cobran sentido durante la audición del texto (Sherman, 1997), pero los estudiantes valoran positivamente poder leer las cuestiones antes de la audición (Sherman, 1997; Buck, 2001). La lectura de las preguntas antes del visionado permite resolver las dudas que puedan generar, si la respuesta de los estudiantes va a ser escrita, la facilita, y permite ampliar la extensión o la forma de la actividad planteada. Por ejemplo, si la actividad implica elegir entre varias frases breves o imágenes, se presentaría en papel la actividad completa, y en la imagen solo aparecería el enunciado de la pregunta. Dicho esto, hay que tener en cuenta que las actividades que se realicen de forma simultánea al visionado deben ser sencillas por la gran dificultad que supone ver, escuchar, leer y escribir al mismo tiempo.

### ÍTEM 10. *Comentarios de los expertos*

Se recomienda que las preguntas sean “sencillas y concretas” y que se utilicen preferiblemente con “niveles bajos de conocimiento de la LE”, que haya congruencia entre las preguntas en papel y las de la pantalla porque “una sola palabra de diferencia puede despistar mucho”, y se alerta de los “conflictos que puede generar la existencia de dos o más planos de trabajo”.

#### III. POSIBLES BENEFICIOS

Frente a un visionado en el que el vídeo se proyecta en una pantalla común y las preguntas están disponibles en papel, la aparición de las preguntas segundos antes del fragmento al que se refieren y durante toda la duración del fragmento en cuestión puede aportar los siguientes beneficios:

a) Ver un vídeo y simultáneamente leer / responder preguntas de comprensión audiovisual en un folio genera un conflicto de atención visual. La integración de los enunciados de las preguntas en la imagen del vídeo no elimina el problema, pero sí reduce el conflicto entre ver y leer las preguntas al presentar dentro del mismo campo visual el vídeo y el enunciado de la pregunta de comprensión relacionada con el fragmento.

TABLA 3. Ítem 11. ¿En qué medida está de acuerdo con la parte subrayada arriba?

	Frecuencia	Porcentaje	Modas
Válidos			
De acuerdo	3	50.0	X
Muy de acuerdo	3	50.0	X
Total	6	100.0	

### ÍTEM 11. *Comentarios de los expertos*

Se considera que reduce el conflicto y “focaliza la atención” en los momentos importantes en relación con la pregunta. Asimismo, se recuerda algo ya señalado en el ítem 8, la posibilidad de que el subtítulo pase parcial o totalmente desapercibido si el estudiante no está concentrado en la imagen.

b) Facilita que los estudiantes orienten su comprensión en función de aquella información que interesa comprender en cada uno de los fragmentos del vídeo, y que no tengan que prestar atención a varias preguntas al mismo tiempo. En otras palabras, permite la concentración en una sola cuestión en cada fragmento. Una de las dificultades de la comprensión auditiva/audiovisual en una LE procede de la necesidad de procesar la información en tiempo real y de forma simultánea a la emisión del *input* (Martín Leralta, 2012). Escuchar, leer y escribir simultáneamente es extremadamente difícil, especialmente en una lengua extranjera (Underwood, 1989). Escuchar, ver, leer y escribir puede serlo todavía más. Favoreciendo que el estudiante solo tenga que prestar atención a una pregunta y no a varias simultáneamente, puede disponer de más memoria de trabajo para la comprensión del texto

audiovisual. Field (2008) explica que los oyentes inexpertos tienen los procesos de descodificación poco automatizados y tienen que dedicar gran parte de la memoria de trabajo a la descodificación. Esta ayuda puede ser especialmente beneficiosa para estudiantes de niveles bajos, ya que son los que pueden ver sobrecargada su memoria de trabajo más rápidamente.

ÍTEM 13b. *Comentarios de los expertos*

Hay muestras de acuerdo (“sí”, “parece lógico”), un experto matiza que es verdadera si “la siguiente pregunta no está vinculada con la anterior”, y otro considera que “la ansiedad creada por la pérdida de información puede perdurar durante toda la audición”.

TABLA 4. 12. ¿En qué medida está de acuerdo con la parte subrayada arriba?

		Frecuencia	Porcentaje	Moda
Válidos	De acuerdo	2	33.3	
	Muy de acuerdo	4	66.7	X
	Total	6	100.0	

ÍTEM 12. *Comentarios de los expertos*

Se muestra acuerdo con la afirmación (“intuitivamente creo que sí”), se considera que la técnica es “provechosa” para niveles bajos, y se observa que el procedimiento puede limitar “el perfil de los ítems”, especialmente el “trabajo de comprensión global y también de interpretación”.

c) En caso de que el alumno pierda la concentración, facilita que continúe la actividad a partir de la aparición de la siguiente pregunta en la pantalla.

TABLA 5. 13. ¿En qué medida está de acuerdo con el punto anterior?

		Frecuencia	Porcentaje	Moda
Válidos	En desacuerdo	1	16.7	
	De acuerdo	2	33.3	
	Muy de acuerdo	3	50.0	X
	Total	6	100.0	

d) Las propiedades acabadas de mencionar (a) reduce el conflicto de atención por la integración de los enunciados de las preguntas dentro del mismo campo visual; b) orienta la comprensión a lo largo del vídeo y dispensa de tener que prestar atención a varias preguntas simultáneamente; y c) facilita continuar la actividad en caso de perder la concentración) pueden contribuir a reducir el posible estrés generado por el visionado y la actividad frente a visionados con las preguntas disponibles en papel, y la disminución del estrés, podría favorecer la comprensión. Elkhafaifi (2005) investiga el efecto de la ansiedad general del aprendizaje de una lengua extranjera y de la ansiedad de la comprensión auditiva/audiovisual en el rendimiento de estudiantes de árabe como LE. A partir de los resultados, el autor sugiere que la ansiedad general del aprendizaje y la ansiedad de la comprensión son fenómenos separados y relacionados, y que están correlacionados negativamente con el rendimiento.

TABLA 6. 14. ¿En qué medida está de acuerdo con la siguiente parte del párrafo anterior: “las propiedades acabadas de mencionar (...) pueden contribuir a reducir el posible estrés generado por el visionado y la actividad frente a visionados con las preguntas disponibles en papel”?

		Frecuencia	Porcentaje	Moda
Válidos	Ni de acuerdo ni en desacuerdo	1	16.7	
	De acuerdo	3	50.0	X
	Muy de acuerdo	2	33.3	
	Total	6	100.0	

ÍTEM 14b. *Comentarios de los expertos*

Se pone de manifiesto que el procedimiento es más efectivo si el estudiante “está avisado de que las preguntas irán apareciendo en pantalla”. También se señala que el estrés puede estar generado por otras razones (“examen final”). Un experto considera que los factores fundamentales para la reducción del estrés radican en que las preguntas aparezcan una detrás de otra y en que permanezcan legibles durante un cierto tiempo.

TABLA 7. 15. ¿En qué medida está de acuerdo con la siguiente parte del párrafo anterior: “la disminución del estrés, podría favorecer la comprensión”?

	Frecuencia	Porcentaje	Moda
De acuerdo	4	66.7	X
Válidos Muy de acuerdo	2	33.3	
Total	6	100.0	

ÍTEM 15b. *Comentarios de los expertos*

Se muestra acuerdo con la afirmación (“el estrés es un obstáculo claro en el proceso de aprendizaje”). Se comenta que plantear preguntas también puede generar estrés, y que las preguntas en la imagen podrían reducirlo al no tener que prestar atención a varias cuestiones simultáneamente y saber reincorporarse.

e) Puede favorecer que los estudiantes presten atención al vídeo durante más tiempo en comparación con visionados que presenten las preguntas solo en papel, debido a que se reduce el conflicto de atención visual a través de la integración en la pantalla, y a que se evita tener que estar pendiente de varias preguntas simultáneamente. Alderson, Clapham y Wall (1995) diseñan un test de comprensión con vídeo y lo pilotan obteniendo resultados que parecen aceptables. Sin embargo, en sucesivas administraciones del test se dan cuenta de que muchos examinandos no miran el vídeo, en su lugar, leen la hoja de preguntas y responden a partir de lo que escuchan.

Wagner (2007) estudia la conducta visual de 36 estudiantes en una prueba de comprensión con vídeo en la que las preguntas se presentaban por escrito y obtiene como resultado que los estudiantes observaron el vídeo un 69 % del tiempo. Wagner sugiere que los informantes interactuaron en gran medida con el vídeo y que los estudiantes percibieron las imágenes como útiles para responder a las preguntas. Sin embargo, se puede hacer otra lectura: durante casi un tercio del tiempo los estudiantes no vieron el vídeo que estaban intentando comprender. Si se propone actividades con vídeo para desarrollar la comprensión auditiva/audiovisual como recomiendan autores como Lynch (2009) y se asume que la comprensión auditiva/audiovisual consiste en la interpretación de información auditiva y visual (Rubin, 1995a; Gruba, 1997), no parece razonable que los estudiantes no observen el vídeo durante una parte considerable del visionado.

TABLA 8. 16. ¿En qué medida está de acuerdo con la parte subrayada arriba?

	Frecuencia	Porcentaje	Modas
De acuerdo	3	50.0	X
Válidos Muy de acuerdo	3	50.0	X
Total	6	100.0	

ÍTEM 16b. *Comentarios de los expertos*

Hay comentarios a favor de este posible beneficio (“en teoría, sí”), y se apunta la posibilidad de que el efecto disminuya o se pierda si se mantiene los dos planos de trabajo (en pantalla y en papel).

f) Permite dividir el visionado en fragmentos. Esto resulta coherente con el principio de segmentación (segmenting principle) de Mayer (2006b: 180) para el aprendizaje multimedia: “people learn more deeply when a multimedia message is presented in user-paced segments rather than as a continuous unit”.

TABLA 9. 17. ¿En qué medida está de acuerdo con la parte subrayada arriba?

	Frecuencia	Porcentaje	Modas
Válidos			
En desacuerdo	1	16.7	
Ni de acuerdo ni en desacuerdo	1	16.7	
De acuerdo	2	33.3	X
Muy de acuerdo	2	33.3	X
Total	6	100.0	100.0

ÍTEM 17b. *Comentarios de los expertos*

Algunos expertos se muestran a favor (“totalmente de acuerdo”), un especialista critica la formulación del principio (“¿Qué significa *user-paced*? La gente es muy diferente (...) ¿qué significa *deeply*? Es algo muy relativo.”), y dos profesores cuestionan la traslación del principio de segmentación al desarrollo de la comprensión auditiva/audiovisual. Se objeta que el principio se concibió para el aprendizaje de una L1 y para “áreas de conocimiento relacionadas con la ciencia y tal vez parcelables”, y que el objetivo no es aprender una “cantidad de informaciones” sino “aprender a focalizar / estructurar la comprensión”.

IV. POSIBLES LIMITACIONES E INTERROGANTES

a) Este tipo de visionado podría no preparar adecuadamente para visionados con preguntas escritas en papel en los que no haya orientaciones sobre el lugar de aparición de la información que interese comprender.

TABLA 10. 18. ¿En qué medida está de acuerdo con la parte subrayada arriba?

	Frecuencia	Porcentaje	Moda
Válidos			
Ni de acuerdo ni en desacuerdo	4	66.7	X
De acuerdo	2	33.3	
Total	6	100.0	

ÍTEM 18b. *Comentarios de los expertos*

El enunciado genera dudas (“no lo sé”, “no creo que sean incompatibles los dos tipos de visionado, sobre todo en niveles altos”). Un experto opina que “el procedimiento es adecuado para una fase en el proceso de aprendizaje” y que después se prescindiría de ellas.

b) Las preguntas que supongan una división en fragmentos de muy pocos segundos pueden limitar el desarrollo de la capacidad de buscar información específica, debido a la corta duración del fragmento donde el estudiante busca la respuesta. Es un elemento que debe tenerse en cuenta si se pretende desarrollar esta capacidad.

TABLA 11. 19. ¿En qué medida está de acuerdo con la parte subrayada arriba?

	Frecuencia	Porcentaje	Moda
Válidos			
En desacuerdo	1	16.7	
Ni de acuerdo ni en desacuerdo	1	16.7	
De acuerdo	3	50.0	X
Muy de acuerdo	1	16.7	
Total	6	100.0	

ÍTEM 19b. *Comentarios de los expertos*

Todos los expertos se muestran de acuerdo (“sí, en teoría, sí”, “pero pueden ser muy útiles si lo que se pretende no afecta a esta capacidad”) salvo uno (“al contrario, ofrece una buena posibilidad de focalizar”). También se señala el efecto “no frustrante” de encontrar la respuesta en poco tiempo, y se añade que nada impide que una fase posterior se llegue “hasta una visión global”.

c) La extensión de las preguntas subtituladas es mucho más limitada que la de aquellas que puedan formularse en papel que los de aquellas preguntas que puedan formularse en papel. Como se ha mencionado, es posible presentar el enunciado de la pregunta en pantalla y la actividad completa en papel. Como también se ha comentado, un principio básico de las actividades escritas que se realicen durante la emisión del *input*, y con más razón si es *input* es audiovisual, es que impliquen poca lectura y escritura.

ÍTEM 20b. *Comentarios de los expertos*

No se ve “inconveniente”, y se recomienda el uso de preguntas breves y concisas.

d) La propia experiencia docente hasta el momento ha sido positiva, pero podría haber un rechazo inicial por parte de los estudiantes no familiarizados con esta práctica, y también por parte de aquellos que consideren que les proporciona una ayuda que no necesitan.

TABLA 12. 21. ¿En qué medida está de acuerdo con la siguiente parte del párrafo anterior: “podría haber un rechazo inicial por parte de los estudiantes no familiarizados con esta práctica”?

	Frecuencia	Porcentaje	Moda	
Válidos	En desacuerdo	3	50.0	X
	Ni de acuerdo ni en desacuerdo	2	33.3	
	Muy de acuerdo	1	16.7	
	Total	6	100.0	

ÍTEM 21b. *Comentarios de los expertos*

Se estima que hay pocas posibilidades de que no sea aceptado (“me parece difícil que alguien rechace este tipo de método docente”, “no debiera comportar rechazo por la simplicidad del procedimiento”). Un experto considera que depende de la forma de guiar a los estudiantes, y otro comenta que “los estudiantes rechazan todos los cambios” basándose en su experiencia docente en asignaturas de informática.

TABLA 13. 22. ¿En qué medida está de acuerdo con la siguiente parte del párrafo anterior: “podría haber un rechazo inicial por parte de aquellos que consideren que les proporciona una ayuda que no necesitan”?

	Frecuencia	Porcentaje	Moda	
Válidos	En desacuerdo	1	16.7	
	Ni de acuerdo ni en desacuerdo	3	50.0	X
	De acuerdo	1	16.7	
	Muy de acuerdo	1	16.7	
	Total	6	100.0	

ÍTEM 22b. *Comentarios de los expertos*

Varios expertos manifiestan sus dudas sobre el enunciado (“no sé hasta que punto puede suceder”, “puede ser, pero sería algo infantil”), uno considera que “puede suponer una interferencia en la atención del fragmento de vídeo”, y otro que el posible rechazo no tendría que ver con el procedimiento sino con el contenido didáctico (“es irritante cuando profesor emplea el tiempo en algo que ya sabes de sobra”).

V. PREGUNTAS GENERALES

TABLA 14. 23. ¿En qué medida está de acuerdo con el siguiente enunciado: el procedimiento puede ser útil para la enseñanza de la comprensión auditiva / audiovisual con grabaciones en vídeo en el aula de LE?

		Frecuencia	Porcentaje	Modas
Válidos	De acuerdo	3	50.0	X
	Muy de acuerdo	3	50.0	X
	Total	6	100.0	

TABLA 15. 24. ¿En qué medida está de acuerdo con el siguiente enunciado: el marco teórico propuesto está fundamentado?

		Frecuencia	Porcentaje	Moda
Válidos	Ni de acuerdo ni en desacuerdo	1	16.7	
	De acuerdo	5	83.3	X
	Total	6	100.0	

A continuación tiene un espacio en el que puede hacer cualquier comentario que considere oportuno sobre el procedimiento descrito o las preguntas que se le han planteado.

*Comentarios de los expertos*

La técnica se valora positivamente (“es un procedimiento mucho más efectivo que el que trabaja con vídeo y papel”, “es muy útil para mejorar la comprensión auditiva”), se sugiere que “el subtítulo activa la atención del alumno”, un experto recomienda la recogida de datos de muchos sujetos para validar las propuestas, y otro manifiesta que aunque el marco está fundamentado intervienen demasiadas variables.

## 5. DISCUSIÓN

### 5.1. ¿EN QUÉ MEDIDA LOS EXPERTOS CONSULTADOS ESTÁN DE ACUERDO CON EL MARCO TEÓRICO?

Existe un alto grado de convergencia entre la información cuantitativa y cualitativa (véase la tabla 16). En relación con los apartados “descripción del procedimiento” y “posibles beneficios”, la moda de las escalas Likert y los comentarios indican que la mayor parte de los especialistas está conforme con los planteamientos. En cambio, hay algunas dudas y aparece “un desacuerdo” en la sección “posibles limitaciones e interrogantes”. En cuanto a las preguntas generales, cabe destacar que los expertos están “de acuerdo” o “muy de acuerdo” con la utilidad del procedimiento para la enseñanza de la comprensión auditiva / audiovisual y que estiman que el marco teórico está fundamentado.



TABLA 16. **Moda de las escalas Likert y ejemplos de comentarios de los expertos**

	Moda <sup>a</sup>	Ejemplos de comentarios
<b>Descripción del procedimiento</b>		
Ítems 2 y 2a (temporalización de la pregunta: pocos segundos antes del fragmento y hasta su finalización)	De acuerdo / muy de acuerdo	“La «desincronización» de las preguntas produciría «ruido» e «interferencias»”
<b>Posibles beneficios (frente a un visionado con las preguntas en papel)</b>		
Ítems 11 y 11b (reducción del conflicto de atención al presentar dentro del mismo campo visual el vídeo y la pregunta)	De acuerdo / muy de acuerdo	“No solo reduce el conflicto, también focaliza la atención”
Ítems 12 y 12b (focalización de la atención en los momentos relevantes)	Muy de acuerdo	“Intuitivamente creo que sí”, “la afirmación es razonable, pero al mismo tiempo limitadora del perfil de los ítems objeto”
Ítems 13 y 13b (en caso de perder la concentración, facilita continuar la actividad a partir de la aparición de la siguiente pregunta subtitulada)	Muy de acuerdo	“Sí”, “parece lógico”
Ítems 14 y 14b (reducción del estrés gracias a las propiedades mencionadas en 11, 12 y 13)	De acuerdo	“Sobre todo si el alumno está avisado de que las preguntas irán apareciendo en pantalla”
Ítems 15 y 15b (la reducción del estrés puede favorecer la comprensión)	De acuerdo	“El estrés es un obstáculo claro en el proceso de aprendizaje”
Ítems 16 y 16b (incremento del tiempo de atención al vídeo)	De acuerdo / muy de acuerdo	“En teoría, sí, pero hay que analizar este aspecto con los estudiantes”

Ítems 17 y 17b (permite dividir el visionado en fragmentos)	De acuerdo / muy de acuerdo	“Totalmente de acuerdo”, “Mayer y Moreno trabajan siempre con L1 y con (...) áreas de conocimiento relacionadas con la ciencia”
<b>Posibles limitaciones e interrogantes</b>		
Ítems 18 y 18b (preparación inadecuada para visionados sin orientaciones sobre el lugar de aparición de la información relevante)	Ni de acuerdo ni en desacuerdo	“No lo sé”, “no creo que sean incompatibles los dos tipos de visionado”
Ítems 19 y 19b (limitación de la capacidad de buscar información específica si los fragmentos son muy breves)	De acuerdo	“Sí, en teoría, sí”, “pero pueden ser muy útiles si lo que se pretende no afecta a esta capacidad”
Ítems 21 y 21b (rechazo inicial por desconocimiento)	En desacuerdo	“Me parece difícil que alguien rechace este tipo de método docente”, “no veo la razón”
Ítems 22 y 22b (rechazo inicial por considerarse una ayuda innecesaria)	Ni de acuerdo ni en desacuerdo	“No sé hasta que punto puede suceder”, “puede ser, pero sería algo infantil”
<b>Preguntas generales<sup>b</sup></b>		
Ítem 23 (útil para la enseñanza de la comprensión auditiva/audiovisual en el aula)	De acuerdo / muy de acuerdo	“Es un procedimiento mucho más efectivo que el que trabaja con vídeo y papel”
Ítem 24 (marco teórico fundamentado)	De acuerdo	“Sí, lo está, pero el territorio es demasiado vasto”

a. Cuando hay dos modas se muestran ambas.

b. Los comentarios provienen de la sección del cuestionario en la que se puede añadir información adicional.

Ciertos aspectos de la configuración de los subtítulos se basan en pautas para la subtitulación en español (según Díaz, 2003). Se ha prescindido de las escalas de actitudes por tratarse de convenciones, pero los comentarios de los expertos sugieren que se valoran positivamente (véase la tabla 17).

TABLA 17. Aspectos del procedimiento basados en convenciones de la subtitulación en español: ejemplos de comentarios de los expertos

Ejemplos de comentarios	
Descripción del procedimiento	
Ítem 3 (número de líneas recomendado para el subtítulo: dos)	"Totalmente de acuerdo", "de acuerdo. Estas convenciones son distintas en otros idiomas"
Ítem 4 (localización de la pregunta: en la parte inferior sobre una franja negra que no invada la imagen)	"Posiblemente sea una buena opción", "de acuerdo"
Ítem 5 (extensión del subtítulo: de veintiocho a cuarenta espacios y caracteres)	"Totalmente de acuerdo", "sí"
Ítem 6 (tipo de letra del subtítulo: Helvética, Arial o Times New Roman)	"Debe ser claramente legible (...) preferiblemente Arial"
Ítem 7 (color del subtítulo: blanco o amarillo)	"Sí, si el fondo es negro", "blanco o amarillo bien"
Ítem 8 (tiempo necesario para la lectura de la pregunta: "regla de los seis segundos" para los subtítulos de dos líneas y competencia de lectura)	"Totalmente de acuerdo", "los estudiantes no siempre prestan atención al <i>input</i> visual (...) la falta de atención (...) puede ocasionar que esa duración de seis segundos no sea real"
Ítem 9 (segmentación de un modo lógico que facilite la lectura)	"Totalmente de acuerdo: no partir palabras y evitar romper la unidad del significado"

En resumen, teniendo en cuenta los datos cuantitativos y cualitativos, se puede afirmar que, globalmente, los expertos están

de acuerdo con la mayor parte de los planteamientos del marco teórico provisional.

## 5.2. ¿QUÉ ASPECTOS DEL MARCO TEÓRICO SE DEBERÍA MODIFICAR, SEGÚN LOS EXPERTOS CONSULTADOS?

Aunque, en conjunto, los expertos están de acuerdo con la mayor parte de las propuestas, sus comentarios revelan puntos susceptibles de mejora. A continuación se examina las observaciones más relevantes:

a) Ítem 2b (temporalización de la pregunta: pocos segundos antes del fragmento y hasta su finalización). Se acepta las propuestas de los expertos. Por un lado, se suscribe la recomendación de que la pregunta subtitulada aparezca en un momento adecuado para la lectura, es decir, en un momento en el que no ocurra nada fundamental, para no generar un conflicto intenso entre el deseo de ver el vídeo y la necesidad de leer la pregunta. Por otro lado, se reconoce que es cuestionable la afirmación –que se había asumido– "in virtually all real-world listening situations, listeners know why they are listening to and what information they need to get from the text" (Buck, 2011: 137), porque hay situaciones en las que el grado de anticipación puede ser bajo o muy bajo, por ejemplo, cuando se ve el telediario o una película, o cuando se conversa con un amigo.

b) Ítem 3 (número de líneas recomendado para el subtítulo: dos). Se comparte la opinión de que si las actividades están mal diseñadas puede ocurrir que se entienda la audición/el vídeo pero no las preguntas de comprensión. Consideramos que las cuestiones no deberían ser más complicadas que el texto que se pretende comprender, ni tampoco deberían dificultar la comprensión. Todo lo contrario, los objetivos principales de las actividades de comprensión

son facilitar la concentración y guiar al estudiante a lo largo del texto (Underwood, 1989).

c) Ítem 6 (tipo de letra del subtítulo: Helvética, Arial o Times New Roman). Se acepta la sugerencia de emplear preferentemente fuentes sin remates. En primer lugar, se utilizan habitualmente en textos breves. En segundo lugar, transmiten alegría y seguridad (Martínez, 2009), y esto podría ayudar a estudiantes en tensión. Y en tercer lugar, son más fáciles de leer en una pantalla porque se adaptan mejor a su resolución (Eguaras, 2014).

d) Ítem 7 (color del subtítulo: blanco o amarillo). Se acepta la recomendación de utilizar el blanco antes que el amarillo sobre la franja negra porque el amarillo es más inestable. Para reforzar la sugerencia se ha consultado la obra de Bright y Cook (2010). Estos autores también advierten de la inestabilidad de este color, y añaden que el blanco es neutral y puede combinarse con cualquier tonalidad. En caso de que el subtítulo invada la imagen, se comparte la necesidad de configurarlo en función de los colores de la fotografía.

e) Ítem 8 (tiempo necesario para la lectura de la pregunta: “regla de los seis segundos” para los subtítulos de dos líneas y competencia de lectura). Estamos de acuerdo con que los estudiantes no siempre mantienen la atención en la imagen y que esto puede reducir el tiempo disponible para leer el subtítulo e incluso ocasionar que pase inadvertido. Para disminuir la posibilidad de que la pregunta pase parcial o totalmente desapercibida, es fundamental explicar su existencia y temporalización. También es necesario que haya tiempo suficiente para responder a cada pregunta antes de que comience la siguiente, bien porque las preguntas aparecen en fragmentos lo suficientemente distantes, bien porque se recurre a la técnica del visionado con pausas, o bien por ambas razones. Y en tercer lugar,

es recomendable que las actividades escritas simultáneas al visionado impliquen poca lectura y escritura.

f) Ítem 10 (compatibilidad de las preguntas subtituladas y en papel). En los casos en los que se presente tanto las preguntas en pantalla como en papel, se comparte la sugerencia de que haya correspondencia total entre los enunciados porque, de lo contrario, sería una fuente de desorientación. También concordamos con que la coexistencia de dos planos de trabajo puede plantear inconvenientes. En este sentido, no se puede asumir, como se había hecho, que la duplicación de las preguntas en papel no mantenga en alguna medida el conflicto de atención visual que precisamente se quiere reducir.

g) Ítem 12b (focalización de la atención en los momentos relevantes). Estamos de acuerdo con que al focalizar en partes del vídeo se limita el “trabajo de comprensión global”. Para practicarla se está experimentando la subtitulación de una pregunta de comprensión general en los últimos segundos de vídeos breves (entre dos y tres minutos). Aunque es cierto que se pierde el carácter focalizador, permite mantener el plano de trabajo en pantalla frente a un visionado con preguntas en papel. También es perfectamente posible aplicar la estrategia docente sugerida por Richards y Burns (2012): efectuar visionados sucesivos con distintos objetivos de comprensión. En ese caso, por ejemplo, un primer visionado sin preguntas podría centrarse en la comprensión global y un segundo visionado con preguntas subtituladas en una comprensión más fina.

h) Ítem 13b (en caso de perder la concentración, facilita continuar la actividad a partir de la aparición de la siguiente pregunta subtitulada). Se comparte la necesidad de que la pregunta siguiente no dependa de la anterior para posibilitar la continuación de la actividad en caso de perder la concentración. Asimismo, se admite

que la ansiedad generada por la pérdida de información puede perdurar durante todo el visionado. Para profundizar sobre este último tema se ha consultado la obra de Greist y Jefferson (2000). Estos autores explican que la ansiedad puede durar desde instantes hasta largos periodos de tiempo.

i) Ítem 17b (permite dividir el visionado en fragmentos). Se reconoce que la aplicación del principio de segmentación (“people learn more deeply when a multimedia message is presented in user-paced segments rather than as a continuous unit”, Mayer, 2006b: 180) al aprendizaje de la comprensión auditiva/audiovisual en el aula de LE plantea varias dudas. En primer lugar, el principio está concebido para el aprendizaje de disciplinas del área de ciencias y tal vez más susceptibles de subdivisión. En segundo lugar, está pensado para el aprendizaje de una L1. Y en tercer lugar, el objetivo principal de las actividades de comprensión no es aprender contenidos sino guiar al estudiante a lo largo del texto.

j) Ítem 18b (preparación inadecuada para visionados sin orientaciones sobre el lugar de aparición de la información relevante). Se comparte la opinión de que el método docente es “adecuado para una fase en el proceso de aprendizaje” y que después se prescindiría de su utilización. No obstante, no se descarta su empleo puntual en otros niveles por dos motivos. Por una parte, permite mantener el plano de trabajo en pantalla (frente a visionados con preguntas en papel). Por otra, la experiencia docente en niveles B1 y B2 ha sido intuitivamente positiva.

k) Ítem 21b (rechazo inicial por desconocimiento). El ítem 21 ha sido el único en el que la respuesta más frecuente ha sido “en desacuerdo”. Varios expertos manifiestan también su disconformidad en el ítem 21b, pero solo se aporta un argumento: la sencillez del procedimiento. Aunque nuestra impresión es que la

metodología ha sido bien recibida por los estudiantes, solo preguntando a los alumnos se podrá saber si existe o no rechazo y cuál es la causa.

l) Ítem 22b (rechazo inicial por considerarse una ayuda innecesaria). Consideramos improbable que la pregunta subtitulada pueda “suponer una interferencia en la atención del fragmento de vídeo” o, al menos, que sea de una intensidad suficiente para distraer, si aparece en un momento adecuado para la lectura y solo es visible poco antes y durante el fragmento al que se refiere. Por un lado, los subtítulos forman parte de la sociedad moderna y, sin duda, son habituales para los estudiantes de lenguas extranjeras. Por otro lado, las preguntas subtituladas son breves e implican poca lectura. Asimismo, si se sigue la recomendación de encajar el vídeo entre dos franjas negras y situar el subtítulo en la banda inferior, se evita invadir y dañar la imagen estética del texto audiovisual.

### 5.3. REFORMULACIÓN DEL MARCO TEÓRICO PROVISIONAL

La consideración de las opiniones de los expertos ha llevado a la modificación de la propuesta inicial. Pasa a consistir en:

#### I. INTRODUCCIÓN

El presente marco teórico describe el uso de preguntas de comprensión audiovisual integradas como subtítulos en la fase de visionado en el aula de LE. Esta propuesta metodológica se ha elaborado a partir de la revisión de la bibliografía, la propia experiencia docente con el procedimiento y las valoraciones y comentarios de seis expertos en didáctica del ELE, lingüística española y/o en el uso de la tecnología. Cabe destacar que a excepción de las observaciones del propio autor (Casañ, 2007,

2012), no se ha encontrado referencias sobre esta técnica ni en la didáctica de la comprensión auditiva/audiovisual ni en la didáctica del vídeo en el aula de LE.

## II. DESCRIPCIÓN DEL PROCEDIMIENTO

Los enunciados de las preguntas de comprensión audiovisual aparecen en pantalla segundos antes de que empiece el fragmento al que se refieren, permanecen visibles mientras dura el fragmento y desaparecen cuando termina. Esta temporalización permite que los estudiantes tengan presente el objetivo de comprensión en cada momento. Si las preguntas fueran visibles mucho antes y/o después, además de durante, la parte del vídeo pertinente, verían disminuido su efecto de facilitar la comprensión a través de la focalización. Si las preguntas no se mostraran durante los fragmentos relevantes, podrían inducir al estudiante a buscar la respuesta en el fragmento equivocado. La temporalización propuesta se supone la más adecuada para facilitar la concentración y poder guiar al estudiante a lo largo del texto, y estos son, según Underwood (1989), los objetivos principales de las actividades de comprensión en la fase de audición / visionado. Para evitar posibles confusiones, se considera necesario informar a los estudiantes del criterio seguido en la colocación de las preguntas subtítuladas.

Ciertos aspectos de la configuración de las preguntas se basan en convenciones establecidas en el uso de subtítulos en español (según Díaz, 2003):

a) Los subtítulos suelen estar compuestos por un máximo de dos líneas para limitar su impacto en la imagen. En el subtítulado intralingüístico para sordos a veces pueden ser tres y hasta cuatro.

Los enunciados de las preguntas subtítuladas serán preferentemente monolineales o bilineales.

b) Los subtítulos aparecen en la parte inferior de la pantalla, salvo cuando haya información visual imprescindible en esa área. Para evitar que los subtítulos dañen la imagen estética del texto audiovisual, algunas distribuidoras encajan la película entre dos franjas negras, una por encima y otra por debajo. En la franja inferior se proyectan los subtítulos. Se recomienda seguir esta estrategia para que la pregunta subtítulada no invada la fotografía.

c) Generalmente los subtítulos constan de veintiocho a cuarenta espacios y caracteres. Un número superior de caracteres suele implicar una reducción del tamaño de letra y una pérdida de legibilidad.

d) El tipo de letra debe ser claramente legible. Se suele utilizar Helvética, Arial o Times New Roman. Se recomienda emplear fuentes sin remates en las preguntas subtítuladas. En primer lugar, este tipo de fuente se usa habitualmente en textos breves. En segundo lugar, según Martínez (2009), transmiten alegría y seguridad y, esto podría beneficiar a los estudiantes en estado de tensión. Y en tercer lugar, son más fáciles de leer en una pantalla porque se adaptan mejor a su resolución (Eguaras, 2014).

e) El color ha de ser fácilmente legible. El blanco y el amarillo son los más utilizados. Según Bright y Cook (2010), el amarillo es un color inestable, mientras que el blanco es neutral y puede combinarse con cualquier tonalidad. Por estas razones, se sugiere la utilización del blanco sobre la franja negra mencionada anteriormente. En caso de que la pregunta subtítulada invada la imagen, es necesario configurarla en función de los colores de la fotografía.

f) A partir de estudios sobre la velocidad de lectura del espectador medio, se recomienda que el subtítulo de una línea, que suele tener siete u ocho palabras, se mantenga en pantalla durante aproximadamente cuatro segundos, y el de dos líneas, que suele tener entre catorce y dieciséis palabras, se mantenga unos seis segundos. En el caso de enunciados superiores a dos líneas se utiliza una equivalencia temporal a partir de la “regla de los seis segundos” para subtítulos de dos líneas. Aunque el estudiante necesitará más o menos tiempo para leer el subtítulo en una determinada LE (o L1 si se utiliza la lengua materna de los estudiantes) en función de su competencia de lectura en esa lengua, esta convención puede servir de orientación para determinar el tiempo necesario para leer cada pregunta subtitulada. Ahora bien, para determinar con cuánta antelación tiene que ser visible la pregunta antes del fragmento, hay que considerar otros dos elementos. En primer lugar, es aconsejable que la pregunta aparezca en un momento propicio para la lectura, es decir, en un momento en el que no ocurra nada fundamental, para no generar un conflicto intenso entre el deseo de ver el vídeo y la necesidad de leer la pregunta. En segundo lugar, los estudiantes no siempre mantienen la atención en la imagen. Algunos autores han investigado cuál es el comportamiento visual de estudiantes de LE cuando efectúan una prueba de comprensión audiovisual. Ockey (2007) no proporciona la media de tiempo que los participantes observan el vídeo, pero a partir de los datos se puede calcular que fue el 44,9 %, Wagner (2007) obtiene el 69 % y Wagner (2010) el 47,9%. Este hecho puede reducir el tiempo disponible para leer el subtítulo e incluso ocasionar que pase inadvertido. Para evitar o reducir esta posibilidad, es fundamental explicar la existencia y la temporalización de las preguntas subtituladas. También es necesario que haya tiempo suficiente para responder a cada pregunta antes de que comience la siguiente, bien porque las preguntas aparecen en fragmentos lo suficientemente distantes, bien porque se recurre a la técnica del visionado con pausas, o bien por ambas razones. Y

en tercer lugar, es recomendable que las actividades escritas simultáneas al visionado impliquen poca lectura y escritura.

g) Se debe segmentar los subtítulos del modo más lógico posible para facilitar la lectura.

#### PREGUNTAS SUBTITULADAS Y PREGUNTAS EN PAPEL

Sherman (1997) sugiere que las preguntas de comprensión auditiva tienen poca significación por sí solas, y que solo cobran sentido durante la audición del texto. No obstante, la entrega de las preguntas antes del visionado comporta algunos beneficios. Por un lado, respeta las preferencias de los aprendices. Según Sherman (1997) y Buck (2001) los estudiantes valoran positivamente poder leer las preguntas antes de la audición. Por otro lado, posibilita aclarar las dudas que puedan plantear. Asimismo, permite ampliar la extensión o la forma de la actividad. Dicho esto, hay que tener en cuenta que las actividades que se realicen de forma simultánea al visionado deben ser sencillas por la gran dificultad que supone ver, escuchar, leer y escribir al mismo tiempo. Asimismo, si las preguntas se presentan tanto en pantalla como en papel, es conveniente que exista total correspondencia entre los enunciados porque una sola palabra de diferencia puede generar confusión.

Aunque se puede presentar de forma simultánea las preguntas en papel y en pantalla durante la emisión del *input*, se desconoce en qué medida la coexistencia de los dos planos de trabajo mantiene el conflicto de atención visual que precisamente se quiere reducir con la subtitulación de las preguntas.

### III. POSIBLES BENEFICIOS

El procedimiento se estima especialmente beneficioso para estudiantes de niveles bajos. Asimismo, *a priori* se considera que esta técnica constituye una "ayuda extra" y que paulatinamente debería abandonarse su uso. Pese a esto, no se descarta su utilización puntual en otros niveles por dos razones. En primer lugar, favorece que se mantenga el plano de trabajo en pantalla frente a visionados con las preguntas en papel. En segundo lugar, la experiencia docente en los niveles B1 y B2 ha sido intuitivamente positiva.

Frente a un visionado en el que el vídeo se proyecta en una pantalla común y las preguntas están disponibles en papel, la aparición de las preguntas subtituladas segundos antes y durante el fragmento al que se refieren puede aportar los siguientes beneficios:

a) Ver un vídeo y simultáneamente leer / responder preguntas de comprensión audiovisual en un folio genera un conflicto de atención visual. Este conflicto contribuye a explicar el grado de (des)atención a la imagen en investigaciones que contabilizan el tiempo que los estudiantes miran en dirección a la pantalla en pruebas de comprensión con vídeo. Como se ha referido arriba, en los estudios de Ockey (2007) y Wagner (2007, 2010) los estudiantes orientan la mirada hacia el vídeo el 44,9 %, el 69 % y el 47,9% del tiempo respectivamente (o, desde otro punto de vista, los estudiantes no observan el vídeo el 55,1 %, 31 % y 52,1 % del tiempo). Si la comprensión auditiva/audiovisual es un proceso de interpretación de información auditiva y visual (Riley, 1979; Martín Peris, 1991; Rubin, 1995a; Lynch, 2012), un elevado grado de desatención puede repercutir negativamente tanto en la comprensión del vídeo como en el desarrollo de esta actividad comunicativa. La subtitulación de las preguntas no elimina el problema, pero sí lo

reduce al aproximar espacialmente las preguntas y los fragmentos relevantes. Gracias a ello, se favorece que los estudiantes presten atención al vídeo durante más tiempo en comparación con visionados con las preguntas en papel.

b) Facilita que los estudiantes orienten su atención en función de aquella información que interesa comprender en cada momento. De esta manera, está en consonancia con uno de los objetivos principales, según Underwood (1989) y Richards y Burns (2012), de las actividades durante el visionado: guiar al estudiante a lo largo del texto.

c) Al favorecer que solo se tenga que prestar atención a una pregunta subtitulada en cada momento, y no a varias preguntas en papel al mismo tiempo, el estudiante puede disponer de más memoria de trabajo para la comprensión del texto audiovisual. La memoria de trabajo se define como "a limited capacity system allowing the temporary storage and manipulation of information necessary for such complex tasks as comprehension, learning and reasoning" (Baddeley, 2000: 418). Para Miller (1956) el límite está en siete elementos, más o menos dos. Cowan (2001) considera que el límite en adultos es de tres a cinco. Como señalan Sweller et al. (2011), desde un punto de vista pedagógico, lo fundamental no es el número exacto, sino que la capacidad es reducidísima. Vandergrift y Goh (2012) alertan de que prestar atención a la imagen, al audio, y a la tarea simultáneamente puede llevar a la saturación de la memoria de trabajo. Los oyentes inexpertos tendrían más probabilidades de llegar a esta sobrecarga porque, de acuerdo con Field (2008), tienen los procesos de decodificación poco automatizados y tienen que dedicar gran parte de la memoria de trabajo a la decodificación. Por lo tanto, este método docente

podría ser especialmente beneficioso para estudiantes de niveles bajos<sup>3</sup>.

d) En caso de que un estudiante pierda la concentración, facilita que continúe la actividad a partir de la aparición de la siguiente pregunta en pantalla. Para que esto sea posible es necesario que la cuestión siguiente no dependa de la anterior. También hay que tener en cuenta que la pérdida de información puede causar ansiedad y que esta, según Greist y Jefferson (2000), puede durar desde pocos segundos hasta un tiempo prolongado.

e) Los posibles beneficios acabados de describir podrían contribuir a reducir el estrés generado por la actividad frente a visionados con las preguntas disponibles en papel, y si fuera así, favorecerían la comprensión. Arnold (2000) cita varios estudios que sugieren que “students in a foreign language classroom are often anxious and that this feeling inhibits their performance in the foreign language” (777). Elkhafaifi (2005) investiga el efecto de la ansiedad general del aprendizaje de una lengua extranjera y de la ansiedad de la comprensión auditiva/audiovisual en el rendimiento de estudiantes de árabe como LE. A partir de los resultados, el autor sugiere que la ansiedad general del aprendizaje y la ansiedad de la comprensión son fenómenos separados y relacionados, y que están correlacionados negativamente con el rendimiento. Cabe reconocer que la existencia de preguntas —ya sea en papel, en pantalla u orales— constituye una fuente de estrés debido a la necesidad de

---

<sup>3</sup> Una cuestión todavía no resuelta es si existen límites independientes para el procesamiento de los estímulos visuales y auditivos, o si hay sistema central de almacenamiento, o si la capacidad de la memoria de trabajo “is set both by task/modality and central, amodal processing limitations” (Fougnie y Marois, 2011: 1338). Si existiera un límite compartido podría haber interferencias cuando se procesa de forma simultánea información visual y auditiva (Cowan y Morey, 2005).

responder. A este respecto, conviene tener presente la recomendación de Richards y Burns (2012: 101): “it’s important to create activities where students feel they can succeed, especially for beginners”.

#### IV. POSIBLES LIMITACIONES E INTERROGANTES

a) Esta metodología podría no preparar adecuadamente para visionados sin orientaciones sobre el lugar de aparición de la información que interese comprender.

b) La focalización puede tener dos efectos secundarios. En primer lugar, puede limitar la práctica de la comprensión global. Para practicarla se está experimentando la subtitulación de una pregunta de comprensión general en los últimos segundos de vídeos breves. Aunque es cierto que se pierde el carácter focalizador, permite mantener el plano de trabajo en pantalla (frente a un visionado con preguntas en papel). En segundo lugar, puede tener el efecto contrario, puede condicionar el desarrollo de la capacidad de buscar información específica debido a la corta duración del fragmento donde el estudiante busca la respuesta. Para practicar distintos tipos de comprensión se puede utilizar la estrategia docente sugerida por Richards y Burns (2012): realizar audiciones / visionados sucesivos con distintos objetivos de comprensión. Por ejemplo, un primer visionado sin preguntas podría centrarse en la comprensión global y un segundo visionado con preguntas subtituladas en una comprensión más fina.

c) La extensión de las preguntas subtituladas es mucho más limitada que la de aquellas que puedan formularse en papel. Como se ha mencionado, es posible presentar los enunciados de las actividades en pantalla y las actividades completas en papel. Sin embargo, un



principio básico de las actividades escritas que se realicen durante la emisión del *input*, y con más razón si es *input* es audiovisual, es que impliquen poca lectura y escritura.

d) Aunque la experiencia docente con el procedimiento ha sido intuitivamente positiva, no se ha investigado cómo valoran los estudiantes este método docente. También se desconoce si los profesores de lenguas extranjeras estarían dispuestos a aprender a subtítular vídeos y a probar esta técnica.

## 6. CONCLUSIÓN

En este estudio mixto de tipo convergente (Creswell y Plano, 2011; Creswell, 2012, 2013), se ha preguntado a seis expertos en didáctica del ELE, lingüística española y/o en el uso de la tecnología de cinco instituciones diferentes qué opinan sobre un marco teórico que describe una metodología novedosa: el empleo de preguntas de comprensión audiovisual integradas en el vídeo como subtítulos en la fase de visionado. Se considera que este método docente podría aportar varios beneficios frente a una explotación en la que el vídeo se proyecta en una pantalla común y los materiales de la actividad se presentan en un folio. Entre ellos, podría reducir el conflicto de atención entre ver el vídeo y realizar la tarea, al aproximar espacialmente las preguntas y los fragmentos relevantes, y podría disminuir las necesidades cognitivas de la actividad al no tener que prestar atención a varias preguntas de forma simultánea. Las respuestas cuantitativas y cualitativas sugieren que, globalmente, los expertos están de acuerdo con la mayor parte de los planteamientos del marco teórico y, al mismo tiempo, revelan puntos susceptibles de mejora. La consideración de las valoraciones y comentarios de los expertos ha llevado a la reformulación de la propuesta inicial.

Este trabajo se contemplaba como la primera actuación en el proceso de validación, refutación o reformulación de los distintos planteamientos del marco teórico. Para continuar este proceso se ha diseñado y pilotado dos actividades de comprensión audiovisual con preguntas subtítuladas para estudiantes de ELE. Nuestra investigación actual se centra en conocer la opinión de estudiantes y profesores de ELE sobre este método, y en averiguar si los profesores estarían dispuestos a aprender a subtítular vídeos y a probar esta técnica. Asimismo, estamos estudiando si hay diferencias estadísticamente significativas entre las puntuaciones de estudiantes de ELE que realizan una prueba de comprensión con preguntas subtítuladas y las puntuaciones de estudiantes de ELE que efectúan el mismo test con las preguntas disponibles solo en papel. También hemos empezado a estudiar la conducta visual de alumnos de ELE con preguntas subtítuladas y con preguntas disponibles solo en papel en la fase de visionado.

### *Agradecimientos*

Quiero agradecer a los expertos su inestimable contribución al desarrollo del marco teórico. También me gustaría dar las gracias a Diana Silver y Robert Grant por su revisión del resumen en inglés. Finalmente, deseo expresar mi gratitud al Banco Santander por la adjudicación de una beca de investigación en España y a la Universidad de Coimbra por la concesión de una ayuda de movilidad Erasmus en Alemania porque buena parte de los datos se recogieron durante estas estancias.

## REFERENCIAS BIBLIOGRÁFICAS

Alderson, J.C., C. Clapham y D. Wall (1995). *Language test construction and evaluation*. Cambridge: CUP.

Allan, M. (1984). "Viewing comprehension in the EFL classroom", en *ELT Journal*, vol. 38, núm. 1, págs. 21-26.

Álvarez-Gayou Jurgenson, J.L. (2003). *Cómo hacer una investigación cualitativa: fundamentos y metodología*. Buenos Aires: Paidós.

Arnold, J. "Seeing Through Listening Comprehension Exam Anxiety", en *Tesol Quarterly*, vol. 34, núm. 4, págs. 777-786.

Baddeley, A.D. (2000). "The episodic buffer: a new component of working memory?", en *Trends in Cognitive Science*, vol. 4, núm. 11, págs. 417-423.

Bright, K y G. Cook (2010). *The Colour, Light and Contrast Manual: Designing and Managing Inclusive Built Environments*. Oxford: Wiley-Blackwell.

Brown, G. y G. Yule (1983). *Teaching the Spoken Language: An approach based on the analysis of conversational English*. Cambridge: CUP.

Buck, G. (2001). *Assessing Listening*. Cambridge: CUP.

Casañ Núñez, J.C. (2012). "Didáctica de las grabaciones audiovisuales para desarrollar la comprensión y el aprendizaje en el aula de L2/LE", en *marcoELE*, núm. 15: <http://marcoele.com/didactica-de-las-grabaciones-audiovisuales/> (consultada el 7 de diciembre de 2014)

Casañ Núñez, J.C. (2009a). *Didáctica de las grabaciones audiovisuales para desarrollar la comprensión oral en el aula de lenguas extranjeras*, en *marcoELE*, núm. 9 (tesina de máster): <http://marcoele.com/suplementos/didactica-de-las-grabaciones-audiovisuales> (consultada el 7 de noviembre de 2014)

Casañ Núñez, J.C. (2009b). "Descripciones", en *Azulejo para el español*, núm. 2, págs. 15-18: <https://sede.educacion.gob.es/publiventa/detalle.action?cod=14146> (consultada el 1 de octubre de 2014)

Casañ Núñez, J.C. (2007). «Dos propuestas de explotación de cortometrajes: "Blanco o negro" y "El sueño de la maestra"», en M. J. Fernández Colomer y M. Albelda Marco, *Actas del III Foro de Profesores de E/LE*. Valencia: Universidad de Valencia, págs. 17-23: [http://www.uv.es/foroel/foro3/Actas\\_III\\_Foro\\_ELE.pdf](http://www.uv.es/foroel/foro3/Actas_III_Foro_ELE.pdf) (consultada el 1 de octubre de 2014)

Chion, M. (1993). *La audiovisión: introducción a un análisis conjunto de la imagen y el sonido*. Trad. de Antonio López Ruiz. Barcelona: Paidós.

Cohen, L, L. Manion y K. Morrison (2011). *Research Methods in Education* (7ª ed.). Nueva York: Routledge.

Coniam, D. (2001). "The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: a case study", en *System*, núm. 29, págs. 1-14.

Consejo de Europa (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Trad. del Instituto Cervantes. Madrid: Anaya.

Cowan, N. (2001). "The magical number 4 in short-term memory: A reconsideration of mental storage capacity", en *Behavioral and Brain Sciences*, vol. 24, núm. 1, págs. 87-114.

Cowan, N. y C. C. Morey (2005). "When Do Visual and Verbal Memories Conflict? The Importance of Working-Memory Load and Retrieval", en *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 31, núm. 4, págs. 703-713.

Creswell, J.W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4ª ed.). Londres: Sage Publications

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Londres: Pearson

Creswell, J.W. y V.L. Plano Clark (2011). *Designing and Conducting Mixed Methods Research* (3ª ed.). Londres: Sage Publications

Díaz Cintas, J. (2003). *Teoría y práctica de la subtitulación inglés-español*. Barcelona: Ariel cine.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: OUP.

Eguaras, M. (2014). *Qué tipografía usar en libros impresos y digitales*. Mariana Eguaras Consultoría Editorial (Edición Kindle).

Elkhafaifi, H. (2005). "Listening Comprehension and Anxiety in the Arabic Language Classroom", en *The Modern Language Journal*, vol. 89, núm. 2, págs. 206-220.

Escobar-Pérez, J. y A. Cuervo-Martínez (2008). "Validez del contenido y juicio de expertos: una aproximación a su utilización", en *Avances en medición*, núm. 6, págs. 27-36.

Fernández Núñez, L. (2006). "¿Cómo analizar datos cualitativos?", en *Butlletí LaRecerca*, ficha 7. Barcelona: Universidad de Barcelona, págs. 1-13

Field, J. (2008) *Listening in the language classroom*. Cambridge: CUP.

Fletcher, J.D. y S. Tobias (2006). "The Multimedia Principle", en R. Mayer (ed.) *The Cambridge handbook of multimedia learning*. Cambridge: CUP, págs. 117-133.

Flowerdew, J. y L. Miller. (2005). *Second language listening*. NY: CUP.

Fougnie, D. y R. Marois (2011). "What Limits Working Memory Capacity? Evidence for Modality-Specific Sources to the Simultaneous Storage of Visual and Auditory Arrays", en *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 37, núm. 6, págs. 1329-1341.

Giovannini, A., E. Martín Peris, M. Rodríguez y T. Simón (1996). *Profesor en acción 3. Destrezas*. Madrid: Edelsa.

Glenn, E. (1989) "A content analysis of fifty definitions of listening", en *Journal of the International Listening Association*, vol. 3, págs. 21-31.

Greist, J. H. y J. W. Jefferson (2000). "Anxiety disorders", en H. H. Goldam (ed.). *Review of General Psychiatry* (5ª ed.). Nueva York: McGraw-Hill Medical, págs. 284-300.

Gruba, P. (1997) "The role of media in listening assessment", en *System*, vol. 25, núm. 3, págs. 335-345.

Gruba, P. (1993). "A Comparison Study of Audio and Video in Language Testing", en *JALT Journal*, vol. 15, núm. 1, págs. 85-88.

Harris, T. (2003). "Listening with Your Eyes: The Importance of Speech-Related Gestures in the Language Classroom", en *Foreign Language Annals*, vol. 36, núm. 2, págs. 180-187.

Lonergan, J. (1984). *Video in Language Teaching*. Cambridge: CUP.

Lynch, T. (2012). "Traditional and Modern Skills. Introduction", en M. Eisenmann y T. Summer (eds.), *Basic issues in EFL teaching and learning*, Heidelberg: Winter, págs. 69-81.

Lynch, T. (2010). "Listening: sources, skills, and strategies", en R. B. Kaplan (ed.), *The Oxford Handbook of Applied Linguistics* (2ª ed.). Oxford: OUP, págs. 74-87.

Lynch, T. (2009). *Teaching Second Language Listening*. Oxford: OUP.

Martínez Castillo, G. (2009). *Presentaciones Digitales y Comunicación Interactiva*. UNISON: México:

<http://clasificaciontipografica.blogspot.pt/>  
(consultada el 20 de enero de 2015)

Martín Leralta, S. (2012). "Escuchar para entender y aprender: un reto didáctico para el profesor de ELE", en *Actas del I encuentro internacional de profesores de ELE: El aprendizaje y la enseñanza de español en contextos multiculturales y multilingües*. Madrid: Instituto Cervantes, págs. 7-17.

Martín Peris, E. (1991). "La didáctica de la comprensión auditiva" en *Cable*, núm. 8. Madrid: Difusión. Rpt. en (2007) *marcoELE*, núm. 5: <http://marcoele.com/la-didactica-de-la-comprension-auditiva/>  
(consultada el 20 de noviembre de 2014)

Mayer, R. (2006a). "Introduction to multimedia learning", en R. Mayer (ed.) *The Cambridge handbook of multimedia learning*. Cambridge: CUP, págs. 1-16.

Mayer, R. (2006b). "Principles for managing essential processing in multimedia learning", en R. Mayer (ed.) *The Cambridge handbook of multimedia learning*. Cambridge: CUP, págs. 169-182.

Mayer, R. (2001). *Multimedia learning*. Cambridge: CUP.

McGurk, H. y MacDonald, J. (1976). "Hearing lips and seeing voices", en *Nature*, vol. 264, págs. 746-748.

Mendelshon, D. (1994). *Learning to Listen: a Strategy Based Approach for the Second Language Learner*. Carlsbad: Dominic Press.

Miller, G. A. (1956). "The magical number seven, plus or minus two: some limits on our capacity for processing information", en *The Psychological Review*, vol. 63, págs. 81-97.

Moliner, M. (2007). *Diccionario de uso del español* (3ª ed.). Madrid: Gredós.

Ockey, G.J. (2007). "Construct implications of including still image or video in computer-based listening tests", en *Language Testing*, vol. 24, núm. 4, págs. 517-537.

Murillo Torrecilla, F.J. (2004). *Cuestionarios y escalas de actitudes*. Madrid: Universidad Autónoma de Madrid:

[http://www.uam.es/personal\\_pdi/stmaria/jmurillo/Metodos/Materiales/Apuntes%20Cuestionario.pdf](http://www.uam.es/personal_pdi/stmaria/jmurillo/Metodos/Materiales/Apuntes%20Cuestionario.pdf)

(Consultada el 15 de enero de 2013)

Plass, J.L. y L.C. Jones (2006). "Multimedia learning in second language acquisition", en R. Mayer (ed.) *The Cambridge handbook of multimedia learning*. Cambridge: CUP, págs. 467-488.

Real Academia Española (2014). *Diccionario de la lengua española* (23ª ed.). Madrid: Espasa.

Richards, J.C. y A. Burns (2012). *Tips for Teaching Listening*. Nueva York: Pearson Education.

Riley, P. (1979). «Viewing comprehension: "L'oeil écoute"», en *Mélanges CRAPEL*, núm. 10, págs. 80-95.

Robson, C. (1993). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford: Blackwell Publishing.

Rodríguez Bravo, A. (1998). *La dimensión sonora del lenguaje audiovisual*. Barcelona: Paidós.

Rost, M. y J.J. Wilson (2013). *Active Listening*. Harlow: Pearson Education.

Rost, M. (2002). *Teaching and researching listening*. Harlow: Pearson Education.

Rubin, J. (1995a). "An Overview to *A Guide for the Teaching of Second Language Listening*", en D. J. Mendelsohn y J. Rubin (eds.) *A Guide for the Teaching of Second Language Listening*. San Diego: Dominie Press, págs. 7-11.

Rubin, J. (1995b). "The Contribution of Video to the Development of Competence in Listening", en D. J. Mendelsohn y J. Rubin (eds.) *A Guide for the Teaching of Second Language Listening*. San Diego: Dominie Press, págs. 151-165.

Sherman, J. (1997). "The effect of question preview in listening comprehension tests", en *Language Testing*, vol. 14, núm. 2, págs. 185-213.

Stempleski, S. y B. Tomalin. (2001). *Film*. Oxford: OUP.

Stoller, F.L. (1992). "Using video in theme-based curricula", en S. Stempleski y P. Arcario (eds.) *Video in second language teaching: using, selecting and producing video for the classroom*. NY: TESOL, págs. 25-46.

Sueyoshi y Hardison (2005). "The Role of Gestures and Facial Cues in Second Language Listening Comprehension", en *Language Learning*, vol. 55, núm. 4, págs. 661-669.

Suvorov, R. (2008). *Context visuals in L2 listening tests: The effectiveness of photographs and video vs. audio-only format*. Ames, Iowa: Iowa State University. (Tesina de máster)

Sweller, J.,P. Ayres y S. Kalyuga (2011). *Cognitive Load Theory*. Londres: Springer

Tomalin, B. (1986). *Video, TV and radio in the English class*. London: Macmillan Publishers Ltd.

Underwood, M. (1989). *Teaching Listening*. Londres: Longman.

Ur, P. (1999). *A Course in Language Teaching: practice and theory*. Cambridge: CUP.

Ur, P. (1984/1994). *Teaching listening comprehension* (12<sup>a</sup> reimp.). Cambridge: CUP.

Vandergrift, L y C. C. M. Goh (2012). *Teaching and Learning Second Language Listening*. Nueva York: Routledge.

Wagner, E. (2007). "Are they watching? Test-taker viewing behavior during an L2 video listening test", en *Language Learning and Technology*, vol. 11, núm. 1, págs. 67-86:

<http://lt.msu.edu/vol11num1/pdf/wagner.pdf>

(consultada el 7 de diciembre de 2014)

Wagner, E. (2010). "Test-takers' interaction with an L2 video listening test", en *System*, vol. 38, núm. 2, págs. 280-291.

Whittaker, S. (1995). "Rethinking video as a technology for interpersonal communications: theory and design implications", en

*International Journal of Human-Computer Studies*, vol. 42, núm. 5, págs. 501-529.

## REFERENCIAS DE SEGUNDA MANO

Gable, R.K. y J.W. Wolf (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings*. Boston: Kluwer Academic, citado en D. McGartland Rubio, M. Berg-Weger, S. Tebb, E. S. Lee, y S. Rauch (2003). "Objectifying content validity: Conducting a content validity study in social work research", en *Social Work Research*, vol. 27, núm. 2, págs. 94-104.

## ANEXO: EXTENSIVE SUMMARY

### *A framework for using audiovisual comprehension questions embedded in the video as subtitles: a mixed methods study*

#### 1. INTRODUCTION

Listening, reading and writing simultaneously is extremely difficult, especially in a foreign language (Underwood, 1989). Listening, viewing, reading and writing can be even more difficult. On the one hand, there is a conflict of visual attention between viewing the image and completing a written activity simultaneously. This conflict helps to explain the degree of (in)attention to the image in studies that researched test-takers' viewing rates during FL video listening tests. Ockey (2007) does not provide the average viewing rate but, from the data it can be calculated that it was 44,9%, Wagner (2007) finds that test-takers make eye contact with the monitor 69% and Wagner (2010) 47,9%. If listening involves the interpretation of auditory and visual information, as suggested by specialists such as Riley (1979), Martin Peris (1991), Rubin (1995a) and Lynch (2012), these degrees of inattention may have a negative effect both on understanding the particular video and on the development of audiovisual comprehension in general. On the other hand, working memory capacity is severely limited (Sweller, Ayres & Kalyuga, 2011) and this should be taken into account when designing while-viewing activities. As Vandergrift and Goh (2012) point out, attention to the image, the audio, and the task simultaneously may cause working memory overload.

The present study forms part of wider research, exploring the use of audiovisual comprehension questions embedded in the image as subtitles and synchronized with the relevant fragments to provide answers (see figure 2). This is apparently an innovative technique, since no references have been found to this procedure in the literature of teaching listening comprehension or of video in the FL classroom, apart from the author's own remarks (Casañ, 2007, 2012). We believe that this teaching technique could provide various benefits in the while-viewing phase over viewings where the activity is available on a sheet. Such benefits could include reducing the conflict of visual attention between viewing the video and doing the task, by spatially approximating the questions and the relevant fragments, and reducing the cognitive requirements of the activity, eliminating the need to pay attention to various questions at the same time.



Fig. 2. Example of a subtitled question. "¿Cómo se llama la chica?" [What is the name of the girl?] From the Spanish film *Ópera prima* by Fernando Trueba.

## 2. STUDY DESIGN

Drawing on the review of the literature and the author's own teaching experience with the procedure a first theoretical framework was produced, describing the teaching method and its potential benefits and limitations. The aim of this mixed methods study was to find out the opinion of a group of experts in teaching Spanish as a foreign language (SFL), Spanish linguistics and/or the use of technology on the provisional framework. As the experts who participated in the study lived in countries different from that of the researcher, the instrument had to be able to gather the data into a single face-to-face meeting, or if necessary to collect the information at a distance. With this in mind a questionnaire was designed with quantitative items (five-point Likert scales) that measured the degree of agreement with various aspects of the theoretical framework, and with qualitative items that elicited comments on the quantitative items and other points. The reason for collecting both quantitative and qualitative data was to merge the two forms of data to gain greater insight into the experts' opinion than would be obtained by either type of data individually. A convergent mixed methods design was used (Creswell & Plano, 2011; Creswell, 2012, 2013): a type of design in which quantitative and qualitative data are collected at the same time, analysed separately, and then merged. This study sought to answer two questions:

- a) To what extent do the experts consulted agree with the theoretical framework?
- b) What aspects of the theoretical framework should be modified, according to the experts consulted?

## 3. A FRAMEWORK FOR USING AUDIOVISUAL COMPREHENSION QUESTIONS EMBEDDED IN THE VIDEO AS SUBTITLES

Quantitative and qualitative results converge in a high degree, suggesting that, overall, the experts agree with most of the framework; results also reveal areas for potential improvement. Consideration of the experts' evaluations and comments led to the reformulation of the initial proposal. The framework is as follows:

### I. Introduction

This framework describes the use of audiovisual comprehension questions embedded as subtitles in the viewing phase in the FL classroom. This methodological proposal draws on the review of the literature, the author's own teaching experience with the procedure, and the opinions of six experts in teaching SFL, Spanish linguistics and/or the use of technology, from five different institutions. It is important to note that apart from the author's own remarks (Casañ, 2007, 2012), no reference to this technique has been found in the literature of listening comprehension in FL classroom or of video in the FL classroom.

### II. Description of the procedure

Audiovisual comprehension questions appear on screen a few seconds before the beginning of the part of the video to which they are related, they remain visible for the duration of the fragment and they disappear when the snatch finishes. This timing allows students to keep in mind the comprehension objectives at all times. If



questions were visible well before and/or after, in addition to during, the relevant part of the video, they would decrease their effect of facilitating understanding through focusing. If questions were not displayed during the relevant fragments, they could lead students to seek the answer in the wrong place. We believe that the timing proposed is the most appropriate to facilitate concentration and to guide students through the text, and these are, according to Underwood (1989), the main aims of while-listening/viewing comprehension activities. To avoid any confusion, it is necessary to inform students of the criteria followed in the placement of subtitled questions.

Certain aspects of the configuration of the questions are based on established conventions in the use of subtitles in Spanish (according to Díaz, 2003):

a) Subtitles are often composed of a maximum of two lines in order to limit their impact on the image. In intralingual subtitling for the deaf, subtitles sometimes have three or even four lines. Subtitled questions will preferably be one or two lines long.

b) Subtitles appear at the bottom of the screen, except when there is essential visual information in that area. To prevent subtitles from damaging the aesthetic image of the audiovisual text, some distributors sandwich the picture between two black stripes, one above and one below. Subtitles are projected in the lower band. This strategy is recommended so that the subtitled question does not encroach on the photography.

c) Generally subtitles contain twenty-eight to forty characters and spaces. A higher number of characters usually involve a reduction of font size and a loss of legibility.

d) The font must be clearly legible: Helvetica, Arial or Times New Roman are typically employed. Sans-serif fonts are recommended for subtitled questions for three main reasons. Firstly they are usually used for short texts; secondly, according to Martinez (2009), they transmit happiness and safety, beneficial for students in a state of tension; and thirdly they are easier to read on a screen because they adjust better to the screen resolution (Eguaras, 2014).

e) The colour must be easily readable: white and yellow are the most commonly used. According to Bright and Cook (2010), yellow is an unstable colour, while white is neutral and can be combined with any colour. It is for these reasons that white on a black stripe has been suggested. If the subtitled question invades the image, it is necessary to re-configure it according to the colours of the photography.

f) On the basis of studies on the reading speed of the average viewer, it is recommended that the subtitle of a line, usually seven or eight words long, stays on screen for about four seconds, and the two-line subtitle, which is usually fourteen to sixteen words long, stays about six seconds. In the case of captions longer than two lines, a temporal equivalence is calculated from the "rule of six seconds" for two-line subtitles. Although a student will need more or less time to read the subtitle in a given FL (or L1 if the mother tongue of the students is used) depending on his/her reading proficiency in that language, this convention can provide an orientation for determining the time needed to read each question. That said, in order to determine how far in advance the question should appear before the fragment, two additional elements should be considered. Firstly, it is advisable that the question becomes visible at a time propitious for reading, that is, at a time when nothing crucial is happening, so as not to generate an intense conflict between the desire to watch the video and the need to read

the question. Secondly, students do not always keep the focus of attention on the image. Some authors have researched test-takers' viewing rates during FL video listening tests. Ockey (2007) does not provide the average viewing rate, but from the data it can be calculated that it was 44.9%; Wagner (2007) finds that test-takers make eye contact with the monitor 69%, and Wagner (2010) 47.9%. This fact may reduce the time available for reading the subtitle and it may even cause the question to pass unnoticed. To avoid or reduce this possibility, it is essential to explain the existence and timing of subtitled questions. Also, there should be enough time to answer each question before the following one starts, either because the questions are placed in sufficiently separated fragments, or because there are short pauses to answer the questions, or both. Lastly, activities carried out while the video is playing should involve little reading and writing.

g) Subtitles should be arranged in the most logical way in order to facilitate reading.

#### SUBTITLED QUESTIONS AND QUESTIONS ON A SHEET OF PAPER

Sherman (1997) suggests that listening comprehension questions have little significance alone, and that they only make sense when listening to the text. Nevertheless, question preview involves some benefits. On the one hand, it respects the preferences of learners. According to Sherman (1997) and Buck (2001) students value positively reading the questions before the listening. On the other hand, possible doubts about the questions can be clarified before playing the video. It also allows the length or form of the activity to be extended. That said, it should be kept in mind that activities carried out in the while-viewing phase should be simple, due to the great difficulty of viewing, listening, reading and writing at the same time. Moreover, if questions are presented both on screen and on

paper, wordings should be the same, because small differences can cause great confusions.

Although the simultaneous availability of questions on paper and on screen in the viewing phase is possible, we do not know to what extent the coexistence of two planes of work might maintain the conflict of visual attention that we are trying to reduce by subtitled the questions.

#### III. Potential benefits

The procedure is considered especially beneficial for low level students. Also, a priori it is regarded as a source of "extra help" and as such its use should be abandoned gradually. In spite of this, occasional use should not be totally rejected at other levels, for two reasons. Firstly, it helps to maintain the focus of work on screen in comparison with viewing with questions on paper. Secondly, the author's teaching experience at B1 and B2 levels has been intuitively positive.

In comparison with a viewing where the video is shown on a communal screen and questions are available on paper, displaying subtitled questions seconds before and during the passage they refer to can provide the following benefits:

a) Watching a video and simultaneously reading/answering audiovisual questions on a sheet of paper creates a conflict of visual attention. This conflict helps to explain the degree of (in)attention to the image in studies that researched test-takers' viewing rates during FL video listening tests. As mentioned above, in Ockey's (2007) and Wagner's (2007, 2010) investigations, students orient towards the video monitor 44.9%, 69% and 47.9% of the time

respectively (or, from another point of view, students do not pay attention to the screen 55.1%, 31% and 52.1 % of the time). If listening implies the interpretation of auditory and visual information (Riley, 1979; Martín Peris, 1991; Rubin, 1995a; Lynch, 2012), a high degree of inattention may have a negative effect both on understanding the video and on the development of audiovisual comprehension. Subtitled questions do not eradicate the problem, but they reduce it by spatially approximating the questions and the relevant scenes. Thanks to that, students may pay more attention to the video as compared to viewings with questions on paper.

b) It helps students to orient their attention towards the listening / viewing objectives at all times. Thus, this is in line with one of the main objectives, according to Underwood (1989) and Richards and Burns (2012), of while-listening / viewing activities: to guide students through the text.

c) If allowed to pay attention to a subtitled question each time, instead of several on- paper questions at the same time, the student may have more working memory available for understanding the audiovisual text. Working memory is defined as “a limited capacity system allowing the temporary storage and manipulation of information necessary for such complex tasks as comprehension, learning and reasoning” (Baddeley, 2000: 418). For Miller (1956) the limit is seven items, plus or minus two. Cowan (2001) reduces the number to three to five items. As noted by Sweller et al. (2011), from a pedagogical perspective, what matters is not the exact figure but that working memory capacity is extremely reduced. Vandergrift and Goh (2012) remark that attention to the image, the audio, and the task simultaneously can lead to working memory saturation. Unskilled listeners may reach this overload faster because, according to Field (2008), they have poorly automatized decoding processes, thus they have to spend a great deal of working memory to

decoding. Therefore, this teaching technique could be especially beneficial for low-level students<sup>4</sup>.

d) In the event that a student loses concentration, he / she is helped to continue the activity when the following question appears on screen. For this possibility to be available, the next question should not be related to the previous one. Also, it should be taken into account that data loss can cause anxiety and that this, according Greist and Jefferson (2000), can last from a very few seconds to an extended period of time.

e) The potential benefits just described could contribute to reduce the stress generated by the activity compared with viewings with questions available on paper, and if so, they would facilitate comprehension. Arnold (2000) cites several studies that suggest that “students in a foreign language classroom are often anxious and that this feeling inhibits their performance in the foreign language” (777). Elkhafaifi (2005) investigates the effect of general foreign language learning anxiety and listening anxiety in the performance of students of Arabic as foreign language. Based on the results, the author suggests that general learning anxiety and listening anxiety are separate but related phenomena, and that they are correlated negatively with achievement. It should be acknowledged that the mere existence of questions — whether on paper, on screen or given orally — is a source of stress, due to the need to respond. In this regard, Richards and Burns (2012: 101) remind us that “it’s

---

<sup>4</sup> It is still unclear whether there are separate working memory subsystems for processing auditory and visual stimuli, or there is a central form of storage, or “working memory capacity is set both by task/modality and central, amodal processing limitations” (Fougnie & Marois, 2011: 1338). If there was a central form of storage, there could be interference between auditory and visual processing (Cowan & Morey, 2005).

important to create activities where students feel they can succeed, especially for beginners”.

#### IV. Potential limitations and unanswered questions

a) This methodology may not adequately prepare students for viewings unless indications are given on where to look for the required information.

b) Focusing on fragments may have two side effects. Firstly, it can limit the practice of global comprehension. In order to practice this, we are experimenting by placing a global understanding question in the last seconds of short videos. Although it is true that the focalising character is lost, it allows maintaining the focus of attention on screen (against viewings with questions on paper). Secondly, it may have the opposite effect: it may condition the development of the ability to search for specific information due to the short duration of the fragment where the student seeks the answer. In any way, it is possible to play the video various times with different purposes, as suggested by Richards and Burns (2012). For example, a first viewing without questions might focus on global understanding and a second viewing with subtitled questions may focus on more specific information.

c) The length of subtitled questions is much more limited than those that may be formulated on paper. As it was already mentioned, it is possible to display the wording of questions on screen and hand out the complete activities on paper. However, a basic principle of written activities during the emission of the input, especially if it is audiovisual, is that they should involve little reading and writing.

d) Although the author’s teaching experience with the procedure has been intuitively positive, the opinion of students regarding this teaching method has not yet been studied. It is also unknown whether foreign language teachers would be willing to learn to subtitle videos and to try this technique.

#### 4. CONCLUSION

In this convergent mixed methods study (Creswell & Plano, 2011; Creswell, 2012, 2013), six experts in teaching SFL, Spanish linguistics and/or the use of technology completed a questionnaire designed to elicit both quantitative and qualitative data. The purpose of the survey form was to find out what experts thought about a framework describing the use of questions embedded in the video as subtitles in the while-viewing phase in the FL classroom. We believe that this teaching technique could provide various benefits over viewings where the activity is available on a sheet. Such benefits include reducing the conflict of visual attention between viewing the video and doing the task, by spatially approximating the questions and the relevant fragments, and reducing the cognitive requirements of the activity, eliminating the need to pay attention to various questions at the same time. Quantitative and qualitative results suggest that, overall, the experts agree with most of the framework; results also reveal areas for potential improvement. The consideration of the comments led to the reformulation of the initial proposal.

This study was regarded as the first action in the process of validating, refuting or reformulating the different aspects of the theoretical framework. To continue this task two audiovisual comprehension activities for students of SFL with subtitled questions

have been designed and piloted. Our current research focuses on finding out what SFL students and teachers think about this methodology, and if teachers would be willing to learn to subtitle videos and to try this technique. Moreover, we are investigating if there is a statistically significant difference between the scores of SFL students taking an audiovisual comprehension test with subtitled questions and the scores of SFL students completing the same test with questions only available on a sheet of paper. In addition to that, we have started to study students' orienting behaviour with subtitled questions and with questions only available on paper in the while viewing phase.

FECHA DE ENVÍO: 21 DE FEBRERO DE 2015