

Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular

An approach to reading subtitles, perception and comprehension of audiovisual pieces using an eye tracking technology

Uma abordagem à leitura de legendas, percepção e compreensão de peças audiovisuais por meio da tecnologia de rastreamento ocular



Anajosé Pardo Díaz
Haide Jineth Briceño Suarez
Judy Costanza Beltrán Rojas
Angelica Mateus Moreno
María Fernanda Lara Díaz



[Pixorena](#)

Photo By/Foto:

Rip
15²

Volumen 15 #2 may-ago
15 Años

Revista Iberoamericana de
Psicología

ISSN-I: 2027-1786 | e-ISSN: 2500-6517
Publicación Cuatrimestral

ID: [10.33881/2027-1786.rip.15202](https://doi.org/10.33881/2027-1786.rip.15202)

Title: An approach to reading subtitles, perception and comprehension of audiovisual pieces using an eye tracking technology

Título: Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular

Titulo: Uma abordagem à leitura de legendas, percepção e compreensão de peças audiovisuais por meio da tecnologia de rastreamento ocular

Alt Title / Título alternativo:

[en]: An approach to reading subtitles, perception and comprehension of audiovisual pieces using an eye tracking technology

[es]: Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular

[pt]: Uma abordagem à leitura de legendas, percepção e compreensão de peças audiovisuais por meio da tecnologia de rastreamento ocular

Author (s) / Autor (es):

Pardo Díaz, Briceño Suarez, Beltrán Rojas, Mateus Moreno & Lara Díaz

Keywords / Palabras Clave:

[en]: Eye-tracking, Reading, Comprehension, Perception, Subtitles, Audiovisual materials

[es]: Seguimiento ocular, Lectura, Comprensión de lectura, Percepción, Subtítulos, Material audiovisual

[pt]: Rastreamento ocular, Leitura, Compreensão, Percepção, Legendas, Materiais audiovisuais

Submitted: 2022-01-16

Accepted: 2022-04-21

Ana José **Pardo Díaz**, [Fon]

ORCID: [0000-0003-1296-0781](https://orcid.org/0000-0003-1296-0781)

Source / Filiación:

Universidad Nacional de Colombia

BIO:

Estudiante fonoaudiología vinculada a Semillero de Investigación

City / Ciudad:

Bogotá DC [co]

e-mail:

apardod@unal.edu.co

Haide Jineth **Briceño Suarez**, [Fon]

ORCID: [0000-0002-7557-6368](https://orcid.org/0000-0002-7557-6368)

Source / Filiación:

Universidad Nacional de Colombia

BIO:

Estudiante fonoaudiología vinculada a Semillero de Investigación

City / Ciudad:

Bogotá DC [co]

e-mail:

hbriceno@unal.edu.co

Resumen

La pandemia por SARS-CoV-2 y las medidas implementadas para la disminución de contagios generó aumento del consumo digital a nivel mundial en distintos idiomas, especialmente en experiencias virtuales de aprendizaje y entretenimiento, como las plataformas de streaming donde hay alto contenido de videos con subtítulos que son vistos por personas de diferentes nacionalidades, por lo que resulta fundamental analizar cómo las personas perciben la información en este tipo de plataformas y el uso y comprensión dado a los subtítulos presentes en este tipo de videos. Esta investigación utilizó un enfoque mixto, cuyo objetivo fue analizar la percepción y comprensión lectora de subtítulos en piezas audiovisuales, durante la proyección de un documental con ayuda del software de seguimiento ocular Sticky y un cuestionario cualitativo. Participaron 64 personas de 18 a 59 años, quienes diligenciaron el consentimiento informado y después de ver los videos, respondieron un cuestionario de validación. Se realizó análisis estadístico de los parámetros de seguimiento ocular e información contenida en el cuestionario. Se concluye que los subtítulos en español como lengua materna (L1) son un apoyo para las personas competentes y no competentes en inglés (L2). En los videos subtítulos en español con audio en inglés (subtítulos interlingüísticos estándar) y español (subtítulos intralingüísticos en L1) se observaron primero a los personajes (con un número mayor de fijaciones) y luego los subtítulos, mientras que en el video doblado y sin subtítulaje sucedió lo contrario; no obstante, la mayoría de los participantes consideraron que se enfocaron más en el texto. Lo anterior tiene implicaciones para el análisis psicolingüístico y el diseño de contenidos.

Abstract

The SARS-CoV-2 pandemic and the measures implemented to reduce infections generated an increase in digital consumption worldwide in different languages, especially in virtual learning and entertainment experiences, such as streaming platforms where there is a high content of videos with subtitles that are seen by people of different nationalities, so it is essential to analyze how people perceive information on these types of platforms and the use and understanding given to the subtitles present in these types of videos. This research used a mixed approach, whose objective was to analyze the perception and reading comprehension of subtitles in audiovisual pieces, during the projection of a documentary with the help of the Sticky eye-tracking software and a qualitative questionnaire. 64 people between 18 and 59 years of age participated, who filled out the informed consent and after watching the videos, answered a validation questionnaire. Statistical analysis of the eye tracking parameters and information contained in the questionnaire was performed. It is concluded that the subtitles in Spanish (L1) are a support for competent and non-English competent people (L2). In the videos subtitled in Spanish with audio in English (standard interlinguistic subtitles) and Spanish (intralinguistic subtitles in L1), the characters were observed first (with a greater number of fixations) and then the subtitles, while in the dubbed video without subtitling the opposite happened; however, most of the participants felt that they focused more on the text. This has implications for psycholinguistic analysis and content design.

Resumo

A pandemia de SARS-CoV-2 e as medidas implementadas para reduzir infecções geraram um aumento no consumo digital em todo o mundo em diferentes idiomas, especialmente em experiências virtuais de aprendizado e entretenimento, como plataformas de streaming onde há alto conteúdo de vídeos com legendas que são vistos por pessoas de diferentes nacionalidades, por isso é fundamental analisar como as pessoas percebem a informação neste tipo de plataforma e o uso e entendimento dado às legendas presentes neste tipo de vídeo. Esta pesquisa utilizou uma abordagem mista, cujo objetivo foi analisar a percepção e compreensão leitora de legendas em peças audiovisuais, durante a projeção de um documentário com o auxílio do software Sticky eye tracking e um questionário qualitativo. Participaram 64 pessoas de 18 a 59 anos, que preencheram o consentimento informado e após assistirem aos vídeos, responderam a um questionário de validação. Foi realizada análise estatística dos parâmetros de eye tracking e informações contidas no questionário. Conclui-se que as legendas em espanhol como língua materna (L1) são um suporte para pessoas competentes e não competentes em inglês (L2). Nos vídeos legendados em espanhol com áudio em inglês (legendas interlínguas padrão) e espanhol (legendas intralínguas em L1) os caracteres foram observados primeiro (com maior número de fixações) e depois as legendas, enquanto no vídeo dublado e sem legenda aconteceu o contrário; no entanto, a maioria dos participantes considerou que se concentrava mais no texto. Isso tem implicações para a análise psicolinguística e design de conteúdo.

Citar como:

Pardo Díaz, A., Briceño Suarez, H. J., Beltrán Rojas, J. C., Mateus Moreno, A., & Lara Díaz, M. F. (2022).

Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular. *Revista Iberoamericana de Psicología*, 15 (2), 13-23.

Obtenido de: <https://reviberopsicologia.iber.edu.co/article/view/2379>

Judy Costanza **Beltrán Rojas**, [Dr]Fon

ORCID: [0000-0002-2896-8852](https://orcid.org/0000-0002-2896-8852)

Source / Filiación:

Universidad Nacional de Colombia

BIO:

Fonoaudióloga, Psicóloga, Candidata a Doctora en Psicología

City / Ciudad:

Bogotá DC [co]

e-mail:

jcbeltranr@unal.edu.co

Angelica **Mateus Moreno**, MSc Fon

ORCID: [0000-0002-8191-4474](https://orcid.org/0000-0002-8191-4474)

Source / Filiación:

Universidad Nacional de Colombia

BIO:

Fonoaudióloga, Magister en Neurociencias

City / Ciudad:

Bogotá DC [co]

e-mail:

amateusm@unal.edu.co

María Fernanda **Lara Díaz**, PhD MSc Fon

ORCID: [0000-0002-7997-5059](https://orcid.org/0000-0002-7997-5059)

Source / Filiación:

Universidad Nacional de Colombia

BIO:

Fonoaudióloga, Magister en Perturbaciones del lenguaje, Doctora en ciencia cognitiva y lenguaje, Docente Universitaria

City / Ciudad:

Bogotá DC [co]

e-mail:

mflarad@unal.edu.co

Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular

An approach to reading subtitles, perception and comprehension of audiovisual pieces using an eye tracking technology

Uma abordagem à leitura de legendas, percepção e compreensão de peças audiovisuais por meio da tecnologia de rastreamento ocular

Anajosé **Pardo Díaz**
Haide Jineth **Briceño Suarez**
Judy Costanza **Beltrán Rojas**

Angelica **Mateus Moreno**
María Fernanda **Lara Díaz**

Introducción

La pandemia por SARS-CoV-2 y las medidas implementadas para la disminución de contagios como las restricciones de movilidad y el aislamiento obligatorio, generaron un aumento del consumo digital, lo cual permitió un mayor desarrollo de entornos y experiencias virtuales en las compras, el entretenimiento y el aprendizaje en el contexto colombiano. A inicios de la emergencia sanitaria en el año 2020 el consumo de televisión en los jóvenes y adultos incrementó aproximadamente un **10%**. Adicionalmente, la plataforma de streaming ‘Netflix’ obtuvo 37 millones de nuevos suscriptores durante el año 2020, mientras que ‘Disney +’ alcanzó un total de 85 millones de usuarios (**BBC News Mundo, 2020; El Tiempo, 2021**).

En consecuencia, el mundo se enfrentó a la consolidación de una sociedad en la que se hace más notorio el intercambio constante de información audiovisual de carácter mundial (**Lavour & Bairstow, 2011; Andreu et al., 2021**), con un aumento en el interés en el contenido de distintos idiomas (**Lavour & Bairstow, 2011**). Es así como los subtítulos juegan un papel fundamental en la comunicación al permitir el acceso al contenido verbal y auditivo del material audiovisual, por parte de distintas audiencias que no reciben el input auditivo o no tienen la posibilidad debido al idioma del diálogo, y por ende, se configuran como apoyo en los procesos de comprensión, aprendizaje y “adquisición de competencias idiomáticas” (**Kruger, Doherty, & Soto, 2017**).

La literatura reporta varios tipos de subtítulos utilizados en las producciones cinematográficas, los más comunes son: interlingüísticos e intralingüísticos. Los primeros son una traducción de un idioma extranjero (L2), que a su vez se dividen en subtítulos interlingüísticos estándar con el audio en L2 y los subtítulos en la lengua materna (L1); mientras que los interlingüísticos inversos presentan el audio en L1 y los subtítulos en L2. Asimismo, el segundo tipo de subtítulos se subdivide en: el captioning o subtítulos intralingüísticos en L1 que contienen la versión escrita y audio en la lengua materna; y los subtítulos intralingüísticos en L2, bimodales o subtítulos didácticos *per se* que presentan la información auditiva y escrita en la lengua extranjera (Díaz, 2012; Szarkowska & Gerber, 2018).

Según un estudio exploratorio de Matielo, de Oliveira & Baretta (2018), se encontró que el contenido subtitulado puede llegar a ser muy enriquecedor, sobre todo para la adquisición de una segunda lengua, en la cual a medida en que esta se desarrolla y se obtienen nuevas habilidades, la dependencia por los subtítulos disminuye. Por otra parte, se señala que la combinación de sonidos e imágenes mejora la relación entre palabras y sus significados, siendo estos a su vez contenido de entretenimiento que logra superar los textos básicos y estáticos, aportando información multisensorial para los distintos tipos de aprendizaje. De lo anterior, los autores concluyen que los subtítulos mejoran la comprensión de lenguaje, no obstante, se resalta que los subtítulos intralingüísticos fueron más efectivos para la comprensión lingüística, en comparación a los interlingüísticos y el no uso de subtítulos (García, 2017).

El subtitulado como técnica de traducción permite al espectador comprender contenido audiovisual con diferentes rasgos lingüísticos a los de su primera lengua (Ávila et al., 2020), por tanto, estos deben ser fáciles de leer y no obstaculizar el disfrute del video. Por ejemplo, en lo referente a la segmentación, los profesionales en el tema sugieren que debe seguir reglas sintácticas, es decir, las unidades lingüísticas deben permanecer juntas en una misma línea dentro del subtítulo, debido a que “los lectores no esperan hasta el final de la oración para interpretarla, sino que intentan encontrarle sentido mientras leen” (Gerber & Szarkowska, 2018).

Ahora bien, el fin último de la lectura es la comprensión, ya que como lo dijo Fonseca et al. (2019 como se cita en Menacho, 2021), la comprensión lectora es proceso mental llevado a cabo con el fin de “construir un significado propio sobre la base del contenido de un texto leído”, es por ello que la lectura involucra diferentes habilidades perceptivas, cognitivas y lingüísticas que se interrelacionan. De manera similar, Borrás & Lafayette (1994, como se citó en Jude, 2019), consideran que los subtítulos actúan como puente para unir la comprensión oral con la comprensión escrita; asimismo dejan en evidencia la relación que hay entre la velocidad a la que se leen los subtítulos, la comprensión y la forma de visualizar material audiovisual.

Las personas que hacen uso del material subtitulado fomentan el aprendizaje de lenguas extranjeras y fortalecen habilidades lingüísticas, como: la comprensión lectora a medida que la velocidad del texto aumenta, además del desarrollo y de la comprensión oral y escrita con la adquisición de nuevo léxico. Este último en conjunto con la escucha interpretativa y el contexto comunicativo real de las piezas audiovisuales permiten mejorar aspectos como la pronunciación (Talaván, 2017; Álvarez, 2018; González & Mañas, 2020; Torralba, 2020). Sin embargo, muchas de las subtitulaciones no cuentan con los efectos de sonido y la información extralingüística diseñada en la pieza audiovisual, fundamental para entender un vídeo en una lengua extranjera que no se domina (Torralba, 2020); perdiendo así parte del mensaje comunicativo, como la esencia de los elementos paralingüísticos y los signos acústicos verbales y no verbales.

De igual manera, esta información se deja de lado, ya que las películas subtituladas requieren de “procesar fuentes adicionales de información verbal y visual” (Kruger et al., 2017), compitiendo por recursos cognitivos atencionales (en especial los de atención visual, entre lo escrito y la imagen), aumentando de esta manera el procesamiento cognitivo tanto para la percepción de los signos visuales no verbales, verbales y auditivos de forma simultánea como la comprensión de la idea a comunicar, puesto que el foco de atención debe variar y no es posible realizar un seguimiento completo, en parte por no tener control alguno sobre la velocidad en la presentación del estímulo (Delabastita, 1989, como se citó en Hajmohammadi, 2004; Kruger & Steyn, 2014).

La lectura de subtítulos utilizados en el cine, al igual que la lectura de un texto estático, implica movimientos oculares como las fijaciones y las sacadas, las cuales han sido investigadas gracias al desarrollo de nuevas herramientas, entre ellas el eye-tracking o técnica de seguimiento ocular. Esta consiste en un conjunto de tecnologías que permiten rastrear, monitorear y registrar el movimiento y la generación de fijaciones oculares (puntos de estabilización de la mirada), en otras palabras, la forma en que un sujeto mira un contenido determinado en la pantalla permite establecer en qué áreas fija su atención (áreas de interés-AOI), el tiempo que permanece en estas y el orden de su exploración visual (Ramírez, 2012; Navarro et al., 2016; Papoutsaki, 2018).

Los movimientos oculares rápidos y amplios (Carballar, Martín, & Gámez, 2017), denominados sacadas tienen como función llevar la visión foveal, o también llamada central, a otra zona del texto o a otro objeto (Fernández, Orozco, Mandolesi, & Agamennoni, 2017) con el fin de captar la información visual allí presentada (Carballar et al., 2017), por tanto, los procesos cognitivos influyen sobre estos movimientos, como es el caso de la atención (Fernández et al., 2017).

Al observar una escena como una imagen estática, una sola fijación bastaría para permitir su comprensión, sin embargo, dada la velocidad y automaticidad de la percepción, el individuo no es totalmente consciente del tiempo que requiere para comprender aquello que observa. Es así como se resaltan dos formas en la que se lleva a cabo la identificación de la escena y sus elementos. Por un lado, pueden identificarse los objetos y luego comprender lo que se observa; y en contraposición, se puede percibir el significado y diseño de la escena, para que finalmente se detallen los objetos particulares (Intraub, 2002).

Teniendo en cuenta los estudios previos relacionados con las temáticas de interés, Ávila et al. (2020) indagaron acerca de los “factores que influyen en el proceso de subtitulación de un contenido audiovisual y que afectan la comprensión de un mensaje en un receptor”, en esta hallaron que la ambigüedad y la adición de palabras en el subtitulaje fueron uno de los errores más comunes hallados en el material observado, es decir, que la información presentada en los subtítulos esta sujeta a variedad de interpretaciones según las subjetividades de los espectadores del material, que para ese caso fueron series y películas de distinta temática y formato. Adicionalmente, estos autores hacen énfasis en la importancia de que el traductor considere los aspectos culturales y se encuentre formado para ello, pues puede incidir en la comprensión del mensaje.

Por su parte, Sueyoshi y Hardison (2005) se enfocaron en la importancia del lenguaje corporal dentro del proceso de comprensión a partir de la presentación de una conferencia, donde los participantes (estudiantes de inglés como L2), que observaron el material con apoyo visual y auditivo presentaron un mejor desempeño en la resolución de preguntas de un test relacionado con la comprensión del contenido. De igual manera, Wagner (2007) determinó la cantidad de tiempo que los participantes mantuvieron la mirada sobre la pantalla donde se proyectó un video en L2, y percibieron la información no verbal. A

partir de ello, el autor sugiere que las imágenes aportaron, según los participantes, para dar respuesta a las interrogantes; sin embargo, el “mirar en dirección a la pantalla no implica necesariamente prestar atención a las imágenes del vídeo” (Wagner, 2007 como se cita en Casañ, 2015).

Con base a lo expuesto, esta investigación buscó analizar la percepción, comprensión de imágenes y comprensión de subtítulos en piezas audiovisuales a partir de la metodología de seguimiento ocular indagando posibles relaciones entre las fijaciones oculares en los subtítulos y las imágenes dinámicas presentadas y su relación con la comprensión del contenido. Esta investigación se considera un aporte a los estudios de lectura mediados por el seguimiento ocular, de igual forma contribuye desde la psicolingüística a los procesos de diseño de piezas audiovisuales, y a la evidencia de la correlación positiva “entre la ejecución de los movimientos oculares sacádicos y el nivel de comprensión lectora [...], lo que implica que a mejor ejecución en los movimientos sacádicos mayor nivel de comprensión lectora” (Carballar et al., 2017).

Método

Diseño del estudio

El presente estudio fue de tipo descriptivo, integrando datos a partir de las medidas de seguimiento ocular y un cuestionario aplicado a los participantes, los cuales conformaron un único grupo de estudio. Todas las mediciones fueron realizadas de forma remota mediante el software *Sticky* de Tobii Pro.

Participantes

81 participantes diligenciaron el consentimiento informado antes del estudio, de los cuales participaron 73 personas en la visualización de los videos (ver Tabla 1). 64 participantes concluyeron con todo el estudio, de los cuales 72% eran mujeres y 28% hombres con edades entre 18-26 años (60.9%). Con respecto a la competencia para comprender el idioma inglés, 35.9% manifestó que lo era, 48.9% tal vez y un 15.6% respondió que no.

Tabla 1. Visualizaciones y datos usables en *Sticky*

	No. de participantes que lo visualizaron	No. de datos usables	Porcentaje de datos usables (%)
Escena 1	76	30	22.8
Escena 2	71	34	24.14
Escena 3	71	32	22.72

Los criterios de inclusión para el grupo de participantes fueron: ser mayor de 18 años, no autoidentificarse con alguna dificultad de tipo cognitivo, neurológico y/o psiquiátrico que comprometa sus habilidades comunicativas y/o comprensivas. Además, la población debía tener visión normal o poder realizar el ejercicio sin gafas, debido a los criterios requeridos para la grabación de los patrones oculares del software utilizado en la presente investigación.

El presente estudio contó con la aprobación del comité de ética de la Facultad de Medicina de la Universidad Nacional de Colombia, en

el Acta 018-197 de 2020, adicionalmente se siguieron todos los protocolos de confidencialidad y protección de los datos personales.

Diseño implementado e instrumentos

Se empleó el software *Sticky* desarrollado por la organización Tobii Pro (2018), el cual integra la tecnología de seguimiento ocular que permite evaluar por medio de la cámara web del computador de cada participante, el seguimiento ocular de largo alcance y aporta el tiempo y porcentaje de visualización, y los puntos de mirada o mapas de calor con el fin de observar cómo se distribuye la mirada sobre los estímulos multimedia evaluados y el foco de atención visual de los participantes (Tobii Pro, 2015). Asimismo, este software, al permitir la recolección de datos de manera remota, facilitó la ejecución de la investigación debido a las restricciones por la emergencia sanitaria (Anthonijsz, 2021); resultando más cómodo y accesible para los participantes e investigadores, lo cual redujo el tiempo, riesgos y costos comparados a la investigación clásica de eye-tracking como lo expone Gibaldi et al. (2017).

Es así que se evaluaron tres piezas audiovisuales del documental “*What the Health*” dirigido por Andersen & Kuhn (2017) que es de carácter naturalista, el cual es un tipo de estímulo que ha contribuido a explicar el apoyo del hipocampo en la segmentación y memorización de eventos cotidianos; que a su vez brinda información sobre los mecanismos cerebrales de atención generados cuando se adoptan perspectivas específicas durante la visualización de este tipo de contenido (Jääskeläinen et al., 2021).

Estas escenas cumplieron con las siguientes características: dos personajes principales (entrevistador y entrevistado) en primer plano, tópico sobre condiciones de salud, sin movimientos drásticos ni cambios de escenario, con la única diferencia entre sí por las condiciones de subtítulos y audio, como se especifica en la Tabla 2.

Tabla 2. Condiciones de los estímulos

	Audio		Subtítulos en español	Tiempo (s)
	Español (L1)	Inglés (L2)		
Escena 1		X	X	41
Escena 2	X		X	46
Escena 3	X			51

*La escena 3 presenta al inicio un pequeño cuadro de texto.

Con el fin de recopilar datos acerca de la comprensión y percepción de los participantes durante la lectura de los subtítulos, se diseñó un cuestionario, el cual fue revisado y aprobado por tres profesionales expertos quienes tuvieron un porcentaje de acuerdo del 90% de los ítems presentados, en este se solicitaba información útil para cumplir el objetivo de esta investigación, con ítems relacionados a los datos de los participantes y el contenido de las escenas:

- Primera sección: tres preguntas de introducción sobre edad, sexo biológico y comprensión del idioma inglés.
- Segunda sección: una pregunta sobre el contenido literal de los subtítulos presentados; dos preguntas sobre contenido inferencial y comprensión global del video; una pregunta subjetiva sobre cómo se sintió respecto a la condición del video (audio-subtítulos).
- Tercera sección: cinco preguntas de conclusión sobre qué aspectos fijó más su atención, qué elementos resalta, en cuál video fue más fácil su percepción y comprensión, el porqué de

su respuesta; por último, si observó algún error entre el audio y el subtítulo que se ubicaba en cada video.

Las respuestas a estas preguntas se encontraban en formato de selección múltiple en su mayoría; algunas de forma abierta, otras respuestas se encontraban en formato de respuesta corta y únicamente había una de verdadero/falso.

Procedimiento

Después de haber diseñado el estudio en la plataforma *Sticky* y haber construido el formulario y consentimiento, se procedió a hacer difusión masiva a los posibles participantes, siguiendo los criterios de inclusión descritos en la sección de participantes.

Las personas diligenciaron el consentimiento informado que les fue enviado. Posteriormente, observaron por única vez cada una de las escenas, lo cual requería que cada sujeto tuviera un computador con cámara web, se ubicara en un ambiente, según su elección con buena iluminación, especialmente frente a su rostro, y, además, no realizara movimientos de cabeza constantes, esto con el fin de tener un mayor porcentaje de respuesta y grabación en el software.

Al finalizar, los participantes respondieron el cuestionario relacionado al contenido de los videos observados.

Análisis de datos

Al alcanzar el registro de 30 datos usables por el software *Sticky*, se realizó el cierre de la visualización de los videos el cierre del estudio y los formularios para la posterior descarga de sus resultados, de los cuales se analizaron en el Programa SPSS 22 teniendo en cuenta la diferencia de medias. Adicionalmente se analizaron los siguientes estadísticos más relevantes según el objetivo del estudio: duración,

número de fijaciones, latencia (tiempo promedio hasta la primera fijación), duración total de la fijación, tiempo visto (tiempo promedio de fijaciones), orden de visualización y la desviación estándar del tiempo de fijación.

Adicionalmente, se realizó un análisis, a partir de la descripción de los principales fenómenos observados en los mapas de calor. A su vez, se tabularon y compararon los valores cuantitativos arrojados por las estadísticas agregadas de Aoi. Finalmente, se analizaron las respuestas arrojadas en el cuestionario de percepción y comprensión de la escena, clasificando los diversos elementos en categorías.

Resultados

A continuación, se presentan los resultados obtenidos en los tres videos utilizados. Primero se describirán los datos del video 1 que contaba con audio en inglés y subtítulos en español, después el video 2 con audio y subtítulos en español y finalmente, el video 3, cuyo audio se encontraba en español y no incluía subtítulos.

Video 1

Se analizaron dos áreas de interés: el personaje principal y los subtítulos del video 1. Para esto, se hizo uso de las métricas brindadas por el software *Sticky* donde se identifica que los participantes observaron más tiempo al personaje principal que los subtítulos, sin embargo no existen diferencias significativas entre estos dos valores. Estos datos pueden ser visualizados en la Tabla 3.

Según los datos aportados por *Sticky* en las estadísticas agregadas de Aoi, se construyó la Tabla 3, la cual consolida los valores relevantes para cada personaje y los subtítulos presentados en el video 1.

Tabla 3. Estadísticas del Área de interés del video 1

AOI	Duración (ms)	Z Duración	Número de fijaciones	Z Número fijaciones	Latencia (tiempo hasta que es notada el Aoi) (ms)	Z Latencia	Duración total de la fijación (ms)	Z Duración	Tiempo promedio de fijaciones (ms)	Z Tiempo de fijación	Orden de visualización o fijación	Desviación estándar del tiempo de fijación
Personaje principal	40664	50389.49*	14017	14200.11*	0.605	8.733*	919.199	931.38*	30.64	31.71*	1	5.362
Subtítulo	30596		6065		11.745		397.977		14.213		2	8.627

* Diferencia estadísticamente significativa (p<0,01)

En el video 1 el mapa de calor muestra que las primeras fijaciones se dan en la camiseta de la señora, posteriormente se enfocan en la cara y los subtítulos ubicados en la parte superior cuando aparecen, como se observa en la Figura 1. En esta, las fijaciones no fueron significativas sobre la zona de los datos de la protagonista y su edad; además, cuando esta aparece traducida, la mirada permanece en otra zona.



Figura 1. Mapa de calor del video 1

Nota: Los colores más cálidos como el rojo y amarillo son los puntos de mayor interés

Asimismo, es notorio que la mirada siempre sigue a los subtítulos, independientemente de la zona en la que son presentados. Se identifica que se enfocan brevemente en un objeto en movimiento ubicado en el fondo de la escena.

En la mayor parte del video, los participantes se enfocan en la protagonista, su rostro y manos, y en los subtítulos. En el entrevistador se resaltan fijaciones rápidas y de baja concentración en el momento en que este habla.

A partir del cuestionario de comprensión, el **45.3%** manifestó que no sabía dar respuesta a la pregunta literal del video sobre la edad de la protagonista y solo un **17.2%** respondió correctamente. Por otra parte, la mayoría de los participantes respondieron correctamente a la pregunta de inferencia del video.

Teniendo en cuenta su percepción sobre las condiciones del video: el **45.3%** expresó que comprendió parcialmente el audio y leyó los subtítulos; el **29.7%** ignoró el audio y se concentró solamente en los subtítulos; un **15.6%** comprendió el audio, sin embargo, leía o se apoyaba en los subtítulos. El **7.8%** no aplicaba para responder a la pregunta, un **1.6%** comprendió parcialmente el audio, pero no leyó los subtítulos; y la opción “comprendí todo el audio y no leí los subtítulos” obtuvo un **0%**.

Video 2

Las estadísticas de las fijaciones sobre la protagonista y los subtítulos presentados se pueden encontrar en la Tabla 4. Se aplicó la prueba Z para la diferencia de medias identificando en cuales son diferentes significativamente. En este video se mantiene un mayor foco atencional al personaje.

Tabla 4. Estadísticas del Área de interés del video 2

AOI	Duración (ms)	Z Duración	Número de fijaciones	Z Número fijaciones	Latencia (tiempo hasta que es notada el Aoi) (ms)	Z Latencia	Duración total de la fijación (ms)	Z Duración	Tiempo promedio de fijaciones (ms)	Z Tiempo de fijación	Orden de visualización o fijación	Desviación estándar del tiempo de fijación
Personaje principal	42348	58787.4*	10344	12043*	4.597	9.77*	678.398	789.7*	20.558	23.93*	1	8.558
Subtítulo	40790		6688		9.225		438.458		13.287		2	9.002

En cuanto al mapa de calor, se observó que las fijaciones inician solamente sobre la protagonista, luego miran brevemente el subtítulo y regresan al personaje principal; sin embargo, luego las fijaciones se dividen y son simultáneas entre la protagonista y los subtítulos.

A pesar que el entrevistador está presente desde el inicio del video, este no recibe atención hasta que habla, puesto que el centro o concentración de las fijaciones permanecen mayor tiempo en la señora y la niña, sin importar el ángulo de estas, lo cual se puede observar en la Figura 2.

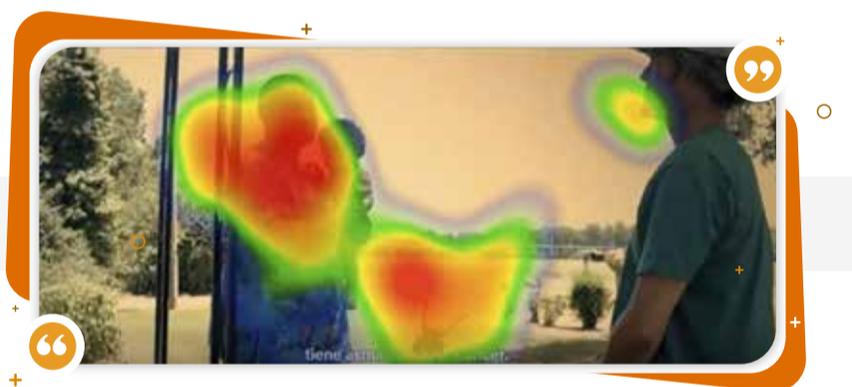


Figura 2. Mapa de calor del video 2

Por otra parte, se identifica que cuando hay un cambio en el plano de la grabación y gracias a esto se muestra un escenario más amplio, los participantes empiezan a concentrar sus miradas allí y duran aproximadamente 7 segundos observando este nuevo fondo. Al finalizar,

los participantes se concentran en los subtítulos, por tanto, disminuye la fijación sobre la protagonista. La mayoría de personas (**43.8%**) respondió correctamente a la pregunta literal del video.

En la pregunta de afirmación, ambas respuestas tuvieron porcentajes similares: Verdadero (**51.6%**), la cual era la respuesta correcta, y Falso (**48.4%**). En cuanto a la pregunta de inferencia, sólo un **25%** respondió correctamente.

Con respecto a cómo se sintieron con el audio y subtítulos en español, considerando que era de respuesta múltiple: a un **40.6%** le sirvieron de apoyo, al **26.6%** les fue indiferente, al **20.3%** le incomodó el audio original (inglés) de fondo, mientras escuchaba el audio en español; el **17.2%** se enfocó más en el subtítulo que en el audio, el **15.6%** al no entender palabras o ser términos diferentes a su vocabulario, se enfocaron en el subtítulo y al **15.6%** le molestó ver los subtítulos, teniendo el audio en el mismo idioma.

Video 3

Los datos estadísticos relevantes sobre los dos personajes y el cuadro de diálogo se consolidan en la Tabla 5. Se administró la prueba Z de diferencia de medias entre el personaje 1 y los subtítulos, se expresa el valor y la significancia en las AOI consideradas. En este caso, el subtítulo fue visto en menor medida tal como se expresa en la tabla en todas las medidas evaluadas.

Tabla 5. Estadísticas del Área de interés del video 3

AOI	Duración	Número de fijaciones	Latencia (ms)	Duración total de la fijación (ms)	Tiempo promedio de fijaciones (ms)	Orden de visualización o fijación	Desviación estándar del tiempo de fijación
Personaje 1	26181**	3214**	28.468**	211.202**	7.283**	3	3.444
Personaje 2	20414	4864	6.421	318.285	10.975	2	5.076
Subtítulo	6890**	569**	1.271**	35.384**	1.415**	1	1.065

** Diferencia estadísticamente significativa ($p < 0.01$)

*El personaje 2 corresponde al sujeto entrevistado Robert Reiner.

Una vez inicia el video, los participantes se enfocan en mirar el texto correspondiente a los datos del entrevistado. Inmediatamente después, pasan a la cara del personaje (primero en los ojos y la boca) y las fijaciones se reparten entre los dos elementos ya mencionados. Pasados cuatro segundos, se concentran solamente en la cara.

Apenas aparece el entrevistador, las fijaciones van netamente al nuevo personaje y luego estas se dirigen sutilmente a los diferentes objetos de fondo, de esta manera, realizando un recorrido ordenado de izquierda a derecha; el cual se vuelve inverso después de que se enfocó al entrevistado y aparece el personaje 1.

Dentro de este mismo recorrido, las fijaciones van de un lado a otro y logran mantenerse entre los personajes como se demuestra en la Figura 3. Además, son notorias breves fijaciones en el movimiento de las manos.



Figura 3. Mapa de calor del video 3

53.1% de la población manifestó que no sabía el nombre del protagonista del video. 35.9% respondió correctamente, a su vez que un 11% seleccionó las opciones de los nombres de las protagonistas de los videos anteriores.

En la pregunta de afirmaciones a partir del contenido, en la cual podían escoger más de una respuesta, las opciones correctas y que mayor porcentaje tuvieron fueron: “la Asociación Americana de la Diabetes no recomienda una dieta específica” (65.6%), “la Asociación Americana de la Diabetes está mejorando la vida de toda la gente afectada por ella (diabetes)” (20.3%), “comer sanamente no revierte la diabetes tipo II” (40.6%).

Un 48.4% anotó que se pudieron enfocar más en los elementos de la escena (objetos, personajes, emociones), un 40.6% se sintió más cómodo frente al hecho de que el video no tenía subtítulos, un 29.7% le fue indiferente, un 6.3% le incomodó el audio original (inglés) de fondo, mientras escuchaba el audio en español; y a un 4.7% le hicieron falta los subtítulos para los términos diferentes a su vocabulario.

A nivel general, la mayoría de los participantes (39.1%) manifestó haberse concentrado más en los subtítulos mientras veía los videos, un 35.9% se enfocó más en la escena y un 20.3% en el audio.

Por otra parte, a partir de la pregunta abierta sobre lo que más les gustó o llamó su atención de los videos, a nivel general se expresó que la temática, en segundo lugar, el contexto o la escenografía, luego los personajes, sus diálogos y actitudes respectivamente. Finalmente, 81.2% de los participantes no identificó errores de traducción o inconsistencias entre las palabras del audio y las del texto, a su vez que un 18.7% sí los notó. Entre estos, la mayoría (9 personas) reportaron que la traducción no era la más adecuada entre el audio y subtítulo durante el video 2, en palabras como bacon, sinusitis, o el uso de sinónimos. En contraparte, 3 personas anotaron que en el video 1 no correspondía la traducción que hubiesen hecho a partir de lo que escucharon o no era completamente exacta.

Discusión

La presente investigación tenía por objetivo analizar la percepción, comprensión de imágenes y comprensión de subtítulos en piezas audiovisuales a través de tecnología de seguimiento ocular, indagando sobre sus posibles relaciones entre las fijaciones oculares en los subtítulos y las imágenes dinámicas presentadas y su relación con la comprensión del contenido.

A partir de las variables analizadas, se encontró que la población objeto del estudio observó más al personaje principal en comparación a los subtítulos; aunque los usuarios manifestaron enfocarse más en el texto. Además, es importante resaltar que, de acuerdo con el tiempo de fijaciones, los datos del video 2 cualitativamente se observan como más dispersos, es decir, los participantes presentaron diferentes duraciones de sus fijaciones en las zonas de interés; lo anterior es posible porque el video cuenta con mayor número de distractores, un escenario más amplio, más información y complejidad de los subtítulos.

En contraste, en el video 3 el tiempo de fijaciones fue similar dentro de la población, ya que este no contaba con subtítulos, el enfoque de los personajes fue más cercano y el fondo era estático con un escenario cerrado. En este, corroborado también por los datos del mapa de calor, el orden de fijación fue diferente respecto al video 1 y 2, ya que el primer objeto visto fue el cuadro de texto y luego los personajes, probablemente porque los participantes presentaron una adaptación previa a los subtítulos vistos anteriormente, lo cual también se identifica en la latencia, puesto que fue aumentado en los personajes y disminuyendo en los subtítulos a medida que se presentaban en cada uno de los estímulos.

Durante la visualización de los videos, es notoria una tendencia a observar e identificar al sujeto que está hablando, lo cual lleva a que las fijaciones se den de manera organizada, creando un recorrido entre los elementos, demostrado en la división de los focos de atención como el personaje principal y los subtítulos, dado que estos son los que realmente aportan información.

Tal recorrido se hizo más evidente en los casos donde la conversación era más fluida, como en el video 3, donde la cara fue una zona de gran relevancia para los participantes, no solo porque los movimientos de la boca al hablar eran notorios, sino que también es considerada como parte de la comunicación no verbal.

De igual forma, los usuarios a través de su percepción reconocieron algunas emociones que predominaban en el protagonista de tal video, por ejemplo, confianza, interés y sorpresa, las cuales pudieron haber surgido de que el personaje demostraba manejo del tema y daba respuesta a la entrevista; aportando mayor credibilidad e información en el mensaje. Lo anterior también sucedió con el movimiento de las manos durante el video 1, ya que ayudaba a ejemplificar visualmente lo que expresaba.

Por otra parte, a medida que se avanzaba en la presentación de los videos, se amplió el campo visual, logrando una mayor cobertura de las escenas, debido a que las condiciones de los videos cambiaban respecto a sus planos de enfoque (video 2 y 3), lo que demostró la dispersión de las fijaciones a otras zonas de la pieza audiovisual. Esto se apoya del estudio de D'Ydewalle & Van Rensbergen (1989), citado por Bisson et al. (2014), a pesar de haberlo desarrollado en niños, encontraron que las fijaciones variaban según el tipo de película, puesto que estas disminuían en el área de subtítulos cuando se presentaban más elementos de interacción y movimiento en las películas de acción.

Asimismo, en el estudio de Bisson et al. (2014) se destaca que los subtítulos son dinámicos, por ello hay una tendencia a leerlos independientemente de ser competente en L2 y no necesitarlos, como sucedió en la presente investigación. No obstante, debido a que el movimiento capta al espectador, es común que al haber cambios de actividad en áreas diferentes a los subtítulos, la mirada se dirija automáticamente hacia allá, fenómeno que fue visible en los mapas de calor de todos los videos y en las estadísticas del número de fijaciones sobre los personajes.

Cabe mencionar que, los participantes realizaron mayor número de fijaciones, bien sea en la imagen o en los subtítulos, no solo por cuestiones de dinamismo, sino también porque seleccionaron uno de los dos inputs visuales disponibles (verbal o no verbal) por razones de eficiencia de procesamiento, como lo señala Matielo et al. (2018) en su estudio.

Sin embargo, las fijaciones en una u otras zonas pudieron haber causado la pérdida de alguna información que se estaba presentando en ese momento en los subtítulos (Bisson et al., 2014), situación que se observa claramente en el video 1, en el caso de la pregunta literal. Además, Muñoz (2017) anota que el público tiende a omitir en mayor medida los subtítulos en L1 en comparación a L2, sugiriendo que se requiere menos esfuerzo respecto a la eficiencia de procesamiento y lectura (Matielo et al., 2018) cuando el texto y la pista de audio son iguales.

Es de gran importancia mencionar que a la mayoría de participantes les fue más fácil la percepción de los detalles (movimientos, personajes, colores, acciones, gestos, lenguaje no verbal) y comprensión del contenido en el video 3, lo cual coincide con lo dicho en la pregunta subjetiva sobre sus actitudes frente a las condiciones de tal video, pues la mayoría anotó que se pudieron enfocar más en los elementos de la escena y se sintieron más cómodos frente al hecho de que el video no tenía subtítulos. Sin embargo, las respuestas no tuvieron un alto porcentaje de acierto en este video, lo cual lleva a considerar que hay otros factores influyentes, como la memoria, siendo esta una de las principales limitaciones en la investigación.

Igualmente, lo anterior lleva a considerar la estrategia en la que se evaluó la comprensión del contenido de los videos, pues si bien se tuvieron en cuenta preguntas literales, de afirmación, inferenciales y subjetivas, la mayoría en modalidad de respuesta cerrada; algunas investigaciones en este campo refieren haber empleado preguntas abiertas de comprensión general tipo “¿Por qué?”, a las cuales posteriormente se les asignaba puntaje. Asimismo, se aplicaron preguntas tipo “Verdadero o Falso” para la comprensión específica (Matielo et al., 2018), metodología que también fue aplicada por Szarkowska & Gerber (2019), quienes presentaron 5 preguntas de tal tipo después de cada video, estrictamente relacionadas con el contenido de los subtítulos, mas no con la imagen. En contraparte, Chen (2012) incluyó un test de lectura de respuesta de selección múltiple.

De igual manera, es necesario evaluar la pertinencia de las preguntas sobre el contenido, de modo que permitan obtener datos sobre la percepción y comprensión, sin la necesidad de depender en gran medida de la memoria u otros factores, para poder realizar un análisis entre ellos y dar respuesta al objetivo. Al igual, el tema de poder evaluar la comprensión se vio afectado debido al no seguimiento de instrucciones, ya que varios participantes respondieron el cuestionario a partir de su opinión personal, mas no de forma objetiva como se pedía en determinadas preguntas.

Como lo ha anotado Kruger & Steyn (2014), las metodologías cualitativas son mayormente utilizadas cuando se consideran mapas de calor y sus muestras son pequeñas en términos de los subtítulos

estudiados y la población, lo cual es evidente en esta aproximación práctica; por ende, se requiere incluir mayor número de participantes en investigaciones de eye-tracking y subtítulos, con el fin de poder generalizar los resultados obtenidos.

Es importante resaltar que, los participantes manifestaron que el estar tan restringidos en sus movimientos, especialmente en la cabeza disminuía su rendimiento, ya que impedía tener la total atención en el contenido de la pieza audiovisual y se concentraban en mantener la postura. También, se expresó que la duración de los videos era muy corta para lograr una adaptación adecuada de lo que se mostraba, se sugiere incluir al menos un ejercicio previo para que los participantes comprendan la metodología de presentación.

Para próximos estudios es importante tener en cuenta el tiempo de presentación de cada uno de los personajes y los subtítulos, ya que estos difieren entre sí en cada uno de los videos, afectando el análisis de los mismos. Además, debe considerarse el formato de los videos para estudiar e incluir diferentes estímulos, como lo son: personajes de grupos etarios diferentes, fondo blanco versus con distractores, escenarios con más elementos, acciones más relevantes, mayor número de personajes y diálogos dinámicos tipo conversaciones más complejas. Asimismo, es necesario incluir dentro de las variables a analizar el nivel de estudios de los participantes, la posible relación con el género de estos y el nivel de exposición previa a piezas audiovisuales con el fin de identificar si existe una influencia de estas condiciones en las respuestas. Sin embargo, se destaca que los tópicos, diálogos y el formato tipo documental utilizado en esta investigación, aportaron información e instruyeron a los participantes.

Conclusiones

Se concluye que los subtítulos en L1 son un apoyo para las personas competentes y no competentes en L2, además, se destaca que los subtítulos son dinámicos y contribuyen al procesamiento de la información presentada, por ello hay una tendencia a leerlos independientemente de su competencia en la lengua extranjera.

Fue evidente que el dominio o manejo de L2 influye en la percepción y comprensión de la pieza audiovisual, manifestado por los propios participantes en cuanto a sus actitudes y la identificación de errores en la estrategia de subtítulaje, como una traducción no adecuada.

Por otro lado, las fijaciones fueron más altas en los personajes y elementos de la escena, debido a la constante actividad que se presenta en estos en comparación a los subtítulos; no obstante, la percepción de los participantes fue lo opuesto, pues la mayoría expresó haber percibido los detalles y comprendido mejor el contenido en el video sin subtítulos.

El orden de visualización de cada uno de los objetos fue diferente en las piezas audiovisuales presentadas, ya que en los videos subtítulados con audio en inglés y español se observaron primero los personajes y luego los subtítulos, mientras que en el video doblado y sin subtítulaje sucedió lo contrario (video 3). Sin embargo, no fue posible identificar si hay una ganancia significativa en la comprensión del contenido cuando hay más fijaciones en la imagen o en los subtítulos.

Es importante resaltar que los movimientos y expresión corporal constituyen un medio de comunicación no verbal (manos, emociones, gestos, etc.) que apoya el mensaje presentado en el input auditivo y visual verbal (subtítulos). Por lo mencionado anteriormente, son nece-

sarios más estudios que incluyan tecnología de eye-tracking y consideren la influencia de otros procesos en la comprensión, percepción y lectura de subtítulos con un posible abordaje multidisciplinar, teniendo en cuenta, la metodología de evaluación respecto a la comprensión y el formato de los estímulos utilizados.

Los hallazgos de esta investigación aportan evidencia para las áreas de diseño de contenidos, usabilidad, psicología del consumidor, psicolingüística, entre otras áreas, dado que se reúne evidencia en torno a la relevancia de la lectura de subtítulos y a marcadores diferenciales en la percepción de imágenes y texto en una misma escena. Se espera profundizar en estos campos en cuanto al estudio de la interacción de la percepción, la atención y la lectura en video.

Referencias

- Álvarez, P. (2018). Una imagen vale más que mil palabras: La subtitulación como recurso didáctico en el aprendizaje de lenguas. *Redit - Revista Electrónica De Didáctica De La Traducción Y La Interpretación*, (11), 16-32. <https://doi.org/10.24310/REDIT.2017.v0i11.3813>
- Andreu, C., Martín M., Gruart, A., & Delgado J. (2021). The Effect of Media Professionalization on Cognitive Neurodynamics During Audiovisual Cuts. *Frontiers in systems neuroscience*, 15. <https://doi.org/10.3389/fnsys.2021.598383>
- Anthonijsz, T. (2021). *Are high-end and webcam eye trackers comparable?* [Tesis de Pregrado, Utrecht University]. <https://studenttheses.uu.nl/bitstream/handle/20.500.12932/40672/Anthonijsz%206642268.pdf?sequence=1#:~:text=High%2Dend%20eye%20trackers%20will,that%20require%20larger%2Dscale%20data>
- Ávila, C., Herrán, K., Muñoz, A., Rodríguez, N., & Torres, D. (2020). La subtitulación audiovisual en un mundo globalizado, ¿cumple su función?. *Comunicación, Cultura Y Política*, 9. <https://doi.org/10.21158/21451494.v9.n0.2018.2729>
- BBC News Mundo. (2020). *Coronavirus: ¿en qué casos es recomendable el autoaislamiento y cómo debe hacerse?*. <https://www.bbc.com/mundo/noticias-51875858>
- Bisson, M., Van Heuven, W., Conklin, K., & Tunney, R. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, 35(2), 399-418. <https://doi.org/10.1017/S0142716412000434>
- Casañ J. (2015). Un marco teórico sobre el uso de preguntas de comprensión audiovisual integradas en el vídeo como subtítulos: un estudio mixto. marcoELE *Revista de Didáctica Español Lengua Extranjera*, (20), 1-45 <https://www.redalyc.org/articulo.oa?id=92152428015>
- Carballar, R., Martín, P., & Gámez, M. (2017). Relación entre habilidades neuropsicológicas y comprensión lectora en Educación Primaria. *Cuadernos de Investigación Educativa*, 8(2), 49-59. <https://doi.org/10.18861/cied.2017.8.2.2686>
- Chen, M. (2012). Effects of the order of reading text or viewing a film and L1/L2 captions on reading comprehension. *Perceptual and motor skills*, 115(1), 18-26. <https://doi.org/10.2466/23.PMS.115.4.18-26>
- Díaz, J. (2012). Los subtítulos y la subtitulación en la clase de lengua extranjera. Abehache: *Revista da Associação Brasileira de Hispanistas*, 2(3), 99. <https://bit.ly/2WxctRR>
- El Tiempo. (2021). *Netflix en cuarentena, así fue el 2020 de la plataforma de «stream»*. El Tiempo. <https://www.eltiempo.com/tecnosfera/apps/netflix-en-cuarentena-asi-fue-el-2020-de-la-plataforma-de-stream-561583>
- Fernández, G., Orozco, D., Mandolesi, P., & Agamennoni, O. (2017). *Evaluación de patologías neurocognitivas mediante los movimientos oculares*. CIC-DIGITAL Repositorio Institucional de la Comisión de Investigaciones Científicas. <https://digital.cic.gba.gob.ar/handle/11746/6650>
- García, B. (2017). Bilingual subtitles for second language acquisition and application to engineering education as learning pills. *Computer applications in engineering education*, 25(3), 468-479. <https://doi.org/10.1002/cae.21814>
- Gerber, O., & Szarkowska, A. (2018). Line breaks in subtitling: an eye tracking study on viewer preferences. *Journal of eye movement research*, 11(3). <https://doi.org/10.16910/jemr.11.3.2>
- Gibaldi, A., Vanegas, M., Bex, P., & Maiello, G. (2017). Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behav Res* 49, 923-946. <https://doi.org/10.3758/s13428-016-0762-9>
- González B., & Mañas, I. (2020). Efectos de los subtítulos intralingüísticos y los subtítulos bilingües aumentados sobre el aprendizaje incidental de vocabulario en español como lengua extranjera. RILEX. *Revista Sobre Investigaciones léxicas*, 3(2), 125-163. <https://doi.org/10.17561/rilex.3.2.5808>
- Hajmohammadi, A. (2004). The Viewer as the Focus of Subtitling Towards a Viewer-oriented Approach. *Translation Journal*, 8(4). <http://translationjournal.net/journal/30subtitling.htm>
- Intraub, H. (2002). *Visual scene perception*. *Encyclopedia of cognitive science*. <https://sites.udel.edu/cas-intraublal/files/2017/10/2002-Intraub-Encyclopedia-of-Cognitive-Sci-1uop5d8.pdf>
- Jääskeläinen, I., Sams, M., Gleran, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage*, 224, 117445. <https://doi.org/10.1016/j.neuroimage.2020.117445>
- Jude, D. (2019). *La subtitulación como estrategia de aprendizaje del español como LE* [Tesis de Maestría, Universidad de Alicante]. https://rua.ua.es/dspace/bitstream/10045/96928/1/La_subtitulacion_como_estrategia_de_aprendizaje_del_espano_Jude_Diana_loana.pdf
- Kruger, J., & Steyn, F. (2014). Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly*, 49(1), 105-120. <https://doi.org/10.1002/rrq.59>
- Kruger, J., Doherty, S., Soto, M. (2017). Subtítulos en lengua original: sus efectos en el espectador nativo y extranjero. *Comunicar*, vol. XXV, núm. 50, enero-marzo, 2017, pp. 23-32. ISSN: 1134-3478. <https://www.redalyc.org/articulo.oa?id=15849613002>
- Lavour, J., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles?. *International Journal of Psychology*, 46: 455-462. <https://doi.org/10.1080/00207594.2011.565343>
- Matielo, R., de Oliveira, R., & Baretta, L. (2018). Intralingual subtitles, interlingual subtitles, and video comprehension: insights from an exploratory study. *Letrônica*, 10(2), 758-774. <https://doi.org/10.15448/1984-4301.2017.2.26370>
- Menacho, L. (2021). Estrategias colaborativas: aprendizaje compartido para el desarrollo de la comprensión lectora en estudiantes de educación primaria. *Praxis educativa*, 25(3), 243-258. <https://dx.doi.org/10.19137/praxiseducativa-2021-250314>
- Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System*, 67, 77-86. <https://doi.org/10.1016/j.system.2017.04.015>
- Navarro, Ó., & Molina, A., & Lacruz, M. (2016). Utilización de Eye Tracking para evaluar el uso de información verbal en materiales multimedia. Pixel-Bit. *Revista de Medios y Educación*, (48), 51-66. ISSN: 1133-8482. <https://www.redalyc.org/articulo.oa?id=36843409005>
- Papoutsaki, A. (2018). *Democratizing Eye Tracking*. <http://cs.brown.edu/people/apapouts/Dissertation.pdf>
- Ramírez, A. (2012). *Eye-tracking: una técnica de seguimiento de la mirada utilizada en la validación de unidades de aprendizaje*. Instituto Tecnológico de Costa Rica. <https://recursos.educoas.org/publicaciones/eye-tracking-una-tcnica-se-seguimiento-de-la-mirada-utilizada-en-la-validaci-n-de>

- Sueyoshi, A. & Hardison, D. (2005). El papel de los gestos y las señales faciales en la comprensión auditiva de un segundo idioma. *Aprendizaje de idiomas*, 55(4), 661–699. <https://doi.org/10.1111/j.0023-8333.2005.00320.x>
- Szarkowska, A., & Gerber, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PLoS ONE* 13(6). <https://doi.org/10.1371/journal.pone.0199331>
- Szarkowska, A., & Gerber, O. (2019). Two or three lines: a mixed-methods study on subtitle processing and preferences, *Perspectives*, 27:1, 144-164, <https://doi.org/10.1080/0907676X.2018.1520267>
- Talaván, N. (2017). Justificación teórico-práctica del uso de los subtítulos en la enseñanza-aprendizaje de lenguas extranjeras. *TRANS. Revista De Traductología*, (16), 23-37. <https://doi.org/10.24310/TRANS.2012.v0i16.3209>
- Tobii Pro. (2015). *How to Work with Heat Maps and Gaze Plots*. <https://www.tobii.com/learn-and-support/learn/steps-in-an-eye-tracking-study/interpret/working-with-heat-maps-and-gaze-plots/>
- Tobii Pro. (2018). *Sticky by Tobii Pro*. <https://www.tobii.com/es/products/sticky-by-tobii-pro/>
- Torralba, G. (2020). The Use of Passive and Active Subtitles in Foreign Language Teaching: A View to their Teaching Potential. Íkala, *Revista de Lenguaje y Cultura*, 25(1), 231-250. <https://doi.org/10.17533/udea.ikala.v25n01a11>
- Wagner, E. (2007). “Are they watching? Test-taker viewing behavior during an L2 video listening test”. *Language Learning and Technology*, 11(1), 67-86. <https://www.lltjournal.org/item/10125-44089/>