

## GRAPE-MARS: una nueva herramienta para el an3lisis multimodal en la investigaci3n sobre segundas lenguas

Inmaculada FORTANET-G3MEZ

Universitat Jaume I

[fortanet@uji.es](mailto:fortanet@uji.es)

<https://orcid.org/0000-0001-8981-565X>

Noelia RUIZ-MADRID

Universitat Jaume I

[madrid@uji.es](mailto:madrid@uji.es)

<https://orcid.org/0000-0001-9800-5624>

Edgar BERNAD-MECH3

[ebernad@uji.es](mailto:ebernad@uji.es)

<https://orcid.org/0000-0002-7271-7735>

Julia VALERIAS-JURADO

[valeiras@uji.es](mailto:valeiras@uji.es)

<https://orcid.org/0000-0001-8487-4968>

**Resumen:** El objetivo de este art3culo es presentar el software GRAPE-MARS (*Multimodal Analysis Research Software*) (<https://www.grape.uji.es/mars>), un programa de c3digo abierto y uso gratuito desarrollado por el grupo de investigaci3n GRAPE (*Group for Research on Academic and Professional English*) para el an3lisis multimodal de v3deos, en el contexto particular de la investigaci3n en segundas lenguas. La investigaci3n multimodal ha avanzado mucho en los 3ltimos a3os, especialmente en el an3lisis del discurso, aunque tambi3n se han realizado estudios relacionados con la traducci3n o la pragm3tica, entre otros. GRAPE-MARS es un programa de interfaz sencilla y usabilidad asequible que pueden utilizar tanto investigadores e investigadoras, como profesores o estudiantado y cuyo objetivo es facilitar la tarea de identificar los elementos o modos que conforman el discurso y facilitar su representaci3n. En una 3poca en que los g3neros digitales, con una multimodalidad inherente, cada vez cobran m3s protagonismo, creemos que GRAPE-MARS puede ser una herramienta esencial para desgranar los elementos que los componen y contribuir a an3lisis multimodales m3s completos que ayuden en la investigaci3n de las segundas lenguas.

**Palabras clave:** an3lisis multimodal, herramientas de investigaci3n, discurso oral, g3neros digitales, segundas lenguas

*Catal3:*

**GRAPE-MARS: una nova eina per a l'an3lisi multimodal en la recerca sobre segones lleng3es**

**Resum:** L'objectiu d'aquest article 3s presentar el programari GRAPE-MARS (*Multimodal Analysis Research Software*) (<https://www.grape.uji.es/mars>), un programa de codi obert i 3s gratu3t desenvolupat pel grup de recerca GRAPE (*Group for Research on Academic and Professional English*) per a l'an3lisi multimodal de v3deos, en el context particular de la recerca en segones lleng3es. La recerca



multimodal ha avançat molt en els últims anys, especialment en l'anàlisi del discurs, encara que també s'han realitzat estudis relacionats amb la traducció o la pragmàtica, entre altres. GRAPE-MARS és un programa d'interfície senzilla i usabilitat assequible que poden utilitzar tant els investigadors i investigadores com els professors o estudiants i l'objectiu del qual és facilitar la tasca d'identificar els elements o modes que conformen el discurs i facilitar la seva representació. En una època en què els gèneres digitals, amb una multimodalitat inherent, cada vegada cobren més protagonisme, creiem que GRAPE-MARS pot ser una eina essencial per a desgranar els elements que els componen i contribuir a anàlisis multimodals més completes que ajudin en la recerca de les segones llengües.

**Paraules clau:** anàlisi multimodal, eines de recerca, discurs oral, gèneres digitals, segones llengües

#### English:

#### **GRAPE-MARS: a new tool for multimodal analysis for research on second languages**

**Abstract** The aim of this article is to present GRAPE-MARS (Multimodal Analysis Research Software) (<https://www.grape.uji.es/mars>), a free open-source software developed by GRAPE (Group for Research on Academic and Professional English) for the multimodal analysis of videos in the context of second language research. Multimodal research has come a long way in recent years, especially in discourse analysis, although studies related to translation, pragmatics, and others have also been carried out. GRAPE-MARS is a programme with a simple interface and user-friendly that can be used by researchers, teachers, and students, and whose aim is to facilitate the task of identifying the elements or modes that make up discourse and to facilitate their representation. At a time when digital genres, with their inherent multimodality, are becoming increasingly prominent, GRAPE-MARS can be an essential tool for unpacking their constituent elements and contributing to more complete multimodal analyses that aid second language research.

**Keywords:** multimodal analysis, research tools, oral discourse, digital genres, L2

## Introducción

En los últimos años se ha producido un auge de los estudios de multimodalidad. En el caso del análisis del discurso, hemos pasado de la investigación sobre textos escritos o transcripciones de textos orales, a tener una visión global que abarca los diferentes elementos que acompañan al texto: tipografías, colores, disposición o imágenes estáticas o dinámicas, en el caso de lenguaje escrito, o proxemia, paralenguaje, gestos y ayudas visuales, en el caso de lenguaje oral (O'Halloran, 2012, Ruiz-Madrid y Fortanet-Gómez, 2015, Hart, 2020).

La multimodalidad se entiende como la representación y construcción del significado a través de diferentes modos comunicativos como puede ser la mirada, los gestos o el lenguaje hablado (Kress, 2010, Manghi, 2011). Según Kress y van Leeuwen (2021), en el panorama semiótico actual, esencialmente multimodal, se están produciendo nuevas relaciones entre lo verbal y lo visual hacia lo que ellos denominan "new visual literacy" (p. 43). Es decir, el significado del lenguaje es el resultado de la combinación de modos, qué se dice junto con cómo se dice.

Se han realizado estudios de multimodalidad en diversos campos, como la semiótica (Wanselin et al., 2022), la traducción (Đorđević y Stamenković, 2023) o la enseñanza de segundas lenguas (Manghi, 2011, Alfonso Lozano y Guiralt Lorenz, 2014, Parodi y Julio, 2017; Ruiz-Madrid y Valeiras-Jurado, 2020; Beltrán-Palanques, 2024), aunque algunos de estos estudios se han centrado más en el discurso escrito. Sin embargo, en el campo del análisis del discurso oral es probablemente donde más se ha trabajado la multimodalidad (Kress y van Leeuwen, 1996; Norris, 2004; O'Halloran, 2004; Jewitt et al., 2016), sobre todo en contextos académicos, centrándose en presentaciones en congresos (Querol-Julían y Fortanet-Gómez, 2012) y clases universitarias (Bernad-Mechó, 2017; Crawford



Camiciottoli, 2015, 2016; Fox y Artemeva, 2012). En algunos casos estos estudios han derivado en aplicaciones de análisis multimodales a la enseñanza y aprendizaje del inglés especializado como segunda lengua desde una perspectiva de estudios de género (por ejemplo, Coccetta, 2018; Crawford Camiciottoli, 2019; Ruiz-Madrid y Valeiras-Jurado, 2020), sugiriendo nuevas pedagogías multimodales para la enseñanza-aprendizaje de géneros (Beltrán-Palanques y Bernad-Mechó, 2024).

Con todo, se hace cada vez más evidente la necesidad de llevar a cabo estudios multimodales de forma práctica y efectiva, que faciliten la rápida aplicación del análisis del discurso oral y de sus resultados en contextos relevantes como la investigación sobre segundas lenguas. Así pues, el objetivo de este artículo es introducir al lector en los estudios de multimodalidad y, en especial, en el uso de GRAPE-MARS (Multimodal Analysis Research Software) para la realización de esos estudios.

## 1. ¿Cómo investigar en multimodalidad en L2?

Para empezar a hacer investigación en multimodalidad, es necesario elegir el enfoque teórico que se le quiere dar al estudio. A continuación, se establecerá un objetivo y se deberá seleccionar cuál va a ser el corpus, es decir, el texto oral o escrito que vamos a analizar; y las variables, es decir, qué pretendemos encontrar en ese corpus. Por ejemplo, el objetivo puede ser comparar el uso de los distintos modos entre un video de YouTube sobre el uso de la IA con muchas visualizaciones y otro del mismo tipo, pero con muy pocas visualizaciones, para identificar cuáles son más efectivos, o cuáles tienen un papel más importante dentro de cada vídeo. En su aplicación a la enseñanza y aprendizaje de la L2, el objetivo puede ser identificar los recursos semióticos que ayudan a la persuasión en la lengua inglesa y que los estudiantes aprenderán a identificar y a utilizar en sus propias producciones de video para sus trabajos académicos o profesionales futuros. Cuando se elija un video se deben tener en cuenta los derechos de autor. Es preferible utilizar un video sin derechos de autor o publicado bajo licencia *Creative Commons*. Si no es así, se debe obtener el consentimiento del autor para utilizarlo para la investigación y, en su caso, para publicar imágenes pertenecientes a dicho vídeo. El objetivo del estudio que estamos ejemplificando determinará el enfoque multimodal que se le quiera conferir (ver sección 1.2. sobre los distintos enfoques y 1.2. sobre conceptos clave). Teniendo en cuenta este objetivo, el corpus consistirá en vídeos de YouTube, aunque hay que determinar antes de emprender el estudio cuánto se quiere abarcar, ya que los análisis multimodales pueden llegar a ser muy detallados y, en esos casos, puede llevar mucho tiempo analizar vídeos largos. Por ello, se recomienda extraer fragmentos de 4-5 minutos comparables entre los dos vídeos según el objetivo de análisis y de investigación. También se puede fragmentar cada vídeo y analizar los extractos por separado. En tercer lugar, se deberán establecer las variables, es decir, los modos que deseamos identificar. En el apartado 1.3. se explican algunas posibilidades de clasificaciones de modos propuestas en estudios anteriores. En cuarto lugar, se deberá elegir



realizar un análisis manual o seleccionar un programa o herramienta digital para realizarlo. Esto se explicará en el apartado 2 de este artículo.

### 1.1 Enfoques multimodales

En primer lugar, debe estar claro que, al contrario de lo que ocurre en la mayoría de los estudios lingüísticos, los enfoques multimodales no se centran en la lengua, sino que la consideran uno más de los modos que contribuyen a la creación de significado. Un análisis multimodal proporciona una visión más completa de cómo tiene lugar una comunicación y qué elementos juegan un papel en ella. Hay tres enfoques básicos para la investigación multimodal del discurso oral (Bernad-Mechó, 2021; Jewitt, 2014): la Semiótica Social Multimodal (SSM) (Kress y van Leeuwen, 1996, 2001), el Análisis del Discurso Multimodal (ADM) (O'Halloran, 2004, 2011), y el Análisis de (Inter)acción Multimodal (AIM) (Norris, 2004). Cada uno de estos enfoques ofrece herramientas analíticas específicas y difiere en sus teorías subyacentes, el énfasis dado al contexto y la interrelación entre modos, así como en la atención prestada al creador de signos (Jewitt, 2014). Por lo tanto, mientras que la SSM muestra un interés particular en el creador de signos en un contexto social, el ADM se centra más en los recursos semióticos como sistemas de elecciones, y el AIM se enfoca principalmente en la interacción de los actores sociales.

El primer enfoque, la SSM, parte de la Semiótica Social de Halliday (1978, 1985) y la Lingüística Sistémico-Funcional para explorar el proceso de creación de significado como una actividad que ocurre en entornos sociales (Kress, 2010). Dentro de este modelo, se destacan dos aspectos principales de la comunicación: el estudio de diferentes modos o recursos semióticos en la construcción de significados y la importancia del contexto social en el que tiene lugar esta construcción. Así, se pone énfasis en el creador de signos y las elecciones semióticas que realiza en un contexto determinado. En otras palabras, la SSM se interesa por cómo los escritores/oradores crean significado mediante el uso de múltiples modos en un mundo social. Adami (2017) resume los objetivos principales de la SSM como la expresión de procesos sociales mediante signos. A partir de un análisis cualitativo minucioso de muestras generalmente pequeñas de textos, este enfoque busca revelar ideologías, valores sociales, roles de poder e identidades expresadas en textos, acercándose así al Análisis Crítico del Discurso Multimodal (ACDM) desarrollado por Machin et al. (2016) o Roderick (2017).

El segundo enfoque, el ADM, se basa en la Lingüística Sistémico-Funcional de Halliday (1985) y se ocupa del estudio del lenguaje como el resultado de la combinación de recursos semióticos. Su objetivo principal es explorar modelos de recursos semióticos, cómo estos interactúan para transmitir significado y cómo pueden cambiar a través de procesos de resemiotización (Iedema, 2003). En el ADM, el lenguaje parece recibir prioridad sobre otros modos; hasta cierto punto, se considera que el lenguaje es el punto de partida que luego se explora en combinación con otros modos. Este enfoque es particularmente relevante en el estudio de las conferencias, ya que el lenguaje hablado se considera el medio fundamental para transmitir contenido en este género (Ruiz-Madrid y Fortanet-Gómez, 2015). En resumen, los



análisis del ADM suelen comenzar con un examen detallado y comprensivo de partes cortas de textos, entendidos como cualquier fuente capaz de transmitir significado, con el objetivo de desarrollar teorías y marcos que expliquen cómo funcionan los recursos semióticos.

El tercer enfoque, el AIM, se fundamenta en la teoría del Análisis de Discurso Mediado de Scollon (2001) y se centra en el concepto de acción mediada. En esencia, Scollon (2001) propone un enfoque para analizar el discurso donde el analista se desvincula del texto — lo que se dice — y se enfoca en las acciones realizadas por los actores sociales — los hablantes. Este enfoque es aplicado por Norris (2004) al estudio multimodal de la comunicación. Así, el AIM se interesa por cómo los actores sociales interactúan en contextos específicos y situados; según Norris (2014a), el trabajo del actor lo es todo.

## 1.2 Conceptos clave del análisis multimodal

El concepto de “modo” es esencial en los estudios multimodales. Kress y van Leeuwen (2001) lo consideran sistemas semióticos con normas y comportamientos regulares. Se trata de imágenes, gestos, lenguaje, música, disposición, modo de escritura, proxémica, postura, entre otros. Van Leeuwen (2005) habla de “recursos semióticos” y los define como

the actions, materials and artifacts we use for communicative purposes, whether produced physiologically – for example, with our vocal apparatus, the muscles we use to make facial expressions and gestures – or technologically – for example, with pen and ink, or computer hardware and software – together with the ways in which these resources can be organized. (van Leeuwen, 2005, p. 285)

Estos dos conceptos se utilizan a menudo como sinónimos en los estudios de multimodalidad. Otros dos conceptos que merecen atención son el *multimodal affordance* o potencial multimodal y el *multimodal ensemble* o conjunto o constructo multimodal. El potencial multimodal alude al hecho de que los modos pueden contribuir a construir la carga de significado discursivo y comunicativo, al mismo tiempo que enfatiza la importancia de los aspectos culturales. Así, subraya la importancia de especificar el contexto y el entorno social como intrínsecamente conectados a los aspectos culturales de un acto comunicativo dado. Como lo expresa Kress (2010), "meanings are socially made, socially agreed and consequently socially and culturally specific" (p. 88). Por otro lado, los conjuntos multimodales se refieren a las combinaciones específicas de modos que ocurren en la comunicación. El estudio de cómo diferentes modos contribuyen a la creación de significado, así como el funcionamiento de esta interrelación en el proceso de creación de significado, es fundamental en el análisis de la SSM.

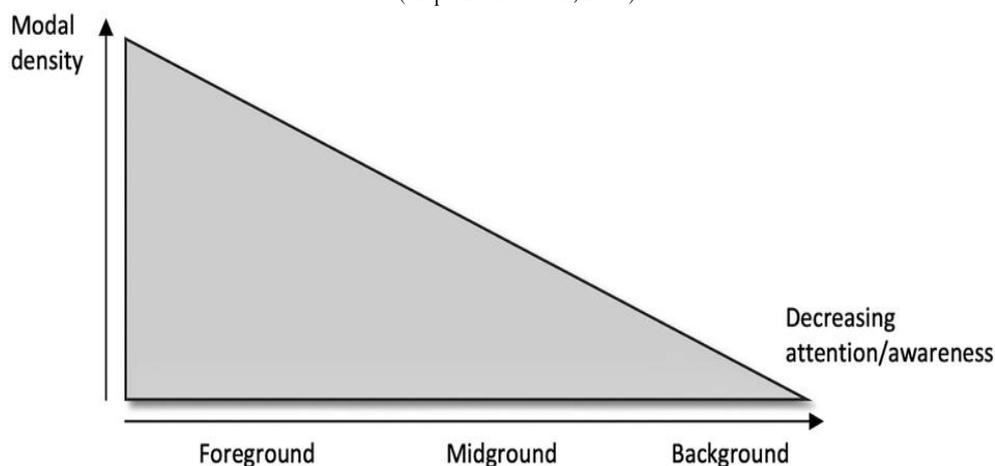
En los estudios de ADM, también se emplean frecuentemente los conceptos de potencial multimodal y constructo multimodal, aunque con enfoques de análisis diferentes en cada paradigma. Mientras que la SSM presta atención al papel del contexto en la comunicación, el ADM se centra en desarrollar marcos que expliquen cómo ciertas combinaciones de modos funcionan en eventos semióticos para identificar patrones recurrentes. Estos análisis suelen realizarse con software de anotación multicapa como ELAN (<https://archive.mpi.nl/tla/elan#intro>) (Wittenburg et



al., 2006), Multimodal Analysis-Video (MMA-Video) (O'Halloran et al., 2012) (<https://multimodal-analysis.com/products/multimodal-analysis-video/index.html>) o GRAPE-MARS (Ruiz-Madrid et al., 2023) (<https://www.grape.uji.es/mars/>). Estas herramientas permiten al analista representar visualmente combinaciones específicas de modos mediante la creación de capas para describir y anotar cada uno de los modos. Con MMA-Video y con GRAPE-MARS, una vez que se ha anotado un archivo de video, el analista también puede obtener datos cuantitativos sobre el uso de recursos semióticos, así como una descripción de la combinación de modos en un momento determinado. Además, como se detalla más adelante, GRAPE-MARS también permite la representación de dichos datos en gráficos y tablas, y fomentar así la visualización de los resultados de investigación (Barros, 2021).

Por último, debemos nombrar el concepto de densidad modal. Este concepto se refiere al hecho de que cuando las acciones reciben más atención, se vuelven modalmente densas. Esto puede ocurrir a través de una mayor complejidad modal, cuando una acción se lleva a cabo a través de muchos modos, o a través de una mayor intensidad modal, cuando uno de los modos asume un papel más intenso. Estos conceptos, a su vez, están relacionados con el de configuraciones modales, que son algo similares a los conjuntos multimodales, con la particularidad de que los modos se representan en orden de importancia, según tengan su papel en la acción. Las acciones más densas están en el primer plano para el hablante y las menos densas en un plano medio o posterior, ya que reciben menos atención (ve Figura 1).

**Figura 1.** Densidad modal en el continuo de primer plano a plano posterior según la atención prestada por el hablante (adaptado de Norris, 2004).



Los tres enfoques para el estudio de la multimodalidad han sido fundamentales para numerosos estudios prácticos con enfoques diversos. La decisión de elegir un enfoque sobre otro en un estudio particular dependerá de los objetivos y las preguntas de investigación planteadas. Aunque las combinaciones explícitas de los tres enfoques son poco comunes, no son mutuamente excluyentes. De hecho, los límites entre ellos no siempre son nítidos y pueden ser reinterpretados y debatidos (Jewitt, 2014), ya que, en algunas ocasiones, una combinación de los tres enfoques puede

proporcionar una comprensión más profunda de las complejidades de la comunicación y ampliará las posibilidades del análisis multimodal (Bernad-Mechó, 2021).

### 1.3 Modos o variables del análisis multimodal

En los últimos años, con el crecimiento de las metodologías multimodales, se ha prestado atención a los elementos no lingüísticos del discurso, como, por ejemplo, los modos no verbales (Bernad-Mechó y Fortanet-Gómez, 2019; Morell, 2018; Thesen, 2016), el uso de recursos visuales (Gunel et al., 2006), o la utilización de pizarras (Fox y Artemeva, 2012) y de recursos multimedia (Tan et al., 2016) en las clases universitarias. Otros estudios han investigado la relación entre elementos lingüísticos y no lingüísticos. Por ejemplo, mientras Fortanet-Gómez y Ruiz-Madrid (2014) exploran el uso multimodal de preguntas en clases impartidas por profesores invitados para facilitar la comprensión en hablantes no nativos de inglés, Crawford Camiciottoli analiza cinco clases universitarias de humanidades con énfasis en cómo se realizan las explicaciones verbales y su relación con el énfasis prosódico, la mirada y los gestos (2015), así como las estrategias de interacción entre conferenciantes y estudiantes y su repercusión en la comprensión (2016). Así mismo, Bernad-Mechó (2017) explora la relación entre el metadiscurso organizativo y los elementos corporales multimodales. En resumen, a través de diferentes enfoques metodológicos, la investigación multimodal brinda nuevas perspectivas sobre las posibilidades de realizar análisis multimodales del discurso, siempre con el objetivo común de describir las estrategias que facilitan la comprensión por parte de los receptores.

Concretando los elementos que frecuentemente constituyen las variables de los estudios de multimodalidad, vamos a introducir dos aportaciones que definen variables para el análisis del lenguaje oral. Por una parte, se pueden aplicar clasificaciones o taxonomías para la parte verbal, que dependerán de aquello que se quiera analizar. Una de las taxonomías más utilizadas es la estructura de movimientos (*moves*) y pasos (*steps*) diseñada por Swales para el análisis de géneros (1990). Por otra parte, aunque existen otras taxonomías utilizadas en estudios multimodales (Kendon, 2004; Valeiras-Jurado y Bernad-Mechó, 2022), en este caso presentamos la de Cestero-Mancera (2006), quien distingue cuatro sistemas de comunicación no verbal: el paralenguaje, la quinésica, proxémica y la cronémica.

El sistema paralingüístico está formado por las cualidades y los modificadores fónicos (tono, timbre, cantidad e intensidad), los indicadores sonoros de reacciones fisiológicas y emocionales (risa, suspiro, grito, tos, llanto castaño de dientes, entre otros), los elementos cuasi-léxicos (interjecciones como “¡Ah!”, “¡Uy!”; onomatopeyas como “Mua”, “Pii-Pii”; y otros sonidos con valor funcional) y las pausas y silencios: Todos estos elementos pueden comunicar por sí mismos o matizar el significados del discurso verbal.

El sistema quinésico abarca movimientos y posturas corporales e incluye también la mirada o el contacto corporal. Cestero-Mancera (2006) reconoce tres categorías básicas de signos quinésicos:

- gestos o movimientos faciales (realizados fundamentalmente con los ojos, las cejas, el entrecejo, la nariz, y la boca) y corporales (realizados con la cabeza, los hombros, los brazos, los dedos, las piernas y los pies);



- posturas, es decir, posiciones estáticas que comunican, de manera activa o pasiva. Por ejemplo, estar de pie, con los brazos arqueados y las manos apoyadas en la cintura (“en jarras”) o sentados, con las piernas ligeramente dobladas y las manos unidas en el regazo puede añadir el significado de desafío, en el primer caso, y de humildad, en el segundo.

El sistema proxémico se refiere a la distribución del espacio y está marcado cultural y socialmente como la intrusión o la violación de la territorialidad y la privacidad. La proxémica interaccional tiene especial importancia, ya que establece la distancia a la que las personas de una comunidad realizan acciones interactivas o comunicativas (consolar, conversar, mantener entrevistas laborales, impartir clases o dar conferencias). Además, aproximarse o distanciarse de una persona al mismo tiempo en que se le habla puede reforzar expresiones como “Venga, ¡vamos juntas!”.

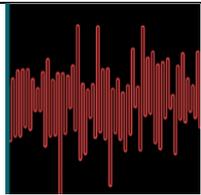
Por último, el sistema cronémico, es decir, relativo al tiempo, también puede complementar el significado de otros elementos verbales y no verbales. Por ejemplo, la duración de entrevistas de trabajo, reuniones y negociaciones comerciales o actividades sociales como puede ser una comida o una cena. También se considera parte de este sistema la mayor o menor duración de los sonidos de algunas palabras, de algún gesto o de las pausas, o la velocidad en la emisión de un enunciado como "no puedes hacerlo así, tienes que hacerlo...", que puede disminuir su efecto crítico. Este sistema se medirá de manera global, a diferencia de los anteriores que responden a acciones más puntuales.

La Tabla 1 resume las principales variables presentadas por Cestero-Mancera (2006), aunque su selección dependerá del tipo de acción de comunicación elegida como corpus (acción social como una conversación entre amigos, clase en que se analiza el lenguaje del profesor, entrevista de trabajo, etc.):

Por otra parte, Valeiras-Jurado y Bernad-Mechó (2022) incorporan una serie de modos a sus análisis de videos ya editados y publicados y que pueden utilizarse como materiales en las clases de segundas lenguas. Se trata de modos filmicos, ya utilizados en otros análisis (Desilla, 2012; Muñoz-Morcillo et al., 2016), y que incluyen tipo de toma, ángulo, puesta en escena, cortes, música, avisos visuales, efectos de sonido y efectos de video. Valeiras-Jurado y Bernad-Mechó (2022) consideran estos modos esenciales para poder entender la comunicación en contextos digitales orales. Por ejemplo, en el caso de las tomas, se distinguen planos cortos, planos medios, plano cowboy, planos medios completos, planos completos o planos largos (Bowen, 2018). El ángulo de la cámara puede ser frontal, lateral, picado, contrapicado, o cenital, entre otros. En la puesta en escena, se puede describir los elementos del fondo: logos, cortinas, mobiliario, paisaje, efecto croma, etc.). La música también se puede describir e indicar cuándo aparece y qué puede aportar, al igual que los cortes, y los efectos de sonido y de video. La Tabla 2 resume estos modos filmicos.



**Tabla 1.** Modos para utilizar como variables en un estudio multimodal  
(las fotografías se han extraído de Internet y tienen licencia Creative Commons)

MODOS		TIPOS	EJEMPLO	REPRESENTACIÓN
Paralenguaje		Reacciones fónicas	Tono Timbre Cantidad Intensidad	
		Reacciones fisiológicas	Tos Risa Suspiro Llanto	
		Elementos quasi-léxicos	“¡Ah!”, “¡Uy!” “Pii-Pii”	
		Pausas y silencios	Pausas entre palabras o entre frases Silencios prolongados (por ejemplo, el gesto de la foto de la fila siguiente iría acompañada de silencio)	
Quinésica		Movimientos faciales	Mirada directa, hacia arriba o hacia abajo. Entrecejo arrugado Labios prominentes	
		Movimientos corporales	Signos con las manos, con los hombros, brazos o con los dedos	
		Posturas	Posiciones estáticas, como tener los brazos “en jarras” o cruzados, entre otros.	
Proxémica		Distancia	Distancia entre personas en un diálogo Distanciamiento acompañando una expresión como “Nos vemos pronto”	

**Tabla 2.** Modos filmicos (las fotografías se han extraído de Internet y tienen licencia Creative Commons)

MODO	TIPO	EJEMPLO	REPRESENTACIÓN
Tipo de toma	Plano corto, medio, de cowboy, etc.	Plano de cowboy	
Ángulo	Frontal, lateral, picado, contrapicado, cenital, etc.	Ángulo frontal	
Puesta en escena	Logos, cortinas, muebles, paisaje, efecto croma, etc.	Fondo mostrando una ciudad. Los personajes parecen estar en un medio de transporte.	
Otros	Música, cortes, efectos de sonido y de vídeo	Efecto de vídeo: incorporación de dos imágenes superpuestas	

Cualquiera que sea el enfoque seleccionado y los modos que se elijan, la mayoría de los investigadores e investigadoras en multimodalidad recomiendan comparar la interpretación de los recursos semióticos proporcionada por diferentes especialistas e, incluso, entrevistar a los y las hablantes, cuando sea posible, para contrastar estas interpretaciones que ellos y ellas mismas les asignan.

## 2. Herramientas digitales para anotar el análisis multimodal: de los antecedentes a GRAPE-MARS

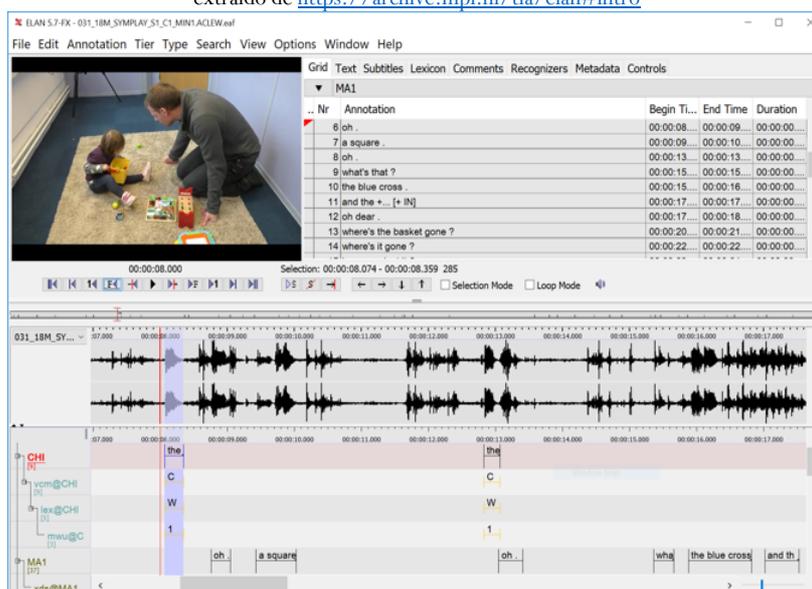
### 2.1. Antecedentes: ELAN y Multimodal Analysis Video

El grupo de investigación GRAPE cuenta con una extensa trayectoria en el análisis multimodal del discurso (cf. <https://www.grape.uji.es>). De hecho, en el marco del grupo se han llevado a cabo proyectos que analizan los recursos semióticos (o modos) empleados en el discurso académico. En muchos de estos proyectos se utilizaron programas ya existentes (principalmente ELAN y Multimodal Analysis Video –MMAV). Aunque existen otros programas, como Atlas.ti (<https://atlasti.com/es>), son estos dos programas los que hemos utilizado para realizar

análisis multimodales para la enseñanza y aprendizaje del inglés como L2 y constituyen el punto de partida para la elaboración de nuestro programa GRAPE-MARS.

ELAN (EUDICO Linguistic Annotator) (<https://archive.mpi.nl/tla/elan#intro>) (ELAN, 2023) es una herramienta de anotación que permite la creación, edición, visualización y búsqueda de anotaciones de datos de video y audio. Se desarrolló en el Instituto Max Planck de Psicolingüística, en Nijmegen, Países Bajos, con el objetivo de ser un punto de referencia tecnológica en la anotación y explotación de las grabaciones multimedia. Este programa está diseñado para analizar el discurso, el lenguaje de signos y los gestos, pero lo pueden utilizar todos aquellos que deseen anotar, analizar y documentar corpus con datos de video y audio (Figura 2). Existen muchos estudios de análisis multimodal realizados con ELAN (Bonsignori, 2019; Ruiz-Madrid y Valeiras-Jurado, 2020; Morell et al., 2022), sin embargo, su interfaz y utilización no son sencillos y requieren un entrenamiento exhaustivo para ser capaz de realizar incluso análisis simples. Por ejemplo, un problema ampliamente detectado es la dificultad para transcribir los videos (el programa no aporta esta tarea) y de sincronizar la transcripción con el video una vez obtenida e incorporada.

Figura 2. Ejemplo de análisis usando ELAN del ACLEW project  
extraído de <https://archive.mpi.nl/tla/elan#intro>



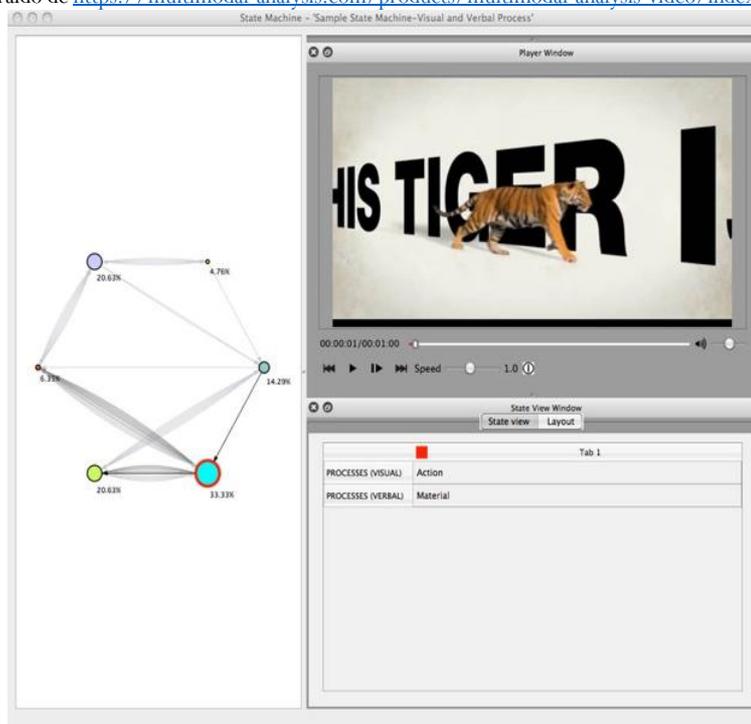
El segundo de los programas que ha servido como antecedente de la creación del GRAPE-MARS es *Multimodal Analysis Video* –MMAV (<https://multimodal-analysis.com/products/multimodal-analysis-video/index.html>) (O'Halloran et al., 2015). Se trata de un programa comercial, es decir, no es gratuito, para la anotación y análisis de videos (Figura 3). Este programa, a diferencia del ELAN, está diseñado para el análisis multimodal del discurso, teniendo en cuenta los recursos lingüísticos, de imagen y de audio. Para facilitar su uso, dispone de ejemplos de análisis

y formatos preparados para su uso. Además, este programa muestra los resultados del análisis cuantitativo de manera gráfica (Figura 4).

**Figura 3.** Ejemplo de la página de análisis usando MMAV  
extraído de <https://multimodal-analysis.com/products/multimodal-analysis-video/index.html>



**Figura 4.** Ejemplo de la página de muestra gráfica del análisis con MMAV  
extraído de <https://multimodal-analysis.com/products/multimodal-analysis-video/index.html>



Este programa resultó útil durante un tiempo para la realización de análisis multimodales en la investigación del grupo GRAPE. Presenta una interfaz más sencilla que el ELAN y está más adaptado específicamente a los estudios multimodales del discurso utilizando video. El MMAV se ha utilizado en un buen número de trabajos sobre multimodalidad en el discurso (Thwaite, 2015; Xia y Hafner, 2021). Sin embargo, en los últimos años se ha dejado de actualizar, lo que crea frecuentes errores al adoptar nuevas versiones de los sistemas operativos. Además, este programa tampoco facilitaba la tarea de conseguir la transcripción y sincronizarla con el vídeo.

## 2.2 GRAPE-MARS: Descripción

Después de una larga trayectoria en el análisis multimodal del discurso oral en inglés, y de haber utilizado y comprobado las limitaciones de los programas anteriormente presentados, el grupo GRAPE se planteó el diseño de un nuevo programa que significara un avance en la optimización del proceso de anotación, lo cual sería un hito en el campo que nos ocupa puesto que daría respuesta a las limitaciones de programas anteriores y facilitaría la realización de estudios cuantitativos con corpus de discurso oral. Por este motivo, junto con el análisis de los programas ya existentes, el grupo GRAPE realizó un análisis de las necesidades actuales mediante un *focus group* con investigadores e investigadoras expertas en el campo. En esta reunión se pusieron de manifiesto una serie de posibles mejoras que podrían implementarse en futuros desarrollos y que se añadirían a las mejoras derivadas de las limitaciones detectadas en los programas anteriores. Entre estas mejoras podemos destacar las siguientes:

- la transcripción automática a través de inteligencia artificial;
- la posibilidad de modificar fácilmente el tamaño de las anotaciones para ajustarlas con precisión a la línea de tiempo;
- la posibilidad de duplicar, copiar y pegar, unir y arrastrar anotaciones de forma rápida y sencilla;
- la posibilidad de duplicar capas de anotación;
- la posibilidad de trabajar con varios videos a la vez;
- la posibilidad de obtener resultados relativos a videos completos o a fragmentos de videos;
- la posibilidad de analizar cuantitativamente el uso de recursos semióticos (por ejemplo, porcentaje, número de veces que aparece) y representar resultados de forma visual con diferentes gráficos;
- y la posibilidad de analizar coocurrencias y patrones de uso.

El GRAPE-MARS (Ruiz-Madrid, et al., 2023) fue desarrollado por el Group for Research on Academic and Professional English (GRAPE) de la Univeersitat Jaume I, con la financiación otorgada por la Generalitat Valenciana (código de la ayuda CIAICO/2021/069). El prototipo se registró como invención en la Oficina de Cooperación a la Investigación y la Tecnología como GRAPE-MARS (Multimodal Analysis Research Software), y consta como Iniciativa de Código Abierto (<https://opensource.org/>). El código abierto aplicado a programas informáticos indica que se puede usar, escribir, modificar y distribuir libremente. Se comparte el código fuente con la comunidad para favorecer los avances y dar la oportunidad a otros investigadores de poder mejorarlo.



El GRAPE-MARS se describe como una herramienta de anotación para el análisis de los recursos semióticos empleados en distintos ámbitos comunicativos (por ejemplo, análisis multimodal). Aunque la interfaz está en inglés en esta primera versión, se puede utilizar en cualquier lengua.

Este programa permite el etiquetado (o anotación) de información gestual, corporal y/o verbal (entre otras) en archivos de vídeo. A partir de esta información etiquetada, se proporcionan herramientas para analizar de manera dinámica y eficiente todos los datos generados. La aplicación cuenta con los siguientes módulos:

- Biblioteca de vídeos, desde donde se selecciona el vídeo o vídeos a analizar.
- Reproductor de vídeos, que permite reproducir hasta dos vídeos simultáneamente (por ejemplo, en el caso de una entrevista, o un docente y los discentes) (Figura 5).
- Transcripciones, donde se muestra una transcripción de grabación generada automáticamente, con opción a editarla (Figura 5).
- Forma de onda, es decir, un gráfico de representación del audio que permite identificar los aspectos paralingüísticos con los otros modos (Figura 5).
- Capas de anotación multimodal (Figura 5).
- Biblioteca en la que crear taxonomías, que se puede editar en cualquier momento y exportar e importar a otros proyectos (Figura 6).
- Analítica: Resultados cuantitativos de los modos (Figura 7)

Figura 5. Pantalla de análisis con la ventana de transcripciones abierta

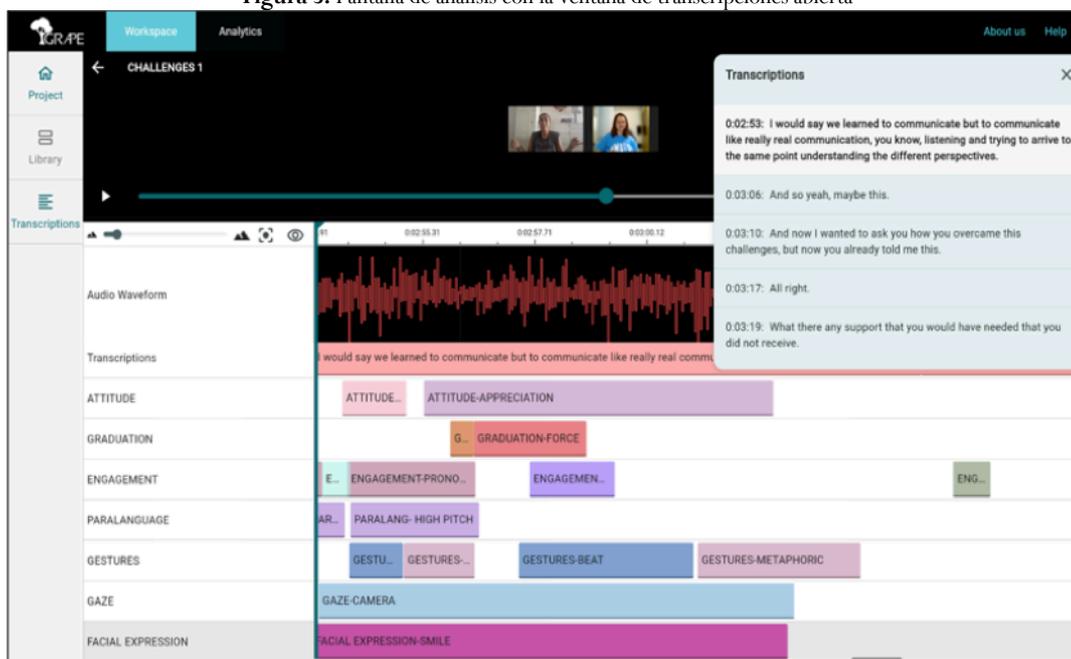


Figura 6. Imagen de la biblioteca

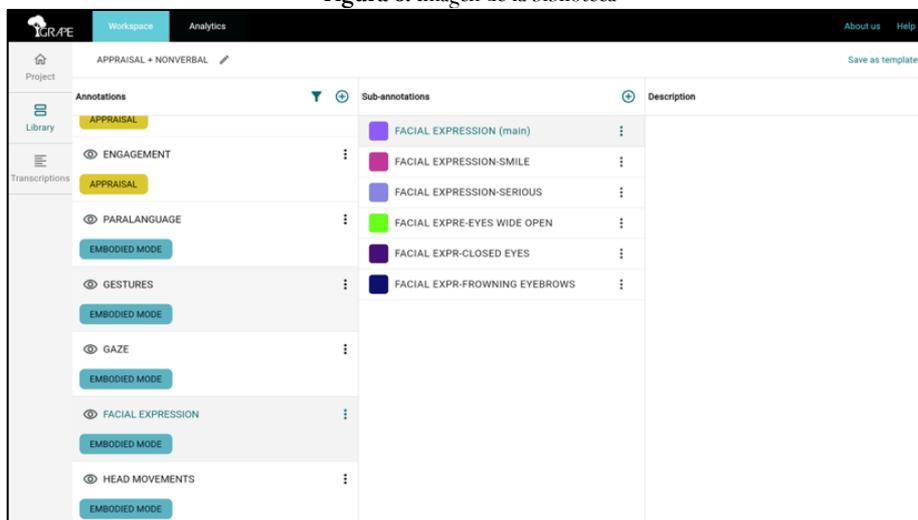
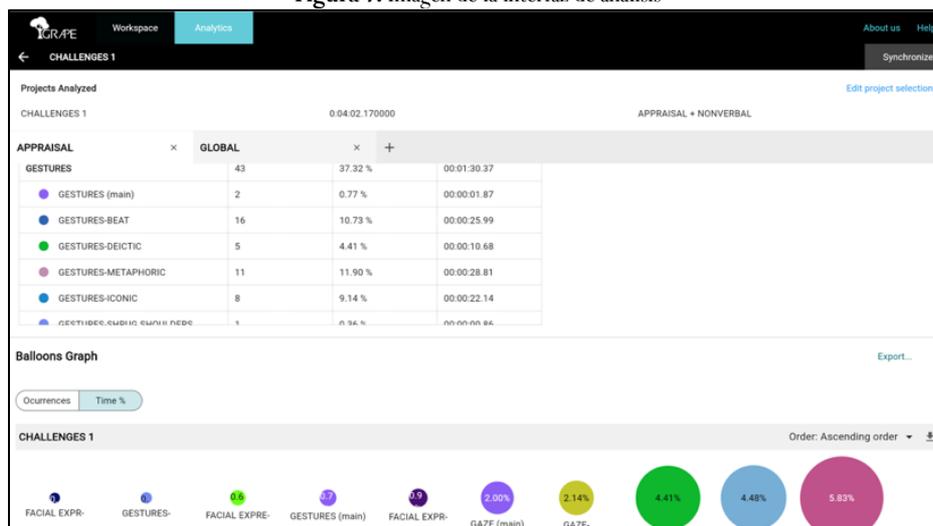


Figura 7. Imagen de la interfaz de análisis



### 2.3 GRAPE-MARS: Utilización

A continuación, se presenta una introducción práctica al programa GRAPE-MARS. En primer lugar, invitamos al lector a localizar el programa en <https://www.grape.uji.es/mars/>. Es importante remarcar que el programa funciona solo con el navegador Google Chrome. Antes de comenzar, se deberá crear una carpeta con el nombre adecuado al proyecto que se quiere realizar. En esta carpeta se guardará el análisis automáticamente.

En la primera pantalla del programa se solicita al investigador que seleccione un directorio (*Select a directory*). Se trata de la carpeta que se acaba de crear. Hay que seleccionarla y, posteriormente, permitir al sistema que edite los archivos. A continuación, hay que hacer un clic en *New Project* y proporcionar un título a este proyecto. Puede haber



más de un proyecto en la misma carpeta. También hay que proporcionar un nombre a la biblioteca, es decir, a los modos o recursos semióticos (ver apartado 1.3) que vamos a introducir en nuestro análisis. También se puede seleccionar una biblioteca creada previamente.

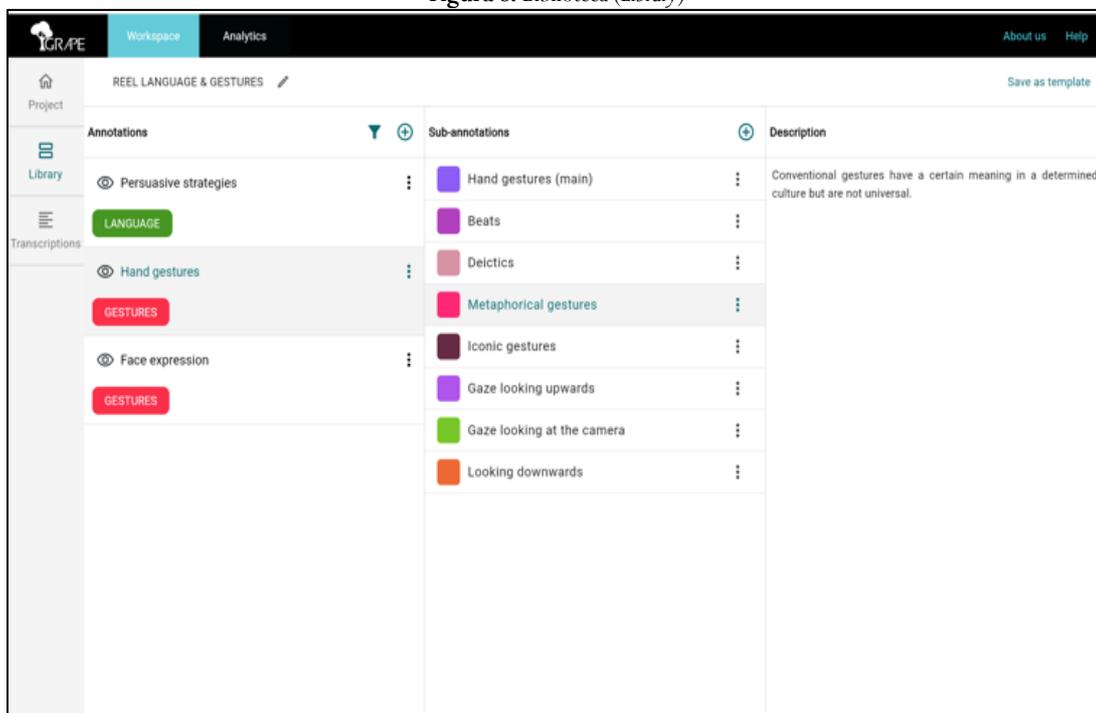
Para facilitar la comprensión del lector, vamos a describir un ejemplo de proyecto, que se podría utilizar en la investigación sobre el inglés como L2, al que hemos denominado *Reel: Why am I like this?*. Los *reels* de Instagram son videos muy cortos, frecuentemente en inglés, en los que los hablantes son muy expresivos y utilizan muchos recursos semióticos. El *reel* utilizado en esta demostración se denomina como el proyecto, *Why am I like this* (50:80 segundos) <https://www.instagram.com/reel/C3sNuFPA9aK/?igsh=MXMyMzUxb2VuYmt4dA%3D%3D>. A la biblioteca se le dio el nombre de *Language and Gestures*.

El paso siguiente consiste en subir el video. Se puede subir uno o dos vídeos para casos en los que es importante analizar a dos hablantes (por ejemplo, una entrevista) o a un hablante y su interlocutor (por ejemplo, en una clase en la que se quiere analizar la efectividad del uso de recursos semióticos para la enseñanza de segundas lenguas). El procesado del video puede tardar unos minutos y dependerá de la longitud del video y del número de personas procesando vídeos al mismo tiempo. Se recomienda seleccionar videos cortos (1-3 minutos) o fragmentar videos más largos.

Una vez se ha procesado el video, vemos ya el proyecto y pulsamos encima de la imagen. Automáticamente el programa nos lleva a la página de trabajo, *Workspace*, y vemos el video, la línea de sonido y la transcripción sincronizada con el video. Es mejor revisar la transcripción, ya que se pueden haber producido errores. Se puede ampliar el gráfico de sonido moviendo el cursor encima de la línea (ver Figura 5). Es el momento de introducir los modos o recursos semióticos, por lo que nos dirigimos a “*Library*” en la columna izquierda de la página. Creamos, pues, anotaciones pulsando en el signo + y les damos nombre. En nuestro ejemplo, la primera anotación de la librería es *Persuasive Language*. También podemos crear *Categories*, que permiten agrupar modos. En este caso, se ha creado y asignado la categoría *Language*. A continuación, se pueden crear sub-anotaciones para las anotaciones, diseñando así una escala jerárquica en la propia taxonomía que responda a diferentes tipos de categorías y modos. Automáticamente, por defecto se crea una sub-anotación con el mismo nombre que la anotación, para casos no contemplados en las sub-anotaciones. Se puede quitar, si no interesa. En este caso, se han creado 4 sub-anotaciones *Questions*, *Use of pronoun ‘you’*, *Use of pronoun ‘I’* y *Use of ‘hedging’*. Las anotaciones y sub-anotaciones dependen de lo que se quiera analizar y del video seleccionado. Se han añadido dos anotaciones más en la biblioteca dentro de la categoría de gestos (*Gestures*): *Hand gestures* y *Face expressions*. Para cada anotación se han añadido algunas sub-anotaciones y se han descrito aquellas que podían resultar menos claras. Esto es una demostración; una investigación real requeriría posiblemente un mayor número de categoría y anotaciones, como la paralingüística, incluyendo la entonación, el énfasis en algunas palabras o los silencios. En la Tabla 3 se puede ver un resumen de las anotaciones de este proyecto y en la Figura 8 se puede ver la pantalla con las anotaciones.

Tabla 3. Taxonomía usada en la biblioteca

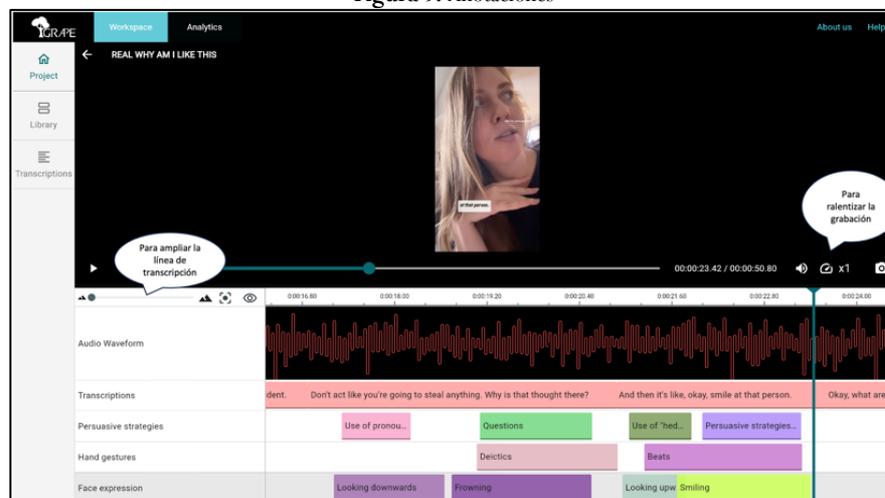
CATEGORIAS	ANOTACIONES	SUBANOTACIONES	DESCRIPCIÓN
LANGUAGE	Persuasive strategies	Persuasive strategies  Questions  Use of pronoun “you”  Use of pronoun “I”  Use of “hedging”	   This pronoun can be addressed to the audience or to an indetermined person.  Talking about herself.  "Seem to" or "like" are examples of "hedging"
GESTURES	Hand gestures	Hand gestures	
		Beats	Movements to mark the rhythm of the speech.
		Deictics	Movements to point at something.
		Metaphorical gestures	Conventional gestures that represent something abstract and tend to be associated to a culture.
		Iconic gestures	They represent something specific and are used to accompany speech to describe what is said.
	Facial gestures	Gaze looking upwards or downwards Gaze looking at the camera,	
	Facial expression	Face expression Smiling Frowning Laughing Opening eyes wide	

Figura 8. Biblioteca (*Library*)

Una vez creada la Biblioteca (*Library*), hay que volver al Proyecto (*Project* en la columna de la izquierda), donde se ver3 el listado detallado de las variables integradas en la taxonomía (anotaciones y sub-anotaciones) en la columna de la izquierda y las capas del análisis debajo del video, la línea de sonido y la transcripción. Es el momento de iniciar el trabajo de anotaci3n.

Se empieza a visionar la grabaci3n. Se puede ralentizar mediante el s3mbolo del reloj x1, subir y bajar el sonido y tomar capturas de fotos del video. Tambi3n se puede agrandar la imagen moviendo hacia abajo la parte inferior. Para anotar, es necesario buscar la parte del video en la que se quiere localizar la anotaci3n y hacer clic en la capa correspondiente a la anotaci3n. Aparecer3 entonces un men3 desplegable para elegir la sub-anotaci3n m3s apropiada. Se puede alargar el espacio de la anotaci3n, moverlo o hacerlo m3s peque1o. Tambi3n se puede cambiar el orden de las capas hacia arriba o hacia abajo arrastr3ndolas. La Figura 9 muestra el video ya anotado.

Figura 9. Anotaciones



Una vez anotado el vídeo, se puede pasar a visualizar los resultados en la analítica (*Analytics* en la parte superior). Se debe crear, dar nombre a un Nuevo análisis y elegir un proyecto (pulsar “Add”). A continuación, se pueden crear diferentes subanálisis. Una vez dentro del subanálisis, se debe elegir el tipo (*General* o *Co-occurrence*, este último para hacer comparativas entre las sub- anotaciones), el rango del análisis (pueden ser solo algunas sub- anotaciones o anotaciones) y el nivel de análisis (categoría, anotación o sub- anotación). En el ejemplo, se ha elegido el tipo general, todo el proyecto y las sub- anotaciones. Se obtiene una tabla con todas las veces que aparecen las anotaciones y sub- anotaciones, y el tiempo que duran en porcentaje sobre el tiempo total y en duración temporal (ver Figura 10). Los datos de la tabla se pueden exportar en formato Excel para ser utilizados con programas más específicos de análisis de datos estadísticos si se desea.

Figura 10. Analítica: tabla

Annotation name	Occurrences	Time %	Total duration
<b>Persuasive strategies</b>	24	57.41 %	00:00:29.15
Persuasive strategies (main)	1	2.53 %	00:00:01.28
Questions	3	11.79 %	00:00:05.98
Use of pronoun "you"	10	21.48 %	00:00:10.90
Use of pronoun "I"	6	13.22 %	00:00:06.71
Use of "hedging"	4	8.38 %	00:00:04.25
<b>Hand gestures</b>	10	39.05 %	00:00:19.83
Hand gestures (main)	1	4.07 %	00:00:02.06
Beats	2	6.20 %	00:00:03.15
Deictics	3	11.24 %	00:00:05.70
Metaphorical gestures	3	15.35 %	00:00:07.79

En segundo lugar, se puede visualizar la representación de los datos en globos de distintos colores, en un gráfico de radar, en un gráfico circular, y en una línea temporal. Todos estos gráficos se pueden exportar en formato png (ver Figuras 11, 12 y 13).

Figura 11. Gráfico de globos

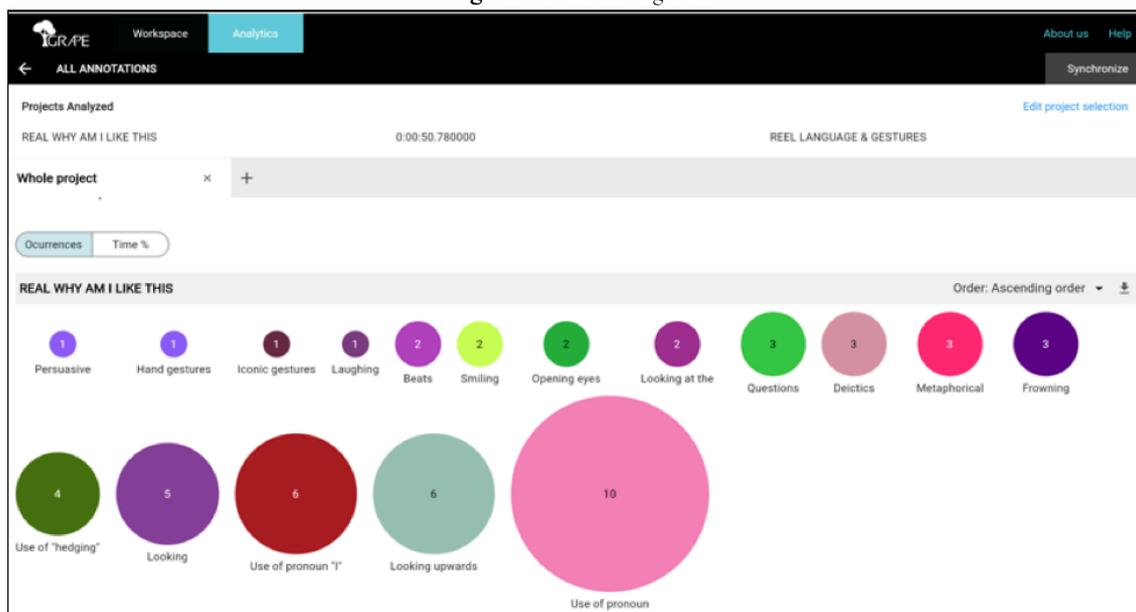


Figura 12. Gráfico de radar

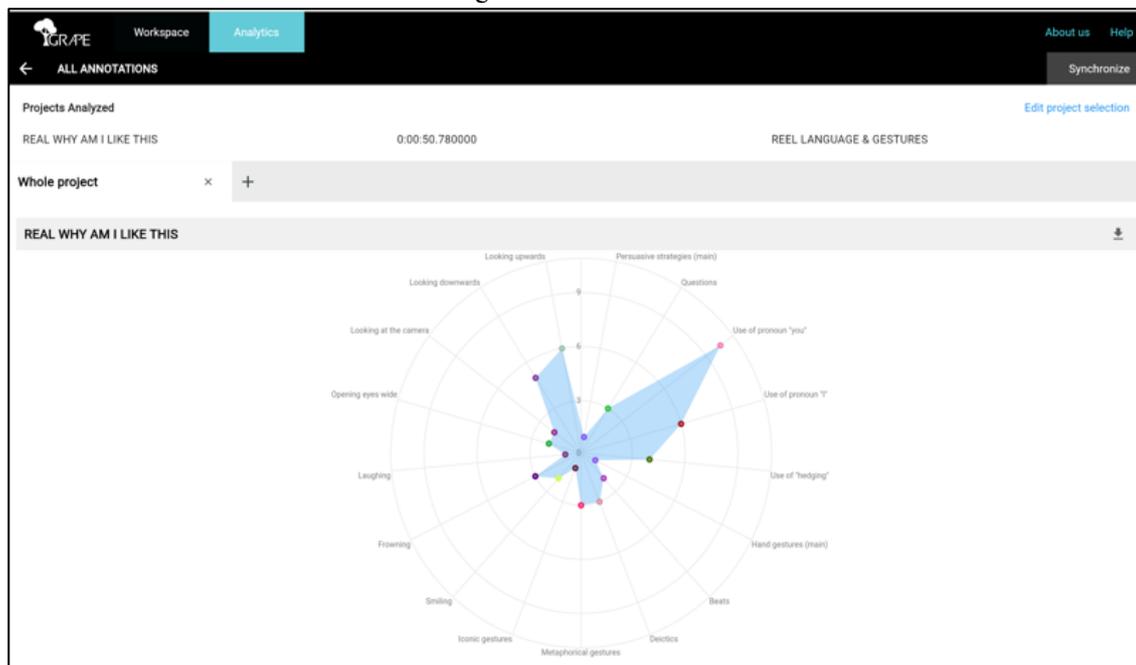
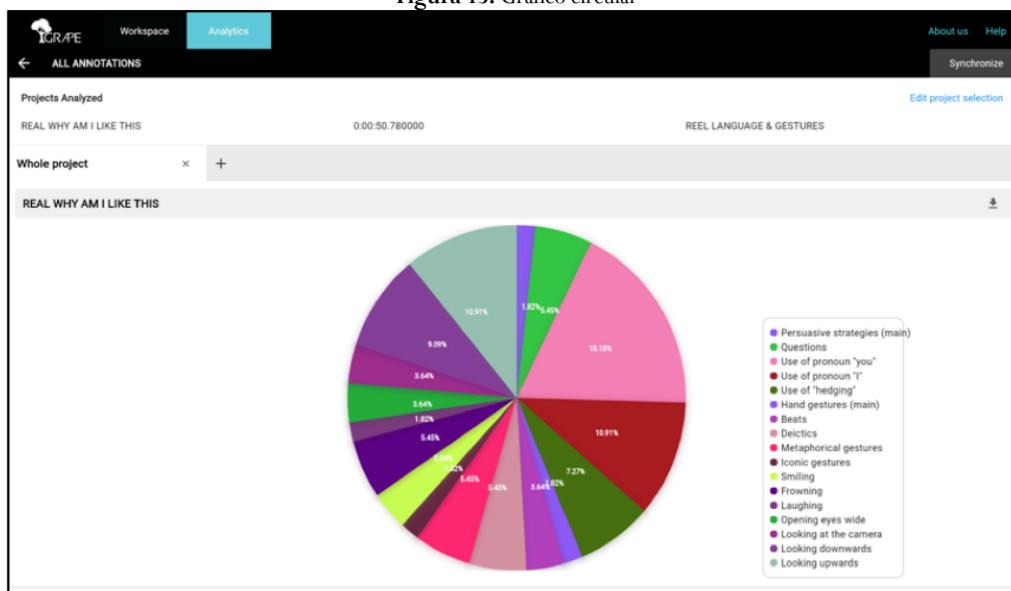


Figura 13. Gráfico circular



Todos los archivos quedarán guardados en la carpeta local que se ha creado y estos siempre pueden ser editados y modificados. Para compartir el análisis, es mejor comprimir la carpeta en zip y enviarla en su totalidad. El receptor también podrá editar los archivos y guardarlos en su ordenador.

### 3. Ejemplos de estudios realizados utilizando el GRAPE-MARS en contextos de investigación sobre L2

GRAPE-MARS es un programa de reciente creación y todavía son pocos los análisis elaborados con esta herramienta. Entre ellos, podemos destacar “New tools for Multimodal Discourse Analysis: Analyzing spoken digital genres with GRAPE-MARS” (Valeiras-Jurado, et al. 2024), en el que se presenta la herramienta y se ejemplifica mediante un fragmento de la serie “The Good Doctor”, como material para la enseñanza del inglés como segunda lengua en el grado de Medicina. Esta serie reproduce ejemplos de diálogo médico-paciente en los que el uso de diversos recursos semióticos es muy relevante.

Se han realizado dos estudios más, utilizando GRAPE-MARS, centrados en la multimodalidad en el aprendizaje de la lengua inglesa en educación secundaria, en ambos casos mediante la introducción de *Instagram reels*: “New pedagogical artefacts for new learning needs: Instagram reels and multimodal literacy in pre-tertiary education” (Mena-Alba, 2024), y “Becoming multimodally literate in the digital era: Instagram reels as pre-tertiary didactic materials” (Serrano-Hernández, 2024). En estos dos estudios se optó por introducir *reels* como alternativa a actividades de comprensión y expresión oral, ya que estos géneros digitales ofrecen una alta densidad de recursos no verbales acompañando al lenguaje y, al mismo tiempo. Los estudiantes conocen bien este tipo de género, pero lo asocian con el entretenimiento más que al aprendizaje de lenguas. Su uso en el aula puede cambiar su visión y aumentar su motivación.

Por último, podemos mencionar el estudio de Cvetkovic y Fortanet-Gómez (2024), en el que se investiga a partir de dos entrevistas a profesoras de primaria tras una experiencia de telecolaboración en la enseñanza del inglés como L2 entre un colegio en España y otro en Alemania. El objetivo es visualizar las diferencias culturales entre ambas profesoras y que coinciden con el mayor reto con el que dicen haberse encontrado, la comunicación intercultural. El estudio se realiza desde tres perspectivas teóricas: el análisis temático, la Teoría de la Apreciación y el análisis multimodal. Para el análisis de las dos últimas se utiliza el GRAPE-MARS. Por último, se ha publicado una reseña del programa (Liu y Xiang, 2024), con comentarios muy positivos sobre su uso.

El GRAPE-MARS es un programa nuevo y su uso está a disposición de cualquier investigador o investigadora.

#### 4. Conclusiones

El objetivo de este artículo era introducir al lector en los estudios de multimodalidad en contextos de L2 y, en concreto, en la utilización de la herramienta GRAPE-MARS desarrollada por miembros del grupo de investigación GRAPE de la Universidad X. En este tipo de estudios se debe elegir el marco teórico para definir después el objetivo, el corpus y las variables a analizar. En este artículo se han presentado una variedad de enfoques teóricos sobre multimodalidad, así como los modos o variables de estudio.

Respecto a las herramientas digitales para la anotación multimodal, se han descrito programas que han servido de antecedente a GRAPE-MARS y, a continuación, se ha presentado de manera práctica este programa. GRAPE-MARS es un programa fácil de utilizar para la anotación de los distintos recursos semióticos de una grabación de video, de manera que se haga evidente el carácter multimodal de la comunicación oral y de los géneros digitales que se están desarrollando hoy en día.

En la investigación multimodal los recursos semióticos habitualmente se complementan, aunque en ocasiones puede haber contradicciones, o puede encontrarse algún recurso aislado. Es la tarea del investigador o la investigadora no solo identificar estos recursos semióticos, sino también interpretarlos. El análisis cuantitativo del uso de los recursos semióticos solo será significativo si va acompañado de un análisis cualitativo, ya que más recursos no siempre significan mejor comunicación. Por otra parte, la interpretación de los datos es esencial en situaciones de interacción en las que es necesario prestar atención a los elementos no verbales en su contexto para poder identificar, por ejemplo, la ironía.

Este tipo de análisis puede resultar muy útil en la investigación sobre la enseñanza y el aprendizaje de segundas lenguas, para demostrar el uso conjunto de distintos modos o recursos semióticos para facilitar la comunicación. Hoy en día cada vez son más asequibles los recursos digitales y concretamente los vídeos se están incorporando en la enseñanza y aprendizaje de lenguas para la comprensión de lenguas y como modelos de producción. Siendo así, el aprendizaje de las habilidades de comprensión y expresión oral no pueden ya basarse en actividades exclusivamente



verbales, sino que es necesario atender a la multimodalidad de los géneros digitales utilizando herramientas como la que proponemos.

Una herramienta como GRAPE-MARS solo puede llegar a su pleno potencial con la colaboración de otros investigadores que la utilicen. Sabemos que, en este momento, se pueden detectar limitaciones, como puede ser el tiempo del procesamiento inicial del video mediante la IA, y esperamos poder mejorar este aspecto y otros que podamos identificar en el periodo más corto posible de tiempo.<sup>1</sup>

## Referencias

- Alfonso Lozano, Raúl; Giralte Lorenz, Marta (2014) Comunicación multimodal y aprendizaje de la lengua oral en E/LE. CAUCE. *Revista Internacional de Filología, Comunicación y sus Didácticas*, 37, 117-133. [https://cvc.cervantes.es/literatura/cauce/pdf/cauce36-37/cauce\\_36-37\\_007.pdf](https://cvc.cervantes.es/literatura/cauce/pdf/cauce36-37/cauce_36-37_007.pdf)
- Beltrán-Palanques, Vicent (2024). Assessing video game narratives: Implications for the assessment of multimodal literacy in ESP. *Assessing Writing*, 60,100809. <https://doi.org/10.1016/j.asw.2024.100809>
- Beltrán-Palanques, Vicent; Bernad-Mechó, Edgar (2024). Preparing doctoral students for conference presentations. Teaching multimodal academic literacy and the role of organizational metadiscourse. En Beltrán-Palanques, Vicent; Bernad-Mechó, Edgar (Eds.). *Current trends in EMI and Multimodality in Higher Education* (pp. 184-201). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003205517-11/preparing-doctoral-students-conference-presentations-vicent-beltr%C3%A1n-palanques-edgar-bernad-mech%C3%B3?context=ubx&refId=7e551001-06f0-44d2-b0e9-6d248ad3660b>
- Bernad-Mechó, Edgar (2017). Metadiscourse and Topic Introductions in an Academic Lecture: A Multimodal Insight. *Multimodal Communication*, 6(1) <https://doi.org/10.1515/mc-2016-0030>
- Bernad-Mechó, Edgar (2021). Combining Multimodal Techniques to Approach the Study of Academic Lectures: A Methodological Reflection. *Atlantis*, 43(1), 178-198. <https://doi.org/10.28914/Atlantis-2021-43.1.10>
- Bernad-Mechó, Edgar; Fortanet-Gómez, Inmaculada (2019). Organizational metadiscourse across lecturing styles. In Sancho Guinda, Carmen (Ed.) *Engagement beyond language* (pp. 321–340). John Benjamins. <https://doi.org/10.1075/pbns.301.17ber>
- Barros García, Benamí (2021) Representar visualmente los resultados de la investigación sobre el español LE/L2. En Cruz Piñol, Mar (Ed.), *e-Research y español LE/L2. Investigar en la era digital* (pp.225-245). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780429433528-13/representar-visualmente-los-resultados-de-la-investigaci%C3%B3n-sobre-el-espa%C3%B1ol-le-l2-benam%C3%AD-barros-garc%C3%ADa?context=ubx&refId=e19290f1-eeb7-4d8f-98bd-1d41fb8280ea>
- Bonsignori, Veronica. (2019). A multimodal analysis of spoken medical English in expert-to-expert interaction in TV programmes. *Ibérica*, 37, 115–140. <https://revistaiberica.org/index.php/iberica/article/view/112>
- Cestero-Mancera, Ana María (2006). La comunicación no verbal y el estudio de su incidencia en fenómenos discursivos como la ironía. *ELUA*, 20, 57-77. <http://dx.doi.org/10.14198/ELUA2006.20.03>
- Cocchetta, Francesca (2018). Developing university students' multimodal communicative competence: Field research into multimodal text studies in English. *System*, 77, 19-17. <https://doi.org/10.1016/j.system.2018.01.004>

<sup>1</sup> Animamos a los investigadores y profesores que utilicen GRAPE-MARS a que se comuniquen con nosotros mediante el correo electrónico [grapemars@uji.es](mailto:grapemars@uji.es) para cualquier duda o pregunta que puedan tener y para informarnos sobre sus publicaciones utilizando esta herramienta.



- Crawford Camiciottoli, Belinda (2016). A multimodal analysis of interaction in academic lectures: a case study. En Bonsignori, Veronica; Crawford Camiciottoli, Belinda (Eds.), *Multimodality Across Communicative Settings, Discourse Domains and Genres* (pp. 65-92). Cambridge Scholars Publishing. <https://dokumen.pub/multimodality-across-communicative-settings-discourse-domains-and-genres-1nbsped-9781443862653-9781443811071.html>
- Crawford Camiciottoli, Belinda (2015). Elaborating explanations during OpenCourseWare humanities lectures: The interplay of verbal and nonverbal strategies. En Crawford Camiciottoli, Belinda; Fortanet-Gómez, Inmaculada (Eds.), *Multimodal analysis in academic settings. From research to teaching* (pp. 144-170). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315738758-10/elaborating-explanations-opencourseware-humanities-lectures-belinda-crawford-camiciottoli?context=ubx&refId=770009fe-efa7-421d-85f9-bba492949eda>
- Crawford Camiciottoli, Belinda (2019). Harnessing multimodal literacy for knowledge dissemination in ESP settings. En Bonsignori, Veronica; Cappelli, Gloria; Mattiello, Elisa (Eds.), *Worlds of Words: Complexity, creativity and conventionality in English language, literature and culture* (Vol. I, pp. 47-52), Pisa University Press. <https://hdl.handle.net/11568/1003973>
- Cvetkovic, Anita; Fortanet-Gómez, Inmaculada (2024). Reflecting on a telecollaboration experience: A thematic and multimodal analysis of online interviews. [Presentación] *41 Congreso de AESLA*, Valencia, 17-19 de abril de 2024.
- Đorđević, Jasmina P.; Stamenković, Dejan (2023). Classification of multimodal translation errors in the entertainment industry: a proposal. *The Translator*, 29(3), 265-280. <https://doi.org/10.1080/13556509.2021.2024654>
- ELAN (2023). Version 6.7. Computer software. Nijmegen: Max Planck Institute for Psycholinguistics, *The Language Archive*. <https://archive.mpi.nl/tla/elan>
- Fortanet-Gómez, Inmaculada; Ruiz-Madrid, Noelia (2014). Multimodality for comprehensive communication in the classroom: Questions in guest lectures. *Ibérica*, 28, 203–224. <https://www.revistaiberica.org/index.php/iberica/article/view/245>
- Fox, Janna; Artemeva, Natasha (2012). The cinematic art of teaching university mathematics: Chalk talk as embodied practice. *Multimodal Communication* 1(1), 83-103. <http://doi.org/10.1515/mc-2012-0007>
- Hart, Christopher (2020). Multimodal discourse analysis. En Hart, Christopher (Ed.), *Researching Discourse* (pp. 119-156). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780367815042-9/multimodal-discourse-analysis-christopher-hart?context=ubx&refId=b0976f21-84a3-4d95-bcbf-7fc696f9cf5d>
- Jewitt, Carey (2014). *The Routledge Handbook of Multimodal Analysis*. Routledge. <https://www.routledge.com/The-Routledge-Handbook-of-Multimodal-Analysis/Jewitt/p/book/9781138245198?srsId=AfmBOoq8qyHKvBtIqQCfD2S6tm4ebr0zEcgnYCHuHhXPrGdjC2LVAso>
- Jewitt, Carey; Bezemer, Jeff; O'Halloran, Kay (2016). *Introducing Multimodality*. Routledge. <https://www.routledge.com/Introducing-Multimodality/Jewitt-Bezemer-OHalloran/p/book/9780415639262?srsId=AfmBOoqSD6RfLBH5Xgrjx50o9FSsaD6Z2GUPSeLsa2DsZsi1hFzvhQU>
- Kress, Gunther (2010) *Multimodality: A social semiotic approach to contemporary communication*. Routledge. [https://www.routledge.com/Multimodality-A-Social-Semiotic-Approach-to-Contemporary-Communication/Kress/p/book/9780415320610?srsId=AfmBOopX\\_EWH707pPJ8QIAhA-sgUyCwDUkDgUfSfG0SBIR9ZVp6K-Zen](https://www.routledge.com/Multimodality-A-Social-Semiotic-Approach-to-Contemporary-Communication/Kress/p/book/9780415320610?srsId=AfmBOopX_EWH707pPJ8QIAhA-sgUyCwDUkDgUfSfG0SBIR9ZVp6K-Zen)



- Kress, Gunther; van Leeuwen, Theo (2021) *Reading Images: The Grammar of Visual Design*. Routledge. [https://www.routledge.com/Reading-Images-The-Grammar-of-Visual-Design/Kress-Leeuwen/p/book/9780415672573?srsId=AfmBOorjwCRDNpjEWQ-88Xs\\_8X6Sxf\\_ymhQClkzIt3\\_UinTZjomUiMXQ](https://www.routledge.com/Reading-Images-The-Grammar-of-Visual-Design/Kress-Leeuwen/p/book/9780415672573?srsId=AfmBOorjwCRDNpjEWQ-88Xs_8X6Sxf_ymhQClkzIt3_UinTZjomUiMXQ)
- Liu, Yingxue; Xiang, Yu (2024). Computer-assisted multimodal analysis research: A technical review of GRAPE-MARS. *Multimodality & Society*, 4(2), 239-246. <https://doi.org/10.1177/26349795241251623>
- Machin, David; Caldas-Coulthard, Carmen. R.; Milani, Tomasso M. (2016). Doing critical multimodality in research on gender, language and discourse. *Gender and Language*, 10(3), 301-308. <https://doi.org/10.1558/genl.v10i3.32037>
- Manghi, Dominique H. (2011) La perspectiva multimodal sobre la comunicación. Desafíos y aportes para la enseñanza en el aula. *Revista Electrónica Diálogos Educativos*, 22, 3-14. <https://revistaschilenas.uchile.cl/handle/2250/49263>
- Mena-Alba, Alejandro. (17-19 de abril de 2024). New pedagogical artefacts for new learning needs: Instagram reels and multimodal literacy in pre-tertiary education. [Presentación] *41 Congreso de AESLA*, Valencia, 17-19 de abril de 2024.
- Morell, Teresa; Beltrán-Palanques, Vicent; Norte, Natalia (2022). A multimodal analysis of pair work engagement episodes: Implications for EMI lecturer training, *Journal of English for Academic Purposes*, 58, 101124. <https://doi.org/10.1016/j.jeap.2022.101124>.
- Muñoz-Morcillo, Jesús; Czurda, Klemens; Robertson-von Trotha, Caroline Y. (2016). Typologies of the popular science web video. *Journal of Science Communication*, 15(4), A02. <https://doi.org/10.22323/2.15040202>
- Norris, Sigrid (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. Routledge. <https://www.routledge.com/Analyzing-Multimodal-Interaction-A-Methodological-Framework/Norris/p/book/9780415328562?srsId=AfmBOoqO4tQaTKKIV37suG7GSgzCmYfydPskFN9bbMld7qHXNXulnmdW>
- O'Halloran, Kay. (2012) Análisis del discurso multimodal. *Revista Latinoamericana de Estudios del Discurso*, 12(1), 75-97. <https://periodicos.unb.br/index.php/raled/article/view/33492/27094>
- O'Halloran, Kay L. (2015) *Multimodal Video Analysis* [programa informático] <https://multimodal-analysis.com/products/multimodal-analysis-video/index.html>
- O'Halloran, Kay L.; Tan, Sabine; y E, Marisa K. L. (2015). Multimodal Analysis for Critical Thinking. *Learning, Media and Technology*, 41 (2). <https://doi.org/10.1080/17439884.2016.1101003>
- O'Halloran, Kay L. (2004). *Multimodal Discourse Analysis: Systemic- Functional Perspectives*. Bloomsbury Publishing. <https://www.bloomsbury.com/us/multimodal-discourse-analysis-9781847142573/>
- Parodi, Giovanni; Julio, Cristóbal (2017). No solo existen palabras en los textos escritos: algunas teorías y modelos de comprensión de textos multimodales o multisemióticos. *Investigaciones Sobre Lectura*, 8, 27-48. [http://www.gianniparodi.cl/gianniparodi/site/artic/20180411/asocfile/20180411220328/parodi\\_julio\\_2017\\_comprensio\\_n\\_multimodal.pdf](http://www.gianniparodi.cl/gianniparodi/site/artic/20180411/asocfile/20180411220328/parodi_julio_2017_comprensio_n_multimodal.pdf)
- Querol-Julián, Mercedes; Fortanet -Gómez, Inmaculada (2012). Multimodal evaluation in academic discussion sessions: How do presenters act and react? *English for Specific Purposes*, 31(4), 271-283. <https://doi.org/10.1016/j.esp.2012.06.001>
- Roderick, Ian (2018). Multimodal critical discourse analysis as ethical praxis. *Critical Discourse Studies*, 15 (2):154-168. <https://www.tandfonline.com/doi/full/10.1080/17405904.2017.1418401#:~:text=https%3A//doi.org/10.1080/17405904.2017.1418401>



- Ruiz-Madrid, Noelia; Fortanet-Gómez, Inmaculada (2015) Contrastive Multimodal Analysis: Conference Plenary Lectures in English and in Spanish. En Belinda Crawford Camiciottoli e Inmaculada Fortanet-Gómez (Eds.) *Multimodal Analysis in Academic Settings. From Research to Teaching*. Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315738758-4/contrastive-multimodal-analysis-noelia-ruiz-madrid-inmaculada-fortanet-g%C3%B3mez?context=ubx&refId=36e921d5-29d2-41b2-a260-48d1e1eae840>
- Ruiz-Madrid, Noelia; Fortanet-Gómez, Inmaculada; Bernad-Mechó, Edgar; Valeiras-Jurado, Julia (2023). *GRAPE-MARS (Multimodal Analysis Research Software)*. [programa informático]. Universitat Jaume I. <https://mars.grape.uji.es>
- Ruiz-Madrid, Noelia; Valeiras-Jurado, Julia (2020). Developing multimodal communicative competence in emerging academic and professional genres. *International Journal of English Studies*, 20 (1), pp. 27–50. <https://doi.org/10.6018/ijes.401481>
- Serrano-Hernández, Elena (2024). Becoming multimodally literate in the digital era: Instagram reels as pre-tertiary didactic materials. [Presentación] *41 Congreso de AESLA*, Valencia, 17-19 de abril de 2024.
- Swales, John M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press. <https://archive.org/details/genreanalysiseng0000swal>
- Thwaite, Anne (2015). Using the "Multimodal Analysis Video" program for register analysis: A preliminary study. *TESOL International Journal*, 10(1), 110-113. <https://www.elejournals.com/tesol-journal/tij-2015/tesol-international-journal-volume-10-issue-1-2015/>
- Valeiras-Jurado, Julia; Bernad-Mecho, Edgar (2022) Modal density and coherence in science dissemination: Orchestrating multimodal ensembles in online TED talks and YouTube science videos. *Journal of English for Academic Purposes*, 58, 101118. <https://doi.org/10.1016/j.jeap.2022.101118>
- Valeiras-Jurado, Julia; Bernad-Mechó, Edgar; Ruiz-Madrid, Noelia (2024). New tools for Multimodal Discourse Analysis: Analyzing spoken digital genres with GRAPE-MARS. [Presentación] *41 Congreso de AESLA*, Valencia, 17-19 de abril de 2024.
- Wanselin, Hanna; Danielsson, Kristina; Wikman, Susanne (2022). Analysing Multimodal Texts in Science—a Social Semiotic Perspective. *Res Sci Educ*, 52, 891–907. <https://doi.org/10.1007/s11165-021-10027-5>
- Xia, Sichen Ada; Hafner, Christoph A. (2021). Engaging the online audience in the digital era: A multimodal analysis of engagement strategies in TED talk videos. *Ibérica*, 42, 33–58. <https://doi.org/10.17398/2340-2784.42.33>

