



LA GRAMÁTICA DE CONSTRUCCIONES MULTIMODAL: POTENCIAL Y DESAFÍOS

Javier VALENZUELA MANZANARES¹ y José Antonio MOMPEÁN GONZÁLEZ²
Universidad de Murcia

Resumen

El presente trabajo trata sobre el encaje de la información multimodal en la gramática de construcciones (GCx). La GCx es una familia de teorías en la Lingüística Cognitiva que sostiene que el conocimiento lingüístico consiste fundamentalmente en construcciones, unidades holísticas dentro de un continuum léxico-gramatical derivadas de los patrones de uso detectados en los actos comunicativos, y que son constantemente actualizadas con cada nuevo acto comunicativo. En principio, la gramática de construcciones representa un marco idóneo para el tratamiento e incorporación de la multimodalidad ya que la teoría no pone límites a qué aspectos de la situación comunicativa pueden ser incorporados a estas estructuras. Si un aspecto dado se repite lo suficiente y tiene poder diagnóstico o predictivo, puede teóricamente formar parte de la información incluida en la construcción, lo que abre la puerta a la inclusión de aspectos multimodales como la prosodia, los gestos co-habla, o incluso otras señales como expresiones faciales, postura corporal y dirección de la mirada. Desde este enfoque, el trabajo describe diversos casos gramaticales en los que aspectos multimodales como la prosodia y gestualidad interactúan y cómo la gramática de construcciones ha modelado estos casos incorporando esta información. El trabajo finaliza con algunos desafíos que la inclusión de la multimodalidad supone para la arquitectura actual de la gramática de construcciones.

Palabras clave: gramática de construcciones; multimodalidad; gestualidad; prosodia

-
1. jvalen@um.es.  <https://orcid.org/0000-0002-0007-7943>
 2. mompean@um.es.  <https://orcid.org/0000-0002-5933-5759>

MULTIMODAL CONSTRUCTION GRAMMAR: POTENTIAL AND CHALLENGES

Abstract

This paper discusses the inclusion of multimodal information in Construction Grammar (CGx). CGx is a family of theories in Cognitive Linguistics that maintains that linguistic knowledge consists fundamentally of constructions, holistic units within a lexical-grammatical continuum, derived from usage patterns detected in communicative acts, that are constantly updated with each new communicative act. In principle, CGx represents an ideal framework for the treatment and incorporation of multimodality, since the theory does not set limits to which aspects of the communicative situation can be incorporated into these structures. If a given aspect is repeated enough and has diagnostic or predictive power, it can theoretically form part of the information included in the construction, which opens the door to the inclusion of multimodal aspects such as prosody, co-speech gestures, or even other cues such as facial expressions, body posture, and gaze direction. From this approach, the paper describes several cases in which multimodal aspects such as prosody and gesturing are seen to interact and how construction grammar has modeled these cases incorporating this information. The paper ends with some challenges that multimodality poses to the current architecture of CGx.

Keywords: construction grammar; multimodality; gesture; prosody

RECIBIDO: 13/02/2022

APROBADO: 08/03/2023

1. INTRODUCCIÓN

A pesar de su relativamente corta trayectoria en comparación con otras ciencias modernas tales como la física o la química, en las que algunos avances han llevado siglos, la lingüística ha experimentado en el último siglo diversos cambios de paradigma que han tenido como consecuencia firmes progresos en la comprensión sobre el lenguaje como fenómeno humano. Entre estos cambios se encuentran dos revoluciones que han tenido lugar en las últimas décadas: el énfasis en el «uso lingüístico» como fuente principal del conocimiento lingüístico y el reconocimiento de la naturaleza multimodal del lenguaje. De estas dos revoluciones en las ciencias del lenguaje, una de ellas, el giro hacia un enfoque basado en el uso, está bastante consolidada. En este sentido, las teorías lingüísticas ya han incorporado diversos aspectos referidos a este nuevo foco, tales como la importancia de la frecuencia

como variable modeladora de las lenguas, o la crucial importancia del contexto, entendido este a muchos niveles, desde el puramente lingüístico al situacional o incluso al social y cultural. Por otro lado, la segunda revolución probablemente esté aún en proceso: el foco en la multimodalidad como aspecto insoslayable del lenguaje. Se reconoce así que el lenguaje no solo consiste en una serie de señales verbales (morfemas, palabras, sintagmas, oraciones...), sino también en otro tipo de . señales procedentes de otras modalidades, como la visual, como ocurre en el caso de los gestos realizados de manera común durante las interacciones comunicativas cara a cara. Como decimos, este giro hacia la multimodalidad está probablemente aún en proceso, puesto que a pesar de la multitud de estudios que demuestran la gran relevancia de esta información multimodal a la hora de conformar y entender las interacciones comunicativas, falta todavía una mayor integración de estos resultados de investigación en las teorías lingüísticas existentes. Con toda probabilidad, ambos giros –el enfoque basado en el uso y el multimodal, muy relacionados entre sí (véase abajo)– son determinantes no solo en la investigación lingüística actual sino que, previsiblemente, seguirán siendo objetivos destacados en las próximas décadas.

El presente trabajo describe el surgimiento de los enfoques del lenguaje basados en el uso, prestando especial atención a uno de ellos, la gramática de construcciones, así como el creciente interés por estudio del lenguaje desde un punto de vista multimodal. Se presentan diversos trabajos en los que esta teoría ha integrado aspectos multimodales de manera puntual y se discuten los desafíos que supone la inclusión más amplia de todo tipo de aspectos multimodales utilizando los mecanismos actuales de la teoría.

2. ENFOQUES BASADOS EN EL USO E INFORMACIÓN MULTIMODAL

Los enfoques del lenguaje «basados en el uso» (*usage-based*, en inglés) (p. ej., Bybee, 2006; Kemmer y Barlow, 2000; Ibbotson, 2013) mantienen que el uso lingüístico es la *única* fuente del conocimiento lingüístico de las personas, que son capaces desde la niñez (p. ej., Tomasello, 2003) de detectar y de extraer aspectos recurrentes en los distintos eventos comunicativos de habla o «actos de uso» (*usage events*, en inglés) o partes del mismo (Langacker, 2008) e incorporarlos al conocimiento que tienen sobre la lengua, como se ilustra en (1). En este ejemplo, procedente del corpus CHILDES (Aguirre, 2003), se presentan tres actos de uso en los que intervienen un niño (N) llamado Magín, de menos de dos años, su madre (M) y su padre (P), en el espacio de un mes y medio y en un contexto de juego en la habitación y el baño de su domicilio.

(1)

Evento de uso 1	Evento de uso 2	Evento de uso 3
M: ven cuco vamos a acostar	P: la lu(na)	M: Magín a acostar, a <i>dormir</i>
M: vamos a <i>dormir</i>	N: la nuna [: luna]	N: <i>momir</i> [: dormir]
M: ¿eh?	P: ¿dónde está?	N: mamá
M: vamos a <i>dormir</i> Magín	N: (l)a nuna	M: ¿qué?
N: xxx, tatín [: calcetín]	...	N: ¡ay!
N: tatín	N: guardar	M: ay, a <i>dormir</i>
M: ¿qué quieres?	P: a guardar	N: narana [: araña]
N: tatín	N: a (dor) <i>mir</i>	P: está durmiendo la araña
M: ¿qué quieres Magín?	P: a <i>dormir</i>	P: a <i>dormir</i> , venga
	N: ino!	

En el ejemplo, se percibe una característica típica del lenguaje de progenitores y cuidadores: la repetición de palabras y otras estructuras en su input, lo cual ayuda al aprendizaje y consolidación de las mismas en el conocimiento lingüístico del menor (Tomasello, 2003). En estos casos, destacamos el uso del verbo *dormir*, cuya repetición por parte de los progenitores se refleja, a nivel producción, en los intentos de Magín de imitar dicho verbo (*..mir... momir*) que, a nivel de producción y con referencia a las normas adultas, todavía tendrá que continuar evolucionando no solo a través de la exposición al input sino a la maduración de las estructuras y habilidades motoras utilizadas en la producción del lenguaje (p. ej., McAllister Byun y Tessier, 2016).

El surgimiento de los enfoques basados en el uso se considera una revolución en la teoría lingüística al romper con décadas de tradición lingüística, en la que el uso de la lengua, que se puede asimilar a los conceptos de «habla» (*parole*) o «actuación» (*performance*) en las tradiciones estructuralista y generativa, respectivamente, era considerado como irrelevante. En el análisis lingüístico, este aspecto se separaba de manera estricta del sistema abstracto e independiente de los hablantes (la *langue* saussuriana) o la facultad y conocimiento innato del lenguaje (la «competencia» chomskyana). (Más tarde esta distinción fue denominada por Chomsky como «Lengua-E», o lengua externa y «Lengua-I» o referida al sistema interno, el objetivo real de la lingüística para este autor; Chomsky, 1986).

Una de las razones por las que el uso de la lengua fue marginado durante gran parte del siglo XX –además de por evocar el espectro del conductismo y el papel de la repetición tan ferozmente criticados por Chomsky– era porque aquel se consideraba como un fenómeno caracterizado por una enorme variación contextual, arbitraria y difícil de predecir. De esta manera, lo que realmente se juzgaba como relevante para averiguar el funcionamiento de la lengua era ese sistema ideal, abstraído y separado de los usos concretos del lenguaje. La actuación,

en cambio, agrupa aquellos aspectos relacionados con el uso del lenguaje en situaciones concretas, entre los que se incluirían aspectos como el conocimiento extralingüístico, la memoria, la atención, los errores idiosincráticos cometidos al hablar, o las vacilaciones, que se estimaban como mucho menos centrales para alcanzar ese objetivo verdadero de la lingüística, que no era otro que el de averiguar el funcionamiento del sistema lingüístico.

En contraposición, los enfoques basados en el uso sostienen que este no es tan impredecible como podría parecer, sino que se puede caracterizar por una serie acotable de regularidades, de aspectos que se repiten y que juegan un papel fundamental como substrato sobre el que, dadas ciertas condiciones, como por ejemplo una exposición mínimamente frecuente a los mismos, se construyen los patrones lingüísticos que los hablantes abstraen e incorporan a su «gramática» (conceptual/mental). El conocimiento lingüístico o gramática, por tanto, representa «la sedimentación de patrones de uso lingüístico» (Thompson, 1998, p. 331) en el discurso, con cada episodio de uso teniendo un impacto en esos patrones, que están constantemente actualizándose. Finalmente, cabe destacar que el uso lingüístico se entiende como un fenómeno «situado» («situated», en inglés) o enraizado en contextos sociales, culturales y de interacción comunicativa, primordialmente en la interacción directa entre personas (Zima y Brône, 2015), por lo que los elementos de esos contextos y prácticas socioculturales son de gran relevancia.

Por otro lado, una segunda revolución más amplia en el estudio del lenguaje desde diversas disciplinas es el foco en la multimodalidad (Goldin-Meadow y Brentari, 2017; Perniss, 2018; Ibarretxe-Antuñano y Valenzuela, 2021), con una especial atención a la relación entre las diversas modalidades en el uso del lenguaje. Esta explosión viene motivada por diversos factores, entre los cuales se hayan el reconocimiento de la necesidad de incorporar la multimodalidad para una comprensión completa de la comunicación humana (que típicamente incluye el uso de la gestualidad), así como el gran incremento experimentado en técnicas y herramientas que permiten el estudio de estos fenómenos. Se asume ahora de manera más explícita que el nicho ecológico principal de la lengua, tanto oral como de signos, lo constituyen los contextos de interacción cara a cara; este es el entorno en el que el lenguaje se usa y se aprende y del que una multitud de elementos, tanto orales como visuales, dependen para la construcción de los enunciados (Perniss, 2018; Vigliocco, Perniss, y Vinson, 2014). Por otro lado, la creciente disponibilidad de corpus multimodales, así como herramientas tecnológicas para su compilación y análisis, ha permitido

revelar patrones entre lenguaje oral y gestualidad anteriormente desconocidos o ignorados (un buen ejemplo es Pagán-Cánovas *et al.*, 2020).

La relevancia de la multimodalidad en el lenguaje humano va más allá del hecho irrefutable de que la comunicación y uso diario de la lengua contenga elementos multimodales: existe cada vez más evidencia de la relevancia del papel de los elementos multimodales en el análisis y la descripción de los distintos componentes tradicionalmente estudiados por la lingüística. Es el caso de la gestualidad, por ejemplo, que mantiene una estrecha relación con el componente fonético-prosódico del lenguaje y con sus usos lingüístico-comunicativos, entre otros. Más adelante (secciones 5 y 6) se describirán algunos trabajos que han explorado estas relaciones entre gestualidad y prosodia.

3. LA GRAMÁTICA DE CONSTRUCCIONES

Bajo las premisas de que tanto el uso del lenguaje como los aspectos multimodales son básicos para entender la naturaleza del conocimiento lingüístico, cabe preguntarse qué herramientas teóricas pueden resultar adecuadas para acomodar dichos componentes. En este sentido, uno de los enfoques basados en el uso que ha ganado especial relevancia en los últimos años es la *gramática de construcciones* o, en su denominación inglesa, *construction grammar* (a menudo abreviada CxG en inglés, GCx en español). Se trata en realidad de una familia de teorías en la Lingüística Cognitiva (p. ej., Croft, 2001; Goldberg, 2006; Kay y Fillmore, 1999; cf. Ibarretxe-Antuñano y Valenzuela, 2021, cap. 3, para un resumen) que, a pesar de su diversidad, comparten la visión central de que el conocimiento lingüístico consiste fundamentalmente en «construcciones», estructuras que se abstraen de las regularidades detectadas en los actos comunicativos, y que son constantemente actualizadas con cada nuevo acto comunicativo. En consonancia con otras aproximaciones basadas en el uso, las construcciones deben entenderse no como unidades estáticas que, una vez abstraídas de los eventos de uso se reutilizan como estructuras preempaquetadas en nuevos eventos, sino como herramientas que se actualizan en esos eventos y se adaptan de forma flexible a las contingencias interaccionales (Zima y Brône, 2015, p. 486).

Las construcciones, que por convención se representan entre corchetes cuadrados [], se pueden definir como combinaciones de elementos formales y semántico-funcionales (p. ej., Croft, 2001; Goldberg, 2006; Langacker, 2008) por lo que evocan en cierto sentido el concepto saussuriano de signo lingüístico. No obstante, difieren

del mismo en diversos aspectos tales como su tamaño, su visión sobre la forma y el contenido, así como su grado de detalle en la representación (o abstracción).

Por un lado, el tamaño de una construcción varía, yendo más allá del de un simple lexema, como muestra el ejemplo (2). Este tamaño oscila desde el morfema ligado (a), pasando por morfemas libres/lexemas (b), y unidades multi-palabra, frecuentemente de carácter idiomático, que van desde las de tipo sintagmático (c), hasta las oracionales simples (d) y compuestas (e). Una consecuencia de esta libertad en el tamaño es que, como se menciona arriba, en vez de funcionar como elementos independientes, el léxico y la sintaxis existen en un continuum en el conocimiento lingüístico.

(2)

(a) [(sab)-er]; [(sab)-ido]

(b) [saber]; [sabelotodo]

(c) [a saber]; [a sabiendas]

(d) [nunca se sabe]; [el saber no ocupa lugar]

(e) [si supiera algo, te lo diría]; [no sabe (o) no contesta]

Por otro lado, la visión sobre los aspectos formales y semántico-funcionales es más amplia que la visión saussureana, ya que incluye elementos de todos los niveles lingüísticos tradicionales. El componente formal de una construcción, por ejemplo, incluye información fonético-fonológica, morfológica, sintáctica y –como vemos más abajo– puede abarcar incluso información multimodal de tipo gestual y, en algunos casos, incluso musical (p. ej., sol–sol–la–sol–do–si en la construcción [cumpleaños feliz]). Por su parte, el componente semántico-funcional incluye información semántica de naturaleza «enciclopédica», es decir, que puede capturar conocimiento amplio y no limitado a semas contrastivos, además de mostrar la «perspectiva» (*construal*, en inglés) (Cienki, 2015) o forma de concebir o representar un evento determinado (p. ej., [me queda la mitad] vs. [me falta la mitad]). Como se muestra en (3), dicha información también puede incluir aspectos como marcadores del discurso en la interacción comunicativa (a), de la estructura global del discurso (b), del registro (c), de una estructura conceptual o guion que describe secuencias apropiadas de eventos en un contexto particular (d), de un contexto situacional (e), del contexto sociocultural (f), del grupo sociolingüístico al que se pertenece (g), o una combinación de dichos aspectos.

(3)

- (a) [*¿sabes?*] / [*¿ves?*] (función fática)
- b) [A: *¿Te importa si...?* + Oración? (estructura petición+respuesta+agradecimiento
B: (respuesta)
A: (*Muchas gracias*)]
- (c) [*compre dos y llévase uno*] (p. ej. en el registro publicitario y ofertas)
- (d) [*la cuenta, por favor*] (al final de una visita al restaurante)
- (e) [*te acompaño en el sentimiento*] (en un velatorio)
- (f) [*quien se fue a Sevilla perdió su silla*] (en la cultura española)
- (g) [*Ok, boomer*] (entre postmilenials, ante un discurso paternalista)

Esta riqueza de información y niveles obedece al hecho de que los hablantes tienen una experiencia holística y multimodal de los actos de uso lingüístico, y la información que extraen de los mismos captura múltiples aspectos que van más allá de lo meramente estructural o mínimamente contrastivo.

Finalmente, el grado de detalle en la representación (o abstracción) de las construcciones es elevado, y puede ir desde representaciones muy específicas con gran detalle fonético segmental y prosódico (p. ej., en construcciones léxicas) basadas en ejemplares (p. ej., Port 2010) hasta representaciones más abstractas derivadas de otras más específicas, como ocurre en las construcciones multi-palabra. Así, mientras algunas están completamente especificadas (p. ej., [*¡lo sabía!*]) otras pueden combinar elementos léxicos específicos y patrones más abstractos, como serían [*¡Qué Nombre más Adjetivo!*] (p. ej., *¡qué artículo más interesante!* *¡qué peli más bonita!*) o [*¿pero es que Oración?*] (p. ej., *¿Pero es que Juan ha venido?* *¿Pero es que tu hermano toca la gaita?*).

4. GRAMÁTICA DE CONSTRUCCIONES Y MULTIMODALIDAD

La gramática de construcciones no pone límites a cuáles son los aspectos de la situación comunicativa que pueden ser incorporados a las estructuras construccionales; si un aspecto dado se repite lo suficiente y tiene poder diagnóstico o predictivo, puede teóricamente pasar a formar parte de la información incluida en la construcción, lo que por supuesto, incluye los aspectos multimodales de los distintos actos de uso (p. ej., Zima y Bergs, 2017; Mompeán y Valenzuela, 2019).

Diversos estudios han tratado de forma explícita el componente multimodal de algunas construcciones, dentro de lo que algunos han venido a llamar *gramática de construcciones multimodal* o, en inglés, *Multimodal Construction Grammar*

(p. ej. Hoffmann, 2021; Schoonjans, 2017; Steen y Turner, 2013; Pagán-Cánovas y Valenzuela, 2016; Zima y Bergs, 2017; entre otros). Por otro lado, este componente multimodal puede constar de varios elementos. En principio, cualquier elemento multimodal puede formar parte de una construcción (véase Ibarretxe-Antuñano y Valenzuela 2022, cap. 5, para una clasificación reciente de tipos de elementos multimodales). Un ejemplo reciente es el análisis de los posts de Instagram como construcciones de tipo abstracto, que incluyen tanto elementos formales como semántico-funcionales y añaden componentes específicos (p. ej., imagen, leyenda, hipervínculos, hashtags), la disposición distintiva de los mismos, y distintas opciones relativas al tipo de color, imagen, etc. (Fischer y Aarestrup, 2021).

A diferencia de construcciones como la anterior, no obstante, la multimodalidad en las construcciones más típicas que se aprenden en el uso de la interacción lingüística cotidiana incluye normalmente un componente *prosódico*, un aspecto del nivel fonético-fonológico que incluye elementos que afectan a la sílaba, palabra o enunciado en su totalidad (p. ej., entonación, acento léxico/rítmico, etc.). Con frecuencia, también se realizan señales relativas a la *gestualidad*, el conjunto de movimientos y postura del cuerpo (en especial del rostro, cabeza y manos) realizados con intención comunicativa. En relación con la prosodia, los acercamientos modulares más clásicos propugnan que la información lingüística se procesa por capas. En este tipo de modelos, la entonación se contempla como una parte añadida o superpuesta a las frases y a las reglas gramaticales. En cambio, en la GCx, se permite que determinados aspectos prosódicos pasen a formar parte indisoluble e integral de algunas construcciones. De manera similar, una diferencia entre la visión de la GCx multimodal y otros enfoques funcionalistas tales como la Lingüística Sistemática Funcional, es que esta última no considera la gestualidad como parte de la gramática, sino como un elemento paralingüístico, como «pequeñas variaciones por medio de las cuales el hablante señala la relevancia de lo que dice» (Halliday, 1985, p. 30). En contraposición, la GCx considera la gestualidad de nuevo como un aspecto completamente integrado en el conjunto de acciones que constituyen la expresión del significado (Kendon, 2004; McNeill, 2008), y por lo tanto susceptible de ser convencionalizado y de formar parte del conocimiento lingüístico o gramatical.

5. LA PROSODIA EN LA GRAMÁTICA DE CONSTRUCCIONES

Existen diversos ejemplos que demuestran que ciertos elementos prosódicos han pasado a formar parte indisoluble de algunas construcciones. Un claro ejemplo es la «construcción de la respuesta de incredulidad» o, en inglés, *Incredulity Response*

Construction (Fillmore, Kay, O'Connor, 1988), utilizada para expresar escepticismo, sorpresa, incredulidad, etc. Mediante un ejemplo icónico de esta construcción en inglés tal como [*what, me worry?*] ('qué, ¿preocuparme yo?'), Lambrecht (1990) muestra algunos de sus rasgos distintivos: el sujeto puede aparecer en acusativo (cuando es un pronombre personal, p. ej., *him, go to the opera?*), el verbo aparece en infinitivo, y existe una división clara en dos unidades tonales.

Esta construcción tiene un equivalente en español; sería el caso ejemplificado por *¿Iraide, cantar tirolés?*. Como es bien sabido, en las oraciones declarativas del español existen unos requisitos morfosintácticos de concordancia entre el sujeto y el verbo: *el niño canta / los niños cantan*. Según esto, la oración *Iraide cantar tirolés* no es gramatical, puesto que el sujeto no concuerda con el verbo principal, que está en infinitivo (cf. *Iraide canta tirolés*). Sin embargo, cuando se asocia esta oración con unos determinados requisitos prosódicos, de manera que rítmicamente se separe el sujeto del verbo principal, y se use una entonación ascendente de tipo interrogativo al final (*¿Iraide, cantar tirolés?*), activamos un tipo de construcción específica: [Sujeto | Verbo infinitivo ↑]; En esta construcción, los parámetros formales morfosintácticos (la inclusión de un verbo en infinitivo que no concuerda con el sujeto, junto con la pausa entre sujeto y verbo así como una entonación ascendente de tipo interrogativo) está indisolublemente asociada con la otra mitad de la construcción, su parte de significado, que activa la expresión de incredulidad o extrañeza por parte del hablante, la sensación de que el evento del que se habla no es habitual o norma. Este tipo de construcción no se puede describir de manera modular: solo la conjunción simultánea de parámetros entonativos y morfosintácticos puede activar la construcción y por tanto su parte de significado.

Otro ejemplo de construcción que muestra una estrecha relación entre prosodia y significado construccional es la construcción [*hay x y x*] en español (Valenzuela y Hilferty, 2007). En esta construcción se reduplica el objeto del verbo *haber* (p. ej., *hay días y días, hay bancos y bancos*). Sin embargo, una mirada más atenta revela que existen en realidad dos subconstrucciones muy similares, que únicamente se distinguen por un parámetro formal: el patrón entonativo. Así, la oración *hay niños y niños* puede activar dos significados distintos. En el primer de ellos, se distingue entre dos tipos de niños, es decir, tiene una función contrastiva. En este caso, este contraste se refleja en los patrones entonativos: el primer *niños* se asocia a una entonación ascendente, y el segundo a una asociación descendente (*hay niños↑ y niños↓*), dando a entender que no todos los niños son iguales. Obsérvese que el aspecto en el que se diferencian los niños no está fijado de antemano; emerge de manera dinámica del

tópico discursivo o de distintos elementos presentes en el contexto, el propósito de la conversación, etc. Por ejemplo, según el contexto, el tipo de contraste en la oración contrastiva *hay libros y libros* puede ser de distintos tipos, como se ve en (4):

- (4)
- (a) *Hay libros* (caros) *y libros* (baratos)
 - (b) *Hay libros* (bien escritos) *y libros* (mal escritos)
 - (c) *Hay libros* (apropiados para regalarle a tu abuela) *y libros* (inapropiados para ello)

Frente a esta construcción contrastiva, existe una construcción muy parecida formalmente, pero distinta en su significado: la construcción enumerativa. Los aspectos morfosintácticos son idénticos (la reduplicación del objeto de *haber*), pero la asociación con una curva entonativa distinta activa otro tipo de semántica. Por ejemplo, si en el enunciado del párrafo anterior se pronuncian ambos nominales con entonación ascendente (*hay niños*↑ *y niños*↑), el significado que se activa no es el de contraste, sino el de enumeración; aproximadamente, el correspondiente a ‘hay muchos niños’. De esta manera, el único parámetro formal que distingue entre ambos tipos de significado (contrastivo frente a enumerativo), es la prosodia.

Valenzuela y Hilferty (2007) comprobaron cómo afectaba esta diferencia prosódica de manera *online* en una tarea de lectura. Estos autores midieron el tiempo que los participantes tardaban en leer una de estas oraciones (p. ej. *hay libros y libros*), pero en dos situaciones distintas: en una de ellas, se daba a los participantes un contexto que sesgaba la interpretación semántica hacia la versión enumerativa. Por ejemplo, tras leer el texto *Pablo dice que no le gusta leer, pero cuando vas a su casa y a continuación hay libros y libros*, la interpretación más coherente es la enumerativa (es decir, que en su casa hay muchos libros). De manera paralela, al leer el texto *¿Pero cómo le regalas ese libro tan subido de tono a tu abuela? ¿Es que no sabes que...?* Seguido de la frase *hay libros y libros*, la interpretación más coherente es la contrastiva: hay distintos tipos de libros y no todos son apropiados para regalarle a tu abuela. La manipulación crucial consistió en insertar entre el contexto y la construcción meta de [*hay x y x*] una frase musical que imitaba la entonación propia de cada una de estas subconstrucciones: una melodía con dos frases ascendentes (imitando la entonación enumerativa) o una melodía con una frase ascendente y otra descendente (imitando la entonación contrastiva). Los resultados mostraron que la concordancia entre sesgo contextual y frase entonativa influía en la velocidad con la que los participantes leían la frase final. Es decir, si el sesgo previo les preparaba para una interpretación contrastiva y escuchaban una melodía contrastiva o les preparaba

para una interpretación enumerativa y escuchaban una melodía enumerativa, tardaban menos tiempo en leer la frase final que en los casos en los que el sesgo y la melodía no coincidían (sesgo contrastivo con melodía enumerativa o sesgo enumerativo con melodía contrastiva). Esto indica que la interpretación final de la construcción [*hay x y x*] requiere de la integración de factores semánticos y prosódicos, una integración que se realiza de manera online, sobre la marcha, al procesar este tipo de construcciones.

6. GESTUALIDAD Y PROSODIA EN LA GRAMÁTICA DE CONSTRUCCIONES

Como hemos apuntado, los diversos elementos formales y semántico-funcionales de las construcciones interactúan en la expresión del significado en el uso lingüístico. De entre todos los elementos formales, no obstante, parece existir una especial relación entre gestualidad y los elementos prosódicos (p. ej., entonación, acento léxico/rítmico, etc.), asociación que ha sido observada desde hace décadas (p. ej., Bolinger, 1986). Esta relación incluye desde el notable alineamiento temporal ambos (Wagner, Malisz, y Kopp, 2014) hasta el hecho de que, al contrario que otros elementos multimodales en el uso del lenguaje (p. ej., la música), tanto gestualidad como prosodia son omnipresentes en el habla conversacional, independientemente de su estructura lingüística u organizativa, lo cual genera la posibilidad de múltiples asociaciones entre ambas.

Existen diversas relaciones entre gesto y prosodia que han sido puestas de relieve en los últimos años. Una de ellas es el denominado «*gestural pitching*», o movimientos arriba y abajo del gesto sincronizados con el contorno entonativo del enunciado (Zappavigna, Clérigh, Dwyer y Martin 2009). Por otro lado, una de las conexiones más mencionadas en la literatura es la coincidencia entre uso correlacionado de los «gestos rítmicos» (*beat gestures*, en inglés) y el pico prosódico de ciertos elementos léxicos en el discurso (McNeill, 2008, p. 41). Los «*beats*» son gestos en los que la forma y movimiento de la mano asemeja un golpe de la misma (hacia arriba y abajo o hacia atrás y adelante) y que tiende a señalar el momento o parte del discurso que el/la hablante percibe como más importante o relevante. Partiendo de una unidad analítica como la frase entonativa con una estructura prosódica que incluye un punto más prominente o núcleo (p. ej., Cruttenden, 1986), podemos señalar que existe una notable correlación entre ejecución del gesto, sílaba tónica y relevancia o importancia de la sílaba o palabra en cuestión (foco o rema) en el

discurso (p. ej., Dimitrova *et al.*, 2016; Jiménez-Bravo y Marrero-Aguilar 2020). No obstante, no debe de entenderse dicha correlación como una necesidad, ya que existen múltiples ejemplos en los que dicha coincidencia es variable, lo cual parece sugerir que los hablantes pueden ajustar tanto prosodia como gesto en tiempo real dependiendo del uso lingüístico comunicativo en un contexto concreto (Kendon 2004, p. 215-126).

Parece interesante también describir no solo relaciones globales, sino específicas entre gestos concretos, patrones prosódicos determinados, y construcciones concretas. Existen múltiples ejemplos de dichas relaciones alrededor de construcciones léxicas simples y frases hechas. En inglés, por ejemplo, las expresiones *uh-uh* [^hʔ_hʔ_h] y *mm* [^hʔ_hʔ_h] se usan para decir ‘no’, con el uso de una oclusiva glotal a principio de cada una de las dos sílabas y un patrón entonativo H*L, que indica un tono alto (H) con acento prosódico (*) en la primera sílaba, seguida de un tono bajo (L) (Cienki, 2015; Ladefoged y Johnson, 2011, p. 61) inapropiados para ello. Asimismo, dicha expresión de negación se suele acompañar, como en otros ejemplos de negación en inglés, con un movimiento lateral de lado a lado de la cabeza, un patrón muy extendido en las lenguas del mundo, pero no universal (p. ej., Andonova y Taylor, 2012), lo que indica que su uso implica un componente de aprendizaje.

Otro ejemplo de relación entre prosodia, gestualidad y significados/funciones concretas en el discurso es el estudiado por Mompeán y Valenzuela (2019) en relación a la construcción tautológica [*Brexit means Brexit*] (‘Brexit significa Brexit’), popularizada a través de los medios en Reino Unido en el año 2016 tras el referéndum sobre la pertenencia de dicho país a la Unión Europea. En un análisis de un corpus de videos que contenían el uso de esta tautología en el discurso político y periodístico, los autores mostraron cómo la expresión adquirió características formales y semántico-funcionales diferentes desde su aparición en junio de 2016 hasta el último dato analizado (noviembre de 2018). El análisis mostraba cómo una construcción no es algo estático, sino que evoluciona a lo largo del tiempo, y se actualiza en nuevos eventos de uso. Así, la tautología fue usada en un primer lugar por partidarios del Brexit (p. ej. Theresa May) como recurso lingüístico para promover la aceptación del resultado del referéndum (salida del Reino Unido de la Unión Europea); más adelante, la tautología comenzó a usarse por comentaristas y críticos en el contexto de su oposición o cuestionamiento de este resultado. El estudio mostró también no solo el uso combinado de acento prosódico y gestos rítmicos (normalmente usando las dos manos, a veces una o incluso el dedo índice)

sino de manera crucial, patrones distintos asociados a estos dos distintos usos discursivos de la tautología.

Como se ejemplifica en (5), en el primer caso la tautología tenía una estructura prosódica de tres pies rítmicos y una duración más larga, uso de aspiración enfática al final de *Brexit*, y tres gestos rítmicos o *beats* (5a). En la segunda acepción, utilizada como burla o crítica al concepto inicial del Brexit, esta estructura se vio reducida a dos pies rítmicos, una duración más corta, ausencia de aspiración y dos únicos pies rítmicos, incluyéndose la sílaba correspondiente a la forma verbal *means* dentro del primer pie (5b).

- (5)
- (a) *Bre-xit means Bre-xit* |'σσ|σ|'σσ|
 (b) *Bre-xit-means Bre-xit* |'σσσ|'σσ|

De esta manera, este estudio mostraba no solo la conjunción de factores multimodales en una estructura construccional, sino cómo estos factores iban modulando su uso, incorporando las nuevas necesidades comunicativas surgidas del contexto socio-político, para así poder diferenciar los dos usos de la construcción.

7. PROBLEMAS DE UNA GRAMÁTICA DE CONSTRUCCIONES MULTIMODAL COMPLETAMENTE INCLUSIVA

El objetivo de incluir información multimodal en una estructura de tipo construccional, es decir, una estructura que agrupe distintos tipos de información heterogénea de manera que esta información se active de manera simultánea y no modular, es algo que se justifica teóricamente de manera muy sólida. Al fin y al cabo, la multimodalidad es claramente información recurrente en distintos actos de habla, por lo que tiene perfecto sentido que este tipo de información sea incluida por los hablantes en sus asociaciones de forma y significado. No obstante, los detalles de cómo se debe producir la integración de la información multimodal en la gramática de construcciones o cualquier otro modelo similar es un asunto que no está carente de problemas de diverso tipo, que pasamos a comentar a continuación.

7.1. *Segmental frente a suprasegmental; analítico frente a holístico*

Uno de los problemas más básicos a los que se enfrenta la gramática de construcciones a la hora de incorporar la multimodalidad es que es un sistema pensado

para el lenguaje verbal, en el que se pueden identificar componentes (o segmentos, si se quiere), que contribuyen con sus peculiaridades e información individual al todo. Una construcción está constituida por la concatenación de unidades discretas (sean morfemas, palabras o categorías gramaticales), que se combinan por procesos de composicionalidad débil o parcial (Langacker, 2008): existen efectos composicionales, pero el significado de las partes no explica por completo el significado de la combinación. Incluso significados más difusos, como puede ser el significado de transferencia unido a la construcción ditransitiva, dependen de la identificación de unidades individuales. Sin embargo, la información multimodal tiende a ser suprasegmental (o prosódica): un patrón entonativo no se corresponde con una palabra o un morfema, sino que se distribuye por una combinación no especificada de ellos. La prosodia, por ejemplo, no es analítica, sino global; lo mismo podríamos decir de determinado tipo de gestos; no hay efectos de composicionalidad claros y la interacción de estos elementos con la parte verbal no está por lo tanto nada clara y no existe un modelo concreto que realice la unión de ambos tipos de señales con un nivel adecuado de granularidad y especificidad descriptiva.

7.2. Heterogeneidad de la información multimodal: distintos tipos de significado gestual

Uno de los grandes problemas de la inclusión de información multimodal en una unidad de tipo construccional es la indefinición del tipo de significado que llevan asociados los distintos tipos de información multimodal. Ciñéndonos únicamente a la información gestual, existen distintos tipos de gestos (Kendon, 2004) y cada uno de ellos se asocia a distintas cargas significativas. Por ejemplo, los «emblemas» son gestos aproximadamente equivalentes a una unidad léxica. El gesto en el que el índice y el pulgar se tocan formando un círculo mientras que el resto de los dedos se despliegan verticalmente, el conocido como signo de OK, es aproximadamente equivalente a una palabra; es un signo lingüístico completo, en el que la forma, en vez de un sonido, es una configuración manual, y el significado es ese acuerdo, que mencionamos. Los emblemas son por lo tanto gestos que tienen asociados un significado completo y, de hecho, no necesitan ningún tipo de asociación con material verbal.

Otros tipos de gestos activan significados diferentes: por ejemplo, los «gestos deícticos», que aparecen cuando señalamos, con el dedo, cabeza o mirada, un objeto de nuestro entorno al tiempo que decimos *eso*, responden a otro tipo de significado distinto, inseparable del contexto físico del acto comunicativo. Estos significados

se parecen por lo tanto a significados deícticos como los asociados a palabras como *ahí*, *aquí* o *ahora* (de hecho, los tiempos verbales, como el pasado o el futuro, son significados deícticos, puesto que dependen de un centro deíctico que corresponde al momento de enunciación).

Otro tipo de gestos distintos, los «gestos co-habla» tienen un significado mucho más difuso; el gesto lateral que hacemos al decir *de Madrid a Valencia* no corresponde de manera única con este significado de desplazamiento espacial; de hecho, estos gestos pueden además ser metafóricos: el gesto lateral que acabamos de mencionar es el mismo que hacemos al decir *de principio a fin*. De hecho, si observamos el gesto de manera independiente no podemos saber con qué significado está asociado (Valenzuela y Alcaraz, en prensa, 2023). Esta es la norma con la gran mayoría de los gestos co-habla: se adaptan de manera flexible al material verbal, pero dependen de él para su interpretación correcta, sin corresponder a un significado fijo.

Aparte de los gestos co-habla, los conocidos como «gestos recurrentes» están a medio camino entre los significados fijos de los emblemas y los significados contextuales de los gestos co-habla. Por ejemplo, un encogimiento de hombros puede asociarse a una variedad de significados (p.ej., *no me importa*, *no lo sé* o *qué más me da*), que no son ni fijos ni con una variabilidad contextual laxa, correspondiendo a un tipo de significado que podría corresponderse con una actitud del hablante que podría definirse como «desvinculación personal de la situación» (Valenzuela, 2017).

Finalmente, los conocidos como «gestos rítmicos», mencionados anteriormente (véase sección 6), se usan para marcar el ritmo de una narración y ayudan a segmentar el lenguaje y enfatizar determinados significados, puntuándolos con un gesto.

Como hemos visto, una construcción es la asociación de unos parámetros de forma con unos parámetros de significado, pero tal y como se ejemplifica en estos distintos tipos gestuales, los gestos se pueden asociar no ya a muchos significados distintos, muchos de ellos de gran variabilidad contextual, sino a distintos niveles de significación. La especificación completa de cómo se unen estos significados heterogéneos al material verbal, por lo tanto, presenta enormes desafíos.

Adicionalmente, la información gestual mantiene distintas relaciones de coocurrencia con el material verbal. Los gestos pueden ser más o menos independientes del elemento verbal con el que se produzcan, sustituyendo complementemente algunos casos a la información verbal o dependiendo de ella para su comprensión. Asimismo, en ocasiones, el gesto puede reforzar lo expresado por la información verbal (como cuando se hace un gesto manual hacia arriba mientras se dice *ha subido al piso de arriba a buscar sus gafas*) o puede dar información no contenida en la expresión

verbal, como cuando se hace el mismo gesto hacia arriba mientras se dice *ha ido a su piso a buscar las gafas*; en este caso, la información de que el piso está arriba la da el gesto, no la frase. Si se dice *ayer por la noche parece que lo pasaron de maravilla* mientras se hace el gesto de beber (haciendo un movimiento manual repetido con el pulgar señalando la boca y el meñique extendido), el mensaje que se transmite es que ayer estuvieron bebiendo, de nuevo información que se expresa gestualmente y no verbalmente.

En resumen, lo que se quiere poner de manifiesto es que la relación entre los significados aportados por el gesto y los significados aportados por el material verbal es una relación muy compleja, no solo por los distintos tipos de significado gestual sino por las distintas posibilidades de distribución informacional entre gesto y mensaje verbal, de manera que el diseño de un mecanismo detallado que permita su integración, dada la variada casuística, en muchos casos, muy dependiente del contexto comunicativo, es un problema de primer orden que todavía ha de resolverse.

7.3. Heterogeneidad de las señales multimodales

Los problemas a los que se enfrenta la integración de la información multimodal en una estructura de tipo construccional se acrecienta si consideramos la diversidad de elementos multimodales que pueden ser potencialmente relevantes. Hasta ahora hemos mencionado información gestual (manual) y prosódica, pero las señales multimodales incluyen otros muchos articuladores (p. ej., expresiones faciales, postura corporal, dirección de la mirada). Por ejemplo, ¿qué debería hacerse con las expresiones faciales? Determinadas expresiones faciales coocurren de manera frecuente con estructuras de tipo construccional, como podría ser la construcción interrogativa. Es el caso de la elevación de las cejas, por ejemplo. Nota y colegas (2022) utilizaron un avatar creado por ordenador para estudiar cómo los signos faciales (en particular el movimiento de las cejas) contribuyen a la rapidez en la identificación de preguntas en el uso lingüístico conversacional. El avatar era una figura humana en la que los diferentes articuladores multimodales –por ejemplo, las diferentes partes de la cara, cejas, ojos, boca– eran manipulables mediante un programa de ordenador. Esta manipulación selectiva de cada uno de los articuladores permitía realizar experimentos muy controlados: por ejemplo, para comprobar el papel que desempeña el levantamiento de cejas en la comunicación, hicieron que el avatar formulara preguntas a los usuarios humanos al tiempo que levantaba las cejas, en unos casos, o sin levantarlas en otros, y midieron el tiempo que tardaban en ambos casos los usuarios en responder (es decir, en reconocer que se les había hecho

una pregunta). Sus resultados mostraron que los usuarios humanos contestaban más rápidamente las preguntas asociadas con un levantamiento de cejas; es decir, este movimiento les ayudaba a interpretar la secuencia verbal como una pregunta. Por lo tanto, parece que esta señal forma parte de la construcción y, por lo tanto, debería incluirse en una descripción completa de la construcción interrogativa.

Otras señales multimodales serían incluso más difíciles de integrar. La dirección de la mirada, por ejemplo, tiene una serie de funciones muy definidas en las interacciones comunicativas: ayuda a identificar cuál es el referente o el foco de la oración al que se refiere el hablante, influyendo por ejemplo en si una situación se describe con voz activa o pasiva (Ibbotson, Lieven y Tomasello, 2013), es básica para el establecimiento de la «atención compartida» (*joint attention*, en inglés; Tomasello, 2003), y es un componente crucial para regular los turnos en las interacciones dialógicas (Deguyte & Astell, 2021) También ayuda a interpretar la intención comunicativa del hablante (una mirada directa a los ojos se asocia a un mayor énfasis en lo que se está diciendo, como una declaración aumentada de veracidad o importancia; Kendon, 1967) y se ha demostrado que la dirección de la mirada es básica en el desarrollo del lenguaje (véase Çetinçelik *et al.*, 2021 para una revisión). La postura corporal, adelantando el tronco hacia el interlocutor, también tiene su función comunicativa; lo mismo podría decirse de mecanismos como la distancia corporal (conocida como «proxémica»), que depende de factores personales y sobre todo culturales: distintas culturas tienen distintas distancias para interacciones íntimas, personales, sociales o públicas (Hall, 1974).

¿Cómo podemos incorporar todas estas señales tan variadas y dependientes en diversos grados del contexto comunicativo en una construcción? Sería necesario introducir una nueva serie de mecanismos de combinación de estas señales; la especificación detallada de cómo las diferentes señales se unen para constreñir el espacio de construcción de significado es de nuevo hartamente complicada y, de hecho, no existe en realidad un modelo completo que lo haya intentado.

7.4. ¿Cómo se podría representar la información multimodal en la gramática de construcciones?

En vista de lo expuesto, parece que la forma de representar la información completa en una unidad debe ser diferente a las estructuras que conocemos en la actualidad como construcciones. Posiblemente podría corresponder a estructuras complejas en las que los distintos tipos de información se presenten de manera

probabilística configurando patrones gestálticos en los que los distintos parámetros se pueden adaptar de manera dinámica a las necesidades comunicativas de un contexto dado. La idea del «construcción» como un repositorio de correspondencias forma-significado que almacenamos en la memoria a largo plazo y traemos a la memoria de trabajo para expresar un significado en la producción lingüística o para identificar un patrón lingüístico en el caso de la comprensión, parece que debe ceder el terreno a un modelo necesariamente más fluido y contextual, en el que las probabilidades de coocurrencia de distintos tipos de información configuran un espacio de sistema dinámico, con determinadas configuraciones construccionales actuando como atractores.

8. CONCLUSIÓN

En general, la necesidad de incluir información multimodal en descripciones completas y exhaustivas de las lenguas es algo cada vez menos discutible. El formalismo que se utilice compartirá de manera necesaria características con una gramática de tipo construccional, que agrupe en una misma estructura información tanto verbal como no-verbal (o multimodal). Existe evidencia de que estas configuraciones, que agrupan distintos tipos de información compleja (que proceden tanto del canal auditivo como del visual), facilitan la comunicación en vez de complicarla. Esto puede parecer contradictorio: si en un intercambio comunicativo, el interlocutor tiene que estar atento no solo a las señales de tipo verbal sino a toda la variedad de señales multimodales (teniendo que discriminar entre, por poner un ejemplo, los movimientos accidentales, como subirse las gafas o rascarse un brazo, y los movimientos con intención comunicativa, como un gesto acercando la mano al oyente) la lógica diría que esta sobrecarga informacional saturaría el sistema de reconocimiento y haría la comunicación más lenta (Valenzuela, 2019).

Sin embargo, como hemos visto en algunos ejemplos explicados en este artículo, lo que ocurre es lo contrario. Esa mayor facilidad de procesamiento al enfrentarse a una serie simultánea de señales variadas y complejas indica claramente que la distinta información multimodal se agrupa en una unidad comunicativa. Por ejemplo, reconocemos más rápido la imagen de un perro si además escuchamos un ladrido; cada una de las partes que conforman el todo sirven para activarlo, de manera que las diversas pistas se suman, confirmando la correcta identificación categorial. De la misma manera, en el caso de la comunicación oral, las distintas pistas multimodales de nuevo se suman, y hacen que el reconocimiento de la

intención comunicativa sea más sencilla, rápida y eficiente. Esto solo es posible si dichas señales no se producen de manera incidental y aislada, sino si todas forman parte de algún tipo de estructura común, ese todo del que cada una de las señales es una parte.

En principio, la teoría cuyos principios la posicionan en mejor lugar para incluir y agrupar informaciones tan heterogéneas es la gramática de construcciones. Existen sin embargo una serie de problemas técnicos que hacen que, con toda posibilidad, una nueva versión de la gramática de construcciones multimodal tenga un sistema de representación de la información distinto de la versión más clásica. En este nuevo sistema, la confluencia de los diferentes tipos de información debe considerarse de manera mucho más probabilística y la descripción de sus interrelaciones a todos los niveles es todavía un problema extremadamente complejo que hay que abordar de manera sistemática y gradual. De esta manera, podría considerarse que la gramática de construcciones multimodal es en estos momentos más una meta a la que dirigirse que una realidad actual. No obstante, es nuestra firme convicción que este es el formato que cualquier descripción completa y de base empírica de los comportamientos lingüísticos deben tener y el destino al que al que con toda probabilidad todos nos dirigimos.

REFERENCIAS

- Aguirre, C. (2003). Early verb development in one Spanish-speaking child, en D. Bittner, W. U. Dressler y M. Kilani-Schoch (eds.), *Development of verb inflection in first language acquisition: A cross-linguistic perspective*, pp. 1-26. Berlín y Nueva York: De Gruyter Mouton
- Andonova, E. y Taylor, H. A. (2012). Nodding in dis/agreement: A tale of two cultures. *Cognitive Processes Suppl 1*:S79-82. DOI: <https://doi.org/10.1007/s10339-012-0472-x>
- Bolinger, Dwight (1986). *Intonation and its parts: Melody in spoken English*. Stanford: Stanford University Press.
- Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language* 82(4), 711-733.
- Chomsky, N. (1986). *Knowledge of language*. Nueva York: Praeger
- Cienki, A. (2015). Spoken language usage events. *Language and Cognition* 7, 499-514.
- Croft, W. (2001). *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford y Nueva York: Oxford University Press.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Çetinelik, M., Rowland, C. F., y Snijders, Tineke M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology* 8;11:589096. DOI: <https://doi.org/10.3389/fpsyg.2020.589096>

- Degutyte, Ziedune y Astell, Arlene (2021). The role of eye gaze in regulating turn taking in conversations: A systematized review of methods and findings. *Frontiers in Psychology* 12:616471. DOI: <https://doi.org/10.3389/fpsyg.2021.616471>
- Dimitrova, D., Chu, M., Wang, L., Özyürek, A. y Hagoort, P. (2016). Beat that word: How listeners integrate beat gesture and focus in multimodal speech discourse. *Journal of Cognitive Neuroscience* 28(9), 1255-1269.
- Fillmore, Ch. J., Kay, P. y O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3), 501-538.
- Fischer, K. y Aarestrup, M. (2021). Relationships between construction grammar(s) and genre: Evidence from an analysis of Instagram posts. *Journal of Pragmatics* 183, 87-104.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford y Nueva York: Oxford University Press.
- Goldin-Meadow, S. y Brentari, D. (2017). Gesture, sign, and language: the coming of age of sign language and gesture studies. *Behavioural Brain Sciences* 40:E46. DOI: <https://doi.org/10.1017/S0140525X15001247>
- Hall, E. T. (1974). *Handbook for proxemic research*. Washington, D. C: Society for the Anthropology of Visual Communication.
- Halliday, M. A. K. (1985). *Spoken and written language*. Victoria, NSW, Australia: Deakin University.
- Hoffmann, T. (2021). Multimodal Construction Grammar: From multimodal constructs to multimodal constructions, en X. Wen y J. R. Taylor (eds.), *The Routledge handbook of Cognitive Linguistics*, pp. 78-92. Londres y Nueva York: Routledge.
- Ibarretxe-Antuñano, I. y Valenzuela, J. (2021). *Lingüística y cognición: el lenguaje desde la lingüística cognitiva*. Madrid: Editorial Síntesis.
- Ibbotson, P. (2013). The scope of usage-based theory. *Frontiers in Psychology* 4(255). DOI: <http://doi.org/10.3389/fpsyg.2013.00255>
- Ibbotson, P., Lieven, E. V. M., y Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics*, vol. 24(3), 457-481. DOI: <https://doi.org/10.1515/cog-2013-0020>
- Jiménez-Bravo, M. y V. Marrero-Aguar (2020). Multimodal perception of prominence in spontaneous speech: A methodological proposal using mixed models and AIC. *Speech Communication* 124, 28-45.
- Kay, P. y Fillmore, Ch. (1999). Grammatical constructions and linguistic generalizations: The what's X doing Y construction. *Language* 75(1), 1-33.
- Kemmer, S. y Barlow, M. (2000). Introduction: A usage-based conception of language, en M. Barlow y S. Kemmer (eds.), *Usage-based Models of Language*, pp. 7-28. Stanford: CSLI Publications.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica* 26, 22-63. DOI: [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4).
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Ladefoged, P. y Johnson, K. (2011). *A course in phonetics*. 6ª edición. Boston, MA: Wadsworth.

- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Lambrecht, K. (1990). What, me worry? — ‘Mad Magazine sentences’ revisited. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society* 16, p. 215-228.
- McAllister Byun, T. y Tessier, A. M. (2016). Motor influences on grammar in an emergentist model of phonology. *Language and Linguistics Compass* 10(9), 431-452.
- McNeill, D. (2008). *Gesture and thought*. Chicago, IL: The University of Chicago Press.
- Mompeán, J. A. y Valenzuela, J. (2019). *Brexit means Brexit: a constructional analysis*. *Complutense Journal of English Studies* 27, 1-37
- Nota, N., Trujillo, J. P., y Holler, J. (2022). Conversational eyebrow frowns facilitate question identification: An online VR study. *PsyArXiv*. DOI: <https://doi.org/10.31234/osf.io/fcj8b>
- Pagán Cánovas, C. y Valenzuela, J. (2017). Timelines and multimodal constructions: Facing new challenges. *Linguistic Vanguard* 3(s1). DOI: <https://doi.org/10.1515/lingvan-2016-0087>
- Pagán Cánovas, C., Valenzuela, J., Alcaraz Carrión, D., Olza, I. y Ramscar, M. (2020). Quantifying the speech-gesture relation with massive multimodal datasets: Informativity in time expressions. *PloS One* 15(6). e0233892.
- Perniss, P. (2018). Why we should study multimodal language. *Frontiers in Psychology* 9(1109). DOI: <https://doi.org/10.3389/fpsyg.2018.01109>
- Port, R. (2010). Rich memory and distributed phonology. *Language Sciences* 32, 43-55.
- Schoonjans, S. (2017). Multimodal Construction Grammar issues are Construction Grammar issues. *Linguistic Vanguard* 3(s1). DOI: <https://doi.org/10.1515/lingvan-2016-0050>
- Steen, F. F. y Turner, M. (2013). Multimodal Construction Grammar, en Michael Borkent, B. Dancygier y J. Hinnell (eds.), *Language and the creative mind*, pp. 255-274. Stanford, CA: CSLI Publications.
- Thompson, S. A. (1998). A discourse explanation for the cross-linguistic differences in the grammar of interrogation and negation, en A. Siewierska y J. Jung Song (eds.), *Case, typology and grammar: In honor of Barry J. Blake*, pp. 307-341. Amsterdam: John Benjamins.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Valenzuela, J. (2019). Señales no verbales en la comunicación: ¿problema o solución?. *Ciencia Cognitiva* 13(3), 71-73.
- Valenzuela, J. (2017). A new multimodal construction? What do I care? *International Cognitive Linguistics Conference (ICLC-14)*. University of Tartu (Estonia). July 11.
- Valenzuela, J. y Hilferty, J. (2007). Music, modularity and syntax. *International Journal of English Studies* 7(2), 101-116.
- Valenzuela, J. y Alcaraz-Carrión, D. (en prensa, 2023). English temporal gestures are temporal gestures.

- Vigliocco, G., Perniss, P. y Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B* 369(1651): 20130292.
DOI: <https://doi.org/10.1098/rstb.2013.0292>
- Wagner, P., Malisz, Z. y S. Kopp (2014). Gesture and speech in interaction: An overview. *Speech Communication* 57, 209-232.
- Zappavigna, M., Cléirigh, Ch., Dwyer, P. y Martin, J. R. (2009). The coupling of gesture and phonology, en M. Bednarek and J.R. Martin [eds.] *New discourse on language: Functional perspectives on multimodality, identity, and affiliation*, pp. 237-266. London: Continuum.
- Zima, E. y Bergs, A. (2017). Multimodality and construction grammar. *Linguistics Vanguard* 3, s1, 2017, pp. 20161006. DOI: <https://doi.org/10.1515/lingvan-2016-1006>
- Zima, E. y B., Geert (2015). Cognitive Linguistics and interactional discourse: Time to enter into dialogue. *Language and Cognition* 7(4), 485-498.

Javier Valenzuela Manzanares
Jose Antonio Mompeán González
Departamento de Filología Inglesa
Campus de la Merced
Universidad de Murcia
Murcia, 30001 (España)

