

Factores que influyen en la comprensión de las cláusulas subordinadas de relativo en español: estudio exploratorio

Factors that influence the comprehension of Spanish relative clauses: exploratory study

Laura M. V. Manoiloff

CIECS-CONICET / U Nacional de Córdoba
Argentina

Pablo E. Requena

University of Montana
United States of America

Ma. Constanza Carando

CIECS-CONICET / U Nacional de Córdoba
Argentina

Cecilia M. Defagó

Universidad Nacional de Córdoba
Argentina

Laura Alonso Alemany

Universidad Nacional de Córdoba
Argentina

Daiana B. Cesaretti

CIECS-CONICET / U Nacional de Córdoba
Argentina

Cecilia Ferrero

CIECS-CONICET / U Nacional de Córdoba
Argentina

Adrian Ramírez

CIECS-CONICET / U Nacional de Córdoba
Argentina

Juan Seguí

CNRS-Paris 5 / EFL Labex
Francia



ONOMÁZEIN 42 (diciembre de 2018): 23-52

DOI: 10.7764/onomazein.42.01 | ISSN: 0718-5758

Laura M. V. Manoiloff, Ma. Constanza Carando, Daiana B. Cesaretti, Cecilia Ferrero y Adrian Ramírez: Equipo de Investigación de Psicología Cognitiva del Lenguaje y Psicolingüística, Laboratorio de Psicología Cognitiva, CIPSI (CIECS-CONICET), Facultad de Psicología, Universidad Nacional de Córdoba, Argentina. | Correo electrónico: lmanoiloff@psyche.unc.edu.ar, constanzacarando@gmail.com, daicesaretti@gmail.com, ferrerochechu@gmail.com y adrianomarramirez@gmail.com

Pablo E. Requena: Department of Modern and Classical Languages and Literatures, College of Humanities and Sciences, University of Montana, United States of America. | Correo electrónico: pablo.requena@umontana.edu

Cecilia M. Defagó: Centro de Investigación María Saleme de Burnichon, Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba, Argentina. | Correo electrónico: ceciliabent@gmail.com

Laura Alonso Alemany: Grupo de Procesamiento del Lenguaje Natural, Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Argentina. | Correo electrónico: lauraalonsoalemany@gmail.com

Juan Seguí: Laboratoire Mémoire et Cognition, CNRS-Paris 5 / EFL Labex, Francia. | Correo electrónico: juan.segui@parisdescartes.fr

Fecha de recepción: noviembre de 2016 | Fecha de aceptación: agosto de 2017

Resumen

La comprensión del significado oracional depende en gran parte de la facilidad con que los hablantes asignan roles temáticos (como Agente o Paciente) a los sintagmas nominales. Este procedimiento es particularmente útil para comprender cláusulas subordinadas de relativo, ya que al procesarlas los hablantes deben discernir si se trata de relativas con extracción de Sujeto (S) ([El joven₁ [que_{1Sujeto} vino] trajo una valija]) o con extracción de Objeto (O) (El joven₁ [que_{1Objeto} yo saludé] trajo una valija). Trabajos previos han mostrado que el primer tipo de extracción resulta más fácil de procesar que el segundo y se ha sugerido que factores como el orden de palabras y la presencia o ausencia de marca de Caso podrían influir en el procesamiento de las relativas de O. El presente estudio exploratorio es el primero que aprovecha la flexibilidad morfosintáctica del español para examinar cuatro estructuras que combinan distintos órdenes de palabras con presencia o ausencia de marca de Caso. Luego de ofrecer un análisis de frecuencia de uso de estas estructuras en un corpus escrito del español, se presentan dos estudios: uno sobre comprensión espontánea (asignación de roles) sin límite de tiempo y otro de elección de opciones con límite de tiempo. Los resultados sugieren que tanto el orden canónico como la marca de Caso benefician la comprensión.

Palabras clave: procesamiento sintáctico; comprensión; cláusulas de relativo en español; orden de palabras; marca de Caso.

Abstract

Most of the comprehension of sentential meaning depends on how easily speakers assign thematic roles (such as Agent or Patient) to noun phrases. This process is particularly useful for comprehending relative clauses given that when processing them speakers need to discern if they are dealing with a Subject (S) relative clause (*El joven₁ [que_{1Sujeto} vino] trajo una valija*, 'The young man₁ [that_{1Subject} came] brought the suitcase') or an Object (O) relative clause (*El joven₁ [que_{1Objeto} yo saludé] trajo una valija*, '[The young man₁ [that_{1Object} greeted] brought the suitcase]'). Previous research has shown that the former type is easier to process than the latter type and it has been suggested that factors such as word order and the presence or absence of Case marking could impact the processing of O relative clauses. The present exploratory study is the first one to exploit the morphosyntactic flexibility of Spanish in order to examine four structures that combine different word orders with the presence or absence

of Case marking. Following the presentation of a frequency analysis of these structures in a written corpus of Spanish, this paper presents two studies: one about spontaneous comprehension (thematic role assignment) without time limit and another one in which responses were timed. Results suggest that both canonical word order and Case marking positively impact comprehension.

Keywords: syntactic processing; comprehension; Spanish relative clauses; word order; Case marking.

1. Introducción

La comprensión del significado oracional implica no solo el reconocimiento de las unidades léxicas (cadenas de palabras), sino también el procesamiento de rasgos funcionales con los cuales se construyen las combinatorias oracionales de características jerárquicas y arbóreas (Chomsky, 1986). El estudio de las cláusulas subordinadas relativas permite apreciar el carácter abstracto de la sintaxis, puesto que el procesamiento de estas implica la combinatoria de elementos que no siempre se hallan explícitos (sino que muchas veces se consideran “implícitos” o elididos). Por eso, estas construcciones han constituido un objeto de estudio privilegiado para los lingüistas y psicolingüistas.

Para la Gramática Generativa, la descripción de dichas oraciones supone que uno de los roles temáticos (Agente y Tema) y uno de los Casos (Nominativo o Acusativo) no se realizan explícitamente a través de una categoría léxica, sino que su información es captada por un pronombre relativo (complementante) que se desplaza de la posición que ocuparía la categoría léxica hasta una posición ocupada por una categoría funcional. Por ejemplo, las oraciones (1) y (2) contienen una cláusula de relativo cada una. Según el referente del pronombre relativo *que* en cada caso, este reemplaza a *el joven* que funciona como Sujeto (S) de la subordinada en (1), pero reemplaza a *la valija* que funciona como Objeto (O) de la subordinada en (2). En consecuencia, por la función sintáctica del pronombre relativo estas construcciones han recibido el nombre de relativa de S y relativa de O, respectivamente¹.

(1) Relativa de S: [El joven₁ [que_{1Sujeto} vino] trajo una valija]

(2) Relativa de O: [La valija₁ [que_{1Objeto} el joven trajo] era grande]

El reconocimiento del tipo de información expresada por el pronombre relativo (Agente y Caso Nominativo, en (1), o Tema y Caso Acusativo, en (2)), es posible gracias a marcas morfosintácticas, semánticas y/o discursivas. Los estudios de procesamiento muestran mayor facilidad en la comprensión de las relativas de S en comparación con las de O para lenguas de la familia indoeuropea (aunque también se ha encontrado en algunas investigaciones que las relativas de S resultan más complejas de procesar que las relativas O en lenguas como el

1 Es importante aclarar que, si bien algunos autores consideran que desde un punto de vista semántico ciertos elementos se eliden o aparecen de forma implícita, esto no ocurre desde las teorías formales. Desde la perspectiva sintáctica el pronombre reemplaza a un elemento, no lo elide. Específicamente, la Gramática Generativa (Chomsky, 1965) no le da el nombre de elisión a este fenómeno, tampoco habla de relativas de S ni O. Sin embargo, esta terminología proveniente del estructuralismo es frecuentemente utilizada en el dominio de la psicolingüística a fin de describir el material lingüístico empleado.

chino —por ej. Hsiao y Gibson, 2003; Lin y Garnsey, 2010—, el japonés —por ej. Ishizuka y otros, 2006— y el vasco —por ej. Carreiras y otros, 2010—.

Las diferencias observadas en el procesamiento de ambos tipos de relativas han sido interpretadas desde distintas perspectivas teóricas. La explicación basada en estructura oracional (por ejemplo, Frazier y Clifton, 1989; Miyamoto y Nakamura, 2003) sostiene que las dificultades de procesamiento de las relativas de O tienen que ver con una preferencia universal por oraciones de relativo de S con independencia de otros factores. Enfoques basados en los recursos de procesamiento, por el contrario, ponen el énfasis en factores funcionales tales como limitaciones de recursos y limitaciones cognitivas de integración. Por ejemplo, la teoría de localidad de la dependencia (DLT) (Gibson, 2000; Grodner y Gibson, 2005) propone que las relativas de O son más difíciles que las relativas de S, porque contienen las primeras un mayor número de dependencias temporalmente incompletas. Estas dependencias restringen tanto los recursos de almacenamiento como de integración de la información (Gibson, 1998, 2000; Lewis y Vasishth, 2005; Lewis y otros, 2006). La complejidad de procesamiento en las relativas de O también se ha explicado como consecuencia de la necesidad del lector de cambiar de perspectiva desde una primera interpretación, que en idiomas con un orden canónico (Sujeto + Verbo + Objeto, en adelante SVO) como el español o el inglés se asienta en la predicción de la presencia de un agente en posición inicial, a una segunda interpretación ante la presencia del orden no canónico de las relativas de O (MacWhinney y Pleh, 1988; Del Río y otros, 2012). En el mismo sentido, autores como Caplan y Waters (1999) lo han descrito como un orden no canónico de los roles temáticos.

Por otro lado, la dificultad de procesamiento de cláusulas de relativo ha sido ligada a la edad de los participantes (en algunos casos considerando la relación del envejecimiento a la capacidad de la memoria operativa y la velocidad de procesamiento) como indican estudios que comparan adultos jóvenes con adultos mayores y ancianos (por ej., Stine-Morrow y otros, 2000; Waters y Caplan, 2001; DeDe y otros, 2004; Caplan y Waters, 2005; Caplan y otros, 2011; en el español: Arancibia y otros, 2014)². Explicaciones basadas en la fuerza de las representaciones lingüísticas, por su parte, proponen que la experiencia del usuario de la lengua y las estadísticas resultantes influyen el modo en que se predicen los constituyentes (Reali y Christiansen, 2007a y 2007b; Wells y otros, 2009; Reali, 2014). En resumen, la evidencia acumulada apunta hacia la conclusión de que las dificultades en el procesamiento de las relativas de O podrían deberse a una combinación de diferentes factores morfosintácticos, funcionales, discursivos, semánticos y basados en la experiencia. El presente estudio se enfoca en los factores morfosintácticos.

2 La literatura proporciona resultados mixtos con respecto a los efectos de la edad en el análisis y la interpretación de las estructuras sintácticas complejas como las relativas de S y O.

Entre los factores sintácticos que se señalan para explicar la ventaja en el procesamiento de las relativas de S aparece la presencia de orden canónico (SVO) y la presencia de una marca morfosintáctica. El presente estudio analiza el rol de la sintaxis en el procesamiento de las cláusulas de relativo mediante el estudio de distintas alternativas en el orden de palabras y la presencia de marca morfosintáctica. Este trabajo es de naturaleza exploratoria, ya que consiste principalmente en tareas *offline* en las que por primera vez se incluyen condiciones no consideradas en estudios previos. Como punto de partida, también se ofrecen datos sobre las frecuencias de uso de las subordinadas relativas en un corpus escrito, lo que permite una aproximación hacia la puesta a prueba de hipótesis que asumen facilitación o dificultad en el procesamiento basadas en la experiencia (patrones de uso).

1.1. Procesamiento de las subordinadas relativas de S y O y los factores morfosintácticos

Los primeros trabajos específicos sobre el procesamiento de las relativas de S y O llegan de la mano de Fodor y otros (1974), y Wanner y Maratsos (1978). Posteriormente, autores como King y Just (1991), utilizando técnicas conductuales, demostraron que en inglés las relativas con extracción de S son más fáciles de procesar que aquellas con extracción de O. Los autores interpretaron este resultado asumiendo que el procesamiento de las primeras requiere una menor carga en la memoria de trabajo u operativa respecto de las segundas. Esto también fue evidenciado por MacWinney y Pleh (1988) en el inglés, y en idiomas cercanos como el alemán (Mak y otros, 2002, 2006).

En francés, y por medio de una presentación auditiva de las oraciones, Frauenfelder y otros (1980) estudiaron subordinadas ‘reversibles’ con idéntico orden superficial [complementante+verbo+sintagma nominal (SN)]. Mientras que la presencia del complementante *qui*, ‘quien’, reemplazaba al sujeto de la subordinada, resultando en el orden canónico SVO en una de las condiciones, el complementante *que*, ‘que’, reemplazaba el objeto directo (OD), resultando en el orden no-canónico OVS en la otra condición. Es decir, los dos órdenes de palabras respondían a las relativas con extracción de S y de O, respectivamente. Los resultados demostraron que las oraciones relativas con extracción de S son más fáciles de comprender que las relativas con extracción de O. Holmes y O'Regan (1981) también hallaron mayores dificultades de lecturas en relativas con extracción de S en comparación con la de extracción de objetos mediante estudios de movimientos oculares, siendo que estas últimas generaron mayores movimientos regresivos.

Resultados similares fueron los obtenidos por Betancort y otros (2009), quienes, aunque su estudio se enfocaba en buscar efectos de animacidad (rasgo semántico) en el procesamiento de relativas, reportaron que las relativas de O causaron más dificultades que las relativas de S. Los autores observaron que el factor animacidad modula esta preferencia. En particular,

el costo de procesamiento de las relativas de O aumenta cuando ambos Sintagmas Nominales (SN) (es decir, dentro y fuera de la relativa) son animados, y disminuye cuando únicamente solo un SN (el de la relativa) es animado. Los trabajos de Frauenfelder y otros (1980) y Betancort y otros (2009) son esclarecedores, porque explotan la flexibilidad en el orden de las palabras de idiomas como el francés y del español, que permiten manipular relativas de S y de O sin alterar el orden superficial [complementante+verbo+SN]. Lo que desambiguaba unas de otras en esos estudios era la marca (el pronombre) de relativo (*qui* o *que*), en el estudio de francés, y la marca diferencial de objeto directo (a o Ø) en el estudio en español.

Otros estudios han aprovechado no solo el solapamiento del orden superficial entre relativas S y O, sino también variantes posibles en el orden superficial, como así también la posibilidad de contar con la presencia de marcas de caso y de concordancia de género y número. Estas marcas morfosintácticas podrían servir para desambiguar el rol de los SN de la subordinada, facilitando así su procesamiento. Por ejemplo, Del Río y López-Higes (2006) contrastaron el procesamiento de relativas de S y de O. Pero, mientras que en las primeras el orden era [complementante+verbo+SN o SVO], en las de O el orden era [complementante+SN+verbo o OSV]. Los autores sostienen que la diferencia entre las relativas que siguen el orden canónico y aquellas que no lo siguen es que, mientras el procesamiento de las primeras sigue un curso incremental (cada palabra nueva es integrada a medida que se va encontrando), en el caso de las segundas es necesario mantener durante un breve espacio de tiempo información sobre el sujeto de la cláusula principal hasta que le pueda ser asignado el papel temático que le corresponda en la cláusula subordinada de relativo. Más recientemente, Del Río y otros (2012) contrastaron relativas de S (con orden canónico SVO) con relativas de O con orden OSV y OVS. Lo interesante de las relativas de O utilizadas es que los pronombres relativos usados en estas eran *el que* o *la que*. Según explica Schroten (1987: 42-3), el español actual ya no permite el uso de estos pronombres relativos, a menos que se introduzcan con la preposición 'a', que es la marca de OD humanos y definidos. Por eso, el ejemplo (3) no es gramatical en el español actual, a diferencia del ejemplo (4) que sí lo es.

(3) *El chico₁ [el que_{1Objeto} saludaste] es vecino mío.

(4) El chico₁ [a₁ que_{1Objeto} saludaste] es vecino mío.

Entonces, junto con el pronombre relativo, Del Río y otros (2012) también agregaron la marca de caso acusativo diferencial ('a') y un artículo definido ('el'), por lo que se podría decir que, a diferencia de estudios previos en los que solo se empleaba el complementante *que*, en su estudio las relativas de O contenían marcas morfosintácticas adicionales. Ejemplos de las relativas de O que su estudio incluyó se presentan en (5) y (6).

(5) El policía₁ [al que_{1Objeto} el comisario criticó] vive en el barrio. (OSV)

(6) El policía₁ [al que_{1Objeto} criticó el comisario] vive en el barrio. (OVS)

Del Río y otros (2012) hallaron tiempos de lectura más largos en el inicio de las relativas de O en el nivel de la marca del pronombre relativo en comparación con las relativas S. Esto fue interpretado como consecuencia del pasaje de una primera perspectiva de orden canónico (SVO) en las primeras a otra de orden no canónico (OVS y OSV) en las segundas. Igualmente, encontraron diferencias en los tiempos de lectura entre estas últimas a nivel del verbo matriz. Las relativas de O con orden OSV produjeron tiempos de lectura más largos que aquellas con orden OVS. Los autores ofrecen hipótesis tentativas para el tiempo de lectura mayor en el verbo matriz luego de relativas de O con orden OSV que tienen que ver con la presencia de dos SN contiguos que podrían acrecentar la interferencia entre ellos al momento de la integración, y la presencia del verbo de la subordinada en posición final, lo cual incurriría en un costo en la asignación de roles temáticos.

El estudio más completo en términos de las variables morfosintácticas se realizó en ruso, dada la flexibilidad de ese idioma. Levy y otros (2013) pudieron manipular no solo el tipo de extracción (S u O), sino también el orden de palabras al interior de ambas condiciones y la presencia de marca morfosintáctica en las de O. Sus resultados indicaron menor porcentaje de acierto para las relativas de O, sin orden canónico y con un tipo de marca de caso sincretizada (*chto*, 'que'). La mayor proporción de acierto en la comprensión se encontró en las relativas de S en ambas condiciones (con y sin orden canónico) y con la marca que indica caso (*kotoryj*, 'quien' —caso Nominativo—, o *kotorogo*, 'a quien' —caso Acusativo—).

A pesar de la literatura previa revisada anteriormente, un estudio que incluya más de dos variantes en la presencia de marcas morfosintácticas aún no existe en español, tal como se pone en evidencia en la tabla 1.

TABLA 1

Comparación de estructuras incluidas en estudios de relativas en español

Estructura (E)	Betancort y otros (2009)	Del Río y López-Higes (2006)	Del Río y otros (2012)	Presente estudio
E1: La niña que mira al hombre cruza la calle. (S)	•	•	•	•
E2: La niña que mira el hombre cruza la calle. (O)	•			•
E3: La niña a la que mira el hombre cruza la calle. (O)			•	•
E4: La niña que el hombre mira cruza la calle. (O)		•		•
E5: La niña a la que el hombre mira cruza la calle. (O)			•	•

Este trabajo se propuso esclarecer el nivel de dificultad en la comprensión de relativas de objeto del español en comparación con las relativas de sujeto en función de dos factores morfosintácticos: el orden de palabras y marca de caso.

El estudio exploratorio que aquí se presenta es el primero que se aproxima a una comparación de la facilidad de procesamiento en relativas de S y O. Lo hace mediante tareas *offline* y manipula dentro de estas relativas de O no solo el orden de palabras, sino también la presencia o ausencia de marcas morfosintácticas adicionales. En la sección siguiente se describen las estructuras seleccionadas y se presentan datos de corpus que dan cuenta de su uso en la lengua escrita. Luego se presentan dos estudios experimentales que buscan proveer una primera aproximación al estudio de todas estas estructuras.

2. Las subordinadas relativas de S y O en español y su frecuencia en el habla escrita

Se ha elegido trabajar con oraciones subordinadas relativas de S y de O de persona (o con el rasgo semántico de +Animacidad), por la variedad de alternativas que estas presentan en la lengua española. En este sentido el español presenta una importante flexibilidad en la estructura superficial que se manifiesta en las diferencias de orden que asumen los distintos constituyentes oracionales.

De la revisión precedente, puede observarse que la mayor parte de los estudios efectuados en este dominio solo analizaron en detalle la oposición entre las relativas de extracción de S y O sin explorar la totalidad de las distintas alternativas de extracción de O que posibilita la flexibilidad sintáctica del español. Por lo tanto, en este trabajo se estudió dicha oposición (extracción de S y O), incluyendo diferentes estructuras oracionales con relativas de extracción de O en las que varían el orden canónico y la marca morfosintáctica. Se tomaron los siguientes tipos de estructuras (en adelante, E):

- E1: La niña **que mira al hombre** cruza la calle. (S)
- E2: La niña **que mira el hombre** cruza la calle. (O)
- E3: La niña **a la que mira el hombre** cruza la calle. (O)
- E4: La niña **que el hombre mira** cruza la calle. (O)
- E5: La niña **a la que el hombre mira** cruza la calle. (O)

Como puede observarse una sola de las estructuras corresponde a la extracción de S (E1), mientras que las otras 4 son de extracción de O (E2-E5), con variación de orden canónico en

español SVO (lo sigue o no lo sigue) y de presencia de marcas morfosintácticas adicionales³ (presencia o ausencia). En todos los casos, dentro de la relativa se elide o reemplaza el componente *la niña* con el pronombre relativo *que*. En algunos casos, *que* corresponde al S de la relativa (E1), mientras que, en otros, al O de la relativa (E2-5). Es importante mencionar que todas las cláusulas de relativo tenían incrustación central, por lo tanto, este no fue considerado como un factor de complejidad (Miller y Chomsky, 1963; Frazier y Rayner, 1988).

Entonces, la extracción de S o de O es uno de los factores que se analizarán en el presente trabajo. La relativa de extracción de S (E1) presenta un orden canónico y no posee marca sintáctica. Su contraparte se presenta en la E4, relativa de extracción de O con orden canónico y sin marca sintáctica. Dentro de las relativas con extracción de O, se conformaron estructuras que presentan un orden canónico (E4 y E5) y otras presentan un orden no-canónico OVS (E2 y E3). Además, E3 y E5 contienen marcas morfosintácticas como la preposición *a* y el artículo *la* que preceden el pronombre relativo, resultando en versiones más marcadas de E2 y E4, respectivamente. La presencia de la marca preposicional (típica de ciertos objetos sintácticos, OD animados y OI) podría desambiguar el tipo de cláusula según su extracción. Además, el artículo especifica el género y número del constituyente elidido (o reemplazado). Es decir, construcciones como *al que*, *a la que* o *a los que* son más informativos que la sola presencia del pronombre relativo *que* presentes en las E2 y E4. De lo que se examinó y presentó precedentemente, emergen hipótesis en relación con los tres principales factores de interés en este trabajo, que pueden ser comprobadas gracias a las cinco E que el español permite.

Hipótesis 1: Extracción de S/O. Replicando estudios previos, las relativas de S (E1) serán más fáciles de procesar que las de O (E2-5), principalmente en la comparación de E1 y E4.

Hipótesis 2: Orden canónico. Dado que el orden canónico del español es SVO, se hipotetiza que las estructuras que presentan este orden (E1, E4 y E5) serán más fáciles de procesar que las que presenten un orden no canónico OVS (E2 y E3).

Hipótesis 3: Marca morfosintáctica. La presencia de marca morfosintáctica facilitará las relativas con extracción de O (E3 y E5).

En resumen, se prevé que la oración con la subordinada de extracción de S tendrá una mayor comprensión que las oraciones con relativas de extracción de O, y dentro de las oraciones que poseen subordinadas de extracción de O, la presencia del orden canónico y de marcas morfosintácticas facilitarán la comprensión. Las propiedades que dificultarían la compre-

3 Existe un tipo de marca sintáctica que no se ha considerado para este trabajo, debido al escaso uso que presenta en el español actual: el relativo “quien”, cuya estructura resultante sería una de tipo: “*La niña quien mira al hombre cruza la calle*”.

sión en las E que tienen extracción de O serían el orden no canónico y la ausencia de marca morfosintáctica. Por lo tanto, se puede suponer un *continuo* de mayor a menor facilidad en la comprensión, en el cual E1 estaría ubicada en el extremo de mayor facilidad y dentro de las relativas de objeto sería: $E_4 > E_5 > E_3 > E_2$, siendo la E2 la más difícil de comprender.

Como se manifestó anteriormente, la contraparte de la E1 es E4, ya que en esta última también se respeta el orden canónico y no posee marca sintáctica. Sin embargo, siguiendo los planteamientos de los trabajos antecedentes en relación a cómo el curso incremental es un factor independiente del orden canónico, E4 sería más compleja que E1.

Con respecto a las relativas con extracción de O, E3 y, E5 el factor que las distingue es el orden canónico (ya que ambas tienen marca sintáctica), por lo que se puede suponer que la estructura que no lo sigue (E3) será más difícil de comprender. Y aquella estructura con extracción de O que no sigue el orden canónico y no posee marca sintáctica, E2, será la de mayor dificultad.

Por lo tanto, se conjetura que el orden de dificultad en la comprensión que se presupone para las cinco E es como sigue (de menor a mayor dificultad): $E_1 > E_4 > E_5 > E_3 > E_2$.

En resumen, basándonos en las diferencias encontradas en el procesamiento de las relativas de S y O en el estudio de Río y otros (2012), el presente trabajo examina la facilidad o dificultad de procesamiento (correcta o incorrecta comprensión) de las cláusulas de relativo de S y O, incluyendo las distintas alternativas de relativa de O que alteran el orden canónico dentro de la subordinada. Esto permite evaluar hipótesis más específicas sobre el rol de factores sintácticos (como la presencia de marca sintáctica), controlando el orden de palabras.

Para conocer si las cinco estructuras conformadas eran encontradas en el uso del lenguaje escrito, y con qué frecuencia, se llevó a cabo un análisis sobre un corpus local (del español de Córdoba, Argentina)⁴ obtenido de la versión *on-line* de La Voz del Interior (un diario local, de mayor venta en la provincia de Córdoba, Argentina, <http://www.lavoz.com.ar/>), de 5 millones de palabras, construido por el Grupo de Procesamiento del Lenguaje Natural de la Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Argentina (<http://pln.famaf.unc.edu.ar/>), y analizado automáticamente mediante el analizador morfosintáctico Freeling 3.1, que tiene un error del 2% para el análisis morfosintáctico (Atserias y otros, 2006). Las cláusulas de relativo fueron buscadas a través de patrones superficiales detectables auto-

4 Estudios similares han sido realizados en otros trabajos sobre el procesamiento de las cláusulas de relativo, como por ejemplo Carreiras y otros (2010).

máticamente sobre el análisis morfológico de Freeling. Se detectaron todos los pronombres relativos en el corpus, descartando los “que” analizados como conjunciones subordinantes. Para cada oración con un pronombre relativo, se examinó que se encontrara alguno de los patrones asociados a alguna de las estructuras del estudio. El 55% de los casos de relativas pudo ser clasificado en alguno de los patrones predeterminados para cada E. No obstante, E1 y E2 quedaron expresadas con la misma secuencia: [PR “que”- V - SN] debido a que no se pudo encontrar un patrón basado en análisis morfosintáctico que pudiera distinguirlos, ya que el *software* utilizado tiene un nivel de error muy alto en el análisis de funciones sintácticas (40%). Más específicamente, las relativas de extracción de S y O del español de Córdoba se pueden hacer casi con las mismas secuencias morfo-sintácticas, como se muestra a continuación: “*La niña que mira el monumento cruza la calle*” (E1); “*La niña que mira el hombre cruza la calle*” (E2).

En este caso, no se puede distinguir automáticamente el núcleo animado del inanimado, que es lo que diferencia la función de O del sintagma nominal “el monumento” del sintagma nominal “el hombre”, que tiene función de S.

En cambio, con la presencia de la preposición “a” que introduce a los objetos animados, sí se pudieron distinguir de forma no ambigua algunas de las oraciones E1. Así, E1, en forma no ambigua, se pudo expresar y buscar de la siguiente manera: [PR “que” - V - SPrep]. E3 quedó expresada en la secuencia [CR “al que/a la que” - V - SN]; E4: [PR “que”- SN - V]; y E5: [CR “al que/a la que” - SN - V]. Las relativas que no pudieron ser clasificadas como pertenecientes a algunas de las estructuras definidas para este trabajo fueron: [PR “que”- SN- sin V](39687 casos: 38%), [PR “que” - SPrep] (2.525 casos: 2,4%); otras estructuras son aquellas en las que PR va seguido de coma (4058 casos: 4%). A partir de esto, se obtuvieron los datos de frecuencia de ocurrencia para cada estructura en el corpus. Dichos datos y las características de las estructuras se muestran en la tabla 2.

La inspección de la tabla 2 muestra que los cinco tipos de estructura se encuentran en el lenguaje escrito, no obstante, la ocurrencia de casos de E2 no resulta clara mediante los datos arrojados. La estructura que tiene mayor frecuencia es E1 (relativa con extracción de S), tanto en combinación con E2 como de manera individual. Luego continúan la E4, E3 y E5, siendo estas dos últimas las que presentan menor ocurrencia de casos.

En relación a las ocurrencias de relativo de S versus O, los datos del corpus muestran un mayor porcentaje de relativas de S (E1: 9%) respecto de las relativas de O (E3, E4 y E5: 8%). Sin embargo, las estructuras ambiguas ocupan el mayor porcentaje (E1 y E2: 38%).

En base a los casos detectados por el analizador morfosintáctico sobre la ocurrencia de las E de relativo en un corpus local, las frecuencias encontradas ($E1 > E4 > E3 > E5$) son muy cercanas al *continuo* planteado inicialmente en relación a la dificultad en la comprensión:

TABLA 2

Secuencia de búsqueda, número y porcentaje (%) de casos de relativas que pudieron ser clasificadas en algunas de las estructuras estudiadas y ejemplos del corpus

Estructura	Marca y orden de palabras	Ejemplos del corpus	N de casos (%)
1 y 2	“que”- V - SN	<i>Me cuesta comprender las posturas políticas que buscan frenar acciones ejecutivas sólo por la especulación electoral.</i>	38.888 (38%)
1 (S)	“que” - V -Sprep	<i>Hay dealers que utilizan a los remises como “pantallas”.</i>	9.692 (9%)
4 (O)	“que”- SN - V	<i>La APE repartía los fondos de acuerdo a la sintonía que los gremios tenían con el Gobierno.</i>	7.229 (7%)
3 (O)	“al/a la/a los/a las que” - V - SN	<i>Veo el maltrato hacia los abuelos por parte de empleados muy mal pagos, a los que apenas les abona 1.400 pesos.</i>	917 (0,8%)
5 (O)	“al/a la/a los/a las que” - SN - V	<i>No tratar de caer simpático con aquellos a los que Kirchner considera enemigos.</i>	198 (0,2%)

$E1 > E4 > E5 > E3 > E2$. La diferencia en el orden de las estructuras se da en las relativas de O, la E3 y E5. Sin embargo, la diferencia es mínima (0,8% y 0,2%, respectivamente), y la E2 no pudo ubicarse, dadas las limitaciones del análisis automático.

Por lo tanto, a través del análisis de corpus se pudo verificar el uso de estas estructuras en el lenguaje escrito, y su frecuencia de ocurrencia reveló estar cercana al continuo hipotético sobre la facilidad/dificultad de comprensión de las E, según la revisión de la literatura.

3. Método

Para examinar la facilidad o dificultad en el procesamiento de la cláusula de relativo del español con elisión de Sujeto y Objeto, incluyendo las distintas alternativas de relativa de O, se realizaron dos estudios (estudio 1 y 2) que incluyeron mediciones con y sin límites de tiempo, ya que los datos obtenidos en estas dos condiciones pueden diferir de manera importante. En el estudio 1, los juicios solicitados a los sujetos son hechos bajo condiciones “ideales” de procesamiento que reducen el impacto de las limitaciones de integración y de memoria de trabajo. En el estudio 2, las medidas son efectuadas en condiciones que exigen del sujeto un procesamiento acelerado de integración y de memorización de la información (sobre este punto, ver Lewis y Phillips, 2015).

En el primer estudio, se buscó analizar la comprensión espontánea sin límite de tiempo y corroborar si los sujetos eran capaces de asignar los roles temáticos de manera correcta. El segundo estudio consistió en medir el nivel de acierto o error en la selección de dos alternativas de interpretación de las oraciones (donde una era correcta y la otra no) con límite de tiempo, es decir, los sujetos debían comprender y elegir la opción correcta lo más rápido posible (elección forzada con límite de tiempo).

En ambos estudios se tomó la cantidad de errores cometidos en la tarea de comprensión como indicador de la facilitación o no. Para identificar una correcta comprensión de la cláusula de relativo, solo es necesario hacerse la pregunta “¿quién hizo qué a quién?”, para la que hay solo dos respuestas posibles y opuestas. Por ejemplo, para la oración “*La niña que mira el hombre cruza la calle*”, las respuestas posibles a esa pregunta son “El hombre mira a la niña” o “La niña mira al hombre”.

3.1. Estudio 1. Comprensión visual espontánea (sin límite de tiempo)

3.1.1. Metodología

3.1.1.1. Participantes

Intervinieron voluntariamente 43 sujetos con nivel terciario y universitario, con un rango de 23 a 68 años ($M = 32$ años)⁵. Todos eran hablantes nativos del español, tenían una visión normal o corregida por el uso de lentes y ninguno tenía historia de problemas o alteraciones del lenguaje.

3.1.1.2. Material

Se crearon 28 oraciones con relativas de extracción de S y O según las 5 estructuras. Los verbos de las relativas fueron transitivos para que ante la pregunta “¿quién hizo qué a quién?” sea posible concebir la idea de que el agente sea uno u otro SN. En la tabla 3 se presenta la descripción.

5 Como ya se mencionó en la introducción, hay trabajo que reporta mayor dificultad de procesamiento de relativas, ya sea como un efecto lineal de la edad de los participantes, es decir, que, a mayor edad, mayor dificultad (por ejemplo, Caplan y otros, 2011), o en sujetos de > 65 años de edad (por ejemplo, Arancibia y otros, 2014). Dado los sujetos que participaron en el presente estudio, hay un solo participante que excede los 60 años y la mayoría ($N=30$) tiene < 60 años, por lo que no se esperan efectos de edad. No obstante, se realizó un análisis que excluyó a los participantes de > 32 años de edad ($N=13$) que arrojó resultados similares a los que se reportan. Por lo tanto, para un acabado conocimiento sobre los efectos de la edad de los participantes en la comprensión de cláusulas de relativo, requeriría una muestra más amplia que la que se tomó para el presente estudio.

TABLA 3

Descripción de cada estructura según las alternativas adoptadas para cada factor

Estructura (E)	Extracción		Orden canónico		Marca sintáctica		Descripción
	S	O	Sí	No	Sí	No	
1	•		•			•	SN que (S elidido) V al SN
2		•		•		•	SN que (OD elidido) V el SN
3		•		•	•		SN a la que (OD elidido) V el SN
4		•	•			•	SN que (OD elidido) SN V
5		•	•		•		SN a la que (OD elidido) SN V

3.1.1.3. Diseño y procedimiento

La condición tipo de estructura y factor fue una variable intrasujetos. La variable dependiente fue el número de errores. A cada uno de los participantes se les presentó en forma escrita (en 4 hojas) cada una de las oraciones (escritas en calibri 11, con un interlineado de 1,5) ordenadas de forma aleatoria. En la prueba aparecía cada oración enumerada de 1 a 28 y, debajo de cada una, la pregunta “¿quién hizo qué a quién?”, seguida de una línea de puntos para completar la respuesta:

1. La niña que mira al hombre cruza la calle. ¿Quién mira a quién?

La consigna fue: “Por favor, lea atentamente cada una de las oraciones y responda la pregunta que le sigue”. Los participantes leyeron cada una de las oraciones silenciosamente (la cantidad de veces que quisieron hasta comprenderlas) e iban completando las respuestas, siguiendo el orden de las oraciones. Realizaron la tarea individualmente, entre 30 y 40 minutos, en una habitación tranquila y con buena iluminación.

3.1.2. Resultados

El análisis descriptivo de los errores mostró un porcentaje total del 4,1% (49 errores) para las 5 estructuras. El gráfico 1 indica el porcentaje de errores para cada una de las E.

Los resultados descriptivos muestran que el mayor porcentaje de errores ocurrió en las oraciones con E2 y, en menor medida, con E3. Las tres estructuras restantes (E1, E4 y E5) no presentaron mayores dificultades para su comprensión. Luego se realizó el mismo análisis, pero discriminado por factor (ver gráfico 2).

GRÁFICO 1

Porcentaje de errores de cada una de las estructuras utilizadas

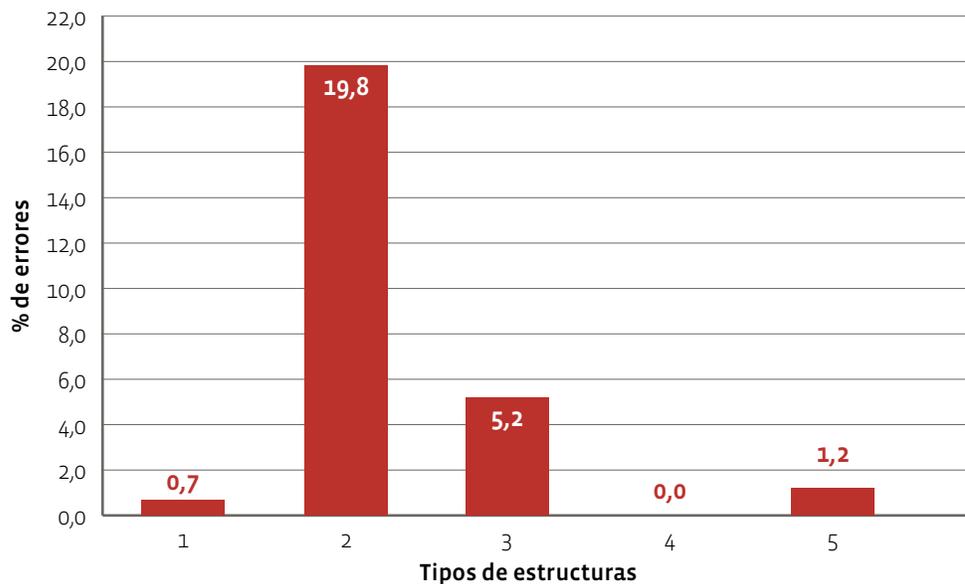
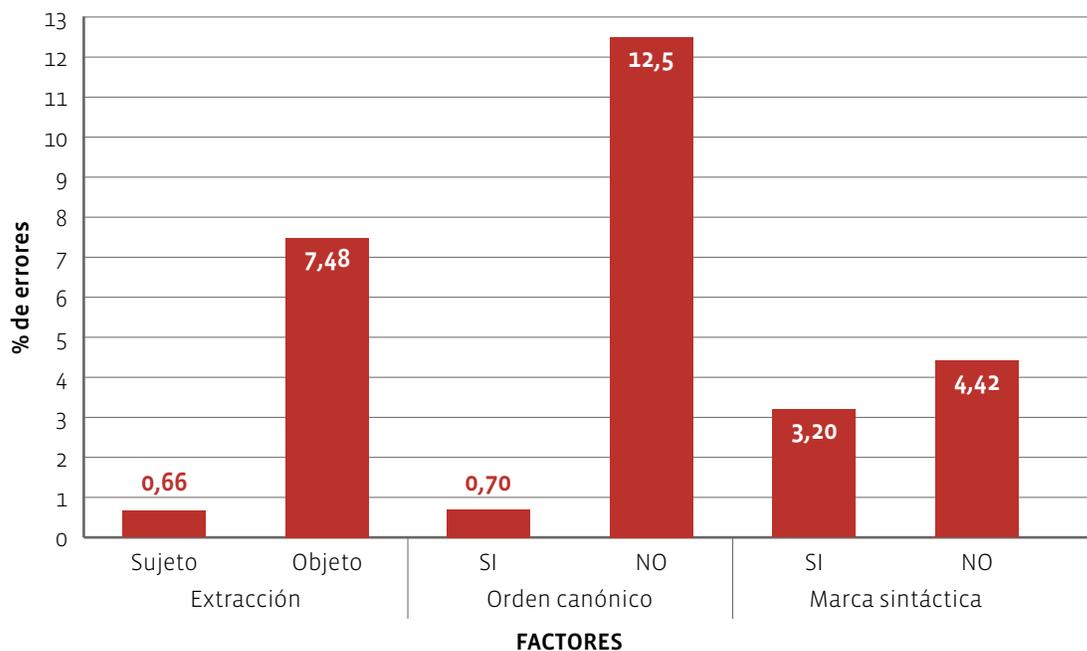


GRÁFICO 2

Porcentaje de errores para cada factor



Como se puede percibir en el gráfico, las oraciones que poseen el factor extracción de O resultan más problemáticas en la tarea de asignar funciones de agente y paciente, en comparación a las relativas de extracción de S. Con respecto al orden canónico, el gráfico muestra una mayor cantidad de errores para las oraciones que no siguen el orden canónico (SVO). Por último, y en relación con la marca sintáctica, se ve un porcentaje de errores apenas mayor para las oraciones que no presentan marca sintáctica. A nivel general, se hallaron las mayores proporciones de errores en aquellas relativas con extracción de O que no siguen el orden canónico, que se manifiestan como las más difíciles de comprender.

Si se examinan los errores de las estructuras con extracción de S (E1) con su contraparte en las oraciones con extracción de O (E4), se puede observar que la diferencia en errores es mínima. Por tanto, cuando solo se varía el factor tipo de extracción pero los demás factores se mantienen constantes, no se encuentra gran diferencia en la dificultad de comprensión, expresada por el porcentaje de errores. Pero cuando se varían otros factores, como el orden canónico y la marca sintáctica, ahí la comprensión pareciera ser más dificultosa para las oraciones con extracción de O en este tipo de pruebas.

Cuando se analizan solamente las estructuras con extracción de O (E2, E3, E4 y E5) se puede percibir que, si agrupamos las estructuras que siguen el orden canónico (E4 y E5) y no (E2 y E3), el no seguir el orden canónico en las relativas con extracción de O hace mucho más difícil su comprensión (1,2% versus 25% de errores). Y si agrupamos las estructuras por marca morfosintáctica (con E3 y E5 y sin E2 y E4), el discernimiento sobre la dificultad en la comprensión se hace menos claro, con menor porcentaje de errores para las estructuras con marca morfosintáctica.

3.1.3. Discusión

Los datos obtenidos en el estudio de comprensión visual espontánea (sin límite de tiempo) permiten establecer el siguiente orden (de menor a mayor dificultad): $E4 > E1 > E5 > E3 > E2$, donde E4 y E1 están invertidas respecto de la hipótesis del trabajo y el porcentaje de frecuencia de ocurrencia obtenido en el estudio de corpus. Los valores de diferencia, no obstante, son poco importantes.

Cabe destacar que las estructuras E1 y E4 presentaron el menor porcentaje de errores (siendo E4 la única estructura que no presentó errores), lo que permite hablar de una sintaxis transparente y directa en estas dos estructuras. Lo más llamativo es que ambas estructuras solo difieren en el tipo de extracción, siendo E1 de S y E4 de O. Por lo tanto, en este tipo de pruebas (que reducen el impacto de las limitaciones de integración y de memoria de trabajo), cuando los otros factores están igualados (orden canónico y marca sintáctica), no parecen

diferenciarse en la complejidad de procesamiento. La E4 introduce un SN explícito inmediatamente, que es diferente al SN antecedente (el de la oración principal), con lo que probablemente se anule la posibilidad de que este último pueda ser el S de dicha relativa. Esto está asegurado por el orden canónico que adopta la relativa.

Hacia el final de la secuencia, se observa también que E2 es la estructura más compleja, con una brecha muy abrupta en comparación con la estructura que le sigue en cantidad de errores (14,6%), mostrando que, cuando se varía el orden canónico y la marca sintáctica en las oraciones con extracción de O, el no seguir el orden canónico y no tener marca sintáctica hacen más compleja su comprensión, incluso en este tipo de pruebas donde el sujeto tiene la posibilidad de relectura de las oraciones hasta llegar a su comprensión, reduciendo el esfuerzo que debe realizar para mantener e integrar la información.

En relación al análisis según factores, efectivamente los factores planteados como facilitadores ofrecieron resultados acordes a lo presumido y los antecedentes explicitados. El orden canónico se muestra como el factor más relevante en la facilitación, con un valor mínimo de errores en las relativas que adoptan el orden canónico (0,7%) versus orden no canónico (12,5%), generando una diferencia considerable (11,8 puntos). La extracción de S versus O también ofrece resultados reveladores en cuanto a su facilitación (0,66% vs. 7,48%, respectivamente), generando una diferencia de 6,82 puntos. Por último, la presencia de marca sintáctica, si bien se presenta como un facilitador, no ofrece una diferencia tan clara (1,22 puntos) respecto de la ausencia de dicha marca.

Para corroborar los datos encontrados en la tarea de comprensión visual sin límite de tiempo, se determinó realizar una segunda tarea con condiciones más semejantes a la comprensión *on-line*. En este segundo estudio, los sujetos debían comprender las oraciones y dar una respuesta mediante la elección de una opción con un límite de tiempo determinado. De esta forma, se evita la relectura de las oraciones y su comparación con respuestas anteriores, lo que permite acercarse a la forma de procesamiento secuencial habitual y observar cómo estas condiciones influyen en la comprensión de las diferentes estructuras.

3.2. Estudio 2. Comprensión visual con elección forzada y límite de tiempo

Con este segundo estudio se buscó estudiar si el mismo patrón de resultados encontrado en el estudio 1 se manifiesta cuando los sujetos deben realizar la comprensión visual lo más rápido posible y mediante la elección forzada sobre dos alternativas, donde se examinó la influencia del factor tiempo sobre la comprensión.

3.2.1. Metodología

3.2.1.1. Participantes

Veintisiete egresados y estudiantes de la Facultad de Psicología, UNC, participaron del experimento. El rango de edad fue de 18 a 32 años ($M= 23$ años). Todos ellos tenían una visión normal o corregida por el uso de lentes, eran hablantes nativos argentinos y ninguno tenía historia de problemas o alteraciones del lenguaje. Participaron voluntariamente en la realización de la prueba y ninguno de ellos participó en el estudio 1.

3.2.1.2. Material

Se utilizaron las mismas 28 oraciones del estudio 1. Para cada oración se crearon dos posibilidades de respuesta que respondían a la pregunta: “¿Quién hizo qué a quién?” Por ejemplo: “El perro al que mira el gato bebe agua”. Respuestas posibles: a) “El perro mira al gato”; b) “El gato mira al perro”.

3.2.1.3. Diseño y procedimiento

La condición tipo de estructura y factor fue una variable intrasujetos. La variable dependiente fue el número de errores. Los participantes realizaron la experiencia en forma individual, en una sesión que duró 20 minutos. El sujeto se sentaba cómodamente frente a la computadora, en una habitación con buena iluminación y aislada de ruidos. Todas las oraciones fueron presentadas en el medio de la pantalla de una computadora personal con una tasa de barrido de 60-Hz. El *software* utilizado para la presentación y detección de los errores fue el DMDX (Forster y Forster, 2003). Las oraciones y sus opciones de respuesta A y B (ubicadas debajo de la oración) se presentaban en un rectángulo blanco de 15,5 cm de ancho y 10 cm de alto, sobre un fondo negro.

Cada oración aparecía en el medio del rectángulo, escrita en minúscula, en calibri tamaño 36 y en estilo negrita. Las opciones de respuesta aparecían debajo de la oración en minúscula en letra calibri tamaño 28. El orden de presentación fue aleatorio. En las opciones de respuesta, se alternó el orden de S y O, es decir, para la mitad de las respuestas, el O como agente de la acción estaba en la opción A, mientras que para la mitad restante se encontraba en la B. Inversamente fue realizado con el S como agente de la acción.

La secuencia experimental fue la siguiente: primero se presentaba un signo numeral (#) en el medio de la pantalla, durante 200 milisegundos (ms). Luego se remplazaba por el rectángulo donde se presentaba la oración con sus dos alternativas de respuesta. Esto permanecía

en la pantalla hasta que el sujeto daba su respuesta o con un tiempo máximo de 20.000 ms. Si el sujeto respondía en ese tiempo, se pasaba a la siguiente secuencia (con un intervalo de 1.000 ms. entre secuencias). La consigna oral dada a los sujetos fue: “Lea atentamente la oración presentada, seleccione la opción correcta, A o B, y presione el botón correspondiente. Siempre lo más rápido posible, pero sin cometer errores”. Se brindó a los sujetos una botonera de respuesta, donde un botón correspondía a la opción A y otro botón a la opción B. La mitad de los sujetos respondió a la opción A con el botón derecho y la opción B con el botón izquierdo, y la otra mitad lo hizo a la inversa.

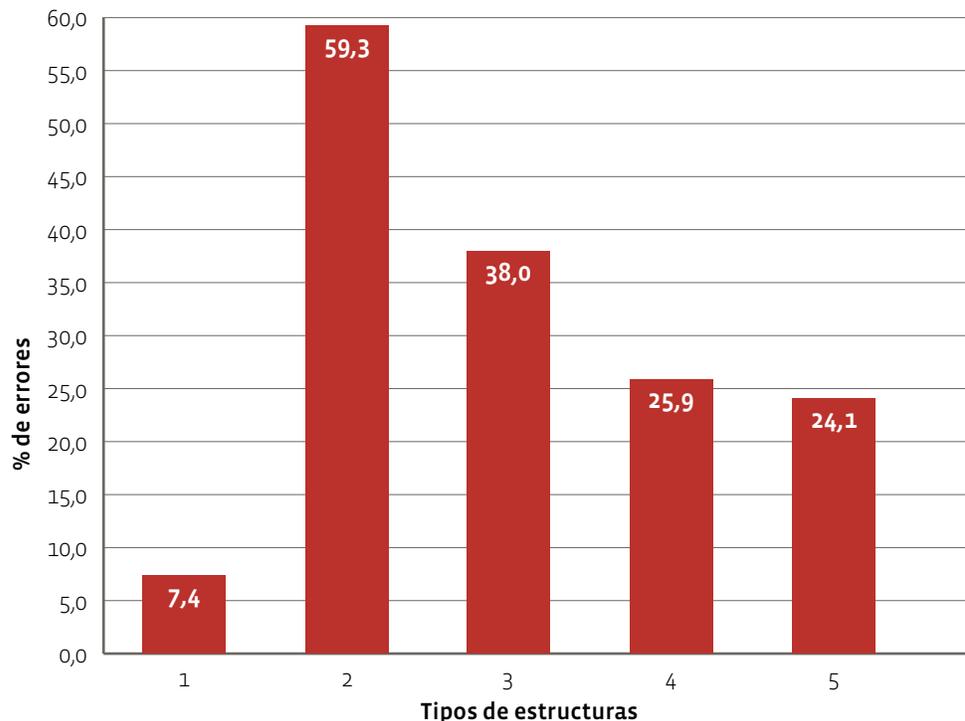
Además, el experimento fue precedido por una fase de práctica que consistía en seis oraciones nuevas con las mismas características del material experimental.

3.2.2. Resultados

Todos los valores superiores a 20 segundos y la elección de la opción incorrecta fueron considerados errores: porcentaje total 23% (173 errores). El gráfico 3 indica el porcentaje de errores correspondiente a cada una de las estructuras examinadas.

GRÁFICO 3

Porcentaje de errores para cada una de las estructuras utilizadas

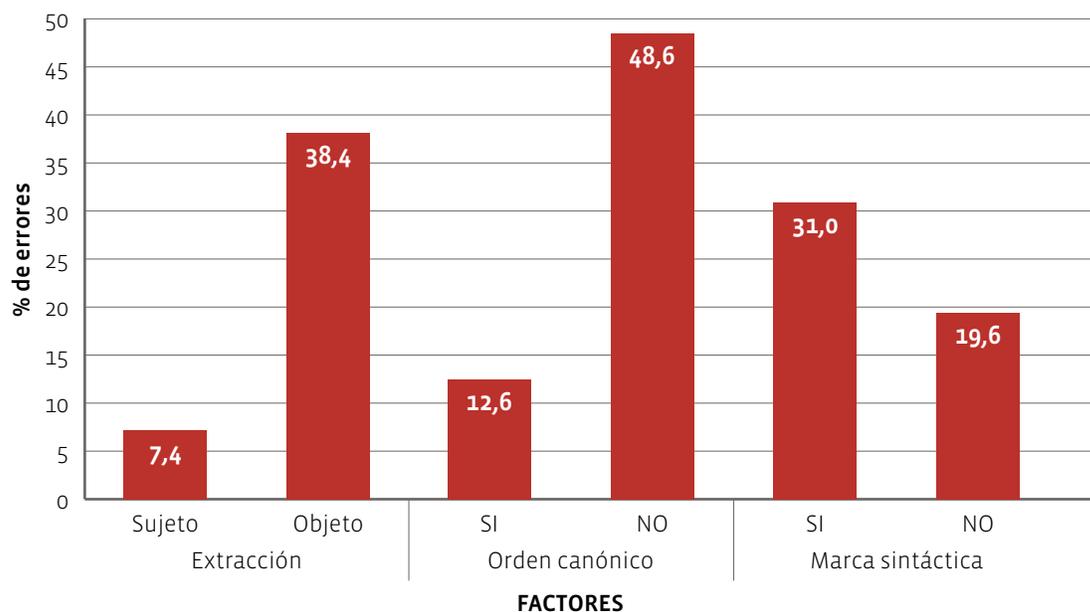


Los resultados descriptivos muestran que el mayor porcentaje de errores fue producido en las oraciones de E2, al igual que en el estudio 1. Es decir, los participantes cometieron más errores para elegir la opción correcta o no pudieron realizarlo en el tiempo estipulado. Luego de E2, la relativa que obtuvo más errores fue E3, luego E4 y E5 con poca diferencia. Y la menor cantidad de errores fue encontrada en E1.

A continuación, se realizó el mismo análisis, pero se discriminó por factor: 1) extracción (S - O), 2) orden canónico (sí - no), y 3) marca sintáctica (sí - no). Esto se presenta en el gráfico 4.

GRÁFICO 4

Porcentaje de errores para cada factor



Del análisis visual del gráfico 4, se desprende que las oraciones que no siguen el orden canónico resultan más difíciles de comprender (con la mayor proporción de errores) en relación con las que siguen el orden canónico. Este mismo patrón de errores fue encontrado en el estudio 1. Con respecto a las oraciones con extracción de S u O, estas últimas concentraron mayor cantidad de errores, al igual que en el estudio 1. Y en relación al factor marca sintáctica, las oraciones con marca presentaron la mayor proporción de errores al ser comparadas con aquellas sin marca. Estos últimos resultados apuntan en la dirección opuesta a los hallados en el estudio 1. A nivel general, se encontraron las mayores proporciones de errores en las relativas con extracción de O que no siguen el orden canónico y con marca sintáctica, que se manifiestan como las más difíciles de comprender con un tiempo limitado de respuesta.

Cuando se examinan los errores de la E1 (con extracción de S) y su contraparte E4 (con extracción de O), se puede observar que con las exigencias de esta prueba (comprensión con límite de tiempo) el factor tipo de extracción marca una diferencia no vista en el estudio 1, siendo las oraciones con extracción de O más difíciles de comprender, según el porcentaje de errores obtenidos en la E4 (25,9%) en comparación con la E1 (7,4%).

Cuando se varían los otros factores como el orden canónico y la marca sintáctica, aumenta la dificultad en la comprensión, al igual que en el estudio 1.

Si se analizan solamente las estructuras con extracción de objetos (E2, E3, E4 y E5) y se las agrupa en función a las estructuras que siguen el orden canónico (E4 y E5) y no (E2 y E3), el no seguir el orden canónico en las relativas con extracción de objeto hace mucho más difícil su comprensión (50% versus 97% de errores), al igual que en el estudio 1. Y si agrupamos las estructuras por marca sintáctica (con E3 y E5 y sin E2 y E4), el discernimiento sobre la dificultad en la comprensión se hace menos claro, con menor porcentaje de errores para las estructuras con marca sintáctica, al igual que en el estudio 1.

3.2.3. Discusión

En relación con este segundo estudio, el orden que se puede establecer en relación a la facilidad de comprensión es: $E1 > E5 > E4 > E3 > E2$. Esto muestra que el tipo de estructura en donde se cometieron menos errores de comprensión fue E1, mientras que en E2 es donde se presenta el mayor número de errores. Con respecto al *continuo* propuesto ($E1 > E4 > E5 > E3 > E2$), los resultados del estudio 2 casi no presentan diferencias, ya que la única alternancia ocurre entre E4 y E5, y dichos porcentajes muestran una diferencia muy pequeña. Por lo tanto, esta prueba coincidiría con los supuestos planteados inicialmente.

A nivel general, el estudio 2 presentó un número mayor de errores que el estudio 1 (estudio 1: 4,1%, 49 errores; estudio 2: 23%, 173 errores), que puede atribuirse al factor tiempo.

Con respecto al estudio 1, es posible advertir que hay coincidencia con las dos estructuras que presentaron mayores dificultades (E3 y E2), aunque en el estudio 2, en el extremo de menor dificultad, se encuentra E1 (la única estructura que elide S). Cuando se comparan la E1 con su contraparte en extracción de O (E4), se observa una importante diferencia en el porcentaje de errores, indicando que el factor tipo de extracción (S u O) tiene una influencia importante en este tipo de pruebas que demanda una respuesta rápida, exigiendo al sujeto un procesamiento acelerado de integración y de memorización de la información para su comprensión. Mientras tanto, en el estudio 1 no se observó tal diferencia, lo que mostraría que cuando el sujeto tiene la opción de relectura y/o comparación con las respuestas anteriores, este factor perdería peso en la comprensión.

Al igual que en el estudio 1, cuando se inspeccionan las oraciones con extracción de O, la falta del orden canónico y de la marca sintáctica parecen hacer más compleja su comprensión; no obstante, la influencia de la marca no es tan clara.

En relación a la frecuencia de ocurrencia de las E de relativo, según el estudio de corpus, coincide la posición de las estructuras en los extremos: E1 y E3 (teniendo en cuenta que E2 no pudo ser discriminada en el análisis del corpus); mientras que E5 cambia de posición. A partir de esto, se podría pensar que la frecuencia de ocurrencia coincide, en alguna medida, con la complejidad en la comprensión de dichas oraciones.

En cuanto al análisis de los factores, la elisión de S aparece como el factor facilitador más importante, ya que posee la menor cantidad de errores, siendo la diferencia entre ambas posibilidades (elisión de S u O) importante (31 puntos). Le sigue el orden canónico con una diferencia de 36 puntos. La marca sintáctica mostró un patrón inverso al encontrado en el estudio 1: se cometieron más errores en las oraciones con marca sintáctica. Sin embargo, en ambos estudios la diferencia entre la proporción de errores para las oraciones con y sin marca sintáctica es numéricamente pequeña (estudio 1, 1,22%, y estudio 2, 11,4%), a diferencia de lo que ocurre con los otros dos factores, donde el porcentaje de errores entre las opciones es de mayor magnitud. El perfil de datos hallados para este último factor hace menos entendible su influencia en la comprensión de este tipo de oraciones.

Para resumir, los datos obtenidos en el estudio 2 son más cercanos a los propuestos en la introducción según el continuo planteado a partir de la revisión de los antecedentes y muestran cierta coincidencia con las frecuencias de ocurrencia de las estructuras halladas en el análisis de un corpus. Se puede hipotetizar que este segundo estudio se acerca más a la comprensión espontánea (al no contar con la opción de relectura), donde se ponen más en juego los procesos de integración y memoria de trabajo.

4. Discusión general

Este trabajo se propuso esclarecer el nivel de dificultad (o facilidad) en la comprensión de relativas de objeto del español en comparación con las de sujeto, en función de dos factores morfosintácticos: el orden de palabras y la marca de caso. Para establecer el nivel de dificultad se seleccionaron estructuras (E) del español en las que los dos factores morfosintácticos están presentes o no. Tomando como punto de partida los estudios previos que (aunque muy informativos) abordaron una selección acotada de E, en el presente estudio se incluyeron cinco de ellas, lo que permitió manipular los factores morfosintácticos de interés y establecer un orden hipotético de facilidad en la comprensión de las E.

El análisis de corpus permitió verificar el uso de estas E en el lenguaje escrito y su frecuencia de ocurrencia, aunque la E2 no pudo ser discriminada de forma clara. La frecuencia de ocurrencia mostró cierta coincidencia con el continuo hipotético sobre la facilidad/dificultad de comprensión de las E, que se estableció según la revisión de la literatura. De este análisis queda por pensar qué impacto tiene la frecuencia de ocurrencia en la comprensión. Es posible que funcione como un facilitador adicional que actúa a la hora de procesar sintaxis, es decir, que la frecuencia de ocurrencia podría tener un efecto facilitador en la comprensión de oraciones con relativas de S y O (Reali y Christiansen, 2007a y 2007b; Wells y otros, 2009; Reali, 2014), o que la relación sea inversa, es decir, que las E más fáciles sean las más utilizadas en el habla escrita. Sin embargo, otros trabajos han encontrado un patrón inverso. Por ejemplo, el trabajo de Carreiras y otros (2010) encontró en sus experimentos 1 y 2 (lectura a ritmo personal e interpretación final) mayor dificultad de procesamiento para las relativas de S sobre las de O en vasco, y su estudio de corpus arrojó un patrón de distribución inverso: mayor frecuencia en las RS y menor frecuencia en las RO.

Responder a esta cuestión excede el propósito del presente trabajo. Sin embargo, cabe resaltar la estrecha relación encontrada en este trabajo entre el lenguaje en uso y su procesamiento, y la utilidad de incluir información sobre el uso de las estructuras analizadas en estudios de procesamiento.

Con respecto al tipo de extracción, se puede decir que se halló una tendencia de que las relativas con extracción de S resultan más fáciles de comprender que aquellas con extracción de O, lo que coincide con los datos generales encontrados en los trabajos antecedentes en diferentes idiomas (Frauenfelder y otros, 1980 y Holmes y O'Regan, 1981, en francés; MacWhinney y Pleh, 1988 y King y Just, 1991, en el inglés; Mak y otros 2002, 2006, en alemán; Levy y otros, 2013, en ruso; Del Río y López-Migues, 2006, Betancort y otros, 2009, Del Río y otros, 2012, y Arancibia y otros, 2014, en español).

Puntualmente, en comparación con los trabajos realizados en español, se ha encontrado que las relativas con extracción de O causaron más dificultades que las relativas de S, cuando se mantuvo el orden superficial de las palabras, aunque se manipularon otros factores como la animacidad (Betancort y otros, 2009), o cuando se alteró el orden superficial de las palabras (Del Río y López-Migues, 2006; y Del Río y otros, 2012).

Es importante advertir que la mayor facilidad de las relativas de S en comparación con las de O (cuando se compararon las E1 y E4, que solo diferían por el tipo de extracción) fue modulada por el tipo de prueba. Cuando la prueba exigió al sujeto un procesamiento acelerado de integración y de memorización de la información para su comprensión (estudio 2), la relativa de O (E4) presentó mayor dificultad que su contraparte de S (E1). Mientras que cuando la prueba no presentó tal exigencia (es más, el sujeto tenía la opción de relectura y/o comparación con las respuestas anteriores), como en el estudio 1, no se observó tal diferen-

cia. Nuevamente, los resultados del estudio 2 son los más semejantes a los estudios previos en español que utilizaron tareas que exigían un procesamiento acelerado de integración y de memorización de la información para su comprensión. Los trabajos de Del Río y López-Migues y Del Río y otros incluyeron tareas de lectura a ritmo personal, con una prueba de verificación de comprensión al final. En Betancort y otros, 2009, los sujetos debieron leer la sentencia; su lectura y tiempo fueron registrados mediante un rastreador de movimientos oculares (*eye-tracker*), pero ninguna prueba de verificación de la comprensión fue dada.

No obstante, este último punto debería ser estudiado en otras condiciones experimentales, donde efectivamente se pueda comparar tareas *offline* (como las de este trabajo) y las *on-line*. Se volverá sobre este punto, al final de este apartado.

El análisis sobre orden de palabras ha mostrado que las cláusulas relativas O con orden canónico (OSV), como en las E4 y E5, exhiben menor dificultad de comprensión que aquellas de orden no canónico (OVS), como en las E2 y E3. Estos resultados se oponen a los que reportan Del Río y otros (2012), quienes encuentran menos errores de comprensión en el orden OVS (*El policía al que criticó el comisario vive en el barrio*) comparado con el orden OSV (*El policía al que el comisario criticó vive en el barrio*). Por el contrario, los resultados de los experimentos 1 y 2 que aquí se presentaron muestran que el orden OSV resultó más fácil de comprender.

Como ya se vislumbró, hay marcadas diferencias entre la metodología empleada en Del Río y otros (2012) y el presente trabajo. Mientras que el trabajo de Del Río y otros consistió en una tarea de lectura a su propio ritmo palabra por palabra que se presentó a los participantes de manera no cumulativa (es decir que, al aparecer una palabra, desaparecía la anterior), en el presente trabajo los participantes veían toda la oración y las dos opciones de respuesta (en el caso del estudio 2) o la pregunta experimental (en el estudio 1). Esta sola diferencia hace que, por ejemplo, la lectura en Del Río y otros resulte de mayor complejidad que en el presente estudio, ya que no permite regresiones o relecturas. La demanda en la memoria de trabajo a la hora de leer de una manera no natural (ya que se sabe gracias a experimentos de *eye-tracking* que la lectura no es lineal) puede ser parte de la explicación de la interferencia o mayor dificultad, debido al cambio de perspectiva en las relativas OSV en Del Río y otros. Otra diferencia entre ese estudio y este se halla en la manera de comprobar la comprensión de la cláusula relativa (lo cual solo se hizo en la mitad de los casos en Del Río y otros (2012), pero en todos los casos en el presente trabajo). Mientras que en ese trabajo, luego de la última palabra de cada oración aparecía una frase como *El comisario criticó al policía* y los participantes presionaban un botón para responder verdadero (si esta frase coincidía con el contenido de la oración) o falso (si no coincidía), en el estudio 1 del presente trabajo los participantes respondieron a la pregunta *¿Quién critica a quién?* Esto se realizó para recabar datos *offline* de comprensión de los roles temáticos y los participantes tenían que escribir

oraciones completas (ej. *El comisario criticó al policía*). En el estudio 2, tanto la oración como dos opciones de respuesta aparecían escritas y permanecían en la pantalla hasta que el participante seleccionara una de ellas. Básicamente la diferencia entre Del Río y otros y el estudio 2 (del presente trabajo) es que el primero requiere un juicio de veracidad (verdadero o falso), mientras que el segundo solo requiere la selección de la opción que mejor representa la oración experimental. Quizá, la naturaleza de la tarea puede haber resultado en mayores o menores puntajes en la tarea de comprensión. Respecto de la marca morfosintáctica, dos estudios previos del español abordaron únicamente relativas de objeto sin marca. Mientras que Del Río y López-Higes (2006) incluyeron la E4, Betancourt y otros (2009) solo incluyeron la E2. De estos dos estudios, solo el primero incluyó datos de comprensión al preguntar a los participantes, por ejemplo, ‘¿Quién agredió al otro? a) el capitán, b) el cadete’. Sus resultados reportan menos del 40% de errores en la comprensión de las relativas. Para esta misma estructura (E4), en el estudio 2 del presente trabajo encontramos 25.9% de errores, lo cual confirma el resultado de Del Río y López-Higes (2006) para la misma estructura. Aunque el trabajo de Betancourt y otros (2009), que incluyó la E2, no reporta tarea de comprensión, los autores encontraron el mayor costo de procesamiento en las relativas O que contenían Sujeto y Objeto animados (A-A). Esa condición del trabajo de Betancourt y otros (2009) es comparable con los datos del presente trabajo, en el que la animacidad se mantuvo constante en las oraciones estímulo. En el estudio 2 del presente trabajo, las relativas con E2 mostraron el mayor costo en la comprensión (59.3%). Entonces, los resultados del presente estudio que muestran gran costo de comprensión de E2 se pueden interpretar a la luz del trabajo de Betancourt y otros (2009), que muestran costo en el procesamiento en una tarea *on-line*.

Otro estudio previo, Del Río y otros (2012), solo incluyó condiciones que contenían marca morfosintáctica y que diferían en el orden de palabras. Los autores las denominaron ORsi y tOR, y en el presente estudio esas condiciones se identificaron como E3 y E5, respectivamente. Mientras que Del Río y otros (2012) reportan más dificultad de comprensión para E5 que para E3, nuestros resultados muestran lo opuesto. Tanto en el estudio 1 como en el estudio 2 del presente trabajo, se hallaron más errores de comprensión de relativas con E3 (5.2% y 38%, respectivamente) que de relativas con estructura E5 (1.2% y 24%, respectivamente). Esta discrepancia en los resultados de E3 y E5, en las que la marca se mantiene constante, pero el orden de palabras cambia, es difícil de explicar de manera taxativa. Mientras que Del Río y otros proponen que la complejidad puede deberse a la dificultad que presenta la presencia de dos SNs contiguos en E5 (ver comentarios anteriores sobre orden de palabras), nuestros resultados podrían reflejar costos que tienen que ver con la coactivación de construcciones análogas en el orden superficial de E3. La E3 (SN_{a la que}-V-SN) es análoga a E2 (SN_{que}-V-SN), pero también a E1 (SN_{que}-V-SN), la cual es relativa de S. El estudio de corpus preliminar que presentamos muestra cómo tanto E1 como E2 son las cláusulas de relativo con mayor frecuencia de uso en el registro escrito. De las que se han podido aislar en el estudio de corpus, se referirá a E1, que es la más frecuente. De acuerdo con las teorías basadas en el uso, la alta frecuencia de

E1 podría hacer de esta estructura un modelo analógico para E3 ya, que ambas contienen similar orden superficial (SN-V-SN). Esto, sumado a la indiscutible evidencia de mayor facilidad de procesamiento y comprensión de las relativas S (E1), podría explicar la mayor dificultad en la comprensión de E3 en el presente estudio. Específicamente, se podría pensar que, frente a la presencia de una relativa con estructura E3, la construcción E1 se activaría provocando que la correcta interpretación dependa de la presencia de la marca morfosintáctica frente al OD. En el caso de E3 la marca se encuentra al comienzo ([**a** la que mira el hombre]), mientras que, en el caso de E1, se encuentra al final ([que mira **a**l hombre]). Esta relación analógica, que podría generar coactivación y dificultar la comprensión, podría explicar los resultados. Por su parte, se podría asumir que, como E5 no mantiene tales similitudes de orden superficial, no cuenta con las ventajas de la coactivación.

Un antecedente del presente estudio es el trabajo sobre el idioma ruso de Levy y otros (2013). En ese estudio se reporta que las relativas O con marca sincretizada (es decir, no marcada por el caso, por ejemplo, 'que') y con orden mezclado resultaron con menor porcentaje de aciertos. De manera similar, el presente estudio ha mostrado que tanto el orden canónico como la marca morfosintáctica benefician la comprensión. Respecto del orden de palabras, E4 y E5 (ambas con orden OSV) resultaron más fáciles de comprender que E2 y E3 (con orden OVS). Respecto de la presencia de marca morfosintáctica, en los dos estudios reportados en el presente trabajo, en las condiciones sin marca (E2 y E4) se observaron más errores de comprensión que en las condiciones con marca de caso (E3 y E5).

El presente estudio constituye un primer paso hacia una mejor comprensión del rol que ejerce el orden de palabras y la presencia de marca morfosintáctica en el procesamiento de las cláusulas de relativo en el español. Las limitaciones del presente trabajo se hallan primordialmente en su naturaleza exploratoria, ya que se hizo foco en la comprensión de las relativas como una manera de abordar la facilidad o dificultad en el procesamiento midiendo la comprensión. Puntualmente, la falta del límite de tiempo en la tarea utilizada en el estudio 1, puede haber generado diferencias no controladas (por ejemplo, en la utilización de diferentes estrategias para la comprensión de las oraciones), añadiendo un factor de incertidumbre sobre los procesos subyacentes a la comprensión. Otra limitación de la metodología utilizada en este trabajo se halla en la falta de control de las variables de la capacidad de memoria operativa y la velocidad de procesamientos de los sujetos en la comprensión de oraciones con cláusula de relativo. Sobre la base del presente trabajo pretendemos estudiar el procesamiento mediante técnicas *on-line*, controlando mayor cantidad de factores (como la memoria operativa, velocidad de procesamiento y edad de los sujetos), superando así las limitaciones antes mencionadas. Una comparación del procesamiento de todos los tipos de relativas incluidos en el presente estudio posibilitará determinar con mayor exactitud el rol del orden de palabras y de la marca morfosintáctica ya no solamente en la comprensión, sino también en el procesamiento en tiempo real. La presente línea de investigación posee potencial para

lograr mayor especificidad en la descripción del impacto entre la variación morfosintáctica en la comprensión.

5. Bibliografía citada

ARANCIBIA GUTIÉRREZ, Beatriz, y otros, 2014: "Procesamiento de cláusulas relativas anidadas, memoria operativa y envejecimiento", *RLA. Revista de lingüística teórica y aplicada* 52 (1), 155-179.

ATSERIAS, Jordi, y otros, 2006: "FreeLing 1.3: Syntactic and semantic services in an open-source NLP library", *Proceedings of LREC* 6.

BETANCORT, Moisés, Manuel CARREIRAS y Patrick STURT, 2009: "The processing of subject and object relative clauses in Spanish: An eye-tracking study", *The Quarterly Journal of Experimental Psychology* 62 (10), 1915-1929.

CAPLAN, David, y Gloria S. WATERS, 1999: "Verbal working memory and sentence comprehension", *Behavioral and brain Sciences* 22 (1), 77-94.

CAPLAN, David, y Gloria S. WATERS, 2005: "The relationship between age, processing speed, working memory capacity, and language comprehension", *Memory* 13 (3-4), 403-413.

CAPLAN, David, y otros, 2011: "Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses", *Psychology and aging* 26 (2), 439-450.

CARREIRAS, Manuel, y otros, 2010: "Subject relative clauses are not universally easier to process: Evidence from Basque", *Cognition* 115 (1), 79-92.

CHOMSKY, Noam, 1965: *Aspects of the Theory of Syntax*, Cambridge, Massachusetts: MIT Press.

CHOMSKY, Noam, 1986: *Knowledge of Language*, New York: Praeger.

DEDE, Gayle, y otros, 2004: "The relationship between age, verbal working memory, and language comprehension", *Psychology and aging* 19 (4), 601-616.

DEL RÍO, David, y Ramón LÓPEZ-MIGUES, 2006: "Efectos de la memoria operativa y de una carga de procesamiento en la comprensión de oraciones", *Psicológica* 27, 79-95.

DEL RÍO, David, Ramón LÓPEZ-HIGES y María Teresa MARTÍN-ARAGONESES, 2012: "Canonical word order and interference-based integration costs during sentence comprehension: The case of Spanish subject and object relative clauses", *Quarterly Journal of Experimental Psychology* 65, 2108-2128.

FODOR, Jerry, Thomas G. BEVER y Merrill F. GARRETT, 1974: *The Psychology of Language*, New York: McGraw-Hill.

FORSTER, Kenneth, y Jonathan FORSTER, 2003: "DMDX: A windows display program with millisecond accuracy", *Behavior Research Methods, Instruments, and Computers* 35, 116-124.

FRAUENFELDER, Ulrich, Juan SEGUI y Jacques MEHLER, 1980: "Monitoring around the relative clause", *Journal of Verbal Learning and Verbal Behavior* 19, 328-337.

FRAZIER, Lyn, y Charles CLIFTON, 1989: "Successive cyclicity in the grammar and the parser", *Language and Cognitive Processes* 4, 93-126.

FRAZIER, Lyn, y Keith RAYNER, 1988: "Parameterizing the language processing system: Left-vs. right-branching within and across languages" en John HAWKINS (ed.): *Explaining language universals*, 247-279.

GIBSON, Edward, 1998: "Linguistic complexity: Locality of syntactic dependencies", *Cognition* 68, 1-76.

GIBSON, Edward, 2000: "The dependency locality theory: A distance-based theory of Linguistic Complexity" en Yasushi MIYASHITA, Alec MARANTZ y Wayne O'NEIL (eds.): *Image, language, brain*, Cambridge, MA: MIT Press, 94-126.

GRODNER, Daniel, y Edward GIBSON, 2005: "Consequences of the serial nature of linguistic input", *Cognitive Science* 29, 261-291.

HOLMES, Virginia M., y J. Kevin O'REGAN, 1981: "Eye fixation patterns during the reading of relative-clause sentences", *Journal of Verbal Learning and Verbal Behavior* 20 (4), 417-430.

HSIAO, Franny, y Edward GIBSON, 2003: "Processing relative clauses in Chinese", *Cognition* 90(1), 3-27.

ISHIZUKA, Tomoko, Kentaro NAKATANI y Edward GIBSON, 2006: "Processing Japanese relative clauses in context", comunicación presentada en el 19th annual CUNY conference on human sentence processing.

KING, Jonathan, y Marcel Adam JUST, 1991: "Individual differences in syntactic processing: The role of working memory", *Journal of Memory and Language* 30, 580-602.

LEVY, Roger, Evelina FEDORENKO y Edward GIBSON, 2013: "The syntactic complexity of Russian relative clauses", *Journal of Memory and Language* 69, 461-495.

LEWIS, Shevaun, y Colin PHILLIPS, 2015: "Aligning grammatical theories and language processing models", *Journal of psycholinguistic research* 44 (1), 27-46.

LEWIS, Richard L., y Shravan VASISHTH, 2005: "An activation-based model of sentence processing as skilled memory retrieval", *Cognitive Science* 29, 375-419.

LEWIS, Richard L., Shravan VASISHTH y Julie A. VAN DYKE, 2006: "Computational principles of working-memory in sentence comprehension", *Trends in Cognitive Sciences* 10 (10), 447-454.

LIN, Yowyu, y Susan M. GARNSEY, 2010: "Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin", *Processing and producing head-final structures*. Springer Netherlands, 241-275.

MACWHINNEY, Brian, y Csaba PLEH, 1988: "The processing of restricted relative clauses in Hungarian", *Cognition* 29, 95-141.

MAK, Willem M., Wietske VONK y Herbert SCHRIEFERS, 2002: "The influence of animacy on relative clause processing", *Journal of Memory and Language* 47 (1), 50-68.

MAK, Willem M., Wietske VONK y Herbert SCHRIEFERS, 2006: "Animacy in processing relative clauses: The hikers that rocks crush", *Journal of Memory and Language* 54 (4), 466-490.

MILLER, George A., y Noam CHOMSKY, 1963: "Finitary models of language users" en Robert Duncan LUCE, Robert R. BUSH y Eugene GALANTER (eds.): *Handbook of Mathematical Psychology*, New York, NY: Wiley, 2-419.

MIYAMOTO, Edson T., y Michiko NAKAMURA, 2003: "Subject/object asymmetries in the processing of relative clauses in Japanese", *Proceedings of 1831 WCCFL* 22, 342-355.

REALI, Florencia, 2014: "Frequency Affects Object Relative Clause Processing: Some Evidence in Favor of Usage-Based Accounts", *Language Learning* 64 (3), 685-714.

REALI, Florencia, y Morten H. CHRISTIANSEN, 2007a: "Word chunk frequencies affect the processing of pronominal object-relative clauses", *The Quarterly Journal of Experimental Psychology* 60 (2), 161-170.

REALI, Florencia, y Morten H. CHRISTIANSEN, 2007b: "Processing of relative clauses is made easier by frequency of occurrence", *Journal of Memory and Language* 57 (1), 1-23.

SCHROTEN, Jan, 1987: "Gramática generativa y gramática estructural en el análisis sintáctico de las cláusulas relativas en español", *Nueva Revista de Filología Hispánica* 35 (1), 37-110.

STINE-MORROW, E. A. L., Sharon RYAN y J. Sandra LEONARD, 2000: "Age differences in on-line syntactic processing", *Experimental Aging Research* 26 (4), 315-322.

WANNER, Eric, y Michael MARATSOS, 1978: "An ATN approach to comprehension" en Morris HALLE, Joan BRESNAN y George A. MILLER (eds.): *Linguistic Theory and Psychological Reality*, Cambridge, Mass: MIT Press, 119-161.

WATERS, Gloria S., y David CAPLAN, 2001: "Age, working memory, and on-line syntactic processing in sentence comprehension", *Psychology and aging* 16 (1), 128-144.

WELLS, Justine B., y otros, 2009: "Experience and sentence processing: Statistical learning and relative clause comprehension", *Cognitive Psychology* 58, 250-271.