

LA FIABILIDAD DE LOS JUICIOS DE ACEPTABILIDAD EN LOS ESTUDIOS DE SINTAXIS DEL ESPAÑOL

THE RELIABILITY OF ACCEPTABILITY JUDGMENTS IN SPANISH SYNTAX¹

IVÁN ORTEGA-SANTOS
University of Memphis
rtgsntos@memphis.edu

Este artículo se centra en la fiabilidad de los juicios de aceptabilidad en la gramática generativa con énfasis en el español. Se presentan las principales diferencias entre, por un lado, la metodología de recogida de datos más extendida, es decir, el uso de juicios de aceptabilidad sin protocolo experimental ni análisis estadístico, y, por otro lado, la sintaxis experimental, que sigue protocolos estándares de recogida de datos y hace uso de análisis estadísticos. Tras discutir las virtudes y limitaciones de ambas metodologías a través de varios ejemplos concretos en los que los datos recogidos difieren, se presentan los resultados de un estudio de fiabilidad de datos del español centrado en una muestra aleatoria de datos aparecidos en *Probus* (2006-2017) y en publicaciones de alto impacto (100+ citas). Dichos datos, recogidos originalmente sin protocolo experimental, son estudiados en un diseño experimental para analizar las divergencias entre ambas metodologías. Se exploran detalladamente el origen de estas divergencias y las limitaciones de los estudios de fiabilidad.

Palabras clave: sintaxis del español, fiabilidad, métodos de recogida de datos, español venezolano

This article focuses on the reliability of acceptability judgments in generative grammar with an emphasis on Spanish. The main differences are presented between, on the one hand, the most widespread data collection method, that is, acceptability judgments gathered without

¹ Este trabajo fue financiado en su totalidad por una beca de investigación de la Facultad de Arte y Ciencias de la Universidad de Memphis. Este apoyo no implica necesariamente que la Universidad avale las conclusiones de la investigación. Por otra parte, la presente investigación se enmarca dentro del Proyecto PGC2018-093774-B-I00 del Ministerio de Ciencia, Innovación y Universidades del Gobierno de España. Agradezco los comentarios, ayuda y/o sugerencias de B. Morgan, J. Sprouse, D. Almeida, J. Uriagereka, Á. Gallego y las audiencias de ALFALito en New York, HLS 2019 y el III Encuentro de Dialectos del Español organizado por la Universidad Alcalá. Finalmente, les agradezco a Andrés Saab, Francisco Ordóñez y Sandra Quarezemin por la invitación a participar en este volumen así como a los revisores anónimos y a Alba Valencia y a Andrés Saab por sus comentarios y ayuda en la revisión del artículo.

following experimental protocols or performing any statistical analysis, and, on the other hand, experimental syntax, which follows standard data collection protocols and performs statistical analyses. The virtues and limitations of both methodologies are discussed through various case studies in which the data collected differ according to the methodology. Furthermore, the results of a data reliability or data assessment study focusing on Spanish using a random sample of data that appeared in *Probus* (2006-2017) and in high impact publications (100+ citations) are presented. The data, originally collected without experimental protocols, are studied in an experimental design to analyze the divergences between both methodologies. The origin of these divergences and the limitations of data reliability studies, as revealed by this sample, are explored in detail.

Keywords: Spanish syntax, reliability, data collection methods, Venezuelan Spanish

Recibido: 11 agosto 2020

Aceptado: 12 octubre 2020

1. INTRODUCCIÓN: EL LUGAR DE LOS JUICIOS DE ACEPTABILIDAD EN LA GRAMÁTICA GENERATIVA

Tradicionalmente, la gramática generativa ha usado los juicios de aceptabilidad, también llamados juicios gramaticales, como fuente de datos. Estos son juicios introspectivos sobre una frase o sintagma para determinar si dicha unidad lingüística es aceptable o inaceptable (i.e., si es parte del español o de la lengua objeto de estudio). Esta noción de aceptabilidad no es normativa; de lo que se trata es de recabar datos para posibilitar la descripción de la gramática mental de un hablante de español o de cualquier otra lengua. Tal gramática debe poder generar las frases aceptables y rechazar las inaceptables de ese idioma. Esta relación entre teorías gramaticales generativas y aceptabilidad supone tratar las frases en cuestión como gramaticales (pueden ser generadas por la gramática) o agramaticales (no pueden ser generadas por la gramática); la diferencia entre aceptabilidad y gramaticalidad reside en que el primer término hace referencia a una percepción, un juicio metalingüístico, mientras que el segundo conlleva la existencia de una teoría gramatical que explique los datos. Ahora bien, la percepción de aceptabilidad puede verse influida por factores independientes de la gramática mental. Esta disociación entre gramaticalidad y aceptabilidad se da, por ejemplo, en las denominadas ilusiones gramaticales, es decir, frases agramaticales –que no son parte de la gramática mental– pero que se perciben como aceptables. Asimismo, también existen frases gramaticales que se perciben como inaceptables. La oración en (1), traducida de un clásico ejemplo de Herman Schultze recogido por Montalbetti (1984:6), ilustra el primer caso: La estructura comparativa no es parte de la gramática del español –no puede ser generada y es considerada, por lo tanto, agramatical– como muestra un examen detallado de su interpretación. El segundo caso se ilustra en (2) con un ejemplo de incrustación central (*center embedding*, Chomsky y Miller 1963):

1. Más gente ha ido a Roma que yo he ido.
2. El perro que el gato que el ratón vio escuchó ladró.

En cuanto a (2), a pesar de que la estructura es gramatical –la gramática puede generarla–, resulta difícil, si no imposible, procesar la oración resultante en cuestión, razón por la cual la percepción de gramaticalidad se ve afectada. Específicamente, esta frase combina las siguientes

estructuras: *El perro ladró / El perro que el gato escuchó / el gato que el ratón vio*. Las denominadas frases *garden path* (Bever 1970) constituyen otro ejemplo de oraciones percibidas como inaceptables por razones de procesamiento, a pesar de ser gramaticales. Es así que, desde los inicios de la gramática generativa, existe una necesidad de estudiar la relación entre aceptabilidad y procesamiento, como parte de la dicotomía entre Competencia y Actuación, presente en este marco teórico. De modo natural, parte de la investigación sobre los juicios de aceptabilidad se ha centrado en entender qué factores facilitan o dificultan el procesamiento (ver Lewis y Vasishth 2005 y Gibson 1998 para ejemplos de teorías de costes de procesamiento). Actualmente, el debate se ha enfocado cada vez más en otro aspecto del diálogo entre gramática generativa y disciplinas tales como la psicolingüística, a saber, en la adopción de estándares de investigación experimental. El propósito de este artículo es proporcionar una introducción a este último debate centrada en el español. Tras introducir el debate en la sección 2, en la sección 3 se presentan varios casos prácticos de la sintaxis del español en los que la adopción de un paradigma de investigación experimental ha dado lugar a nuevas generalizaciones. La sección 4, basada en Ortega-Santos *et al.* (en evaluación) presenta los resultados de dos estudios que comparan los datos de publicaciones basadas en datos introspectivos recogidos mediante protocolos de recopilación de datos informales (o descritos de modo insuficiente) con datos recogidos experimentalmente. El presente estudio complementa la discusión original de Ortega-Santos *et al.* (en evaluación) no solo por su carácter de introducción al tema de la fiabilidad de los datos –requerido por el presente volumen– sino por el énfasis específico en los aspectos cualitativos de la investigación. Específicamente, se discuten detalladamente los casos de divergencia entre ambas metodologías, con el fin de facilitar un acercamiento crítico a los estudios de fiabilidad, más allá del análisis predominantemente cuantitativo de Ortega-Santos *et al.* (en evaluación). En los casos pertinentes, se menciona la posición de la *Nueva Gramática de la Lengua Española* (Real Academia Española, RAE, 2010) con respecto a los datos objeto de discusión.

2. MÉTODOS DE RECOPIACIÓN DE LOS JUICIOS DE ACEPTABILIDAD

El uso generalizado de los juicios de aceptabilidad en la sintaxis generativa ha sido criticado por las siguientes razones (ver Edelman y Christiansen 2003, Wasow y Arnold 2005, Ferreira 2005, Gibson y Fedorenko 2010a, 2010b, entre otros, y la discusión reciente en Ortega-Santos en prensa y Muñoz Pérez 2014)²:

- Ausencia de información sobre el origen de los juicios (número de hablantes que fueron consultados, su perfil lingüístico, su formación en sintaxis teórica y/o potenciales conflictos de interés)
- Ausencia de información sobre cómo se recogieron los datos (número de lexicalizaciones presentadas a los hablantes, presencia/ausencia de datos de relleno para que los hablantes no sean conscientes del aspecto concreto que se está estudiando y

² Obvia decir que la lista no es exhaustiva y que las críticas no se aplican uniformemente a las publicaciones de sintaxis generativa.

evitar que los juicios se vean afectados por la sobreexposición a un tipo de estructura; ver Snyder 2000 para discusión de este último fenómeno)

- Ausencia de análisis estadístico de los resultados

A continuación, se incluyen varios ejemplos en los que el uso de metodología experimental ha permitido refinar los resultados recogidos sin metodología experimental (sección 3). En la sección 4, se lleva a cabo una comparación sistemática entre los resultados de ambas metodologías de recopilación de datos en una muestra reciente aparecida en *Probus* (2006-2017) y en publicaciones de alto impacto.

3. ¿CUÁNDO SON ÚTILES LOS MÉTODOS EXPERIMENTALES DE RECOPIACIÓN DE DATOS?

El objetivo de esta sección es presentar varios estudios de caso en los que una metodología experimental ha permitido llegar a nuevas generalizaciones sobre la sintaxis del español como paso necesario para entender la relevancia de los estudios de fiabilidad. En particular, se discuten las islas-Qu, el orden VOS y la microvariación en las estructuras comparativas.

3.1. Islas sintácticas

Existe debate sobre si las comúnmente denominadas islas sintácticas son realmente el resultado de restricciones de la gramática o la consecuencia de la dificultad para procesar ciertas estructuras (ver Ortega-Santos 2011 para una introducción centrada en el español). Se trata, por lo tanto, de una línea de investigación relevante para el estudio de los juicios de aceptabilidad y su relación con la gramática. A nivel teórico, las restricciones al movimiento-Qu y su hipotética universalidad han sido de gran importancia para el desarrollo de la gramática generativa (consultar, por ejemplo, Chomsky 1964, 1973 y Ross 1967, así como mucha de la bibliografía posterior que estas obras inspiraron). Las cláusulas subordinadas con un elemento-Qu en el sintagma complementante son una configuración sintáctica de la cual se prohíbe la extracción, si bien hay diferencias dependiendo de la lengua de que se trate (note que la aceptabilidad de los datos del español en (3) se clarifica en esta sección posteriormente):

3. ¿Quién_i no sabes por qué _{t_i} escribió el informe?

4. *Who don't you know why _{t_i} wrote the report?

Inglés

Mucha de la bibliografía tradicional sobre islas en el español asume, al menos desde Torrego (1984), que, a diferencia del inglés, las islas-Qu no están presentes en el castellano (salvo que haya inversión verbo-sujeto obligatoria en la cláusula subordinada). Obviamente, de ser correcto, esto pondría en cuestión la universalidad de esta restricción de localidad. Es aquí donde un diseño experimental resulta útil, tal y como señalan Sprouse *et al.* (2016) (ver también Phillips 2009: 56), pues permite comparar de modo independiente el efecto en la percepción de aceptabilidad de al menos los siguientes factores: (i.) la presencia/ausencia de la isla-Qu, (ii.) el movimiento-Qu (corto o largo) y (iii.) la violación de la isla (violación presente o ausente). De acuerdo a Sprouse *et al.* (2016: 333), para determinar si las islas-Qu están activas en una lengua, es necesario determinar si el efecto de la violación de la isla en la percepción de aceptabilidad es

independiente del efecto de los otros componentes (la presencia de una dependencia de larga distancia y la presencia de la hipotética isla-Qu). Este paso es necesario pues dichos componentes crean dificultades de procesamiento susceptibles de afectar la percepción de aceptabilidad. Según esta lógica, Ortega-Santos *et al.* (2019), siguiendo a Sprouse *et al.* (2016) diseñaron una serie de datos que contemplan los parámetros mencionados. Así, las oraciones en (5a) y (5c) ejemplifican el movimiento-Qu corto. La cláusula subordinada es una isla en el segundo caso (hipotéticamente hablando, por la presencia de un elemento-Qu en el sintagma complementante subordinado), aunque el movimiento-Qu no cruza esa hipotética isla. Por su parte, (5b) y (5d) ejemplifican el movimiento-Qu de larga distancia. Solo en el segundo caso la cláusula subordinada constituye una isla (hipotética):

5. a. *Movimiento-Qu corto, estructura sin isla sintáctica*
¿Quién_i t_i piensa que Pablo escribió el informe?
- b. *Movimiento-Qu de larga distancia, estructura sin isla sintáctica*
¿Quién_i piensas que t_i escribió el informe?
- c. *Movimiento-Qu corto, isla sintáctica (sin violación)*
¿Quién_i no t_i sabe por qué Juan escribió el informe?
- d. *Movimiento-Qu de larga distancia, isla sintáctica (con violación hipotética)*
¿Quién_i no sabes por qué t_i escribió el informe?

En caso de que exista un efecto de la violación de la isla-Qu, tendríamos el resultado del Gráfico 1, con líneas no paralelas de la evaluación promedio (puntuación z), por ejemplo, de una escala Likert del 1 al 7 (Sprouse *et al.* 2016: 333):

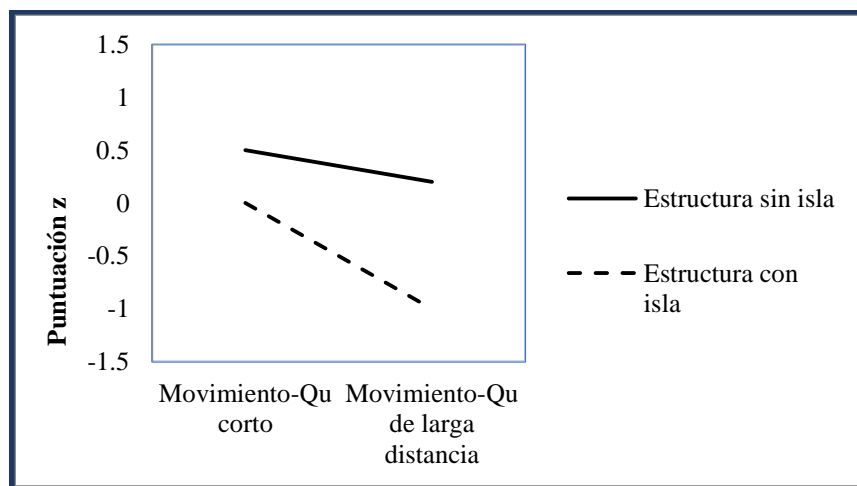


Gráfico 1. Evidencia a favor del efecto de isla sintáctica

En contraste, la ausencia de efectos de la violación de isla se correspondería con el escenario del Gráfico 2, con líneas paralelas:

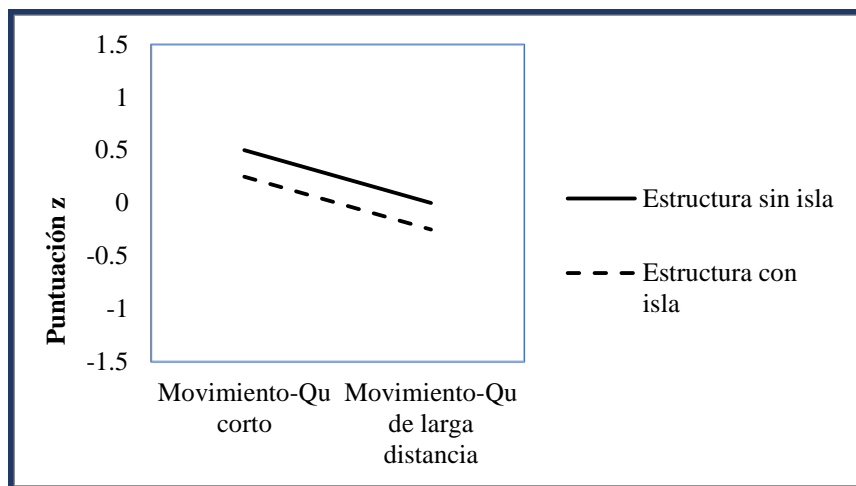


Gráfico 2. Evidencia en contra del efecto de isla sintáctica

Usando este diseño experimental, Ortega-Santos *et al.* (2019) demostraron que el castellano muestra efectos de islas-Qu (ver también López Sancio 2015 y Pañeda *et al.* 2020) independientemente de si el elemento-Qu en el sintagma complementante subordinado fuerza la inversión verbo-sujeto, en contraste con la descripción tradicional que, explícita o implícitamente, se basa en la comparación entre estructuras como (5b) y (5d), sin considerar el resto de los datos³. Cabe mencionar que aunque incluir el paradigma completo de (5) es importante, el contraste es sutil y, por lo tanto, la inclusión de un número no trivial de participantes –necesaria en un diseño experimental⁴–, así como el uso de la estadística, jugaron un papel fundamental para clarificar los resultados.

Una situación similar se da, por ejemplo, en el estudio de los efectos de Superioridad en español. De acuerdo a Vicente (2013), si bien la bibliografía generativista considera ambas frases en (6) aceptables (Ordóñez 1997, siguiendo a Jaeggli 1982), el estudio experimental de estos datos por medio de juicios de aceptabilidad en una escala Likert de 7 puntos provee evidencia a favor de un contraste estadísticamente significativo, en el que la oración en (b) se evalúa como peor.

6. a. ¿Quién compró qué?
- b. ¿Qué compró quién?

En concreto, Vicente analizó la valoración de estructuras sintácticas análogas a (6) y determinó que hay un efecto de penalización en el caso del movimiento del objeto, (6b). Es

³ Para una discusión de los efectos de localidad y su relación con la inversión verbo-sujeto en Torrego (1984) en contextos diferentes a las islas-Qu, consultar Gallego (2007, 133-141)

⁴ El número de participantes es determinado por factores tales como el hipotético tamaño del efecto objeto de estudio. Para efectos pequeños se necesita un número alto de participantes, a diferencia de los efectos grandes. El lector puede consultar los capítulos 4 y 7 de Lowie y Seton (2013) y Cowart (1997), respectivamente, así como las obras allí citadas, para conocer cómo calcular el número de participantes necesario al diseñar un experimento. En el caso de Ortega-Santos *et al.* (2019), 37 hablantes de español como lengua materna participaron en el experimento.

decir, ese ejemplo fue valorado como menos aceptable que (6a). Asimismo, Vicente comparó el tamaño de dicha penalización con el efecto de ese mismo movimiento del objeto en preguntas sencillas (con tan solo un elemento-Qu). La similitud en el efecto de la penalización en ambos contextos (preguntas con dos elementos-Qu y preguntas con uno solo) le permitió concluir que el origen del contraste en (6) es la consecuencia de la penalización asociada a mover el objeto, no de la hipotética violación de Superioridad⁵.

Por lo demás, Ortega-Santos *et al.* (2019) discuten otro factor importante: Los estudios de adquisición de segunda lengua son relevantes para el debate sobre la fiabilidad de los datos de la sintaxis teórica. Esto se debe a la presencia de grupos de control nativo-parlantes en dichas investigaciones. Ortega-Santos *et al.* (2019), de hecho, investigan la adquisición de islas en inglés y en español como segunda lengua, pero los resultados de ese trabajo, como hemos visto, son relevantes para el estudio del español como primera lengua. Esta hipotética aportación de los estudios de adquisición de la sintaxis de una segunda lengua al debate sobre los métodos de recogida de datos en la gramática generativa no ha sido explorada sistemáticamente y constituye una posibilidad de investigación para el futuro.

3.2. Estructura informativa

En el caso de la estructura informativa, en particular, en cuanto a las alternancias en el orden de las palabras determinadas por la focalización del sujeto (SVO / VOS), ha habido un debate generado por la tensión entre los resultados obtenidos por los juicios de aceptabilidad introspectivos y los resultados obtenidos en los estudios de corpus y de metodología experimental. En concreto, los generativistas, a partir especialmente de Zubizarreta (1998), han afirmado que hay una tendencia a que los sujetos focalizados aparezcan a la derecha (VOS).

7. A: ¿Quién compró un libro ayer?

B: Ayer compró un libro PEDRO.

Estos datos, sin embargo, han sido cuestionados desde la sintaxis experimental. Así, diversos investigadores han descubierto una *preferencia* por el orden de palabras canónico en este contexto (ver Gabriel 2007, Muntendam 2009, Leal Méndez y Slabakova 2011, Hoot 2012 y Jiménez-Fernández 2015 para una discusión relevante centrada en varios dialectos, desde el español de Argentina al mexicano o al peninsular; ver Ortega-Santos 2016 para un resumen reciente):

8. B': Ayer PEDRO compró un libro.

Tal y como señala Ortega-Santos (2016), puede ser que la descripción hecha desde la gramática generativa sea incorrecta, que haya variación entre los hablantes o que haya factores que interfieren (por ejemplo, la falta de naturalidad de la repetición de los argumentos en el orden VO_[F]S], donde _F representa el proceso de focalización, en comparación a una frase con el objeto topicalizado o pronominalizado). Lo cierto es que la aceptabilidad del orden [F]S]VO es

⁵ El caso del movimiento-Qu también ilustra la importancia de los estudios de corpus o del uso de la *web* para recoger datos. Tradicionalmente, la gramática generativa asume que el movimiento-Qu múltiple presente, por ejemplo, en lenguas como el serbocroata (Bošković 1997, entre otros) es inaceptable en español. Esta es la postura defendida también por RAE (2010, 3173). Bazaco (2014), sin embargo, recoge datos que muestran la existencia de movimiento-Qu múltiple en castellano.

consistente con la existencia del fenómeno de desacentuación (*destraining*) de la parte de la oración presupuesta en español (Gabriel *et al.* 2010 y Labastía 2006 en el caso del dialecto argentino; ver también Face y D'Imperio 2005 y Domínguez 2004); precisamente, esa era la propiedad que Zubizarreta (1998) señalaba como responsable del orden [F_S]VO en lenguas como el inglés.

Por supuesto, un/a investigador/a es susceptible de equivocarse y los avances investigativos se basan muchas veces en rectificar o cualificar observaciones previas. Sin embargo, estos datos son relevantes en el sentido de que los juicios de aceptabilidad introspectivos recogidos sin metodología experimental podrían no ser lo suficientemente sutiles como herramienta para describir esta área de la sintaxis o, simplemente, para resistir la influencia de generalizaciones anteriores, como, por ejemplo, la recién comentada de Zubizarreta (1998). Desde este punto de vista, es interesante destacar los trabajos de corpus desarrollados por Ocampo (2010) centrados en la distribución de los sujetos postverbales, en los que ha señalado que una situación comunicativa idealizada –sin mayor contexto– como la asumida por los juicios introspectivos no es lo suficientemente rica para capturar la complejidad de factores que determinan el orden de las palabras. En un contexto de este tipo, un paradigma experimental puede contribuir a esclarecer los datos, particularmente cuando el contraste de (7)B y (8)B', de acuerdo a los estudios experimentales, se basa en preferencias y no en el rechazo taxativo de ejemplos como (7)B.

Dicho esto, varias investigaciones no experimentales también han señalado que el orden SVO con el sujeto focalizado como respuesta a (7)A es posible (consultar, por ejemplo, Gutiérrez Ordóñez 1997, Casielles-Suárez 2004, Olarrea 2012, Zagona 2002 y Ortega-Santos 2008)⁶. De hecho, suponiendo que la perspectiva de Ocampo sea correcta, un experimento de aceptabilidad, a pesar de aplicar criterios rigurosos de recogida de los datos y presentar un análisis estadístico, seguiría dependiendo de los juicios introspectivos en un contexto altamente idealizado. Es por eso que, dependiendo de la pregunta de investigación, resulta importante considerar la adopción de otros métodos de recogida de datos. En este contexto, serían particularmente relevantes los estudios de corpus o el uso de herramientas de investigación experimental diferentes a los juicios de aceptabilidad como, por ejemplo, el uso del rastreo ocular (*eye-tracking*), la lectura autoguiada (*self-paced reading*) o las tareas de finalización de frases (*sentence completion tasks*).

3.3 Estructuras comparativas

Las estructuras comparativas permiten ejemplificar otra área en la que los experimentos resultan útiles. Esta utilidad no se deriva de las propiedades de la estructura en sí, sino de que algunas de sus variantes están estigmatizadas. Por ejemplo, Ortega-Santos (2013) discute los siguientes aspectos sobre las comparativas sintagmáticas de desigualdad en español, como las de (9): (i.) la cuestión en torno al tipo de análisis que debe adoptarse: elipsis clausal vs. el

⁶ Como señala un revisor anónimo, el orden de palabras que encontramos en (8) se encuentra en los corpus pues *Pedro*, a pesar de estar focalizado, es también el tópico de la frase en el sentido de que se trata de una oración sobre *Pedro* (ver el estudio de una variedad de lenguas desarrollado por Gernsbacher y Hargreaves 1992). Cabe mencionar que el concepto de tópico en este contexto no se corresponde a una noción sintáctica, sino que sería análogo a los juicios categóricos de Kuroda (1972) (para discusión relevante, consultar Raposo y Uriagereka 1995, Gallego 2007 y Ortega-Santos 2016). De acuerdo al revisor anónimo, no son, por lo tanto, incompatibles la función de tópico –definida en esos términos– y la de foco. Agradezco al revisor la oportuna apreciación, incluyendo la relevancia del trabajo de Gernsbacher y Hargreaves (1992).

análisis preposicional (ver Gutiérrez Ordóñez 1994, Brucart 2003 y Sáez del Álamo 1992, 1999 para el español y Napoli 1983, Hankamer 1973 o Bhatt y Takahashi 2011 para otras lenguas); (ii.) la cuestión de la evolución histórica en la sintaxis del español desde estructuras preposicionales al uso del *que* (Romero Cambrón 1998); (iii.) la cuestión de la ausencia de caso preposicional en el pronombre en (9) como factor inesperado para el análisis preposicional (Piera 1983, Plann 1984, Price 1990 o Reglero 2007).

9. Pedro es más inteligente que yo.

Tras documentar la existencia de estructuras comparativas como las de (10) en el español de Chile, Ortega-Santos (2013) enfrentó los siguientes problemas: (a) la presencia extremadamente limitada de la variante tanto en corpus como en internet en general; (b) un nivel alto de estigmatización –el hablante a quien el investigador escuchó esta variante negó usarla posteriormente con la consiguiente imposibilidad de poder recoger suficientes juicios para su análisis.

10. Pedro es más inteligente que mí.

Por estas razones, y ante la dificultad de localizar hablantes que pudieran proporcionar suficientes datos, Ortega-Santos (2013) diseñó un experimento en el que los participantes finalizaban oraciones susceptibles de usar cualquiera de las dos variables así como un experimento de juicios de aceptabilidad indirectos. En este segundo experimento, los participantes juzgan no la aceptabilidad de las frases en su propio idiolecto sino su naturalidad en una conversación contextualmente situada. La idea es que, bajo esas condiciones, los hablantes emitirían juicios menos normativos (ver Labov 1972 y Barbiers y Cornips 2000). De 30 participantes, 9 aceptaron la variante estigmatizada y proporcionaron datos que avalan el análisis preposicional de *que* en (10) –debido a la morfología del pronombre– frente a un análisis de elipsis clausal, (11). Un hipotético análisis alternativo en el cual *mí* recibe caso por defecto fue descartado por no encontrarse dicho caso preposicional en otros contextos sintácticos, (12).

11. Ana es más inteligente que yo / **mí soy inteligente*.

12. a. Ana y yo / **mí* ...

Coordinación

b. Yo / **mí*, parece que tengo la culpa.

Dislocación a la izquierda

c. Ella está ansiosa por conocerlos, y yo / **mí* también. *Vaciado*

Resulta pertinente señalar que la metodología experimental fue necesaria dada la estigmatización ya mencionada y la dificultad de localizar hablantes que usaran la variante objeto de estudio. En contraste, los hablantes de herencia de EEUU o incluso los hablantes de español panameño aceptan la estructura objeto de estudio sin tales niveles de estigmatización, razón por la cual sería plausible llevar a cabo una investigación en esas variedades sin la necesidad de un diseño experimental⁷.

⁷ RAE (2010, 3393) documenta, aunque no analiza, el uso de pronombre oblicuo en variedades del español ibérico y en ‘el lenguaje de los niños a edad temprana’, haciendo notar su estigmatización.

4. ESTUDIO EMPÍRICO DE LA FIABILIDAD DE LOS DATOS DEL ESPAÑOL LATINOAMERICANO

La discusión sobre la fiabilidad de los juicios de aceptabilidad basada en el estudio de casos concretos, como la discutida en la sección 3, si bien es relevante, tiene también limitaciones. Específicamente, la elección del paradigma objeto de estudio puede inclinar la balanza a favor o en contra de una metodología de recopilación de datos, pues no necesariamente todas las áreas de la sintaxis muestran los mismos problemas de fiabilidad en los datos. Por esta razón, aparecieron los primeros estudios a gran escala comparando los datos publicados que se recopilaron de modo informal con datos recopilados con metodología experimental, en particular, con la elección aleatoria de dichos datos. En Sprouse *et al.* (2013) se realizó este trabajo, centrado en el inglés, con una muestra aleatoria proveniente de artículos aparecidos en *Linguistic Inquiry* en el período 2001-2010 y se reveló un 93% de convergencia entre los datos de ambas metodologías (escala de Likert, prueba de hipótesis de una cola; ver el trabajo citado para los resultados de análisis estadísticos y diseños experimentales alternativos). Por lo demás, Sprouse y Almeida (2012) encontraron un 98% de convergencia en el caso de los datos de un libro de texto de sintaxis del inglés (estimación de las magnitudes y juicios de aceptabilidad con respuesta *sí-no*). Finalmente, en Ortega-Santos (en prensa), se presentan los primeros resultados de un proyecto de evaluación de la fiabilidad de los datos publicados en sintaxis hispánica sin uso de metodología experimental. En concreto, Ortega-Santos (en prensa) se centra en una muestra aleatoria de publicaciones de sintaxis aparecida en *Probus* (2006-2017). En el presente artículo, se comparan esos resultados con los correspondientes a publicaciones de gran impacto (100 citas o más; excluyendo libros y tesis), aparecidas entre el año 1971 y el 2009 (fecha de publicación promedio: 1990). El objetivo es (i.) comparar la fiabilidad de los datos del castellano con los datos del inglés (Sprouse *et al.* 2013), y (ii.) comparar el efecto del número de citas y/o la fecha de publicación en la fiabilidad de los datos. El diseño de estas investigaciones se basa en la idea de que ambas metodologías difieren en diversos aspectos y es necesario esclarecer cuáles de ellos (no) causan los hipotéticos problemas de fiabilidad. Por lo tanto, cada uno de ellos deberá ser estudiado por separado (Sprouse *et al.* 2013). Por ejemplo, se ha sugerido que en el caso de los datos de lenguas diferentes al inglés, la dificultad para encontrar revisores cualificados en el proceso de revisión por pares suscita mayores problemas de fiabilidad (Haider 2016 y Linzen y Oseki 2018). Para posibilitar la comparación de los estudios del español con los estudios del inglés, se siguió metodológicamente la investigación de Sprouse *et al.* (2013). Por lo demás, hay desacuerdo sobre (a) si los datos de publicaciones recientes son más sutiles que los datos de etapas anteriores y, posiblemente, más adecuados para el estudio experimental; (b) si el impacto de un artículo conlleva una mayor fiabilidad. Se remite al lector a Ortega-Santos (en prensa) y Ortega-Santos *et al.* (en evaluación) para una discusión detallada de estas cuestiones, de la metodología seguida y del análisis estadístico. Dichos aspectos son incluidos de forma resumida aquí, pues el objetivo del presente estudio es discutir en detalle los casos en los que no hubo convergencia entre ambas metodologías. A continuación, se incluye una lista no exhaustiva de los temas presentes en cada experimento:

- *Probus* (2006-2017): recomplementación, sintaxis de los clíticos de objeto, el marcado diferencial del objeto, objetos nulos, *se* espurio, preguntas eco con elemento interrogativo *in situ*, sintagmas nominales sin determinación, dislocación a la izquierda con clítico, operadores de grado y sujetos de infinitivo.

- Publicaciones de gran impacto: recomplementación, inversión verbo-sujeto, alternancia entre inergatividad e inacusatividad, huecos parasíticos, marcado diferencial del objeto, complementos de dirección, subida de clíticos, sintagmas nominales sin determinación, oraciones copulativas, el modo subjuntivo y el SE arbitrario.

La lista de obras incluidas en los experimentos (Experimento 1 centrado en *Probus* y Experimento 2 centrado en las publicaciones de gran impacto) se encuentra en el apéndice. Por su parte, la lista de datos incluidos en los experimentos y el análisis estadístico se pueden consultar en la página de *ResearchGate* del autor.

En cuanto a la metodología, se eligieron aleatoriamente pares mínimos para un experimento de juicios de aceptabilidad con una escala de Likert del 1 al 7. Siguiendo a Sprouse *et al.* (2013), se eligieron al azar frases inaceptables, dado que es más probable que sirvan para desarrollar un argumento en contraposición a frases aceptables que podrían simplemente ejemplificar una propiedad. En caso de que una frase inaceptable seleccionada no tuviera una versión aceptable, esta fue creada a partir de la discusión en el texto. Los datos se recogieron por medio de Amazon Mechanical Turk (MTurk) y se pagó a cada participante \$5 estadounidenses (ver Ortega-Santos 2019 para una introducción al uso de esta plataforma). Siguiendo las recomendaciones de Ortega-Santos (2019) para el estudio de lingüística hispánica en MTurk en poblaciones sin exposición significativa al inglés, se recogieron datos del español de Venezuela. En total, cada experimento consistió en cincuenta (50) pares mínimos, si bien no todos se incluyeron en el estudio de fiabilidad por no calificar. Para calificar, los datos tenían que ser recogidos de forma no experimental, ser parte del español moderno como primera lengua y no tener propiedades prosódicas específicas. Solo si se cumplían estas condiciones podían incluirse en los experimentos. Si el dialecto era consistente con el español de Venezuela (por ejemplo, se describía como español latinoamericano o simplemente español), podía ser incluido en el estudio de fiabilidad. Si se correspondía, por ejemplo, al español ibérico, podía ser incluido en el experimento, pero no en el estudio de fiabilidad. Esto es porque los experimentos originales estaban diseñados para contestar otras preguntas relacionadas tales como hasta qué punto la bibliografía es representativa del español de Venezuela, una cuestión ligeramente diferente a la de la fiabilidad⁸.

Los resultados de la direccionalidad de las respuestas (dirección prevista / dirección opuesta basada en las diferencias del promedio de los juicios de aceptabilidad de cada par mínimo) fueron los siguientes:

	Porcentaje en la dirección prevista	Dirección prevista N	Dirección opuesta N
<i>Probus</i> (2006-2017)	94.117%	32	2
Publicaciones de gran impacto	95.348%	41	2

Tabla 1. Direccionalidad de las respuestas basada en las diferencias entre el promedio de los juicios de aceptabilidad de cada par mínimo

⁸ En caso necesario, se adaptaron el léxico y la morfología de los datos al español venezolano. En cuanto a la muestra del experimento de *Probus* (Ortega-Santos en prensa), el cuestionario fue adaptado también para su uso en el español de Chile y de Puerto Rico con el objetivo de medir la distancia entre estos dialectos y el dialecto venezolano.

Se llevó a cabo una prueba de *t* de muestras pareadas (*paired-samples t-test*) para cada par mínimo siguiendo a Sprouse *et al.* (2013), para poder comparar los resultados⁹.

	<i>Probus</i> (2006-2017)		Publicaciones de gran impacto	
	N	Porcentaje	N	Porcentaje
Significativo en la dirección opuesta	2	5.882%	0	
Marginal en la dirección opuesta	0		0	
No-significativo en la dirección opuesta	0	0	2	4.651%
No-significativo en la dirección prevista	3	8.823%	5	11.627%
Marginal en la dirección prevista	1	2.941%	2	4.651%
Significativo en la dirección prevista	28	82.352%	34	79.069%

Tabla 2. Resultados por categorías. Valor *p* definido como $p < .05$ y $p \leq .1$ para significación estadística y marginalidad respectivamente en cada dirección.

4.1. Comentarios generales sobre los resultados

Si bien ambas muestras de Venezuela presentan porcentajes aproximadamente similares de resultados significativos en la dirección prevista, solo en el caso de la muestra más reciente, de *Probus*, se encuentran resultados significativos en la dirección opuesta. Por lo demás, varias razones explican la diferencia de resultados en el español y el inglés. Por ejemplo, tal y como se ha señalado, la elección del dialecto del estudio en castellano, el español venezolano, se determinó por la facilidad para recoger datos de esta variedad por medio de la plataforma MTurk. Las publicaciones cuyos datos fueron incluidos en los experimentos se presentan, por lo general, como ‘español’ o ‘español latinoamericano’; son consistentes en principio con el dialecto y, por lo tanto, fueron incluidas en el estudio de fiabilidad. No se puede excluir, en cualquier caso, la existencia de variación dialectal (ver la discusión detallada en Ortega-Santos *et al.* en evaluación sobre este y otros temas). En este sentido, el presente estudio constituye una aportación relevante al estudio de la fiabilidad de los datos de la sintaxis del español, pero para lograr una visión más completa, sería necesario controlar ese factor. El presente trabajo se centra en discutir detalladamente los datos no replicados, con el objeto de evidenciar otros factores causantes de este resultado (sección 4.2).

⁹ En el caso de la muestra de *Probus*, Ortega-Santos (en prensa) incluye un análisis de los resultados del estudio de fiabilidad en los dialectos de Venezuela, Chile y Puerto Rico en términos de 3×2 ANOVAs.

4.2. Comentarios sobre datos no replicados

4.2.1. Datos de *Probus* (2006-2017)

La tabla 3 incluye los datos para los que el contraste en la valoración de la aceptabilidad no llegó a la significación estadística en la dirección prevista.

	<u>Identificador</u>
Significativo en la dirección opuesta	
13. a. *He leído el libro increíblemente interesante de Rómulo Gallegos.	Pastor08.57b
b. He leído el libro poco interesante de Rómulo Gallegos.	Pastor08.35b
14. a. *Juan es como yo de feo para ser modelo.	Pastor10.P47
b. Ese hombre es un montón de gordo para su estatura y edad.	Pastor10.8
Marginal en la dirección opuesta	
N/A	
No-significativo en la dirección opuesta	
N/A	
No-significativo en la dirección prevista	
15. a. *Científicos estudian el asunto.	MacDonald16.23b
b. Los científicos estudian el asunto.	MacDonald16.23a
16. a. *Y tú no sabes quién llegó cómo?	Reglero07.18b
b. Y tú no sabes cómo llegó quién?	Reglero07.18a
17. a. **/?Y tú le diste a quién la guitarra?	Reglero07.5b
b. Y tú le diste la guitarra a quién?	Reglero07.5a
Marginal en la dirección prevista	
18. a. *Pedro tocó.	Armstrong16.47a
b. Pedro tocó la mesa.	Armstrong16.47a

Tabla 3. Datos para los que la diferencia en la puntuación de los juicios de aceptabilidad no alcanzó significación estadística en la dirección prevista. El indicador hace referencia al primer autor del artículo, año y página o número de frase.

En el caso de los contrastes en (13) y (14) estos son los únicos que alcanzaron la significación estadística en la dirección opuesta. El primer caso muestra la dificultad de investigar matices sutiles relacionados con la expresividad y el énfasis retórico. Debido a esa sutileza, la recopilación de datos por medio de experimentos puede ser necesaria. Dicho esto, el experimento incluyó otros datos provenientes del mismo artículo y fueron replicados. En cuanto a (14), este resultado parece relacionado con el requerimiento de *montón* de tener un argumento contable, a diferencia de lo que los datos sugieren (ver Saab 2010 para discusión de las propiedades de *montón*).

En el caso de (15)15, cuyos resultados son no-significativos en la dirección prevista, este contraste ejemplifica la variación interdialectal y/o la controversia concerniente a la distribución de los sintagmas nominales sin determinante o escuetos (ver Benedicto 1998), en contra de lo tradicionalmente asumido en el campo¹⁰. De hecho, hay variación dialectal en las propiedades de la posición de sujeto en español, lo que también puede afectar la valoración de los datos. En cuanto a (16) y (17), cuyos resultados también son no-significativos en la dirección prevista, presentan características particularmente interesantes: Son preguntas eco que desde el punto de vista pragmático dependen de un contexto precedente. La ausencia de ese contexto en el experimento –sobrentendido en el artículo correspondiente (Reglero 2007: 268, n. 2)–, junto con el hecho de que este tipo de frases se corresponden a un registro coloquial, puede explicar los resultados¹¹. En general, ambos factores son relevantes para comprender las limitaciones de los estudios de fiabilidad: En las investigaciones publicadas se pueden sobrentender factores o características (p.ej., el contexto pragmático de la frase). En cambio, los participantes de los experimentos no tienen acceso a dicha información. La inclusión de un contexto podría salvar el problema del empobrecimiento de los datos en el experimento; sin embargo, en casos donde dicha riqueza va más allá del contexto lingüístico (por ejemplo, detalles de la derivación de la frase), hay una pérdida de información que posiblemente no sea compensable, lo que introduciría una asimetría importante entre la publicación y el experimento de fiabilidad. Por lo demás, los experimentos con juicios de aceptabilidad no necesariamente son la metodología más apropiada para registros coloquiales o estigmatizados (Cornips y Poletto 2005). De esto se deduce que parte de los resultados de los estudios de fiabilidad son determinados por la metodología de recopilación de datos. Finalmente, en lo que respecta a (18), que alcanzó la marginalidad estadística en la dirección correcta, la valoración de este contraste puede estar relacionada con un factor no controlado: La existencia de una interpretación gramatical para la versión (a), en concreto, con la interpretación de ‘Pedro tocó un concierto’.

4.2.2. Datos de las publicaciones de alto impacto

La tabla 4 incluye los datos para los que el contraste en la valoración de la aceptabilidad no llegó a la significación estadística en la dirección prevista en el caso de la muestra de las publicaciones de alto impacto en el español de Venezuela.

¹⁰ RAE (2010, 2530) afirma que los grupos nominales escuetos “no ejercen en español la función de sujeto preverbal fuera del registro particular de los titulares de prensa (...) y de algunos otros contextos restringidos”, como grupos nominales escuetos coordinados o modificados por sintagmas preposicionales. Menciona también otros factores relevantes que pudieran afectar los resultados: (i.) los grupos nominales escuetos en titulares de prensa están más extendidos en Latinoamérica, siendo que el presente estudio se enfoca en una variedad lingüística latinoamericana; (ii.) la focalización del sintagma nominal preverbal sin determinante mejora la aceptabilidad de la frase. Pudiera ser que los hablantes asuman dicho proceso de focalización al juzgar la frase (15).

¹¹ De acuerdo a Reglero, la conjunción inicial y es opcional; en el experimento fue incluida. Para una discusión general sobre las preguntas de eco, el lector puede consultar RAE (2010, 3182-3188).

	<u>Identificador</u>
Significativo en la dirección opuesta	
N/A	
Marginal en la dirección opuesta	
N/A	
No-significativo en la dirección opuesta	
19. a. *Alrededor del parque falta césped.	FernándezSoriano 92.19c
b. En este pueblo falta césped alrededor del parque.	FernándezSoriano 92.20d
20. a. *No le parece que Lola habla palabra de francés.	Rivero71.35b
b. No le parece que Lola hable palabra de francés.	Rivero71.35
No-significativo en la dirección prevista	
21. a. *?Esta es la señal que hizo a varios hombres el alcalde.	Ordóñez98.51b
b. Esta es la señal que les hizo a varios hombres el alcalde.	Ordóñez98.51b
22. a. *Me contestó por qué lo quería saber.	Plann82.24d
b. Me contestó que por qué lo quería saber.	Plann82.25d
23. a. ?*Esa respuesta de que por qué lo tenía que hacer ella le costó el puesto.	Plann82.fn4ii
b. Esa respuesta de por qué lo tenía que hacer ella le costó el puesto.	Plann82.fn4ii
24. a. *Juan juega constantemente limpio a las cartas.	Suñer94.14d
b. Juan juega limpio a las cartas constantemente.	Created from the context
25. a. *Un edificio de dos pisos propuso la alcaldía que la urbanización aprobara que los albañiles derrumbaran.	Torrego84.27b
b. Un edificio de dos pisos propuso la alcaldía que aprobara la urbanización que derrumbaran los albañiles.	Torrego84.27a
Marginal en la dirección prevista	
26. a. *Café y leche son caros.	Contreras86.84
b. Viejos y niños escuchan con atención sus palabras.	Contreras86.83
27. a. *Se reunieron a los miembros de la junta.	Otero72.6a
b. Se persigue a los que disienten.	Otero72.6a

Tabla 4. Datos para los que la diferencia en la puntuación de los juicios de aceptabilidad no alcanzó significación estadística en la dirección prevista en el caso de la muestra de las publicaciones de alto impacto en el español de Venezuela

En el caso de (19), que no es significativo en la dirección opuesta, no está claro qué determina los resultados, si bien (b) es en principio más compleja que (a), debido a la presencia de un sintagma preposicional adicional. Por lo que respecta a (20), que es también no significativa en la dirección opuesta, la propia investigadora (Rivero 1971) menciona la existencia de variación dialectal en los juicios. Un estudio de la fiabilidad como el presente no es lo suficientemente detallado como para verificar que estemos ciñéndonos a la variedad / sociolecto correcto al estudiar los datos. Tampoco las investigaciones teóricas incluyen descripciones lo suficientemente detalladas para hacerlo, a diferencia de la investigación en el campo de la denominada sociosintaxis. Dicha subdisciplina podría prestar atención sistemáticamente a las variables relevantes y contribuir a clarificar el estatus de los datos. Por otra parte, RAE (2010: 1884) incluye la noción de probabilidad como parte de los contextos en los que aparece el subjuntivo, tal y como se ve en (20).

Para (21), no significativa en la dirección prevista, no está claro qué determina la falta de convergencia. (22) y (23), no significativas en la dirección prevista, son estructuras de recomplementación restringidas a un registro oral o coloquial. Tal y como se mencionó en la sección 4.2.1, ese tipo de datos no son los más apropiados para experimentos de aceptabilidad que pueden ser percibidos como un examen. Por su parte, (23) pudiera haber sido interpretado como aceptable por algunos participantes; en concreto, es posible que algunos de ellos ignoraran la naturaleza interrogativa de *por qué* y lo interpretaran como *porque* (causal) debido a la similitud ortográfica (“Él me respondió porque quería saberlo”). Los resultados de (24), no significativos en la dirección prevista, pueden estar relacionados con el hecho de que para que *constantemente* aparezca en la periferia derecha de la frase en (b), tendría que estar focalizado (ver Zubizarreta 1998 y Ortega-Santos 2016). No se proporciona contexto alguno para tal focalización; como consecuencia, la frase puede resultar pragmáticamente inadecuada. Por lo demás, el ejemplo (a) es especial en el sentido que fue creado por el autor para el experimento, lo que introduce un factor extra, más allá de los considerados en Suñer (1994). El caso de (25), no significativo en la dirección prevista, constituye parte de la evidencia de Torrego (1984) a favor de la naturaleza cíclica del movimiento-Qu, evidencia que ha sido considerada sutil en otros estudios (ver, por ejemplo, Gallego 2007). Por su parte, la oración en (26), marginal en la dirección prevista, se basa en la distribución de los sintagmas nominales sin determinante explícito. Tal y como se mencionó con respecto a (15) en la muestra de *Probus* (sección 4.2.1), este resultado podría estar relacionado con la existencia de variación dialectal en dicha distribución (ver Benedicto 1998 y la nota al pie número 10 en este artículo; para discusión de los grupos nominales escuetos coordinados, ver RAE 2010, 2531 y 2638). Finalmente, en el caso de (27), la gramática prescriptiva coincide con Otero (1972) en que ese contraste existe. Sin embargo, se ha demostrado que la concordancia en plural con el objeto de la preposición, (27b), está presente en variedades del español (ver Gallego 2019). En consecuencia, pudiera tratarse de un caso de variación dialectal, no documentado en Otero (1972).

5. LIMITACIONES DE LOS ESTUDIOS DE RÉPLICA O REPETICIÓN

En la sección anterior, se han señalado varios factores que limitan la relevancia de los estudios experimentales de fiabilidad, como por ejemplo la hipotética imposibilidad de incluir en los experimentos el contexto rico de la discusión presente en los trabajos publicados. También se ha señalado la dificultad de recopilar datos de variantes o construcciones

estigmatizadas. Dicho esto, en Ortega-Santos (en prensa) se estudian los datos de la muestra de *Probus* en otros dos dialectos adicionales (el español de Chile y el de Puerto Rico) y se observa que los resultados de las construcciones estigmatizadas presentan divergencias entre dialectos. En este sentido, es importante mencionar que dichos factores no llevan consistentemente a diferencias entre las metodologías de recopilación de datos, aunque pueden constituir áreas problemáticas. De hecho, un mismo fenómeno, como por ejemplo la recomplementación, se encuentra incluido varias veces en la muestra y varía en sus resultados (convergentes/no convergentes). Por lo demás, la variación dialectal –con frecuencia gran desconocida en el caso de gran parte de los datos de la sintaxis teórica– puede haber afectado los resultados.

Asimismo, cabe también mencionar que un análisis en términos de valores *p* –adoptado de la bibliografía relevante para poder contextualizar los resultados– presenta un grado de arbitrariedad (ver Ortega-Santos *et al.* en evaluación) para discusión. Otros resultados como por ejemplo el tamaño del efecto Cohen son relevantes en este contexto.

6. RECOMENDACIONES

Con el objetivo de evitar críticas a aspectos metodológicos y/o hipotéticos problemas de replicación o fiabilidad en los estudios de sintaxis no experimental, se recomienda claridad en cuanto al origen de los datos (número de hablantes consultados, origen dialectal, etc.). El análisis del corpus objeto de estudio muestra la magnitud del problema: De la totalidad de los artículos incluidos en el estudio de fiabilidad (ver el apéndice para consultar la lista) el origen de los juicios de aceptabilidad solo se presenta (i.) en el caso de citas de datos de estudios previos, (ii.) cuando un/a revisor/a o un colega ha proporcionado un juicio concreto; (iii.) cuando se complementa la discusión con datos extraídos de corpus u obras literarias, etc. (p. ej., Suñer 1988 y Suñer 1994 en el caso de obras literarias y Fábregas 2007 y Suñer 1998 en el caso de corpus); (iv.) en los agradecimientos (Aissen y Perlmutter 1976, Jaeggli 1986, Suñer 1988, Suñer 2006, Reglero 2007). Las siguientes citas ilustran el caso de (iv.) (traducción y énfasis míos):

- I would also like to thank Hagit Borer, Robin Cooper, John Hawkins, Nina Hyams, David Lebeaux, Mario Montalbetti, Tim Stowell, and three anonymous reviewers for NLLT for helpful comments, criticism, suggestions, and *data*. (Jaeggli 1986: 43)
(También quisiera agradecer a Hagit Borer, Robin Cooper, John Hawkins, Nina Hyams, David Lebeaux, Mario Montalbetti, Tim Stowell y tres revisores anónimos de NLLT por sus útiles comentarios, críticas, sugerencias y *datos*.)
- I am indebted to two anonymous reviewers and to Ignacio Bosque, Vicky Escandell-Vidal, Olga Fernández Soriano, Klaus von Heusinger, Daniel Jacob, Georg Kaiser, Udo Klein, Louise McNally, Teresa Parodi and Cristina Sánchez for *judgements*, suggestions and ideas. (Leonetti 2008: 33, n. 1)
(Estoy en deuda con dos revisores anónimos y con Ignacio Bosque, Vicky Escandell-Vidal, Olga Fernández Soriano, Klaus von Heusinger, Daniel Jacob, Georg Kaiser, Udo Klein, Louise McNally, Teresa Parodi y Cristina Sánchez por sus *juicios*, sugerencias e ideas.)
- A great thanks also to my *Spanish-speaking informants*: Fernando Sánchez Miret, Rosa Ana Martín Vegas, and Marietta Calderón Tichy. (Pöll 2007: 93, n. 1)
(Muchas gracias también a mis *informantes hispano-hablantes*: Fernando Sánchez Miret, Rosa Ana Martín Vegas, and Marietta Calderón Tichy.)

El problema de incluir el origen de los juicios en los agradecimientos es que, a diferencia de los estudios experimentales, no es una práctica estandarizada y, por lo tanto, no es lo suficientemente rigurosa. Específicamente, faltan detalles sobre la formación en lingüística teórica o posibles conflictos de interés de las personas que proporcionaron juicios, su perfil lingüístico (qué dialecto hablan o si son bilingües), si juzgaron todos los datos del estudio o solo parte, si fueron expuestos a varias frases (lexicalizaciones) a la hora de juzgar una propiedad concreta o si hubo distractores, etc. Ser sistemáticos en estos detalles, acercándose hasta cierto punto a las prácticas generalizadas en los estudios experimentales, ayudaría a evitar críticas innecesarias y a poder evaluar la calidad de las investigaciones.

Este problema resulta también evidente en el siguiente contexto: Cuando se atestigua variación en los juicios de aceptabilidad, es frecuente encontrar referencias a los informantes, pero dichas referencias no son lo suficientemente explícitas en el sentido de que se tiende a mencionar los juicios de *algunos* o *ciertos hablantes*, sin detalles del número de hablantes consultados, su sociolecto o su dialecto. Los siguientes ejemplos son representativos de esta característica relativamente frecuente (mi traducción y énfasis):

- However, both searches in corpora and native speaker's intuitions show that there is nothing unusual in this construction in Spanish. *Some speakers* also allow... (Fábregas 2007: 169)
(Sin embargo, tanto las búsquedas en corpus como las intuiciones de hablantes nativos muestran que no hay nada inusual en esta construcción en español. *Algunos hablantes* también permiten...)
- As pointed out by an anonymous reviewer, the following example appears possible for *some speakers* of Spanish. (Ordóñez y Treviño 1999: 43, n. 4)
(Tal y como señala un revisor anónimo, el siguiente ejemplo parece ser posible para *algunos hablantes* de español.)
- ...illustrates that, without Asp_{SE}, an *in*-adverbial or a *for*-adverbial is grammatical (although the *for*-adverbial is marked for *some speakers*). (MacDonald 2016: 75)
(...ejemplifica que, sin Asp_{SE}, una expresión adverbial con (la preposición) *en* o con la preposición *por* es gramatical (aunque en el caso de *por* para *algunos hablantes* la estructura se considera marcada)

Resulta claro que los/as sintactistas de la muestra tratan sistemáticamente de proporcionar datos relevantes sobre la validez de los juicios. Sería útil para el campo que se proveyeran detalles más explícitos como el mencionado sobre el número de hablantes o el dialecto abordado.

7. CONCLUSIÓN

Este artículo trató de proporcionar una visión general sobre la investigación de la fiabilidad de los juicios de aceptabilidad en la gramática generativa tal y como se da en el español. Específicamente, se discutió la metodología de recogida de datos y se planteó cómo la adopción de estándares científicos de otras disciplinas como la psicolingüística abre la puerta al estudio de las divergencias en los resultados de investigaciones que no usan protocolos experimentales de recogida y análisis de datos frente a investigaciones experimentales. Varios casos concretos ilustraron esas diferencias (el estudio de islas-Qu, el orden de las palabras en la focalización en

el sujeto y la microvariación en estructuras comparativas). Finalmente, se presentaron los resultados de dos experimentos de fiabilidad en los que se estudiaban datos elegidos al azar de artículos de sintaxis aparecidos en *Probus* (2006-2017) y en publicaciones de alto impacto (100+ citas). En concreto, dichos datos, recopilados originalmente sin protocolos experimentales y estudiados sin análisis estadísticos, fueron comparados con los resultados de su estudio con protocolos experimentales. Se discutieron los factores que determinan convergencia o divergencia entre ambos métodos de recopilación de datos, así como las limitaciones de los estudios de fiabilidad. Finalmente, se incluyeron varias recomendaciones para las investigaciones sintácticas no experimentales.

REFERENCIAS BIBLIOGRÁFICAS

- Armstrong, Grant. 2016. Spanish unspecified objects as null incorporated nouns, en *Probus*, 28: 165-229.
- Barbiers, Sjef y Leonie Cornips. 2000. Introduction to syntactic microvariation, en Sjef Barbiers, Leonie Cornips y Susanne van der Kleij (eds.), *Syntactic microvariation*, Meertens Institute Electronic Publications in Linguistics (MIEPiL): 1-12. <http://www.meertens.knaw.nl/books/synmic/>.
- Bazaco, Carmelo. 2014. Multiple Wh questions in Spanish. A new interpretation. Ponencia presentada en el *Hispanic Linguistics Symposium 2014*, organizado por la Universidad de Purdue, 13 al 16 de noviembre de 2014.
- Benedicto, Elena E. 1998. Verb movement and its effects on determinerless plural subjects, en Armin Schwegler, Bernard Tranel y Myriam Uribe-Etxebarria (eds.), *Romance linguistics: Theoretical perspectives*, Amsterdam, John Benjamins: 25-40.
- Bever, Thomas G. 1970. The cognitive basis for linguistic structures, en John R. Hayes (ed.), *Cognition and the Development of Language*, New York, NY, Wiley: 279-362.
- Bhatt, Rajesh y Soichi Takahashi. 2011. Reduced and unreduced phrasal comparatives, en *Natural Language and Linguistic Theory*, 29: 581-620.
- Bošković, Željko. 1997. Superiority effects with multiple *wh*-fronting in Serbo-Croatian, en *Lingua*, 102: 1-20.
- Brucart, José María. 2003. Adición, sustracción y comparación, en Fernando Miret Sánchez (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*, Tübingen, Max Niemeyer: 11-60.
- Casielles-Suárez, Eugenia. 2004. *The Syntax-Information Structure Interface*, London: Routledge.
- Chomsky, Noam y George A. Miller. 1963. Introduction to the formal analysis of natural languages, en R. Duncan Luce, Robert R. Bush y Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2, New York, NY, John Wiley: 269-321.
- Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*, The Hague, The Netherlands, Mouton.
- Chomsky, Noam. 1973. Conditions on Transformations, en Stephen R. Anderson y Paul Kiparsky (eds.), *Festschrift for Morris Halle*, New York, Holt, Rinehart y Winston: 232-286.
- Contreras, Heles. 1986. Spanish bare NP's and the ECP, en Ivonne Bordelais, Heles Contreras y Karen Zagona (eds.), *Generative Studies in Spanish Syntax*, Dordrecht, Foris Publications: 25-49.
- Cornips, Leonie y Cecilia Poletto, C. 2005. On standardising syntactic elicitation techniques, *Lingua*, 115: 939-957.
- Cowart, Wayne. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*, London, Sage Publications.
- Domínguez, Laura. 2004. *Mapping Focus: The Syntax and Prosody of Focus in Spanish*. Tesis de Doctorado, Boston University, Boston. Inédita.
- Edelman, Shimon y Morten H. Christiansen, 2003. How seriously should we take Minimalist syntax? en *Trends in Cognitive Sciences*, 7: 60-61.
- Face, Timothy L. y Mariapaola D'Imperio. 2005. Reconsidering a focal typology: evidence from Spanish and Italian, en *Rivista di Linguistica* 17: 271-289.
- Fernández-Soriano, Olga. 1999. Two Types of Impersonal Sentences in Spanish: Locative and Dative Subjects, en *Syntax*, 2: 101-140.
- Ferreira, Fernanda. 2005. Psycholinguistics, formal grammars, and cognitive science, en *Linguistic Review*, 22: 365-380.
- Gabriel, Christoph, Iñigo Feldhausen, Andrea Pešková, Laura Colantoni, Su-Ah Lee, Valeria Arana y Leopoldo Labastía. 2010. Argentinian Spanish Intonation, en Pilar Prieto, Paolo Roseano (eds.), *Transcription of Intonation of the Spanish Language*, Munich. Lincom: 285-387.

- Gabriel, Christoph. 2007. *Fokus im Spannungsfeld von Phonologie und Syntax*, Frankfurt, Vervuert Verlag.
- Gallego, Ángel J. 2007. Phase Theory and parametric variation. Tesis de Doctorado, Universitat Autònoma de Barcelona, Barcelona. [Publicada como Gallego, Ángel J. 2010. *Phase Theory*, Amsterdam, John Benjamins].
- Gallego, Ángel J. 2019. Long Distance Agreement in Spanish Dialects, en Ludovico Franco, Mihaela Marchis Moreno y Matthew Reeve (eds.), *Agreement, case and locality in the nominal and verbal domains*, Berlin, Language Science Press: 85-108.
- Gernsbacher Morton Ann y David Hargreaves. 1992. The privilege of primacy: Experimental data and cognitive explanations, en Doris L. Payne (ed.), *Pragmatics of word order flexibility*, Philadelphia, John Benjamins: 83-116.
- Gibson, Edward y Evelina Fedorenko. 2010a. Weak quantitative standards in linguistics research, en *Trends in Cognitive Sciences*, 14: 233-234.
- Gibson, Edward y Evelina Fedorenko. 2010b. The need for quantitative methods in syntax and semantics research, en *Language and Cognitive Processes*, 28: 88-124.
- Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies, en *Cognition*, 68: 1-76.
- Gutiérrez Ordóñez, Salvador. 1994. *Estructuras comparativas*, Madrid, Arco Libros.
- Gutiérrez Ordóñez, Salvador. 1997. *Temas, remas, focos, tópicos y comentarios*, Madrid, Arco Libros.
- Haider, Hubert. 2016. *Incredible syntax – between Cognitive Science and imposture*. Manuscrito, University of Salzburg.
- Hankamer, Jorge. 1973. Why there are two than's in English, en Claudia Corum, Cedric T. Smith-Stark y Ann Weiser (eds.), *Proceedings of the CLS 9*, Chicago, Chicago Linguistics Society: 179-191.
- Hoot, Bradley. 2012. *Presentational focus in heritage and monolingual Spanish*. Tesis de Doctorado, University of Illinois, Chicago. Inédita.
- Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*, Dordrecht, Foris.
- Jiménez-Fernández, Ángel. 2015. Towards a typology of focus: Subject position and microvariation at the discourse-syntax interface, en *Ampersand: An International Journal of General and Applied Linguistics*, 2: 49-60.
- Kuroda, Sige-Yuki. 1972. The categorial and the thetic judgment, en *Foundations of Language*, 9: 153-185.
- Labastía, Leopoldo. 2006. Prosodic Prominence in Argentinian Spanish, en *Journal of Pragmatics*, 38: 1677-1705.
- Labov, William. 1972. Negative attraction and negate concord in English grammar, en *Language*, 48: 773-818.
- Leal Méndez, Tania y Roumyana Slabakova. 2011. Pragmatic Consequences of P-movement and Focus Fronting in L2 Spanish: Unraveling the Syntax- Discourse Interface, en Julia Herschensohn y Darren Tanner (eds.), *Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011)*, Somerville, MA, Cascadilla Proceedings Project: 63-75.
- Lewis, Richard L. y Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval, en *Cognitive-Science*, 29: 1- 45.
- Linzen, Tal y Yohei Oseki. 2018. The reliability of acceptability judgments across languages, en *Glossa*, 3(1): 100-125.
- López Sancho, Sergio. 2015. *Testing syntactic islands in Spanish*. Tesis de Máster, Universidad del País Vasco/Euskal Herriko Unibertsitatea, Vitoria-Gasteiz. Inédita.
- Lowie, Wander y Bregtje Seton. 2013. *Essential Statistics for Applied Linguistics*, New York, Palgrave Macmillan.
- MacDonald, Jonathan. 2016. Spanish aspectual *se* as an indirect object reflexive: The import of atelicity, bare nouns, and léista PCC repairs, en *Probus*, 28: 73-117.
- Montalbetti, Mario. 1984. *On the interpretation of pronouns*. Tesis de Doctorado, MIT, Cambridge. Inédita.
- Muntendam, Antje. 2009. *Linguistic transfer in Andean Spanish: Syntax or pragmatics?* Tesis de Doctorado, University of Illinois, Urbana-Champaign. Inédita.
- Muñoz Pérez, Carlos. 2014. Una nota acerca del uso de juicios den teoría gramatical, en *Signo y Señal*, 26: 107-120.
- Napoli, Donna-Jo. 1983. Comparative ellipsis: a phrase structure account, en *Linguistic Inquiry*, 14: 675-694.
- Ocampo, Francisco. 2010. The place of conversational data in Spanish syntax: Topic, focus and word order, en *Studies in Hispanic and Lusophone Linguistics*, 3: 533-543.
- Olarrea, Antxon. 2012. Word order and information structure. José Ignacio Hualde, Antxon Olarrea y Erin O'Rourke (eds.), *Handbook of Hispanic Linguistics*, Oxford, Wiley-Blackwell: 603-629.
- Ordóñez, Francisco. 1997. *Word order and clause structure in Spanish and other Romance languages*. Tesis Doctoral, CUNY Graduate Center. [Publicada como Ordóñez, Francisco. 2000. *The clausal structure of Spanish: a comparative perspective*, New York, Garland].
- Ordóñez, Francisco. 1998. Post-Verbal Asymmetries in Spanish, en *Natural Language and Linguistic Theory*, 16: 313-346.
- Ormazabal, Javier y Juan Romero. 2013. Object Clitics, Agreement and Dialectal Variation, en *Probus*, 25: 301-344.

- Ortega-Santos, Iván. 2008. *Projecting Subjects in Spanish and English*. Tesis de Doctorado, College Park, University of Maryland. Inédita.
- Ortega-Santos, Iván. 2013. Microvariation in Spanish Comparatives, en *Catalan Journal of Linguistics*, 12: 175-192.
- Ortega-Santos, Iván. 2016. *Focus-related Operations at the Right Edge in Spanish: Subjects and Ellipsis*, Amsterdam, John Benjamins.
- Ortega-Santos, Iván. 2019. Crowdsourcing for Hispanic Linguistics: Amazon's Mechanical Turk as a source of Spanish data, en *Borealis: An International Journal of Hispanic Linguistics*, 8: 187-215.
- Ortega-Santos, Iván. En prensa. Dialect distance and data assessment in Chilean, Venezuelan and Puerto Rican Spanish, en Brandon M.A. Rogers y Mauricio A. Figueroa Candia (eds.), *Lingüística del castellano chileno: estudios sobre variación, innovación, contacto e identidad*, Wilmington, Vernon Press.
- Ortega-Santos, Iván, Juan Uriagereka y Ángel J. Gallego. En evaluación. Reliability in theoretical syntax: recent vs. classic high-impact data and language choice.
- Ortega-Santos, Iván, Lara Reglero y Jon Franco. 2019. Wh-Islands in L2 Spanish and L2 English: Between Poverty of the Stimulus and Data Assessment, en *Fontes Lingvæ Vasconvm stvdia et documenta*, 126, 435-471.
- Otero, Carlos. 1972. Acceptable Ungrammatical Sentences in Spanish, en *Linguistic Inquiry*, 3: 233-242.
- Pañeda, Claudia, Sol Lago, Elena Vares, João Veríssimo y Claudia Felser. 2020. Island effects in Spanish comprehension, en *Glossa: A Journal of General Linguistics*, 5: 1-30.
- Pastor, Alberto. 2008. Split analysis of gradable adjectives in Spanish, en *Probus*, 20: 257-299.
- Pastor, Alberto. 2010. Predicative degree constructions in Spanish, en *Probus*, 22: 27-71.
- Piera, Carlos. 1983. Spanish comparatives, deletion operations and the ECP, en *Cornell Working Papers in Linguistics*, 4: 185-199.
- Plann, Susan. 1982. Indirect Questions in Spanish, en *Linguistic Inquiry*, 13: 297-312.
- Plann, Susan. 1984. The syntax and semantics of más/menos... que versus más/ menos... de in comparatives of inequality, en *Hispanic Linguistics*, 1: 191-213.
- Price, Susan. 1990. *Comparative constructions in Spanish and French syntax*, London, Routledge.
- Raposo, Eduardo P. y Juan Uriagereka. 1995. Two Types of small clauses, en Anna Cardinaletti y Maria Teresa Guasti (eds.), *Small Clauses*, New York, Academic Press: 179-206.
- Real Academia Española. 2010. *Nueva gramática de la lengua española*, 2 vols., Madrid, Espasa Calpe.
- Reglero, Lara. 2007. On Spanish comparative subdeletion constructions, en *Studia Lingüística*, 61: 130-169.
- Reglero, Lara. 2007. Wh-in-situ interrogatives in Spanish, en *Probus*, 19: 267-297.
- Rivero, María-Luisa. 1971. Mood and Preposition in Spanish, en *Foundations of Language*, 7: 305-336.
- Romero Cambrón, Ángeles. 1998. *Historia sintáctica de las construcciones comparativas de desigualdad*, Cuenca, Ediciones de la Universidad de Castilla-La Mancha.
- Ross, John R. 1967. *Constraints on Variables in Syntax*. Tesis de Doctorado, MIT, Cambridge. [Publicada como Ross, John, R. 1986. *Infinite syntax!*, Norwood, NJ, Ablex.]
- Saab, Andrés. 2010. Problemas actuales de gramática española: microvariación y estructuras. Manuscrito, Linguistics, Leiden University, Netherlands.
- Sáez del Álamo, Luis. 1992. La elipsis verbal en las construcciones comparativas, en Carlos Martín Vide (ed.), *Lenguajes naturales y lenguajes formales: actas del VII congreso de lenguajes naturales y lenguajes formales*, Barcelona, Universidad de Barcelona: 561-568.
- Sáez del Álamo, Luis. 1999. Los cuantificadores: las construcciones comparativas y superlativas, en Ignacio Bosque y, Violeta Demonte, Violeta (dirs.), *Gramática descriptiva de la lengua española*, Madrid, Espasa Calpe: 1129-1188.
- Snyder, William. 2000. An Experimental Investigation of Syntactic Satiation Effects, en *Linguistic Inquiry*, 31: 575-582.
- Sprouse, Jon y Diogo Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's Core Syntax, en *Journal of Linguistics*, 48: 609-652.
- Sprouse, Jon, Carson T. Schütze y Diogo Almeida. 2013. Assessing the reliability of journal data in syntax: Linguistic Inquiry 2001-2010, en *Lingua*, 134: 219-248.
- Sprouse, Jon, Ivan Caponigro, Ciro Greco y Carlo Cechetto. 2016. Experimental syntax and the variation of island effects in Italian and English, en *Natural Language and Linguistic Theory*, 34: 307-344.
- Suñer, Margarita. 1994. V-Movement and the Licensing of Argumental Wh-Phrases in Spanish, en *Natural Language and Linguistic Theory*, 12: 335-372.
- Torrego, Esther. 1984. On Inversion in Spanish and Some of Its Effects, en *Linguistic Inquiry*, 15: 103-129.
- Vicente, Luis. 2013. On the causes of Superiority effects in Spanish: preliminary results and prospect, en *Catalan Journal of Linguistics*, 12: 283-300.
- Wasow, Thomas y Jennifer Arnold. 2005. Intuitions in linguistic argumentation, en *Lingua*, 115: 1481-1496.

Zagona, Karen. 2000. *The syntax of Spanish*, Cambridge, Cambridge University Press.
 Zubizarreta, Maria Luisa. 1998. *Prosody, Focus and Word Order*, Cambridge, MA, MIT Press.

APÉNDICE
LISTA DE OBRAS INCLUIDAS EN LOS EXPERIMENTOS¹²

1.1 Artículos de sintaxis española publicados en *Probus* (2006-2017) incluidos en el estudio de fiabilidad de los datos

1. Armstrong, Grant. 2016. Spanish unspecified objects as null incorporated nouns, en *Probus*, 28: 165-229.
2. Demonte, Violeta y Olga. Fernández-Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system, en *Probus* 21: 23-49.
3. Leonetti, Manuel. 2008. Specificity in Clitic Doubling and in Differential Object Marking, en *Probus*, 20, 33-66.
4. MacDonald, Jonathan. 2016. Spanish aspectual *se* as an indirect object reflexive: The import of atelicity, bare nouns, and leísta PCC repairs, en *Probus*, 28: 73-117.
5. Ormazabal, Javier y Juan Romero. 2013. Object Clitics, Agreement and Dialectal Variation, en *Probus*, 25: 301-344.
6. Pastor, Alberto. 2008. Split analysis of gradable adjectives in Spanish, en *Probus*, 20: 257-299.
7. Pastor, Alberto. 2010. Predicative degree constructions in Spanish, en *Probus*, 22: 27-71.
8. Pöll, Bernhard. 2007. On the licensing of overt subjects in Spanish infinitival clauses, en *Probus*, 19: 93-120.
9. Reglero, Lara. 2007. Wh-in-situ interrogatives in Spanish, en *Probus*, 19: 267-297.
10. Suñer, Margarita. 2006. Left dislocations with and without epithets, en *Probus*, 18: 127-158.
11. Villa-García, Julio. 2012. Recomplementation and locality of movement in Spanish, en *Probus*, 24: 257-314.

1.2. Artículos de sintaxis española de alto impacto incluidos en el estudio de fiabilidad de los datos

1. Aissen, Judith y David. M. Perlmutter. 1976. Clause Reduction in Spanish. *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistics Society*, 2: 1-30.
2. Campos, Héctor. 1986. Indefinite Object Drop, en *Linguistic Inquiry*, 17: 354-359.
3. Contreras, Heles. 1986. Spanish bare NP's and the ECP, en Ivonne Bordelois, Heles Contreras y Karen Zagona (eds.), *Generative Studies in Spanish Syntax*, Dordrecht, Foris Publications: 25-49.
4. Contreras, Heles. Small Clauses in Spanish and English, en *Natural Language and Linguistic Theory*, 5.2 (May 1987): 225-243.
5. Contreras, Heles. 1991. On the Position of Subjects, en *Syntax and Semantics*, 25: 63-79.
6. Fábregas, Antonio. 2007. An Exhaustive Lexicalisation Account of Directional Complements, en *Tromsø Working Papers on Language and Linguistics*, 34: 165-199.
7. Fernández-Soriano, Olga. 1999. Two Types of Impersonal Sentences in Spanish: Locative and Dative Subjects, en *Syntax*, 2: 101-140.
8. Jaeggli, Osvaldo. 1986. Arbitrary Plural Pronominals, en *Natural Language and Linguistic Theory*, 4: 43-76.
9. Leonetti, Manuel. 2004. Specificity and Differential Object Marking in Spanish, en *Catalan Journal of Linguistics*, 3: 75-114.
10. Ordóñez, Francisco. 1998. Post-Verbal Asymmetries in Spanish, en *Natural Language and Linguistic Theory*, 16: 313-346.
11. Ordóñez, Francisco y Esthela Treviño. 1999. Left dislocated subjects and the pro-drop parameter: A case study of Spanish, en *Lingua*, 107: 39-68.
12. Otero, Carlos. 1972. Acceptable Ungrammatical Sentences in Spanish, en *Linguistic Inquiry*, 3: 233-242.
13. Plann, Susan. 1982. Indirect Questions in Spanish, en *Linguistic Inquiry*, 13: 297-312.
14. Rivero, María-Luisa. 1971. Mood and Preposition in Spanish, en *Foundations of Language*, 7: 305-336.
15. Rivero, María-Luisa. 1994. Negation, imperatives and Wackernagel effects, en *Rivista di Linguistica*, 6: 39-66.
16. Rodrigues, Cilene, Andrew Nevins y Luis Vicente. 2009. Cleaving the interactions between sluicing and preposition stranding, en Danièle Tork y W. Leo Wetzels (eds.), *Romance Languages and Linguistic Theory 2006*, John Benjamins, Amsterdam: 175-198.

¹² Los datos de los experimentos así como las evaluaciones de los participantes y el análisis estadístico se encuentran disponibles en la página de ResearchGate del autor.

17. Suñer, Margarita. 1988. The Role of Agreement in Clitic-Doubled Constructions, en *Natural Language and Linguistic Theory*, 6: 391-434.
18. Suñer, Margarita. 1994. V-Movement and the Licensing of Argumental Wh-Phrases in Spanish, en *Natural Language and Linguistic Theory*, 12: 335-372.
19. Suñer, Margarita. 1998. Resumptive Restrictive Relatives: A Crosslinguistic Perspective, en *Language*, 74: 335-364.
20. Torrego, Esther. 1984. On Inversion in Spanish and Some of Its Effects, en *Linguistic Inquiry*, 15: 103-129.
21. Torrego, Esther. 1989. Unergative-Unaccusative Alternations in Spanish, en *MIT Working Papers in Linguistics*, 10: 253-272.
22. Zagona, Karen. 1988. Proper Government of Antecedentless VP in English and Spanish, en *Natural Language and Linguistic Theory*, 6: 95-128.