

ARTICLE

# Predictive processing can override perceptual information: evidence from Spanish object relative clauses

Sara Fernández Santos<sup>1</sup> , Miquel Llompart<sup>1,2</sup>  and Ewa Dąbrowska<sup>1,3</sup>

<sup>1</sup>Department of English and American Studies, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, Germany; <sup>2</sup>Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain and <sup>3</sup>Department of English Language and Applied Linguistics, University of Birmingham, Birmingham, UK

**Corresponding author:** Sara Fernández Santos; Email: [sara.fernandez@fau.de](mailto:sara.fernandez@fau.de)

(Received 17 January 2025; Revised 17 June 2025; Accepted 25 June 2025)

## Abstract

This study examines the role of the timing of obligatory disambiguating information – obligatory cues – and presence/absence of optional morphological markers in resolving temporary syntactic ambiguity in Spanish object relative clauses. Native adult comprehension (Study 1) reveals similar accuracy for clauses with relatively early obligatory cues, regardless of the presence/absence of additional markers, and those with late obligatory cues with additional markers, but reduced accuracy for those with late obligatory cues without additional markers. Given the phonetic resemblance of the late-disambiguated variant with its corresponding subject relative, we conduct two follow-up perceptual identification tasks with the whole relative clause, including the head (Study 2), and relative clause fragments (Study 3). The identification tasks show that, when instructed to attend to the form of the structures, participants perceive acoustic differences but retain a bias towards subject-relative interpretations. Our results suggest that additional markers aid comprehension of non-canonical structures when obligatory cues occur relatively late within the structure and highlight the dominance of predictive processing over perceptual information in such cases of late disambiguation.

**Keywords:** good-enough processing; language comprehension; predictive processing; relative clauses; Spanish

## 1. Introduction

Local ambiguity is common in natural language (Dąbrowska, 2004; Wasow et al., 2005), particularly in dialogue, where comprehenders construct grammatical representations in real time (Ferreira, 2003). Re-analysis models of ambiguity resolution suggest that, in the temporary absence of disambiguating information, the most



plausible syntactic interpretation is initially activated following ‘good-enough’ shallow processing heuristics and reanalysed only when contradictory linguistic evidence arises, e.g. in non-canonical structures (Van Gompel et al., 2005). Predictive processing approaches complement this by proposing that probabilistic information is used to anticipate upcoming linguistic input, streamlining comprehension when correct and hindering it when predictions fail (Ferreira & Lowder, 2016). For instance, in the sentence *the reporter saw her friend was not succeeding*, it is common to initially analyse the second NP as a direct object driven by statistical information of the verb *to see*, and a subsequent re-analysis is triggered by the verb in the subordinate clause (Hillert, 1997). This is illustrated in a continuing syntactic decision task where participants produced longer reaction times to the second NP in sentences where the verb prefers a direct object, as above, compared to sentences where the verb prefers a complement (e.g., *to doubt*; Hillert, 1997). Similarly, research using physiological measures has shown that syntactically globally ambiguous sentences (never disambiguated) are easier to process than locally ambiguous ones (disambiguated), plausibly due to the additional processing cost of re-analysis (Van Gompel et al., 2005).

Grammaticality judgments and response time measures show that the cost of syntactic re-analysis increases as more information is encountered before the disambiguating cue, i.e., the ‘head position effect’ (Ferreira, 2003; Ferreira & Henderson, 1991), highlighting the importance of disambiguation timing in processing non-canonical structures. In this sense, the processing of the above sentence is facilitated by the presence of earlier disambiguating information, i.e., *the reporter saw that her friend was not succeeding*. However, by including supplementary information, a resulting processing facilitation could be expected not only based on disambiguation timing but also on the presence of redundant information. Potential disambiguation timing effects may be better isolated with sentence pairs that differ in timing of disambiguation without additional cues, for example, when the presence/absence of a cue (like the word *that* in the example above) may be found in different positions as a function of varying word order. Thus, the effects of additional morphological markers and those of disambiguation timing in non-canonical sentence processing can be distinguished by examining structures with varying word order and optional markers. A case in point, as elaborated below, are Spanish object relative (OR) clauses. In this study, we investigate the interplay between subject–verb position and optional object marking with the preposition *a* in the comprehension of Spanish ORs. In the next two sections, we first explore the different types of Spanish relative clauses and then introduce the present investigation.

### 1.1 Spanish relatives

Relative clauses (RCs) are subordinate clauses that modify a preceding noun in the main clause. This noun can function as the subject, forming a subject relative (SR) clause, e.g., *the girl that kicks the boy*, or as the object, forming an OR clause, e.g., *the girl that the boy kicks*. In Spanish, restrictive RCs – those that constrain the interpretation of the head noun to the specifications within the RC – are introduced by the relativiser *que* ‘that’. The basic and more frequent word order in Spanish is SVO (Leonetti, 2017), although alternative word orders (with topicalised objects or postposed subjects) are also relatively frequent compared to languages with a fixed word order, such as English. Accordingly, Spanish SRs are predominantly

**Table 1.** Examples of the 6 Spanish relative clause (RC) structures used in Study 1

Structure number	Type of RC	Verb position	Variant (OR)	Example
1	Subject relative	Verb-medial	–	<i>El niño que abraza al abuelo</i> the boy that hug:3s to:the grandpa 'the boy that hugs the grandpa'
2	Subject relative	Verb-final	–	<i>El niño que al abuelo abraza</i> the boy that to:the grandpa hug:3s 'the boy that hugs the grandpa'
3	Object relative	Verb-medial	Plain-variant	<i>El niño que abraza el abuelo</i> the boy that hug:3s the grandpa 'the boy that the grandpa is hugging'
4	Object relative	Verb-final	Plain-variant	<i>El niño que el abuelo abraza</i> the boy that the grandpa hug:3s 'the boy that the grandpa is hugging'
5	Object relative	Verb-medial	A-variant	<i>El niño al que abraza el abuelo</i> the boy to:the that hug:3s the grandpa 'the boy that the grandpa is hugging'
6	Object relative	Verb-final	A-variant	<i>El niño al que el abuelo abraza</i> the boy to:the that the grandpa hug:3s 'the boy that the grandpa is hugging'

verb-medial (see Table 1, structure number 1). Verb-final SRs are possible but much less common (see Table 1, structure number 2). Indeed, Realí (2014) searched a 212,400-word corpus of spoken peninsular Spanish and found 345 instances of transitive non-pronominal SRs. Of these, none was verb-final (Realí, 2014). In elicitation tasks, both children and adults produced only verb-medial SRs (Ezeizabarrena, 2012), and comprehension of verb-final SRs tends to be low, even among adult native speakers (Sánchez Walker & Montrul, 2020). By contrast, object relatives show a more flexible surface structure. Verb-medial ORs are more frequent (around 80% of ORs with full NPs) than verb-final ORs, but both are attested in corpus data (Ezeizabarrena, 2012; Realí, 2014). Given these distributions, surface word order may guide preferences for interpreting RCs as SRs or ORs. Nonetheless, since both word orders are possible for both SRs and ORs, word order can only be regarded as a probabilistic factor and not as a categorical determiner of RC structure.

Spanish is a language with differential object marking (DOM), requiring the preposition *a*, at a minimum, before human, specific direct objects<sup>1</sup>, as exemplified in SRs with transitive verbs. However, in ORs with such direct objects, *a* can be optionally used after the antecedent of the RC (i.e., before the relativiser *que*) followed by the definite article in agreement with the antecedent, e.g., *el niño (al) que el abuelo abraza*. This results in two OR variants, which we will refer to as plain-variant (without *a* + article) and *a*-variant. Table 1 introduces a summary with examples of the RCs discussed. Note that when *a* is followed by the masculine singular definite article *el*, the two are reduced to the form *al* (*a* + *el*). This reduction does not happen with the rest of the definite articles (*a la*, *a los*, *a las*).

Listeners use different kinds of linguistic information to differentiate between OR and SR syntactic interpretations of a RC, and this information can be either deterministic or

<sup>1</sup>Note that the distribution of DOM in Spanish is much more complex than presently discussed, but such discussion lies beyond the scope of this paper (see e.g., Fábregas, 2013; Leonetti, 2004).

probabilistic. Probabilistic information – like preference for interpretations of canonical word order – is not 100% reliable, and here, we will refer to this kind of information as processing heuristics, in line with good-enough processing approaches (Ferreira et al., 2009). In contrast, we will henceforth use the word ‘cue’ to signal deterministic information that unequivocally disambiguates SR and OR interpretations. The ORs in Table 1 present two different types of such cues. Firstly, since subject relatives require DOM in the second NP when it is animate and specific, the absence of this marking can be considered a (deterministic) cue to an OR interpretation (Betancort et al., 2009). Hence, we refer to the absence of DOM as the obligatory cue for OR disambiguation, since this applies to all OR types. However, the absence of a function word is arguably a minimally perceivable cue which could augment its processing cost (Bates & MacWhinney, 1989). As in ORs this cue is encountered immediately before the subject of the RC, it appears earlier in verb-final than in verb-medial ORs. This is most relevant in the plain-variant, given that it does not contain any other disambiguating cue. Secondly, the *a*-variant presents an additional, earlier, and arguably more salient cue: the preposition *a* after the head noun, which marks it as an object. Hence, *a*-variant ORs may be identified as early as the beginning of the RC thanks to the preceding preposition *a*. Note, however, that the absence of this additional marker does not constitute a cue since both ORs and SRs can appear without this marker. Therefore, only the *a*-variant contains an additional cue to an OR interpretation.

In Spanish, SVO is considered the basic and presumably most frequent word order, where the initial element of a structure is typically the subject, which often has the role of agent (Leonetti, 2017). Therefore, Spanish sentences are often processed using the heuristic that the first NP corresponds to the subject (Ferreira et al., 2009), leading to a bias towards interpreting initial NPs as agents unless contradicted. This bias may contribute to the relative difficulty of processing ORs compared to SRs<sup>2</sup>, as ORs require re-interpretation upon disambiguation (Betancort et al., 2009), with processing costs during OR reading arising at the point of deviation from canonical structure (del Río et al., 2012). This relates to what has been reported as canonicity effects in OR-SR processing asymmetries, although there are alternative explanations that lie beyond the scope of this study, for which we refer the interested reader to Lau and Tanaka’s (2021) review. As previously mentioned, the cost of such syntactic re-analysis in non-canonical structures increases with the amount of information preceding disambiguation (Ferreira & Henderson, 1991). This is relevant for the case of Spanish relatives, since the obligatory cue to disambiguate between SRs and ORs – the presence or absence of DOM – may appear earlier (in verb-final RCs) or later (in verb-medial RCs). Hence, while verb position cannot distinguish SRs from ORs on its own, presence or absence of DOM with the subordinate NP does distinguish between the two analyses, and verb position determines the *timing* of this cue. In addition, disambiguation may be affected by the optional earlier cue of morphological marking through *a* (+ article) in *a*-variant ORs. Building on this, in this study, we investigate the role of verb position<sup>3</sup> and

<sup>2</sup>Note that we focus on the Spanish OR-SR processing asymmetry. While this argument may work for other Indo-European languages, it becomes problematic when considering languages with typologically different RC constructions (see Lau and Tanaka, 2021).

<sup>3</sup>Note that we use the term ‘verb position’ here to avoid the potentially problematic term ‘word order’, which can be interpreted to mean that we assume a preference for SV over VS as in previous studies (e.g., Presotto & Torregrossa, 2024), and because ‘verb position’ works for both ORs and SRs. Nonetheless, since the obligatory cue is attached to the NP, it is the position of the NP in relation to the verb that modulates obligatory cue timing.

additional morphological markers in adult native speakers' comprehension of Spanish ORs, explored further in the next subsection.

## 1.2 The present investigation

There are only a few studies that have compared the processing of verb-medial and verb-final versions of the two variants that arise from the optionality of the redundant object marker *a* in Spanish ORs (Llompert et al., 2025; Murujosa et al., 2024; Presotto & Torregrossa, 2024). Presotto and Torregrossa (2024) tested children's comprehension of these four structures and found that verb-final ORs produced better comprehension rates than verb-medial ORs. Within the verb-final position, the plain-variant produced higher accuracies than the *a*-variant, whereas within the verb-medial position, the *a*-variant produced higher accuracies than the plain-variant (Presotto & Torregrossa, 2024). These results could be explained by considering that verb-medial structures are more similar to the canonical SVO structure in Spanish and may thus be more prone to misinterpretation. That is, the only difference between verb-medial SRs (Table 1, structure number 1) and verb-medial plain-variant ORs (Table 1, structure number 3) is the presence of the preposition *a* 'to' in the subject relative, making the verb-medial plain-variant OR potentially ambiguous in speech given the preference for SVO word orders. This difference could be even smaller in sentences with feminine referents, since present-tense third person singular verbs in Spanish predominantly end in *-a*, which is thought to fuse in speech with the following preposition *a* when present, as in the SR in 1a (vs the OR in 1b). It follows from this that the additional marker in the *a*-variant may be more useful in verb-medial ORs because of the potential ambiguity and late disambiguation of the verb-medial plain-variant.

- (1) a) *La niña que abraza a la abuela*  
       'the girl that hugs the grandma'  
       b) *La niña que abraza la abuela*  
       'the girl that the grandma is hugging'

Llompert et al. (2025) compared children and adult comprehension of the two variants with verb-final position. While children consistently performed better with the more frequent plain-variant, as in Presotto and Torregrossa (2024), Llompert et al. (2025) found no significant differences in adult comprehension of the two variants, which suggests that the presence of the additional marker did not facilitate comprehension when the verb appeared in final position. Similarly, Presotto and Torregrossa (2024) reported as supplementary materials additional results from adult comprehension of these structures, which further illustrate a less accurate performance with verb-medial plain-variant ORs (59%) in comparison to verb-medial *a*-variant ORs (97%) and verb-final ORs (plain-variant: 95%; *a*-variant: 96%). This may be because in verb-final ORs the distance between the additional marker (*a* + article) and the obligatory cue (absence of DOM) is relatively short, with only the relativiser *que* separating them. Consequently, when the additional marker is absent – i.e., in verb-final plain-variant ORs – the disambiguation timing is similar to that in the *a*-variant, as opposed to the late disambiguation of verb-medial plain-variant ORs. Presotto and Torregrossa's (2024) adult results, together with their

finding that children's comprehension of verb-medial ORs is worse than verb-final could also be taken to suggest that late disambiguation may hinder comprehension, in line with syntactic re-analysis (Ferreira & Lowder, 2016).

In this article, we examine the relative role of additional morphological marking and verb position, in so far as the latter reflects the timing of the obligatory disambiguating cue, on adult's processing of Spanish ORs and interpret our results in relation to predictive processing accounts and good-enough processing heuristics. In Study 1, we compared adult native speaker's comprehension accuracy and reaction time to the four types of OR presented above using a picture selection task (PST). Subsequently, motivated by the results of Study 1, we tested whether, when there is potential SR/OR ambiguity, listeners could correctly identify the clause type that had been produced after hearing the whole structure (Study 2), or just a fragment of it (Study 3) to focus the task more clearly on phonetic form.

## 2. Study 1

The aim of this study was to assess the comprehension of the four types of Spanish OR sentences described above. Using a PST, we tested adult native speakers' comprehension accuracy and reaction time to these sentences. Firstly, we expected *a*-variant ORs to produce higher accuracy and shorter reaction times overall than plain-variant ORs, given their additional morphological marker. Moreover, since this additional cue appears at the onset of the RC, it follows that initial biases towards an SR interpretation will be avoided in these structures, i.e., even if the first noun (antecedent) is initially regarded as a subject, this interpretation will be rejected before the relativiser is encountered. In a similar manner, we expected the two possible verb positions in the *a*-variant to exhibit similar reaction times and comprehension rates, since the early appearance of the morphological marker should render the timing of the subsequent obligatory cue irrelevant. Additionally, although not the focus of our study, we expected SRs to produce high accuracy across the board (as in Llompart et al., 2025), although verb-medial SRs could produce higher accuracy rates and faster reaction times than verb-final SRs given the latter's marginal occurrence in spoken Spanish (Reali, 2014)<sup>4</sup>.

Going back to ORs, if morphological marking facilitates comprehension beyond verb position, we would expect both *a*-variant ORs to produce higher accuracy rates than verb-final plain-variant ORs, since the three sentence types show similar disambiguation timing<sup>5</sup> but *a*-variant ORs have an additional morphological marker. Conversely, if both *a*-variant ORs produce similar accuracy rates to verb-final plain-variant ORs, this would highlight the role of verb position over additional morphological marking, suggesting that, when an earlier disambiguation is provided,

<sup>4</sup>An anonymous reviewer pointed out that one might similarly expect differences in accuracy or reaction time of ORs with verb-medial vs verb-final positions on the basis of frequency differences. However, previous data has found no such difference in adults, suggesting that frequency effects for ORs may reach asymptote by adulthood (Llompart et al., 2025).

<sup>5</sup>Note that in the case of *a*-variant ORs, disambiguation appears slightly earlier than in verb-final plain-variant ORs (just before the relativiser vs just after the relativiser). However, previous data shows that this difference does not affect comprehension (Llompart et al., 2025; Presotto & Torregrossa, 2024). Moreover, the point still holds that disambiguation timing is relatively late in verb-medial plain-variant ORs in comparison to the other three OR structures under investigation.

additional marking does not facilitate processing further. Finally, and most importantly, we expected verb-medial plain-variant ORs to exhibit lower comprehension accuracy and longer reaction times than the rest of OR types given their high resemblance to canonical SRs (see 1 above) and the late appearance of the disambiguating cue, which may slow down and hinder re-interpretation (Ferreira, 2003; Ferreira & Henderson, 1991).

## 2.1 Methods Study 1

### 2.1.1 Participants

A total of 42 adults were recruited for this study ( $M_{age} = 32.9$ ,  $SD_{age} = 9.5$ ; female = 21) in exchange for financial compensation. Participants had spent a mean of 19 years in formal education ( $SD = 4.6$ ), with a doctorate being the highest educational achievement ( $N = 1$ ), followed by master's degree ( $N = 13$ ), bachelor's degree ( $N = 16$ ), vocational training ( $N = 5$ ), advanced secondary education ( $N = 4$ ), obligatory secondary education ( $N = 1$ ) and primary education ( $N = 2$ ). Our sample was recruited using the online recruitment platform Prolific ([www.prolific.co](http://www.prolific.co)) and participation was limited to adult native speakers of Spanish born in Spain. Participants were randomly assigned to one of two conditions regarding verb position so that 20 of them were exposed to verb-final RCs ( $M_{age} = 32.4$ ,  $SD_{age} = 7.38$ ; female = 10) and the remaining 22 to verb-medial RCs ( $M_{age} = 33.9$ ,  $SD_{age} = 11$ ; female = 11). A background questionnaire was administered to gauge participants' demographics, educational levels and self-reported reading habits. Informed consent was obtained from the participants, and the study was conducted in accordance with the Declaration of Helsinki.

### 2.1.2 Materials and procedure

Participants took part in the study online through Gorilla Experiment Builder ([www.gorilla.sc](http://www.gorilla.sc)). Here, they completed a background questionnaire, a picture selection task and two unrelated tasks. The stimuli and tasks used in this and all other experiments in this paper can be found at <https://app.gorilla.sc/openmaterials/929742>. The total duration of the experimental session was approximately 20 min.

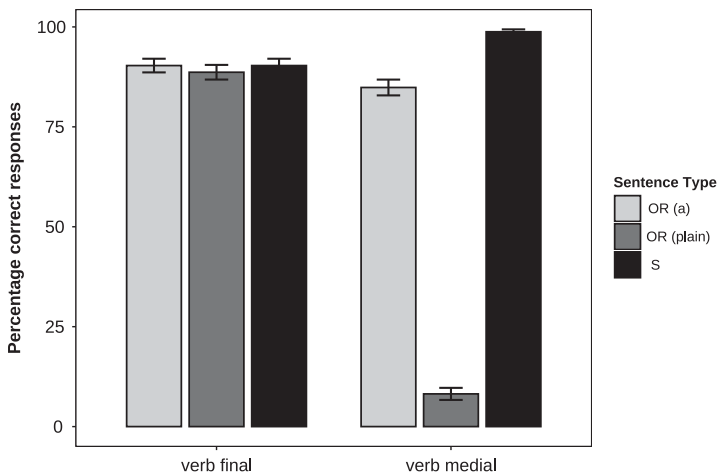
The picture selection task (PST) was modelled after that used by Llopart et al. (2025). The materials included 16 pairs of pictures (32 pictures in total) and 96 audio recordings of Spanish sentences, recorded by a native Spanish speaker in a neutral voice tone and a comfortable speaking rate. The picture pairs showed two human characters (e.g., a girl and a grandma) and differed in who carried out a transitive action (e.g., a girl drawing a grandma versus a grandma drawing a girl). Across the stimuli, there were four pairs of characters: (i) a boy and a grandpa, (ii) a girl and a grandma, (iii) a teenage boy and a man and (iv) a teenage girl and a woman. Each participant pair was shown performing 4 actions, amounting to a total of 16 actions. For each action (e.g., hugging), 6 sentences were recorded: 2 SRs (1 verb-medial and 1 verb-final), 2 plain-variant ORs (1 verb-medial and 1 verb-final) and 2 *a*-variant ORs (1 verb-medial and 1 verb-final). See Table 1 for examples of all RC types included as stimuli.

The task consisted of 48 trials for each participant, with 16 SRs, 16 plain-variant ORs and 16 *a*-variant ORs. Note that for all analyses we removed one item from all conditions for consistency with Study 2 and Study 3. This item used the verb *persigue*, the only verb ending in *-e*, which affects how SRs and ORs differ perceptually in

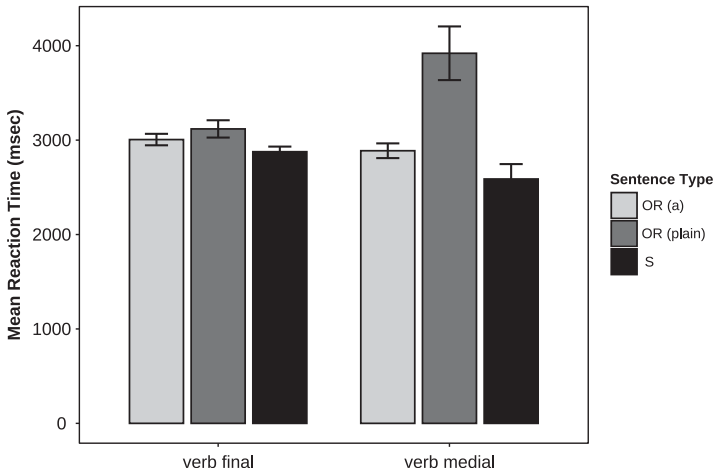
comparison to the other items. The verb position of the RCs that participants heard was determined by the condition to which they were assigned: verb-final or verb-medial. We decided to make verb position a between-subjects variable in order to avoid presenting participants with too many items involving the same characters engaged in the same action. An additional advantage of keeping verb position constant within the task is that participants were strongly discouraged by the design to base their decisions merely on verb position, e.g. judging all verb-medial sentences as SRs, while we could still compare the two verb positions between subjects. Sentence presentation was pseudorandomised: sentences with the same action were at least two trials apart, and no more than two consecutive sentences shared the same structure or participants. After reading instructions and completing two practice trials with simple transitive sentences, participants proceeded to the experimental trials. Each trial began with a 500 ms fixation screen, followed by a 1 s picture preview to eliminate visual processing delays and then the automatic playback of a sentence. Participants responded by clicking on one of the pictures.

## 2.2 Results Study 1

All the datasets and the code scripts used for all the analyses in this article are available at <https://osf.io/4x6pz/> (Open Science Framework). Figure 1 illustrates the percentage of correct responses for each OR variant and SRs with each verb position. In the first place, a visual inspection of this figure suggests that comprehension of ORs and verb-final SRs (mean % correct = 90.33 ( $SD = 29.60$ ), count correct = 271/300), even in a native adult population, is not at ceiling level, i.e., these sentences still pose some level of difficulty (vs verb-medial SRs: mean % correct = 98.79 ( $SD = 10.96$ ), count correct = 326/330). Secondly, within verb-final ORs, accuracy seems to be similar for the *a*-variant (mean % correct = 90.33 ( $SD = 29.60$ ), count correct = 271/300) and the plain-variant (mean % correct = 88.67 ( $SD = 31.75$ ), count correct = 266/300). Within verb-medial ORs, the *a*-variant produced similar accuracies to those from verb-final



**Figure 1.** Percentage of correct responses by condition (verb-final and verb-medial) and sentence type (*a*-variant OR, plain-variant OR and SR).



**Figure 2.** Reaction times (ms) by condition (verb-final and verb-medial) and sentence type (*a*-variant OR, plain-variant OR and SR).

ORs (mean % correct = 84.84 ( $SD = 35.90$ ), count correct = 280/330), while accuracy of responses to verb-medial plain-variant ORs was remarkably lower (mean % correct = 8.18 ( $SD = 27.45$ ), count correct = 27/330). Reaction times for correct trials, illustrated in Figure 2, seem to mirror this pattern, whereby reaction times for SRs are shorter for verb-medial ( $M = 2588$  ms,  $SD = 2851$ ) than verb-final ( $M = 2877$  ms,  $SD = 892$ ) and reaction times for ORs are similar for the verb-final *a*-variant ( $M = 3006$  ms;  $SD = 1000$ ), the verb-final plain-variant ( $M = 3119$  ms;  $SD = 1496$ ) and the verb-medial *a*-variant ( $M = 2888$  ms;  $SD = 1305$ ), whereas the verb-medial plain-variant produced longer reaction times ( $M = 3920$  ms;  $SD = 1477$ ).

We ran mixed-effects regression models on trial-by-trial accuracy and RT data using the lme4 package (version 1.1-23, Bates et al., 2015) in R (R Core Team, 2017; version 4.2.2). In all models, Variant and Verb Position were effect coded with the *a*-variant and verb-final position as 0.5 and the plain-variant and verb-medial position as  $-0.5$ . Johnson's (2014) conditional and marginal pseudo- $R^2$  was calculated for logistic regressions (accuracy models) using the r.squaredGLMM function from the MuMIn package (version 1.47.5; Barton & Barton, 2015). Nakagawa's conditional and marginal  $R^2$  (Nakagawa & Schielzeth, 2013) were calculated for linear models (reaction time models) using the tab\_model function from the sjPlot package (version 2.8.11; Lüdtke, 2021). Odds ratios and their 95% confidence intervals for the fixed effects of all models were computed also using the tab\_model function.

Before analysing reaction times, the data were adjusted as follows: Firstly, only trials in which participants provided the correct response (76.24% of trials) were included. For each participant, RTs that were above or below 3 absolute deviations from their individual median were eliminated as outliers (Leys et al., 2013), which removed a total of 54 correct trials (3.75%). After a visual inspection of the distribution of reaction times, the data were transformed with a logarithm function and 6 additional datapoints with RTs over 6000 ms (0.43% of remaining data) were deleted after being identified as outliers.

We analysed SRs and ORs separately in order to keep OR variant as a predictor in the OR models, since SRs do not exhibit this variability. Firstly, given the apparent difference between verb-medial and verb-final SR comprehension, we explored potential differences in accuracy and reaction time with only SR data. For the accuracy analysis, we ran a generalised linear mixed-effects model (glmer) with a logit linking function using the optimiser ‘bobyqa’ in the glmer control options. The model included Response as the binary dependent variable (1 = correct, 0 = incorrect) and Verb Position as the predictor of interest. The random-effects structure of the model included random intercepts by Participant and by Item. Random slopes for Verb Position over Participant were not included because each participant viewed only one of the two verb positions. Random slopes for Verb Position over Item were not included because item numbers varied for each RC type (as in Table 1)<sup>6</sup>. The results for this model showed a significant effect of Verb Position in SR response accuracy ( $b = 2.26$ ,  $SE = 0.90$ ,  $t = 2.52$ ,  $p < .05$ , marginal pseudo- $R^2 = .16$ , conditional pseudo- $R^2 = .59$ ).

To analyse reaction times to SRs, a linear mixed-effects regression model (lmer) was run with a logit linking function using the optimiser ‘bobyqa’ in the control options. The model included log-transformed reaction time (RT) as the numeric dependent variable and Verb Position as the predictor of interest. The random-effects structure of the model included random intercepts by Participant and by Item. The results for this model showed a significant effect of Verb Position in SR response time ( $b = -0.18$ ,  $SE = 0.06$ ,  $t = -2.78$ ,  $p < .01$ , marginal  $R^2 = .08$ , conditional  $R^2 = .45$ )<sup>7</sup>.

Turning to ORs, for the accuracy analysis, we ran a glmer with Response as the binary dependent variable (1 = correct, 0 = incorrect) and OR Variant, Verb Position and their 2-way interaction as predictors. The random-effects structure of the model included random intercepts by Participant and by Item and random slopes for OR Variant over Participant. The results for this model are summarised in Table 2,

**Table 2.** Results of the glmer for accuracy to object relative clauses on the PST

Variable	$\beta$	SE	z	p	Odds ratios (CI)
(Intercept)	1.47	0.29	5.05	<.001	4.34 (2.45–7.67)
OR variant	3.35	0.46	7.25	<.001	28.48 (0.01–0.09)
Verb position	3.27	0.58	5.67	<.001	26.24 (0.01–0.12)
OR variant: Verb position	−5.36	0.89	−6.01	<.001	0.01 (0.00–0.03)
Random effects	Variance		SD		
Item	0.56		0.75		
Participant	1.98		1.41		
Participant   OR variant	2.43		1.56		
Marginal pseudo- $R^2$	0.54				
Conditional pseudo- $R^2$	0.76				

<sup>6</sup>Since items are variants of the same visual target, which limits the random-effects structure, we ran additional analyses assigning an item number to each distinct verb and including additional slopes and obtained the same results.

<sup>7</sup>Note that reaction time data were collected via mouse-clicks, which are known to introduce greater variability compared to other methods (like button presses) because of the larger motor component involved. While this adds noise to the reaction time data, any sizeable effects observed under these conditions can be taken as particularly robust, since they emerged despite this increased measurement variability.

**Table 3.** Results of the glmer for reaction time to object relative clauses on the PST

Variable	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>CI</i>
(Intercept)	7.95	0.03	227.67	<.001	7.88–8.02
OR variant	−0.10	0.04	−2.18	<.05	−0.18–0.01
Verb position	−0.06	0.07	−0.85	0.40	−0.20–0.08
OR variant: verb position	0.21	0.09	2.42	<.05	0.04–0.39
Random effects	Variance		SD		
Item	0.00		0.07		
Participant	0.03		0.18		
Participant   OR variant	0.01		0.11		
Marginal <i>R</i> <sup>2</sup>	0.01				
Conditional <i>R</i> <sup>2</sup>	0.39				

revealing a significant effect of OR Variant, a significant effect of Verb Position and a significant interaction between OR Variant and Verb Position.

In order to follow-up the interaction, the data were split by Verb Position and two additional glmer models were run, one with the verb-final ORs and the other with the verb-medial ORs. Both models included Response as the binary dependent variable (1 = correct, 0 = incorrect) and OR Variant as the predictor. The random-effects structure of both follow-up models included random intercepts by Participant and by Item and Random slopes for OR Variant over Participant. The results indicated that there was no significant difference in the accuracy of responses to the two variants in verb-final position ( $b = 0.09$ ,  $SE = 0.47$ ,  $z = 0.21$ ,  $p = 0.83$ ), whereas in the verb-medial position, accuracy of responses differed significantly between the two OR variants ( $b = 6.51$ ,  $SE = 0.89$ ,  $z = 7.30$ ,  $p < .001$ ). This indicates that the interaction above was driven by responses to the verb-medial plain-variant being significantly less accurate than responses to the other ORs.

For the reaction time analysis, we ran an lmer model with log-transformed RT as the numeric dependent variable and OR Variant, Verb Position and their 2-way interaction as predictors. The random-effects structure of the model included random intercepts by Participant and by Item and random slopes for OR Variant over Participant. The results, summarised in Table 3, reveal a significant effect of OR Variant and a significant interaction between OR Variant and Verb Position.

The interaction between OR Variant and Verb Position was followed up in a similar way as the accuracy analysis. After splitting the data by Verb Position, we ran two additional lmer models, one with the verb-final ORs and the other with the verb-medial ORs. Both models included log-transformed RT as the numeric dependent variable and OR Variant as the predictor. The random-effects structure of both models included random intercepts by Participant and by Item and random slopes for OR Variant over Participant. Once again, the results showed that there was no significant difference in the reaction times to the two variants in the verb-final ORs ( $b = 0.01$ ,  $SE = 0.04$ ,  $t = 0.26$ ,  $p = 0.80$ ). On the other hand, in the verb-medial ORs, reaction times for the plain-variant were significantly longer than for the *a*-variant ( $b = -0.18$ ,  $SE = 0.06$ ,  $t = -2.88$ ,  $p < .01$ ).

### 2.3 Discussion Study 1

Study 1 revealed asymmetries in SR and OR comprehension. Verb-medial SRs achieved ceiling accuracy, consistent with Llompart et al. (2025), whereas verb-final

SRs showed reduced accuracy and slower response times. This fits with the patterns reported in Sánchez Walker and Montrul (2020) and is likely due to the rarity of these SRs in spoken Spanish (Reali, 2014). Regarding ORs, participants performed equally well with both verb-final OR variants, as the adults in Llompart et al. (2025), but unlike children (Llompart et al., 2025; Presotto & Torregrossa, 2024). This difference between children and adult performance may stem from the need to develop a specialised function-form mapping for *a* + article in ORs that differs from common uses of DOM. That is, while *a* + article is predominantly used to signal that the following element functions as a direct object, in ORs it signals that the *preceding* element functions as an object within the relative clause. Hence, children, lacking sufficient experience with *a*-variant ORs, may interpret the optional *a* as marking the following noun as the object, which would be consistent with their experience of DOM rules but result in an incorrect analysis of *a*-variant ORs (e.g., *\*al que el abuelo = al abuelo*). This argument would be in line with the findings of Presotto and Torregrossa (2024) that children with better scores in common usage of DOM performed worse with the *a*-variant ORs (vs plain-variant) in the verb-final position, i.e., when the NP immediately followed the marker and relativiser.

In verb-medial ORs, the *a*-variant produced similar accuracy and response times to both types of verb-final ORs. In contrast, verb-medial plain-variant ORs were misinterpreted as SRs over 90% of the time and resulted in significantly longer response times. While the accuracy pattern with the different types of ORs mirrors that of previous studies (Murujosa et al., 2024; Presotto & Torregrossa, 2024), the present study produced overall lower accuracy, especially for verb-medial plain-variant ORs. In the case of Presotto and Torregrossa (2024), since the RCs they used are directly comparable to ours, we speculate that differences in RC articulation, e.g. hyperarticulation, could have enhanced comprehension in comparison to our study, although we lack access to their spoken stimuli to confirm this. In the case of Murujosa et al. (2024), their use of RCs in the past tense is likely what results in a higher accuracy compared to our data, since verbs in the past tense have a different vowel ending and therefore eliminate the potential ambiguity of verb-medial plain-variant ORs in speech (see Table 4). Murujosa et al. (2024) also used stimuli with additional prosodic material following the RCs, which could influence the acceptability of verb-medial ORs (Gutiérrez-Bravo, 2005). Given this variant's similarity to an SR, this suggests that perceptual cues distinguishing verb-medial plain-variant ORs from SRs in the present study may have been often missed. However, note that if the differentiating cues were not perceived at all, one would expect performance to be at chance for both verb-medial plain-variant ORs and verb-medial SRs. Instead, we observed a strong bias towards an SR interpretation for both types of items (>90% correct in verb-medial SRs and <10% correct in plain-variant verb-medial ORs). This bias may stem from the 'head position effect' in shallow processing (Ferreira & Henderson, 1991), that is, since this cue appears later in the sentence than in other variants, re-evaluating the initial SR judgement may be more difficult.

Accordingly, our findings seem to support the predicted relevance of disambiguation timing in comprehension of ORs, in line with theories of syntactic re-analysis. What is unclear, though, is whether this bias stems from the similarity between the two sentence types (even if they are discriminable) or from their complete indiscriminability. The first scenario would suggest that ORs' high phonetic similarity with SRs could lead to predictive processing overriding perception. That is, since at the point of the last NP an object is expected, one is more likely to interpret what they

**Table 4.** Examples of the eight types of RC presented in the auditory identification task

Type of RC	Tense	Gender of NPs	Example
Subject relative	Present	Feminine	( <i>La chica que</i> ) <i>peina a la mujer</i> the girl that comb:3s to:the woman 'the girl that combs the woman'
Object relative	Present	Feminine	( <i>La chica que</i> ) <i>peina la mujer</i> the girl that comb:3s the woman 'the girl that the woman combs'
Subject relative	Past	Feminine	( <i>La chica que</i> ) <i>peinó a la mujer</i> the girl that comb:3s-PAST to:the woman 'the girl that combed the woman'
Object relative	Past	Feminine	( <i>La chica que</i> ) <i>peinó la mujer</i> the girl that comb:3s-PAST the woman 'the girl that the woman combed'
Subject relative	Present	Masculine	( <i>El niño que</i> ) <i>abraz a al abuelo</i> the boy that hug:3s to:the grandpa 'the boy that hugs the grandpa'
Object relative	Present	Masculine	( <i>El niño que</i> ) <i>abraz a el abuelo</i> the boy that hug:3s the grandpa 'the boy that the grandpa is hugging'
Subject relative	Past	Masculine	( <i>El niño que</i> ) <i>abrazó al abuelo</i> the boy that hug:3s-PAST to:the grandpa 'the boy that hugged the grandpa'
Object relative	Past	Masculine	( <i>El niño que</i> ) <i>abrazó el abuelo</i> the boy that hug:3s-PAST the grandpa 'the boy that the grandpa hugged'

Note: The parts of the RC in parentheses were not included in Study 3 (see below).

hear as *al/a la* rather than *el/la* even when no initial [a] is heard. In contrast, if participants are not able to discriminate between the two sets of stimuli, this would suggest that predictive processing is used in the face of complete ambiguity. To tease these two possibilities apart, we first conducted an acoustic analysis of the stimuli, followed by an identification task (Study 2).

### 3. Acoustic analysis

We ran an acoustic analysis to quantify the differences between the SRs and the verb-medial plain-variant ORs used in Study 1, since these two structures are highly similar. This similarity is even greater in sentences with feminine nouns (see examples in 1 in the introduction). Indeed, it has been argued that in natural language use there is no real phonological difference between speech strings such as *abraz a la* and *abraz a la* (Colina, 2009), which is the only distinction between verb-medial plain-variant ORs and SRs with feminine referents. Note, however, that debates regarding the duration of reduced sounds such as the conjoined vowels in *abraz a* persist, with recent research reporting differences in duration of native speakers' productions between cases previously assumed to be acoustically equal (e.g., *la salas* vs *las alas*; Andreu Rascón, 2024).

Since acoustic differences vary for RCs with masculine and feminine referents, we consider gender as one of the relevant factors in the acoustic analysis as well as in the analyses of identification data in Studies 2 and 3. For the same reason, we decided to revisit data from Study 1 in order to investigate whether comprehension performance with verb-medial plain-variant ORs was affected by gender. Comprehension of ORs

with masculine nouns produced an accuracy of 8.44% ( $SD = 27.9$ , count correct = 13/154) and comprehension of ORs with feminine nouns produced an accuracy of 7.95% ( $SD = 27.1$ , count correct = 14/176). A glmer run on trial-by-trial data from the verb-medial plain-variant ORs in Study 1 showed that the effect of gender was not significant ( $b = 0.27$ ,  $SE = 0.93$ ,  $z = 0.29$ ,  $p = 0.77$ ).

The acoustic analysis was carried out on the segments of the audio recordings that differentiate SRs from verb-medial plain-variant ORs. In the case of sentences with feminine nouns, the difference is between the vocalic portion corresponding to the verb-final *a* in ORs and the combination of the verb-final *a* and the preposition *a* in SRs. In the case of sentences with masculine nouns, the difference is between the combination of the vocalic portion corresponding to the verb-final *a* and the initial *e* of the following determiner in ORs and the combination of the verb-final *a* and the preposition *a* in SRs.

The onset and offset of each vowel or vowel combination were identified within the sentence by using the information provided by both the waveform and the spectrogram as generated by Praat (Boersma & Weenink, 2018); temporal landmarks were manually marked on a time-aligned text file, and landmarks were always placed at upward zero-crossings. Vowel/vowel combination durations (in milliseconds), and F1 and F2 values (in Hertz) at the first, second and third quartiles of their duration were extracted using a custom Praat script. Then, F1 and F2 values were averaged across the time points and we subtracted the F1 to the F2 (i.e.,  $F2 - F1$ ) to obtain just one value per token. Following previous work with similar vowel pairs (e.g., Llompart & Reinisch, 2017),  $F2 - F1$  values were used to assess potential differences in vowel quality because /e/ is expected to have a lower F1 and a higher F2 than /a/. Therefore, if present, these differences should manifest themselves in our stimuli in that the  $F2 - F1$  values for *a + e* combinations in ORs with masculine referents should be higher than those for the *a + a* sequences in their SR counterparts. In terms of duration, the question was whether the vocalic portion in SRs (e.g., *ayuda al/ayuda a la*) is longer than that in ORs (e.g., *ayuda el/ayuda la*). The results showed that in sentences with feminine nouns there was indeed a difference in duration (SRs:  $M = 96$  ms,  $SD = 13$ ; ORs:  $M = 58$  ms,  $SD = 9$ ). In the sentences with masculine nouns, there were differences in both duration (SRs:  $M = 93$  ms,  $SD = 13$ ; ORs:  $M = 69$  ms,  $SD = 10$ ) and vowel quality (SRs:  $M = 1101$  Hz,  $SD = 259$ ; ORs:  $M = 1357$  Hz;  $SD = 239$ ).

Therefore, the acoustic analysis indeed revealed some differences between the vocalic portions of verb-medial SRs and verb-medial plain-variant ORs. Among the clauses with feminine referents, we found a vowel duration difference of 38 ms, which exceeds the reported thresholds for a just noticeable difference in vowel duration ( $\geq 25$  ms; Klatt & Cooper, 1975). Among the clauses with masculine referents, we found a just noticeable difference in vowel duration (24 ms) and a clear difference in vowel quality pointing towards a raising of the later part of the vocalic portion from [a] to [e] as expected. Overall, the acoustic analysis suggests that the difference between these two types of clause in Study 1 should be, in principle, perceivable. This could be taken to suggest that duration differences in the case of feminine referents were enhanced by the recording conditions (laboratory speech (hyperarticulation) and sentence reading) and the stimuli's purpose (clarity enhancement for comprehension), assuming that natural speech tends to show no such distinction (Colina, 2009). Alternatively, it could be the case that there is indeed a difference in vowel duration in natural speech, even if phonological accounts tend not to predict one

(see Andreu Rascón, 2024), although teasing these two possibilities apart would go beyond the purpose of this study.

## 4. Study 2

The aim of this study was to determine whether the differences between two types of auditive stimuli used in Study 1, verb-medial plain-variant OR (e.g., *el niño que abraza el abuelo*) and SR (e.g., *el niño que abraza al abuelo*) sentence recordings, are perceivable by Spanish native speakers. The two structures are highly similar, which can lead to potential ambiguity in speech. Nonetheless, our analysis (Section 3) indicates that the relevant acoustic differences between them in our stimuli are large enough to be perceivable. To assess this, we examined whether adult native speakers of Spanish could differentiate between the two spoken forms when instructed to do so. With this purpose, an auditory identification task was designed that eliminated the need to interpret the meaning of the sentences and focused very clearly on the formal distinction between the two types of RC.

### 4.1 Methods Study 2

#### 4.1.1 Participants

A total of 20 adults were recruited for this study ( $M_{age} = 28.4$ ;  $SD_{age} = 7.97$ ; female = 10) via Prolific in exchange for financial compensation. Participation was limited to adult native speakers of Spanish born in Spain and excluded all participants who had taken part in Study 1. Participants spent a mean of 19.5 years in formal education ( $SD = 3.46$ ), with doctorate being the highest educational achievement ( $N = 1$ ), followed by master's degree ( $N = 7$ ), bachelor's degree ( $N = 6$ ), vocational training ( $N = 1$ ) and advanced secondary education ( $N = 5$ ). The study was conducted in accordance with the Declaration of Helsinki and participants provided informed consent before taking part.

#### 4.1.2 Materials and procedure

Participants took part in the study online through Gorilla Experiment Builder ([www.gorilla.sc](http://www.gorilla.sc)), where they completed the same background questionnaire as the one administered in Study 1 and an auditory identification task (AIT). The total duration of the experiment session was approximately 15 minutes.

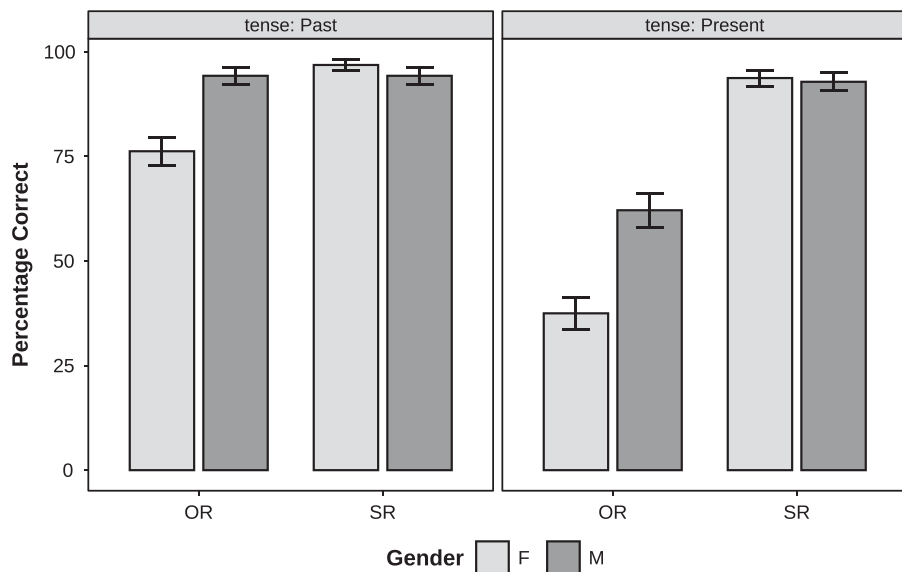
In the AIT, participants read two Spanish sentences, a verb-medial SR and a verb-medial plain-variant OR and then presented with an audio recording; their task was to decide which of the sentences was being spoken. The materials encompassed 60 sentences and 60 corresponding audio recordings. The sentences included 15 SRs and 15 verb-medial plain-variant ORs from Study 1 (see Table 1, structure numbers 1 and 3, respectively) and the 30 past-tense versions of these sentences (e.g., *la chica que peinó (a) la mujer*). In the past-tense versions, the difference between the verb's final vowel (–ó) and the preposition *a* (when present) makes the preposition easier to perceive than in present-tense versions, hence reducing the potential ambiguity regarding its presence or absence. Accordingly, the difference between verb-medial plain-variant ORs and SRs should be easier to perceive. Each OR was paired with their SR counterpart, differing only in the presence or absence of *a* before the determiner on the last NP (e.g., *la chica que peina (a) la mujer*). Each pair of sentences

was presented twice, once with the SR recording and once with the OR recording. No fillers were included to ensure participants focused on the stimuli's form, allowing assessment of their distinguishability when the focus was on auditory identification rather than comprehension. Examples of every type of RC presented in this task are presented in Table 4.

The task consisted of 60 trials, and sentences were presented in a pseudorandom order with the same constraints as the PST in Study 1 and the additional constraint that there could be no more than two consecutive sentences of the same tense. Each trial consisted of a 500 ms fixation screen, followed by a 3 s written sentence preview, during which participants were instructed to read both options. After this, a play button appeared, and the sentence recording played when participants clicked on it. When the recording finished playing, participants could select one of the written sentences with a click. The written sentences remained on the screen until a response was recorded.

#### 4.2 Results Study 2

Figure 3 depicts response accuracy for both tenses (past and present), separated by the gender of nouns (masculine and feminine) and structure type (verb-medial plain-variant ORs and SRs). Participants showed high accuracy for all SRs: present-tense clauses with masculine nouns (mean % correct = 92.86 ( $SD = 25.85$ ), count correct = 130/140) and feminine nouns (mean % correct = 93.75 ( $SD = 24.28$ ), count correct = 150/160), past-tense clauses with masculine nouns (mean % correct = 94.29 ( $SD = 23.29$ ), count correct = 132/140) and feminine nouns (mean % correct = 96.87 ( $SD = 17.45$ ), count correct = 155/160). In contrast, OR identification was more



**Figure 3.** Percentage of correct responses by tense (past and present), gender of the referents in the sentence (feminine and masculine) and sentence type (OR = verb-medial plain-variant OR and SR = verb-medial SR).

challenging, particularly for present-tense clauses with feminine nouns (mean % correct = 37.50 ( $SD = 48.56$ ), count correct = 60/160), while those with masculine nouns showed higher accuracy (mean % correct = 62.14 ( $SD = 48.68$ ), count correct = 87/140). Surprisingly, past-tense ORs with feminine nouns also showed reduced accuracy (mean % correct = 76.25 ( $SD = 42.69$ ), count correct = 122/160), though performance for masculine-referent ORs in past tense was comparable to SRs (mean % correct = 94.29 ( $SD = 23.29$ ), count correct = 132/140). Delta prime scores for structures in present tense were significantly greater than zero for both clauses with feminine ( $p < .001$ ) and with masculine referents ( $p < .001$ ), indicating that participants responded differently to the ORs than to the SRs.

We ran a glmer on trial-by-trial accuracy data from all sentence types. Gender, Tense and Type were effect coded with ORs, present-tense clauses and those with feminine referents as  $-0.5$  and SRs, past-tense clauses and those with masculine referents as  $0.5$ . The model included Response as the binary dependent variable (1 = correct, 0 = incorrect) and Gender, Tense and Type as predictors as well as their 3-way interaction. The random-effects structure of the model included random intercepts by Participant and by Item and random slopes for Type over Participant. Random slopes for Gender over Participant and Tense over Participant were not included because neither of them improved the fit of the model (Gender over Participant:  $\chi^2(4) = 2.62$ ,  $p = 0.62$ ; Tense over Participant:  $\chi^2(4) = 3.18$ ,  $p = 0.53$ ). The results for this model are summarised in Table 5, revealing a significant effect of Type and Tense on response accuracy and significant interactions between Type and Tense and between Type and Gender.

The interactions were followed up by splitting the data by Type and running two additional glmer models, one with the OR data and the other with the SR data. Both models included Response as the binary dependent variable (1 = correct, 0 = incorrect) and Tense, Gender and their interaction as predictors. The random-effects structure of both models included random intercepts by Participant and by Item. The results of the follow-up models show that accuracy of identification of ORs was significantly worse for present-tense than past-tense versions ( $b = 2.57$ ,  $SE = 0.50$ ,  $z = 5.14$ ,  $p < .001$ ) and significantly worse for sentences with feminine nouns than those with masculine nouns ( $b = 1.47$ ,  $SE = 0.97$ ,  $z = 3.01$ ,  $p < .01$ ), but there was no significant

**Table 5.** Results of the glmer for accuracy to relative clauses on the AIT

Variable	$\beta$	$SE$	$z$	$p$	Odds ratios (CI)
(Intercept)	2.39	0.24	9.88	<.001	10.94 (6.81–17.59)
Type	2.23	0.47	4.70	<.001	9.28 (3.66–23.49)
Tense	1.58	0.37	4.24	<.001	4.86 (2.34–10.10)
Gender	0.53	0.37	1.45	0.15	1.70 (0.83–3.48)
Type: Tense	−1.90	0.73	−2.60	<.01	0.15 (0.04–0.62)
Type: Gender	−1.86	0.73	−2.55	<.05	0.15 (0.04–0.65)
Tense: Gender	−0.07	0.73	−0.10	0.92	0.93 (0.22–3.90)
Type:Tense: Gender	−0.83	1.46	−0.57	0.57	0.44 (0.02–7.63)
<i>Random effects</i>					
	<i>Variance</i>				
Item		1.03	1.02		
Participant		0.26	0.51		
Participant   Type		1.27	1.13		
Marginal pseudo- $R^2$		0.33			
Conditional pseudo- $R^2$		0.55			

interaction between the tense and gender of the referents. In contrast, there were no significant differences in accuracy of identification of SRs between past- and present-tense versions ( $b = 0.58$ ,  $SE = 0.51$ ,  $z = 1.12$ ,  $p = 0.26$ ) and sentences with masculine or feminine referents ( $b = -0.40$ ,  $SE = 0.50$ ,  $z = -0.78$ ,  $p = 0.43$ ).

### 4.3 Discussion Study 2

The results of Study 2 showed that the identification of SRs was near ceiling (mirroring SR comprehension accuracy in Study 1) and not affected by tense. In contrast, accuracy with past-tense ORs was higher than with present-tense ORs, and the former is comparable to the accuracies reported in Murujosa et al.'s (2024) comprehension study of these past-tense RCs. This seems to confirm the potential ambiguity of present-tense ORs in speech. In the remainder of this section, we focus on the results from present-tense stimuli.

Identification of SRs did not differ by gender of the referents in the clause, whereas OR identification showed higher accuracy for RCs with masculine (vs feminine) referents. The lower accuracy in identification of ORs with feminine nouns as opposed to those with masculine nouns suggests that the bias towards an SR interpretation is stronger when there is only a vowel duration difference (vs vowel duration and quality). Moreover, as such difference may not reliably occur in natural speech given vowel deletion (Colina, 2009), it is likely that participants do not rely much on this information when identifying the clauses.

Importantly, performance for ORs with masculine referents was clearly above chance, and for ORs with feminine referents, performances also appear to be considerably higher than what was observed for comprehension in Study 1. This suggests that participants were to an extent able to identify the differences between verb-medial plain-variant ORs and SRs when explicitly instructed to pay attention to the form of the speech stimuli. Nonetheless, clause identification still showed a strong bias towards an SR interpretation for these structures. Since performance in identification of ORs was still relatively low, we considered it possible that participants still relied on the meaning of the sentences to some extent and thus favoured the more canonical pattern (SR). This is why, in Study 3, we ran the AIT again with only a smaller portion of the relative clauses in order to reduce the context and therefore prevent listeners' use of sentence meaning to identify the correct form to a larger extent than in Study 2.

## 5. Study 3

The aim of this study was to test the identification of the critical fragments of spoken verb-medial plain-variant ORs and SRs used in Studies 1 and 2 (e.g., *abrazo el/al abuelo*) when the rest of the clause is not presented. With this we aimed to increase listeners' focus on form even further and limit their access to potentially biased interpretations derived from the analyses of the larger clausal structures in the two previous studies.

## 5.1 Methods Study 3

### 5.1.1 Participants

A total of 20 adults were recruited for this study ( $M_{age} = 33$ ;  $SD_{age} = 10.9$ ; female = 9) via Prolific in exchange for financial compensation. Participation was limited to adult native speakers of Spanish born in Spain and excluding participants who had taken part in Study 1 and Study 2. Participants spent a mean of 17.7 years in formal education ( $SD = 4.77$ ), with doctorate being the highest educational achievement ( $N = 1$ ), followed by master's degree ( $N = 4$ ), bachelor's degree ( $N = 7$ ), vocational training ( $N = 4$ ), advanced secondary education ( $N = 2$ ), primary education ( $N = 1$ ) and no formal education ( $N = 1$ ). Informed consent was obtained from participants, and the study was conducted in accordance with the Declaration of Helsinki.

### 5.1.2 Materials and procedure

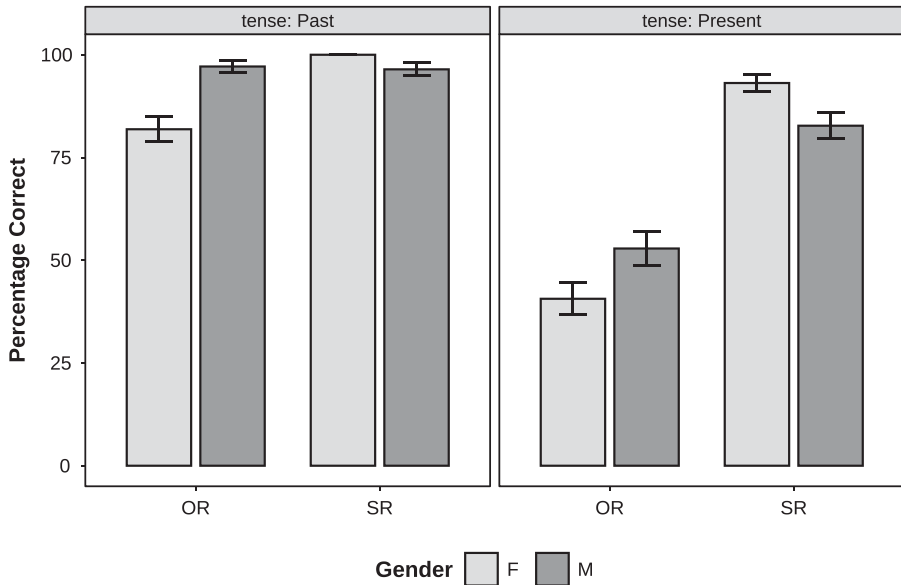
Participants took part in the study online through Gorilla Experiment Builder ([www.gorilla.sc](http://www.gorilla.sc)), where they completed the same background questionnaire as the one administered in Study 1 and Study 2 and the AIT from Study 2 with only a fragment of the clauses previously used. The total duration of the experiment session was approximately 13 minutes.

The identification task was the same as that described in Study 2, but including only clause fragments from the verb onwards. For example, where in Study 2 participants heard *el niño que abraza el abuelo*, here we only presented the fragment *abraza el abuelo* (see Table 4). This trimming process was done in both the written sentences and the audio recordings. Due to the reduction in the written sentences, the sentence preview time in each trial was changed to 1.5 s. Everything else was identical to the AIT in Study 2.

## 5.2 Results Study 3

Figure 4 illustrates response accuracy for both tenses (past and present), separated by gender of the nouns in the structure (masculine and feminine), for the two structure types (fragments of verb-medial plain-variant ORs and SRs), revealing very similar results to those in Study 2. Participants were generally accurate in identifying stimuli trimmed from the original SRs. Nonetheless, note that identification of present-tense fragments with masculine nouns was slightly lower (mean % correct = 82.73 ( $SD = 37.93$ ), count correct = 115/139) than the other SR fragments: present-tense fragments with feminine nouns (mean % correct = 93.12 ( $SD = 25.38$ ), count correct = 149/160), past-tense fragments with masculine nouns (mean % correct = 96.45 ( $SD = 18.56$ ), count correct = 136/141) and past-tense fragments with feminine nouns (mean % correct = 100.00 ( $SD = 0.00$ ), count correct = 160/160). Identification of OR fragments followed the same pattern as ORs in Study 2, with lower identification rates especially for present-tense fragments with feminine nouns (mean % correct = 40.62 ( $SD = 49.27$ ), count correct = 65/160), followed by present-tense fragments with masculine nouns (mean % correct = 52.86 ( $SD = 50.10$ ), count correct = 74/140), past-tense fragments with feminine nouns (mean % correct = 81.87 ( $SD = 38.64$ ), count correct = 131/160) and past-tense fragments with masculine nouns (mean % correct = 97.14 ( $SD = 16.72$ ), count correct = 136/140).

Once again, we ran a glmer on trial-by-trial accuracy data from all structure types. Gender, Tense and Type were effect coded in the same way as Study 2. The model



**Figure 4.** Percentage of correct responses by tense (past and present), gender of the nouns in the structure (feminine and masculine) and structure type (fragments of verb-medial plain-variant ORs and SRs).

included Response as the binary dependent variable (1 = correct, 0 = incorrect) and Gender, Tense and Type as predictors as well as their 3-way interaction. The random-effects structure of the model included random intercepts by Participant and by Item and random slopes for Type over Participant. Random slopes for Tense over Participant and for Gender over Participant were not included because neither improved the fit of the model (Tense over Participant:  $\chi^2(4) = 2.67, p = 0.61$ ; Gender over Participant:  $\chi^2(4) = 5.48, p = 0.24$ ). The 100% accuracy on fragments of feminine SRs in the past tense resulted in a perfect correlation of the fixed effects, which rendered the model uninterpretable; therefore, we changed one random data point from this group from correct to incorrect, which fixed the problem. The results for this model are summarised in Table 6, revealing a significant effect of Type and Tense on response accuracy and a significant interaction between Type and Gender.

### 5.3 Discussion Study 3

The findings of this study were very similar to those of Study 2: above-chance performance for masculine-referent clause fragments and seemingly improved performance for feminine-referent clause fragments compared to Study 1. This suggests that the bias towards an (S)VO interpretation remains even when the head of the relative clause and the relativiser are removed. This could be explained by predictive processing, in this case by transitional probabilities impacting speech segmentation (Saffran et al., 1996). Specifically, the verb may have triggered predictions about the subsequent element for two distinct reasons. First, the processing system tends to anticipate simple and frequent word orders (Ferreira, 2003), in this case favouring the canonical (S)VO (vs (O)VS). Second, the verb forms used in this study

**Table 6.** Results of the glmer for accuracy to relative clauses on the AIT with trimmed clauses

Variable	$\beta$	SE	z	p	Odds ratios (CI)
(Intercept)	2.56	0.25	10.06	<.001	12.97 (7.87–21.38)
Type	2.11	0.52	4.10	<.001	8.27 (3.01–22.74)
Tense	2.78	0.44	6.27	<.001	16.08 (6.75–38.32)
Gender	–0.06	0.43	–0.13	0.89	0.94 (0.41–2.20)
Type: Tense	–1.18	0.87	–1.36	0.17	0.31 (0.06–1.68)
Type: Gender	–2.74	0.86	–3.18	<.01	1.31 (0.24–7.11)
Tense: Gender	0.27	0.86	0.32	0.75	1.31 (0.24–7.11)
Type: Tense: Gender	–2.19	1.72	–1.27	0.20	0.11 (0.00–3.29)
Random effects	Variance		SD		
Item	0.98		0.99		
Participant	0.15		0.39		
Participant   Type	1.29		1.13		
Marginal pseudo- $R^2$	0.44				
Conditional pseudo- $R^2$	0.61				

(present-tense, 3rd person singular) also coincide with imperative forms. As a result, participants may have interpreted these structures as imperatives upon hearing the verb, consequently predicting the next element to be an object, as transitive verbs in imperative constructions typically require one. Either explanation would suggest that the bias originates at least partly from syntactic expectations driven from sheer frequency, thus supporting the ‘good enough’ processing account.

### 6. General discussion

This article looked at OR and SR processing in Spanish to assess the roles of verb position and additional morphological markers in ORs. With this purpose, Study 1 tested comprehension of these structures with a picture selection task. Study 2 and Study 3 tested the identification of SR and verb-medial plain-variant OR clauses and OR clause fragments, respectively, to gauge participants’ ability to discriminate between them. Study 1 showed that verb-medial plain-variant ORs were consistently misinterpreted, with a strong bias in favour of an SR interpretation. In addition, we found no facilitation effects of the additional marker in the comprehension of verb-final ORs. Study 2 showed that, when participants are explicitly asked to identify the form of verb-medial plain-variant ORs and SRs, the acoustic differences are not fully neutralised (overall identification is above chance), suggesting that the acoustic cues differentiating verb-medial plain-variant ORs and SRs are perceivable, but interpretations are still biased towards the latter. Study 3 showed that this bias is still present even after further reducing the interpretability of the semantic content of the spoken stimuli by omitting the initial part of the relative clause. As discussed in Section 5.3 right above, this suggests that the observed bias could be driven by predictions generated by transitional probabilities.

Taken together, the studies suggest that the comprehension of the relatively late-disambiguated verb-medial plain-variant ORs is driven by shallow processing. In addition, their acoustic similarity to canonical SRs could lead to predictive processing overriding perceptual information, that is, at the point of the last NP, an object is expected, making it more likely to interpret the acoustic signal as *al/a la* rather than

*el/la*<sup>8</sup>. This argument is indirectly supported by evidence from vowel identification shifts towards grammatical gender agreement depending on the preceding noun (Martin et al., 2017). Specifically, adjective-final vowels in Spanish were more likely to be identified as /o/ when preceded by a grammatically masculine noun and as /a/ when preceded by a grammatically feminine noun. These shifts were most pronounced when the vowel's acoustic properties were highly ambiguous (midway between /a/ and /o/), suggesting that listeners use predictive processing to resolve uncertain signals and that reliance on prediction increases with uncertainty. In this study, this would explain why participants performed above chance when given both form options in advance (Study 2) but consistently misinterpreted them when no information about form was provided (Study 1), which increases uncertainty. This also fits in with our finding that the bias towards an SR analysis in form identification was stronger for stimuli that were more similar to the predicted SR, i.e., OR clauses with feminine referents versus those with masculine referents.

The current results are in line with previous findings supporting the initial shallow processing and subsequent re-analysis of Spanish OR clauses (del Rio et al., 2012). This paper expands on previous research, which found no effect of the additional morphological marker in adult comprehension of verb-final OR clauses (Llompart et al., 2025), by finding that this marker facilitates comprehension within verb-medial ORs, which are otherwise highly prone to misinterpretation. Hence, additional morphological markers seem to be specifically important when the obligatory disambiguating cue appears relatively late within the structure.

## 7. Conclusion

In short, this article suggests that additional morphological markers may not facilitate comprehension when disambiguation occurs early, but they do when the obligatory cue appears relatively late in the structure. We interpret this as an indication that the cost of re-interpretation in non-canonical structures is dependent on the timing of the disambiguating cue, in line with the 'head position effect' (Ferreira & Henderson, 1991). Since late-disambiguated structures entail more information that is congruent with the erroneous initial interpretation before a conflicting cue is encountered, such information can be used to confirm the misinterpretation and hence hinder re-interpretation. Moreover, our findings suggest that in such cases of strengthened expectations, these predictions can even override perceptual information. Further research could explore whether this phenomenon extends to other locally ambiguous structures and languages. Another interesting question that arises from this research is whether different morphosyntactic systems interact differently with disambiguation timing, for example, whether more salient markers facilitate the processing of relatively early-disambiguated structures and not only late-disambiguated ones.

**Data availability statement.** The dataset analysed in this article and the code to reproduce the analyses reported are available at <https://osf.io/4x6pz/> (Open Science Framework).

<sup>8</sup> An alternative explanation could be that listeners accept sentences without DOM in the NP within the RC because the marker is only required for a subset of direct objects and hence interpret ORs as SRs when the subject is post-verbal and there is no additional object marking, e.g., *el abuelo que abraza el niño*. However, comprehension data from Murujosa (2024) using past-tense RCs, which removes the potential auditory ambiguity, reveals higher accuracies for this condition, which suggests that listeners are not likely to accept lack of DOM across the board.

**Competing interests.** The authors declare none.

## References

- Andreu Rascón, I. (2024). Segmenting speech: The role of Resyllabification in Spanish phonology. *Language*, 9(11), 346.
- Barton, K., & Barton, M. K. (2015). *Package 'mumin'*. R package version 1.47.5.
- Bates, D., Mächler, M., Bolker, B., Walker, S., Christensen, R. H., Singmann, H., & Dai, B. (2015). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Cambridge University Press.
- Betancort, M., Carreiras, M., & Sturt, P. (2009). Short article: The processing of subject and object relative clauses in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology*, 62(10), 1915–1929. <https://doi.org/10.1080/17470210902866672>
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program].
- Colina, S. (2009). *Spanish phonology: A syllabic perspective*. Georgetown University Press.
- Dąbrowska, E. (2004). *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Edinburgh University Press.
- del Río, D., López-Higes, R., & Martín-Aragoneses, M. (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(11), 2108–2128. <https://doi.org/10.1080/17470218.2012.674951>
- Ezeizabarrena, M. J. (2012). Children do not substitute object relatives with subject relatives in every romance language: The case of Spanish. *Revue Roumaine de Linguistique*, 57, 161–181.
- Fábregas, A. (2013). Differential object marking in Spanish: State of the art. *Borealis: An International Journal of Hispanic Linguistics*, 2(2), 1–80. <https://doi.org/10.7557/1.2.2.2603>
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164–203. [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7)
- Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good enough language processing: A satisficing approach. In N. Taatgen, H. Rijn, J. Nerbonne & L. Schomaker (Eds.), *Proceedings of the 31st Annual conference of the Cognitive Science Society* (vol. 1, pp. 413–418). Cognitive Science Society.
- Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of 'garden-path sentences. *Journal of Memory and Language*, 30(6), 725–745.
- Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In B. H. Ross (Ed.), *Psychology of learning and motivation* (vol. 65, pp. 217–247). Academic Press.
- Gutierrez-Bravo, R. (2005). Subject inversion in Spanish relative clauses. A Case of prosody-induced word order variation without narrow focus. In T. Geerts & H. Jacobs (Eds.), *Romance Languages and Linguistic Theory 2003*. Amsterdam: John Benjamins.
- Hillert, D. (1997). Language in time: Lexical and structural ambiguity resolution. In M. Stamenov (Ed.), *Language structure, discourse and the access to consciousness* (pp. 77–112). John Benjamins Publishing.
- Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's  $R^2_{GLMM}$  to random slopes models. *Methods in Ecology and Evolution*, 5, 44–946.
- Klatt, D. H., & Cooper, W. E. (1975). Perception of segment duration in sentence contexts. In A. Cohen & S. G. Neeboom (Eds.), *Structure and process in speech perception: Proceedings of the symposium on dynamic aspects of speech perception* (pp. 69–89). Springer.
- Lau, E., & Tanaka, N. (2021). The subject advantage in relative clauses: A review. *Glossa: A Journal of General Linguistics*, 6(1). <https://doi.org/10.5334/gjgl.1343>
- Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics*, 3, 75–114. <https://raco.cat/index.php/CatalanJournal/article/view/309007>
- Leonetti, M. (2017). 24. Basic constituent orders. In A. Dufter & E. Stark (Eds.), *Manual of romance morphosyntax and syntax* (pp. 887–932). De Gruyter. [10.1515/9783110377088-024](https://doi.org/10.1515/9783110377088-024).

- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766.
- Llompart, M., Fernández Santos, S., & Dąbrowska, E. (2025). Comprehension of object relatives in Spanish: the role of frequency and transparency in acquisition and adult grammar. *Cognitive Linguistics*, 36(1), 31–57. <https://doi.org/10.1515/cog-2024-0016>
- Llompart, M., & Reinisch, E. (2017). Articulatory information helps encode lexical contrasts in a second language. *Journal of Experimental Psychology: Human Perception and Performance*, 43(5), 1040.
- Lüdtke, D. (2021). *sjPlot: Data visualization for statistics in social science*. R package version 2.8.11. R Foundation for Statistical Computing.
- Martin, A. E., Monahan, P. J., & Samuel, A. G. (2017). Prediction of agreement and phonetic overlap shape sublexical identification. *Language and speech*, 60(3), 356–376. <https://doi.org/10.1177/0023830916650714>
- Murujosa, M., Shalóm, E. D. & Sevilla Y. (2024, September 5–7). *Word order and case marking during the comprehension of object relatives in Spanish* [Conference poster presentation]. The 30th Architectures and Mechanisms for Language Processing conference. The University of Edinburgh, Scotland.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142.
- Presotto, G., & Torregrossa, J. (2024). Intervention and amelioration effects in the acquisition of Spanish object relative clauses: The role of word order and DOM. *Glossa: A Journal of General Linguistics*, 9(1). <https://doi.org/10.16995/glossa.11254>
- R core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Real, F. (2014). Frequency affects object relative clause processing: Some evidence in Favor of usage-based accounts: Frequency affects object relative clause processing. *Language Learning*, 64(3), 685–714. <https://doi.org/10.1111/lang.12066>
- Saffran, J. R., Newport, E. L., and Aslin, R. N.. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.
- Sánchez Walker, N., & Montrul, S. (2020). Language experience affects comprehension of Spanish passive clauses: A study of heritage speakers and second language learners. *Language*, 6(1), 2. <https://doi.org/10.3390/languages6010002>
- Van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, 52(2), 284–307. <https://doi.org/10.1016/j.jml.2004.11.003>
- Wasow, T., Perfors, A., & Beaver, D. (2005). The puzzle of ambiguity. In O. C. Orhan & P. Sells (Eds.), *Morphology web grammar: Essays in memory of Steven G. Lapointe* (pp. 265–282). The University of Chicago Press.

---

**Cite this article:** Fernández Santos, S., Llompart, M., & Dąbrowska, E. (2025). Predictive processing can override perceptual information: evidence from Spanish object relative clauses, *Language and Cognition*, 17, e59, 1–24. <https://doi.org/10.1017/langcog.2025.10015>