

Article

Variable Agreement Constructions in Spanish: Between Perception Modalities and Conceptual Foregrounding

Renata Enghels ^{1,*}  and Mariia Baltais ² ¹ Department of Linguistics, Ghent University, 9000 Ghent, Belgium² Department of Experimental Psychology, Ghent University, 9000 Ghent, Belgium; mariia.baltais@ugent.be

* Correspondence: renata.enghels@ugent.be

Abstract

This article investigates how cognitive and grammatical mechanisms shape variable singular–plural agreement in Spanish perception–verb constructions, a domain where speakers alternate between agreement with the postverbal NP₂ and agreement with the infinitival complement. Building on usage-based and cognitive linguistics approaches, this study examines whether factors related to perceptual modality and conceptual salience underlie these alternations. A corpus analysis of pronominal infinitive constructions with *ver* and *oír* reveals divergent patterns across modalities, with visual perception favoring plural agreement and auditory perception favoring singular agreement. To evaluate whether these tendencies reflect deeper linguistic preferences, an acceptability-rating task systematically manipulated modality, agreement, and animacy. The results show no overall interaction between modality and agreement, but they identify a robust effect of animacy: sentences with human referents received higher ratings than those with inanimate referents. Moreover, animacy modulated the influence of modality and agreement in opposite directions, suggesting that speakers’ evaluations are sensitive to the ontological nature of the perceived stimulus. Together, the findings show that agreement variation reflects flexible conceptual construal and that corpus and experimental evidence offer complementary insights into the interface between morphosyntax, perception and salience in Spanish.

Keywords: Spanish verbal agreement; perception verbs; infinitive complement; animacy; salience; corpus linguistics; acceptability ratings; constructional alternations; Cognitive Grammar

1. Introduction

The central goal of this paper is to show how both grammatical and cognitive mechanisms involved in language processing can account for variable singular–plural agreement phenomena that, from a normative point of view, are labeled as ‘non-standard’. More specifically, we explore how the conceptualization of perceived referents leaves visible traces in the morphosyntax of Spanish. In concrete terms, in several Spanish constructions, there is a possible mismatch between the verbal agreement actually produced by speakers and the agreement form prescribed by the grammatical norm. The widespread nature of this phenomenon can be illustrated by a series of agreement patterns (1–3) that diverge, to varying degrees, from the prescriptive norm. Although some of these are traditionally classified as deviations, they often display systematic behavior that warrants closer examination¹.



Academic Editor: John Lipski

Received: 19 December 2025

Revised: 11 February 2026

Accepted: 12 February 2026

Published: 27 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

- | | | | | | | |
|-----|---|---------------------------------|--------------------------------|------------------------------------|---|--|
| (1) | <i>Habían</i>
exist-3PL.PST
'There were many people at the party.' | <i>muchas</i>
many-F.PL | <i>personas</i>
person-F.PL | <i>en</i>
in | <i>la</i>
ART.F.SG | <i>fiesta.</i>
party |
| (2) | <i>Me</i>
1SG.DAT
'My legs hurt.' | <i>dolieron</i>
hurt-3PL.PST | <i>las</i>
ART.F.PL | <i>piernas.</i>
leg-F.PL | | |
| (3) | <i>Un</i>
INDEF.M.SG
'A group of students walked along the avenue.' | <i>grupo</i>
group | <i>de</i>
of | <i>estudiantes</i>
student-M.PL | <i>recorrió(-ieron)</i>
walk-3SG(PL).PST | <i>la</i>
ART.F.SG
<i>avenida.</i>
avenue |

In example (1), the impersonal verb *haber* ‘there is’, which is considered invariable in standard Spanish, appears with plural morphology under the influence of the plural postverbal noun phrase (NP) *muchas personas* ‘many persons’. This pattern can be readily found in spontaneous speech and in corpus data, although the grammatical norm would classify it as non-standard. So, the plural agreement morphology reflects the way speakers conceptualize the referent set, rather than the formal grammatical features of the syntactic head (e.g., Claes, 2015). In (2), the verb *dolieron* ‘hurt’ agrees with the plural NP *las piernas* ‘the feet’, which functions as the grammatical subject, while the dative *me* ‘to me’ designates the experiencer (e.g., Melis, 2022). Although normative grammar fully accepts this construction, it illustrates a different kind of mismatch: the experiencer functions as the conceptual subject of the event, while the grammatical agreement is driven by the postposed, yet semantically salient, body-part referent. Finally, the case in (3) displays variable singular or plural agreement patterns depending on how speakers choose to construe the referent of the clause including a collective NP. Although both variants are acceptable in contemporary Spanish, the alternation between singular and plural agreement reflects a conceptual choice: whether the speaker treats the group as a unified collective entity (*recorrió*) or focuses on the plurality of its members (*recorrieron*) (e.g., Martínez, 1999). This leads to a comparable alternating set of constructions that constitute the empirical focus of this study, namely those in which perception verbs are followed by an infinitival clause (4–5).

- | | | | | | | | | | |
|-----|---|------------------------|----------------------------|---------------------------|-----------------------------------|---------------------|-------------------------------------|------------------------|--------------------------------|
| (4) | <i>A</i>
to | <i>lo</i>
ART.N.SG | <i>lejos</i>
far | <i>se</i>
REFL | <i>oye(n)</i>
hear-3SG(PL).PRS | <i>volar</i>
fly | <i>palomas.</i>
pigeon-F.PL | | |
| | 'In the distance one hears pigeons flying.' | | | | | | | | |
| | | | | | | | | | |
| (5) | <i>En</i>
in | <i>las</i>
ART.F.PL | <i>aguas</i>
water-F.PL | <i>del</i>
of-ART.M.SG | <i>mar</i>
sea | <i>se</i>
REFL | <i>ve(n)</i>
see-
3SG(PL).PRS | <i>flotar</i>
float | <i>cañas.</i>
reed-
F.PL |
| | 'In the waters of the sea one sees reeds floating.' | | | | | | | | |

In these constructions, the finite perception verb may appear either in singular or in plural. Formally, this variation reflects two possible agreement targets: the speaker may establish agreement either with the postposed NP (in concrete terms, *palomas* ‘doves’ and *cañas* ‘reeds’) or with the infinitival complement representing the perceived event as a whole (that is, *volar palomas* ‘palomas fly’ and *flotar cañas* ‘reeds float’). What remains to be determined is whether this alternation correlates with differences in the speaker’s perceptual construal of the witnessed scene. In the literature, examples such as these are often described as instances of ‘agreement *ad sensum*’, that is, agreement motivated by meaning rather than by formal morphosyntactic features (e.g., Corbett, 2003). In such cases, the speaker’s cognitive representation of the referent, such as its animacy or perceptual salience, overrides the syntactic configuration that would otherwise determine agreement.

The objectives of this study are as follows: The first is to investigate which cognitive mechanisms underlie these cases of variable agreement in Spanish and more specifically to examine the role of ‘salience’, understood as the relative cognitive prominence that speakers assign to elements within a perceived/conceptualized scene (Langacker, 1987,

2008; Talmy, 2000; Croft & Cruse, 2004). Salience has long been recognized as a fundamental factor in linguistic construal, governing whether speakers foreground an event as a holistic occurrence or highlight a particular participant as the Figure/Trajector against a more backgrounded Ground/Landmark. From this perspective, agreement alternations in perception–verb constructions may reflect speakers’ shifting attentional focus: at times privileging the perceived event as a single unit, and at others foregrounding the entity that initiates the perceived action. The second objective is to analyze how specific properties of the perception act or the referent (such as the overall perceptual modality and animacy of the stimulus) correlate with these grammatical options. More broadly, the aim of this study is to show why such variable agreement phenomena are relevant for linguistic theory: they offer an empirical window onto the interface between grammar and cognition and reveal how speakers negotiate between formal and conceptual constraints when producing language (see also Acuña-Fariña et al., 2014, p. 120).

Addressing these questions requires a methodological approach that combines usage-based and experimental evidence. Corpus data are indispensable for identifying actual distributional tendencies, but they cannot on their own reveal the cognitive representations that guide speakers’ grammatical choices (e.g., Schütze, 2016; Granvik et al., 2025, among many others). For this reason, corpus-based observations are complemented here with an acceptability-rating task, which allows us to probe whether observed distributional patterns align with speakers’ internal evaluations under controlled conditions. Similar corpus–experiment triangulation has proved fruitful in related domains such as argument structure alternations, as earlier research has shown (e.g., Bresnan et al., 2007).

The remainder of this paper is structured as follows: Section 2 examines standard and non-standard agreement patterns in Spanish, situating the phenomenon within broader theoretical discussions on the syntax–semantics interface and the role of meaning in morphosyntactic variation. Section 3 introduces the infinitival construction with perception verbs, which will serve as the testing ground for the present study. Sections 4 and 5 present, respectively, the results of a corpus analysis and an experimental study designed to capture both the distributional tendencies and the acceptability of singular and plural agreement in these constructions. Finally, Section 6 brings together the findings and discusses their implications for our understanding of agreement in Spanish as a phenomenon at the interface between grammatical constraints, perceptual modalities, and conceptual foregrounding.

2. Variable Verbal Agreement in Spanish

First, it is necessary to establish what the grammatical norm of Spanish defines as standard agreement. In general terms, (verbal) agreement is a formal mechanism through which elements within a clause share morphosyntactic information. As a consequence, agreement ensures internal harmony among syntactically related constituents (Bello, 1964; Avendaño, 2007). Classical examples include, at the level of the NP, *las casas blancas* (‘the white houses’, gender and number agreement between a determiner, the noun and adjective) and, at the level of the VP, *los chicos bailan* (‘the kids dance’, number and person agreement between the subject and verb). Among these, subject–verb agreement has been described as the most widespread agreement mechanism cross-linguistically, playing a central role in encoding grammatical relations (Mallinson & Blake, 1981; Vigliocco et al., 1996; Mare & Pato, 2018; Sánchez et al., 2014; Avendaño, 2007; Haskell & MacDonald, 2003). Indeed, from a typological perspective, three main formal mechanisms are generally recognized for identifying the syntactic function of a constituent: (i) word order, (ii) case marking, and (iii) agreement. Of these, agreement is considered the most consistent and reliable indicator across languages (Givón, 2001).

Still, agreement is not universal, nor does it operate in the same way in all languages (Corbett, 2006). Some languages lack morphological agreement entirely and rely on other devices to encode syntactic relations. Japanese, for instance, does not inflect the verb for person or number, and grammatical relations are marked instead by particles alongside relatively fixed word order. Swedish, though inflectional, also displays minimal verbal agreement (e.g., *Jag talar* ‘I speak’, *du talar* ‘you speak’, *han talar* ‘s/he speaks’ all show an invariant verbal form). Spanish, by contrast, has a rich agreement system that allows for considerable syntactic flexibility. Compared to English, which exhibits highly restricted subject–verb agreement (e.g., *s/he runs* vs. *you run* vs. *they run*), Spanish encodes number and person systematically, making it possible to omit the subject (e.g., *corre* ‘s/he runs’, *corres* ‘you run’, *corren* ‘they run’) or vary word order without loss of grammatical information.

The canonical rule in Spanish prescribes that “los rasgos de número y persona de los verbos conjugados constituyen el reflejo gramatical de los de su sujeto” (‘the number and person features of conjugated verbs constitute the grammatical reflection of those of their subject’) (RAE & ASALE, 2009, p. 2559). In practice, however, this idealized pattern is often complicated by semantic, discourse-related, and cognitive factors that introduce variation.

One key observation is that verbal agreement does not always align with the canonical subject, even when a clear subject is present in the structure. Numerous cases can be found in which the verb agrees instead with another nearby NP with the syntactic function of a direct object (DO) (6a)², an adverbial adjunct (6b), a secondary predicate (6c) or even the subject of a subordinate clause (6d). In these examples the verb does not agree with the grammatical subject but with another nearby constituent instead.

- (6a) *Se han ido citando a varios guardia civiles.*
REFL AUX.PRS.3PL go-PTCP cite-PTCP to several guard civil-PL
‘Several members of the Civil Guard have been summoned.’
(Gómez Torrego, 1992, p. 24)
- (6b) *Para llegar a Madrid sólo se tardan diez minutos.*
for arrive-INF to Madrid only REFL take-3PL.PRS ten minute-PL
‘It only takes ten minutes to get to Madrid.’
- (6c) *Miningas es como se llaman allí a las muchachas.*
Miningas be-3SG.PRS how REFL call-3PL.PRS there to ART.F.PL girl-F.PL
‘Miningas is what girls are called there.’
(Goytisolo, J., *Fiestas*, p. 149)
- (6d) *No se saben cuántos alumnos habrán asistido a clase.*
not REFL know-3PL.PRS how.many student-M.PL AUX.FUT.3PL attend-PTCP to class
‘It is not known how many students will have attended class.’

In the literature, mismatches such as these have been accounted for in terms of agreement with an ‘attractor’. The attractor is an intervening constituent situated between the grammatical subject and the verb, which may ‘attract’ agreement features (Vigliocco et al., 1996; Wagers et al., 2009; Sánchez et al., 2014). This phenomenon typically arises when the distance between the subject and the verb increases or the subject is not explicitly mentioned, and the syntactic or semantic salience of a nearby constituent exerts a stronger influence. Similar attraction effects have been widely attested in psycholinguistic studies of agreement production (cf. Bock & Miller, 1991)³. A second explanatory factor underlying certain cases of agreement mismatch, and interfering with the grammatical one, involves ‘notional’ or ‘conceptual agreement’. As early as Gili Gaya (1980), reference was made to *concordancia mentada* ‘intended agreement’, that is, agreement that responds to the speaker’s expressive or communicative intentions rather than to purely grammatical constraints. This type of agreement operates as a principle of textual cohesion, where morphological concord serves as a mechanism that is a “guía de interpretación y especificación de los distintos elementos del discurso” (‘a guide of the interpretation and specification of the various

elements of discourse’) (Avendaño, 2007, p. 207). From this perspective, agreement fulfills not only a grammatical but also a discursive and cognitive function, as an interpretive strategy through which speakers activate particular ‘construals’ in their interlocutors. The principle can also be understood under a broader principle: the so-called agreement *ad sensum*, or agreement according to the overall meaning of the clause (Corbett, 2003; Haskell & MacDonald, 2003; Sánchez et al., 2013; Mare & Pato, 2018; Enghels, 2019).

This approach aligns closely with cognitively and functionally oriented hypotheses about language. From the perspective of Cognitive Grammar (e.g., Langacker, 1987, 2008), grammatical patterns emerge from the way speakers conceptualize a scene. In this sense, agreement tends to be established with the element that is most ‘salient’ or mentally accessible at the moment of production. This principle is also connected to notions of ‘perspective’ and ‘evidentiality’: speakers may align agreement with the entity they perceive first, or about which they have the most immediate information (cf. *infra* Section 5.2).

Naturally, the idea that language and perception are closely interrelated has been expressed before. As early as the 1970s, Miller and Johnson-Laird observed that “the impression that perception and language are closely related may stem from a feeling that people use language primarily to talk about the world they perceive” (Miller & Johnson-Laird, 1976, p. 119). A few years later, Jackendoff (1983, p. 3) reformulated this as a guiding question: “What does the grammatical structure of natural language reveal about the nature of perception and cognition?”

From this perspective, it is particularly interesting to interpret the variable agreement phenomena within the framework of Construction Grammar (e.g., Goldberg, 2006, amongst others). This approach conceives grammatical constructions as complex symbolic units, that is, conventionalized pairings of form and meaning. According to the Principle of No Synonymy, two constructions that differ formally must also differ in their meaning or pragmatic function (Bolinger, 1968, p. 127). Consequently, this view that a difference in syntactic form spells a difference in meaning implies that the choice between singular or plural agreement in contexts of variability is not neutral either. Rather, it may produce subtle effects on focus or perspective of the utterance. In this sense, variable agreement in perception–verb constructions offers an ideal testing ground to explore how morphosyntactic alternation reflects differences in conceptualization, as further explained in the next section.

3. Testing Ground: The Infinitive Construction with Perception Verbs

In most languages, the semantic nature of perception verbs is determined by two fundamental parameters: the degree of agentivity of the perceiver and the sensory modality involved (Viberg, 1984; Sweetser, 1990; Evans & Wilkins, 2000, among others). In this study, we will focus on four types of verbs that combine these two dimensions: *ver* ‘to see’ and *mirar* ‘to look at’ for visual perception, and *oír* ‘to hear’ and *escuchar* ‘to listen to’ for auditory perception. Among the various possible complementation patterns that perception verbs allow, one of the most distinctive is the infinitival construction. This construction represents the type of complement that encodes an act of direct perception of an event by an experiencer subject. The perceived event, expressed by a subordinate infinitive complement, includes a second participant, here referred to as the subordinate participant or NP₂, who is responsible for the action expressed by the infinitive. Example (7) illustrates this type of construction.

(7)	perceiver NP₁	perception verb PV	perceived event Inf	NP₂
	Ø	<i>Veía</i>	<i>agrandarse</i>	<i>las bocas de aquellas dos mujeres.</i>
	I	see-1SG.IPFV	enlarge-INF.REFL	ART.F.PL mouth-F.PL of those two woman-F.PL

‘I saw the mouths of those two women grow.’

(Muñoz Molina, A., *El jinete polaco*, p. 353 [CREA])

Since the infinitive is a non-finite form, it cannot display verbal agreement with a subject. To observe agreement marking within this domain, we must therefore turn to the pronominalized infinitival construction, where the perception verb itself becomes the locus of agreement (8a–b). In this type of construction, the presence of the reflexive *se* eliminates the explicit perceiver subject and turns the structure into an impersonal or passive configuration (e.g., Mendikoetxea, 1999); the perception verb may occur either in singular or plural.

(8a)	PV	Inf	NP₂
<i>Se</i>	<i>oye(n)</i>	<i>sonar</i>	<i>las campanas.</i>
REFL	hear-3SG(PL).PRS	ring-INF	ART.F.PL bell-F.PL
‘One hears the bells ring.’ / ‘The bells are heard ringing.’			
(8b)	PV	Inf	NP₂
<i>Se</i>	<i>ve(n)</i>	<i>volar</i>	<i>los pájaros.</i>
REFL	see-3SG(PL).PRS	fly-INF	ART.M.PL bird-M.PL
‘One sees the birds fly.’ / ‘The birds are seen flying.’			

In the descriptive and theoretical literature, Spanish *se*-constructions are commonly discussed as a domain where closely related form-meaning pairings cluster together and partially overlap. In concrete, three relevant construals have been distinguished—passive *se*, impersonal *se*, and middle(-like) *se*—although actual tokens tend to occupy intermediate positions rather than neatly instantiate one discrete type (Maldonado, 1999; Mendikoetxea, 1999).

First, in passive *se* constructions, the clause lacks an overt external argument and instead favors a construal in which the postverbal NP is interpreted as subject-like (as a promoted internal argument) and may control agreement on the finite verb. The classical illustration is *Se venden casas* ‘Houses are sold’. In the perception verb + infinitive configuration, a passive-like construal can be defined as especially compatible with plural agreement, as in *Se ven volar los pájaros*, where the perceived entities are foregrounded as the salient stimulus and can be paraphrased in a subject-prominent way (‘The birds are seen flying’). Second, impersonal *se* is usually defined as a structure without an explicit subject in which *se* introduces a non-specific human perspective (‘one/people’), while the postverbal NP retains a more object-like status; the finite verb then tends to appear in singular. A standard example is *Aquí se trabaja bien* ‘People work well here’ (Sánchez López, 2002). In perception verb + infinitive clauses, singular morphology has therefore often been associated with an impersonal analysis. Third, middle(-like) *se* constructions are typically characterized by participant defocusing and a shift in attention toward the event or situation as such. What is salient is not who initiates the event nor the affected participant, but rather the manifestation of the situation (Maldonado, 1999). In the perception verb + infinitive construction, this middle-like region is relevant because singular agreement readily supports an event-centered construal: *Se oye volar las moscas* naturally profiles the perceptual scene (‘the flying is audible/one can hear the flying’), with reduced prominence of the participants.

Traditional grammars, including the *Nueva Gramática de la lengua española*, frequently treat agreement as a key diagnostic for distinguishing passive vs. impersonal *se* and have often attributed singular agreement to the presence of *se* itself. However, the impersonal

analysis has been questioned because *se*, as a clitic, does not behave like a canonical subject in the ordinary sense (NGLE, §41.12e–l). Moreover, real usage does not reveal a clean passive-impersonal classification as the same *se* pattern can occur in contexts compatible with either interpretation, and agreement and other cues do not consistently align with one category or the other. So, closely related sentences can exhibit mismatches between agreement, Differential Object Marking, and interpretive intuitions (e.g., the contrast between *Se busca a los culpables* ‘one looks for the guilty ones’ and *Se buscan soluciones* ‘solutions are sought’), suggesting that agreement morphology does not transparently encode a categorical construction choice.

This paper therefore goes beyond the diagnostic tradition and treats the pronominal (infinitive) construction as a constructional family in which agreement participates in the profiling of different construals. Crucially, in the complex perception verb + infinitive configuration, the hearer needs not first compute a fully articulated impersonal agent interpretation and only then assign agreement. Rather, singular vs. plural agreement functions as a cue toward an event-centered interpretation (often middle-like) or an entity-centered one (often passive-like). The empirical goal of the paper is to identify which features most strongly condition these interpretations in usage and acceptability patterns.

Importantly, the focus on the infinitival construction is motivated by the considerable controversy it has generated in the literature regarding its internal structure. The debate centers above all on the hybrid status of the second participant, the so-called NP₂, which occupies an intermediate position between the perception verb and the infinitive or can be postposed to the complex predicate. On the one hand, NP₂ clearly functions as the semantic subject of the event expressed by the infinitive, that is, the participant performing the action being perceived. At the same time, the infinitive, by its very nature, lacks the capacity to assign case to its arguments. For this reason, NP₂ often appears morphologically marked as the DO of the perception verb. In the grammatical tradition, this configuration is therefore referred to as *accusativus cum infinitivo* (e.g., [Del Rey Quesada, 2022](#)).

Without going into detail, this functional ambiguity has given rise to a large debate as to whether we are dealing with a simple construction, involving a DO and a predicative complement, or rather with a complex construction, in which the infinitive functions as an independent subordinate predicate. A further question concerns the nature of this subordinate unit: if we accept that the infinitive forms part of a subordinate structure, does it represent a propositional unit, that is, a complete clause (e.g., [Felser, 1999](#)), or rather a non-propositional structure (e.g., [Rodríguez Espiñeira, 2000](#))? These issues relate directly to the type of syntactic link established, on the one hand, between the perception verb and NP₂, and on the other, between that NP₂ and the infinitive (see [Engels \(2007\)](#) for an overview of these different analyses and the arguments raised).

Consequently, this construction offers a particularly rich window into the interaction between syntactic and conceptual dimensions in structures that do not fit into traditional grammatical categories. From a constructional perspective, this raises a further question: how many argument slots does a construction such as *ver a los niños jugar* ‘to see the children play’ or *oír a los niños cantar* ‘to hear the children play’ actually contain? Are we dealing with a three-slot structure, comprising subject, verb, and a complex complement that combines an NP₂ and an infinitive, or rather with a construction that includes four clearly differentiated elements?

These observations lead to the question whether a homogeneous analysis of all infinitive constructions following perception verbs is in fact possible and necessary. An alternative hypothesis would be that the internal organization of the construction depends on the semantics of the perception verb itself and on the nature of the perceived stimulus. In other words, the variability observed across verbs such as *ver*, *oír*, *mirar* and *escuchar*

may be motivated by differences in how the perceptual experience is construed. From this perspective, the alternation between singular and plural agreement in *se*-constructions could be seen as a form of agreement *ad sensum*, that is, agreement guided by the speaker's conceptualization of the perceptual scene rather than by purely syntactic constraints.

It is generally known that, under the label of 'perception verbs', we find verbal forms that, while sharing certain general characteristics, follow distinct semantic dynamics depending on the perceptual modality they encode (Ibarretxe-Antuñano, 1999; Enghels, 2007, 2019). Crucially, these differences are not arbitrary but reflect a well-attested hierarchy of perceptual modalities (Viberg, 1984), according to which vision occupies a privileged position in Western societies as a source of reliable and objective information, followed by auditory perception and at a lower scale, tactile, gustative, and olfactory perception. This perceptual hierarchy leaves clear traces in language: visual perceptions tend to be encoded by the highest number of lexically specific verbs, whereas less dominant modalities such as touch, taste, or smell are more often expressed through semantically underspecified or multifunctional verbs (e.g., *sentir* 'to feel', 'to smell', sometimes 'to hear'). In addition, visual perception verbs stand out for their greater syntactic flexibility, allowing a wider range of complementation patterns. Finally, these verbs also display a higher degree of semantic flexibility, giving rise to extensive metaphorical extensions and polysemy (Sweetser, 1990; Evans & Wilkins, 2000).

What is most relevant for the present study is the difference in the nature of the perceived stimulus. In the case of visual perception, it is sufficient for the object to be present in order to be perceived. In contrast, auditory perception is not activated merely by the presence of a stimulus: it requires the emission of a sound effect. This implies that auditory stimuli are inherently dynamic, since they must produce sound. Thus, when we say *oigo al niño* 'I hear the child', what we actually perceive is the auditory result of some activity carried out by the child. This contrasts with visual perception, where the child may simply be present and serve as the object of perception. This distinction is crucial, as it significantly conditions the relationship between the verb, the type of stimulus, and the grammatical structure that accompanies it, as further analyzed through a corpus study.

4. Variable Agreement: Results from a Previous Corpus Study

The first case study examines empirical data from a corpus. This section provides a concise summary of Enghels (2019), from which only the most relevant conclusions are retained, serving as a starting point for the subsequent experimental analysis. To empirically determine which factors condition singular or plural agreement in the pronominal infinitive construction, a corpus was compiled containing all occurrences of this construction with the four perception verbs under investigation: *ver*, *oír*, *mirar*, and *escuchar*. The data were extracted from a range of sources, including written press (e.g., *El País*, *El Mundo*), literary texts, and electronic Spanish databases (e.g., Corpus de Referencia del Español Actual (CREA), Corpus del Español (CDE)) in order to cover different genres and registers. The final dataset comprised approximately 4000 infinitive constructions, of which 748 correspond to pronominal cases.

The collected constructions were classified according to the formal agreement relationship between the perception verb and the subordinate NP₂ participant. The first group comprises those cases in which both the verb and the participant are singular. In a second group, the subordinate participant NP₂ is plural while the verb appears in singular (e.g., *Se oye_{SG} entrechocar [botellas]_{PL}* 'One hears bottles clinking together'). A third category contains examples showing plural agreement, i.e., both the subordinate participant NP₂ and the verb appear in plural (e.g., *Se veían_{PL} bajar [reatas de mulas]_{PL}* 'Strings of mules were seen coming down'). The analysis only concentrated on the second and third patterns, as these are the

only contexts that allow a contrastive examination of whether a postverbal plural NP₂ participant triggers formal agreement in the perception verb. Notably, these represented a relatively small but balanced subset of the total corpus (singular agreement n = 88; plural agreement n = 90).

Figure 1 displays the correlation between perceptual modality and the type of verbal agreement: singular agreement is represented in blue, and plural agreement in contrasting color (see also Enghels, 2019, p. 119). Based on these data, the statistical analysis reveals a clear contrast between visual and auditory perception verbs⁴. The visual ones show a marked preference for plural agreement: in 59.3% of the analyzed cases, the plural subordinate participant triggers plural formal agreement on the main verb (that is, pattern 3 described above). In contrast, the auditory verbs display a strong preference for singular agreement, even when the subordinate NP₂ participant is plural, namely in 63.8% of the analyzed cases (that is, pattern 2 described above).

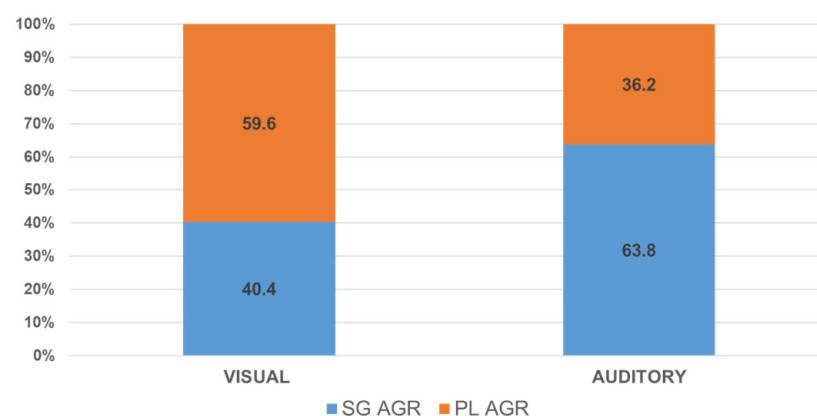


Figure 1. Distribution of singular and plural agreement according to perception modality.

In the examples involving visual perception verbs (9a–b), the second participant typically follows the infinitive, which creates a certain structural distance between the main verb and the subordinate participant. Despite this distance, however, formal agreement systematically occurs between the two constituents. In (9b), the structural distance is even greater, as the main verb and the infinitive are separated by a prepositional complement (*con dirección al trabajo*), yet plural agreement is still maintained. This observation suggests that such agreement cannot be accounted for by purely grammatical factors and instead calls for a conceptual interpretation. A plausible hypothesis is that the speaker perceives a strong connection between the perceptive act and the initiator of the perceived event. By contrast, the examples with auditory verbs in (10a–b) systematically lack agreement between the main verb and the subordinate NP₂ participant, regardless of whether the latter appears immediately after the verb or at a greater distance. In the literature, this behavior has been interpreted as reflecting the more impersonal nature of auditory constructions (Mendikoetxea, 1999; Enghels, 2007). In sentences such as (10a) *se oye a los grillos beber*, the interpretation does not involve a specific perceiver but rather conveys a generic perception, as in ‘it is heard that the crickets are drinking’.

- (9a) *Más allá se ven pasar los transeúntes.*
 farther.away SE see-3PL pass-INF the passers-PL
 ‘Farther away, one sees passers-by going by.’
 (Romero, L., *La noria*, 1952 [Spanish Online])

- (9b) [...] *se miraban pasar con dirección al trabajo*
 cuando SE watch-3PL-IPFV pass-INF with direction to.the work
a los empleados y funcionarios del Banco Nacional.
 DOM the employees and officials of.the Bank National.PL
 ‘When one watched the employees and officials of the National Bank on their way to work, wearing their impeccable shirts, going by.’
 (*La Prensa*, December 5, 1997 [CREA])
- (10a) *Antes de salir el sol se oye a los grillos beber.*
 before of rise-INF SE hear-3SG DOM the crickets.PL drink-INF
 the sun
 ‘Before sunrise, one hears the crickets drinking.’
 (Ruy Sánchez, A., *Los jardines secretos de Mogador*, 2002 [CDE])
- (10b) *Dicen los campesinos que en las noches oscuras aún se escucha gemir*
 say-3PL the peasants that in the nights dark still SE hear-3SG moan-INF
a las almas sin sepultura.
 DOM the souls without burial.PL
 ‘Peasants say that on dark nights one can still hear the souls without burial moaning.’
 (Cuví, P., Ecuador. *Paso a paso*, 1994 [CREA])

According to the hypothesis proposed, the higher proportion of plural agreement observed with *ver* and *mirar* can be explained by the object-centered nature of visual perception (*supra* Section 3). Visual perception tends to focus on and foreground the participant who initiates the perceived event, thereby favoring an interpretation in which this second participant is placed in the foreground. In such cases, visual verbs select infinitival complements of a non-clausal type, corresponding to a 4-slot type of the whole construction, and the speaker establishes a direct semantic link between the subordinate participant and the perception verb itself. This promotes the interpretation of that participant as the (grammatical) subject of the main predicate, which in turn triggers plural agreement. The prototypical constructional configuration for visual perception verbs corresponds to a non-clausal analysis of the construction and can be represented as:

$$SE\ ve(n)/mira(n)_{SG/PL}\ [NP_2]_{SG/PL}\ [Inf]$$

(agreement → salience of the perceived entity as stimulus)

In contrast, the more event-oriented nature of auditory perception explains the predominance of singular agreement in pronominal infinitive constructions with *oír* and *escuchar*. These verbs preferentially select infinitival complements that denote dynamic events, where the speaker first establishes a semantic relationship between the subordinate participant NP_2 and the infinitive, and only subsequently between that complex and the perception verb. In this configuration, the second participant is no longer the core perceived entity; rather, the entire subordinate clause constitutes the perceptual focus. Consequently, agreement is in singular. Hence the prototypical constructional configuration for auditory perception verbs, corresponding to a clausal analysis of the construction, is:

$$SE\ oye/escucha_{SG}\ [(NP_2\ energy\ source)_{SG/PL} + INF]_{SG}$$

(SG agreement → salience of the event as stimulus)

Thus, the observed alternation in the corpus data shows a tendency toward semantic agreement, guided by conceptual salience of the stimulus rather than by strictly grammatical rules. In this view, variation reflects subtle cognitive adjustments during linguistic production.

However, it was also observed that these configurations do not represent fixed patterns but rather prototypical tendencies. As the corpus data reveal, instances of singular agreement also occur with visual perception verbs (11a), and plural agreement occasionally

5. Variable Agreement: An Acceptability-Rating Task

This section presents the experimental phase of the study, which systematically manipulates the factors identified in the corpus (perceptual modality and agreement type), along with additional variables related to the notion of salience (such as the animacy of the perceived participant), in order to test their effect on acceptability by Spanish speakers. Section 5.1 describes the design of the acceptability-rating task, including the selection and manipulation of stimuli, participant characteristics, and the procedure used to elicit speakers' ratings. Section 5.2 outlines the hypotheses guiding the experiment, derived from the main factors identified in the corpus data, and formulates predictions concerning their expected influence on acceptability. Section 5.3 reports the results of the analysis, beginning with an overall inspection of the ratings and followed by a multivariate statistical model testing the effects and interactions of the experimental variables.

5.1. Methods: Participants, Materials and Procedure

We recruited 110 native speakers of Peninsular Spanish through the Prolific platform (Palan & Schitter, 2018; <https://www.prolific.co/>, accessed on 19 December 2025). The task was restricted to monolingual participants of Spanish nationality with normal or corrected-to-normal vision, and who reported no history of language-related, neurological, or diagnosed mental health disorders. Ten participants were excluded from the dataset due to low response accuracy (<80% correct responses) on the yes/no comprehension questions, suggesting insufficient attention to the task. Consequently, the final sample included 100 participants (38 women, 62 men; mean age = 30.4 years, SD = 10.6, range = 19–57), whose acceptability ratings were retained for statistical analysis⁶.

In total, the materials for the acceptability-rating experiment included 200 stimuli: 40 critical sentences and 160 fillers. All sentences are based on authentic examples from the corpus referred to in Section 4 above but were systematically manipulated to have a homogeneous structure. The materials were all simple clauses with predicates in either the Present Indicative or the Preterit. Sentence length varied between six and twelve words and was therefore included as a control variable in the statistical analysis. Word order was also controlled by adopting a fixed constituent sequence: (adverbial) [Se] [PV] [INF] [NP₂]. This structure ensured comparability across stimuli and isolated the factors of interest. In order to be able to measure the effect of variable agreement with the verb, all target sentences contained a plural NP₂.

In concrete, three independent variables were manipulated:

1. The modality of the perception verb, contrasting visual and auditory perception;
2. Agreement, which could be singular or plural between the verb and NP₂;
3. The animacy of NP₂, which could refer to a human or an inanimate entity, to further operationalize the concept of salience (cf. *infra* Section 5.2).

Two perception verbs were used, *ver* 'see' and *oír* 'hear'⁷, with twenty sentences per verb, half combined with a singular verb and half with a plural verb. This agreement factor was further crossed with the semantic nature of the referent: half of the examples featured a human NP₂, and the other half an inanimate one. In addition, we also considered the presence or absence of Differential Object Marking (DOM). As will be further developed below (cf. *infra* Section 5.2), this phenomenon refers to the use of the preposition *a* in Spanish (and other (Romance) languages) to mark DOs that are typically human (or animate) and specific. Its presence increases the syntactic visibility of the object and signals its high referentiality or prominence in the discourse (e.g., Torrego Salcedo, 1999). Following this grammatical norm and the most common patterns observed in the corpus, sentences with singular agreement of the perception verb contained a human NP₂ marked by DOM (e.g., *se ve retroceder a los policías* 'one sees the police officers move back'). The presence of DOM

marks NP₂ unequivocally as the DO of the perception verb, thereby preventing structural ambiguities that could have otherwise influenced participants' ratings. This was not the case for (a) sentences with plural agreement (e.g., *se oyen reír sus compañeros* 'one hears their companions laughing'), where agreement already disambiguates the structure for the rater with NP₂ marked as the subject of the perception verb, nor (b) sentences with inanimate NP₂s (e.g., *se ven transitar autos* 'one sees cars going by') that do not license DOM. Table 1 summarizes how the critical sentences were distributed across the variables.

Table 1. Distribution of critical sentences across independent variables.

Perception Verb	Agreement	Human	Inanimate	Total
<i>ver</i>	PL	5	5	10
	SG	5	5	10
<i>oír</i>	PL	5	5	10
	SG	5	5	10
Total		20	20	40

To illustrate the experimental materials, a selection of target sentences is presented under (13).

- (13)
- | | |
|-------------------------|---|
| ver—SG—human | <i>En muchas situaciones peligrosas se ve retroceder a los policías.</i>
'In many dangerous situations, one sees the police officers move back.' |
| ver—SG—inanimate | <i>En las aguas del mar se ve flotar cañas.</i>
'In the waters of the sea, one sees reeds floating.' |
| ver—PL—human | <i>Por la noche se ven pasar patrullas militares.</i>
'At night, one sees military patrols passing by.' |
| ver—PL—inanimate | <i>En las calles se ven transitar autos.</i>
'In the streets, one sees cars going by.' |
| oír—SG—human | <i>En la sala de al lado se oye hablar a las mujeres.</i>
'In the room next door, one hears the women talking.' |
| oír—SG—inanimate | <i>En el club se oye chocar fichas de dominó.</i>
'In the club, one hears domino tiles clacking together.' |
| oír—PL—human | <i>A lo lejos se oyen reír sus compañeros.</i>
'In the distance, one hears their companions laughing.' |
| oír—PL—inanimate | <i>A un costado se oyen crujir ramas pisoteadas.</i>
'Off to the side, one hears trampled branches cracking.' |

The fillers were designed to elicit variable acceptability ratings based on previous findings, in order to encourage participants to use the entire rating scale. Specifically, some fillers contained the Spanish inchoative construction (serving as critical stimuli for [Baltais et al., 2026](#)), some featured the Spanish factitive construction ([Roegiest & Enghels, 2008](#)), and others instantiated the rare (cf. note 7) pronominal infinitive constructions with *escuchar* ‘to listen’ and *mirar* ‘to look’ ([Enghels, 2019](#)). The practice phase included 10 filler sentences. The remaining 190 stimuli were distributed over ten experimental blocks, which were presented in random order. Within each block, the order of stimuli was pseudorandomized: each contained four critical sentences, and no two critical items appeared consecutively. Finally, 21 ‘yes/no’ comprehension questions were included throughout the experiment as attention checks.

All participants completed the experiment online via LimeSurvey (<https://www.limesurvey.org/>, accessed on 19 December 2025; average completion time: 51 min). At the start, they were informed that the study had been approved by the Ethics Committee of the Faculty of Psychology and Educational Sciences at Ghent University, and they provided their informed consent to participate. Next, they filled out a sociodemographic

questionnaire, followed by the acceptability-rating task. Finally, they completed the Spanish version of the Big Five Inventory-2 (BFI-2: Soto & John, 2017).

For the main task, participants were instructed to rate each sentence on a 7-point scale according to how ‘acceptable’ it sounded. An acceptable sentence was defined as one that would seem natural for a native speaker of Spanish to say or to write. They were also informed that some sentences would be followed by a content-related ‘yes/no’ comprehension question. The sentences were presented one at a time on the screen, preventing participants from directly comparing them or revising previous ratings. The task began with the practice block, followed by ten experimental blocks, during which participants were offered two opportunities to take a short break. After completing the rating task, participants answered an open question about what type of sentences they believed the researchers were interested in. The responses indicated that all participants remained unaware of the study’s purpose. All participants received payment for their participation (hourly rate: 10 €).

5.2. Hypotheses and Predictions

5.2.1. Hypothesis 1

As described in Section 5.1, the experimental design crossed the factors of perceptual modality (*ver* vs. *oír*) and verb–NP₂ agreement (singular vs. plural). This was motivated by previous corpus-based observations (see Section 4), which suggested that variation in agreement patterns may depend on the conceptualization of the perceptual event.

First, building on the earlier definition of the perception modalities (see Section 3), it is reasonable to hypothesize that the dominance of visual perception and its strong grounding in embodied experience may affect acceptability ratings more generally. Constructions involving visual perception verbs may benefit from greater processing fluency and conceptual accessibility, which could translate into higher acceptability ratings overall, independently of strictly grammatical constraints. Importantly, this does not imply that visual constructions are inherently more grammatical, but rather that they may be perceived as more natural or self-evident due to the privileged status of vision in human perception and cognition. This leads to prediction (H1a) including an overall perception modality dominance effect.

Prediction (H1a): *ver* > *oír*

A follow-up prediction (H1b) concerns the interaction between perceptual modality and agreement. It was formerly stated that the two perception verbs differ in how they shape the mental representation of events. The visual modality, associated with *ver*, tends to promote a discrete construal of object referents: it isolates individual entities within the visual field, enhancing their perceptual distinctness. This conceptualization makes plural NP₂ referents potentially more salient and is thus expected to favor plural agreement. In contrast, the auditory modality, associated with *oír*, evokes a unitary construal in which the perceived stimulus is processed as a single, continuous auditory event, favoring singular agreement. If these modality-specific construals influence speakers’ grammatical evaluations, they should be reflected in the acceptability ratings. Specifically, we expect that sentences with *ver* in plural will receive the highest ratings, followed by *ver* in singular. On the contrary, we expect higher ratings for the sentences with *oír* in singular, and lower ones for *oír* in plural, which represents the least natural conceptual match.

Prediction (H1b): *ver* PL ≥ *ver* SG ↔ *oír* SG > *oír* PL

5.2.2. Hypothesis 2

The second hypothesis is more explorative, as it focuses on the role of animacy in determining agreement preferences as a consequence of different effects of cognitive and perceptual salience. In Cognitive Grammar (e.g., Langacker, 1987, 2008), ‘salience’ refers to the relative cognitive prominence of elements within a conceptual scene, determined by factors such as animacy, agency, and perceptual distinctness, which shape how speakers construe and encode linguistic relations. As such, the notion refers to the relative prominence accorded to specific elements in a conceptual whole, often described through the notions of Figure vs. Ground, Trajector vs. Landmark, or Profile vs. Base (Langacker, 2008, p. 66). More specifically, salience can be the result of multiple interacting factors, including (a) perceptual properties (e.g., more noticeable elements because of being involved in movement, or a higher level of brightness or proximity), (b) discourse-cognitive factors (e.g., elements that are more accessible or predictable in context because of being introduced before), or (c) overall conceptual properties (e.g., entities that are more central in an event). In this framework, we hypothesize that animacy may function as a source of conceptual salience: human and animate entities, being capable of action, volition, and perception, are expected to attract greater attentional prominence and thus become more strongly foregrounded in mental representations.

We therefore expect that animate (plural) referents will display a higher degree of conceptual salience than inanimate ones, which leads us to two predictions. First, we hypothesize that sentences containing animate, particularly human, plural referents will generally receive higher acceptability ratings than sentences with inanimate plural referents. This prediction (H2a) targets a general animacy effect, motivated by the fundamentally anthropocentric orientation of human language (e.g., Duranti, 1997). Because language is produced by and for human interlocutors, it systematically privileges animate, intentional, and agentive beings, which are more readily construed as actors and perspective holders in discourse.

Prediction (H2a): *human > inanimate*

Additionally, higher salience should make plural agreement appear more natural for speakers rating the sentences in the experiment, leading to higher acceptability ratings when the verb morphologically reflects the plurality of a human referent. Building on this reasoning, our prediction (H2b) is that human NP₂s will favor plural agreement more strongly than inanimate NP₂s. This preference is formulated as a greater-than-or-equal relation, since, following predictions (H1a) and (H1b), human referents are also expected to receive higher acceptability ratings overall, independently of agreement, potentially giving rise to a ceiling effect. Conversely, when the referent is inanimate, its conceptual salience is expected to be lower. Inanimate entities are less likely to be construed as central participants and are instead perceived as part of a homogeneous or backgrounded setting, being the perceived event as a whole. As a result, plural agreement in these cases may appear less natural, since the grammatical marking of plurality is not supported by a strong conceptual cue.

Prediction (H2b): *human PL ≥ human SG ↔ inanimate SG > inanimate PL*

5.3. Results

5.3.1. Visual Inspection of the Data

To begin, as shown in Figure 2, the distribution of acceptability ratings of the critical stimuli is clearly skewed toward the higher end of the scale. This suggests that, overall, pronominal infinitive constructions are not perceived by speakers as odd or ungrammatical,

but rather as generally acceptable forms within the language, although there is considerable variability across items (mean rating = 5.17, SD = 1.90, range: 1–7).

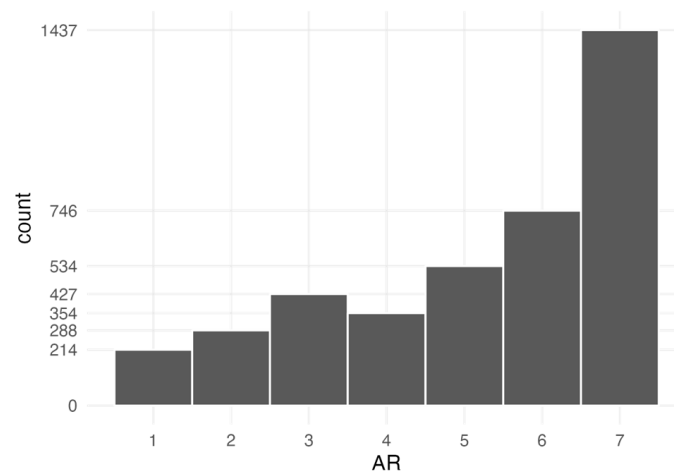


Figure 2. Distribution of acceptability ratings for the critical stimuli.

Next, Figure 3 shows that no significant differences were found between the acceptability ratings of constructions with singular and plural agreement. This result is particularly relevant, as it suggests that both variants are cognitively available to Spanish speakers and perceived as grammatically valid options. In other words, agreement with a whole infinitive constituent (singular) and agreement with a specific participant NP₂ (plural) appear to be equally represented in speakers' internal grammatical competence. This finding reinforces the view that these patterns correspond to competing constructions rather than errors or performance deviations, in line with a (constructionist) perspective in which multiple schemas may coexist and be selectively activated depending on semantic, discourse, or processing factors (e.g., Hilpert, 2014).

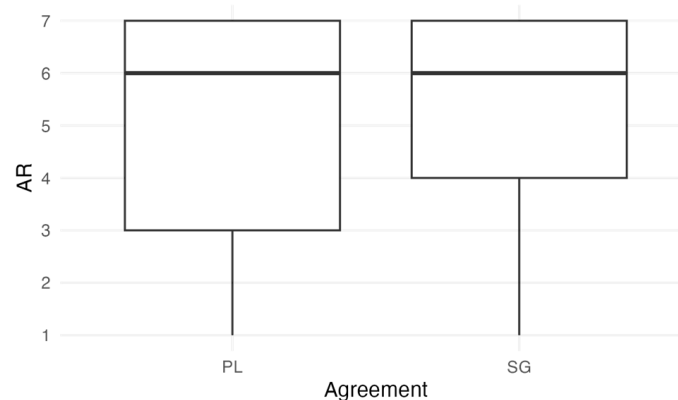


Figure 3. Distribution of acceptability ratings per singular or plural pattern.

Figure 4 compares the acceptability ratings for constructions with visual perception verbs (*ver*) and auditory perception verbs (*oír*). As shown, there are no substantial differences either in the median or in the overall distribution of ratings across the two perceptual modalities. This absence of significant differences indicates that these constructions are not prototypically associated with the visual domain, despite its status of dominant perception modality. Consequently, the perception dominance hypothesis (H1a) is not supported by the experimental data.

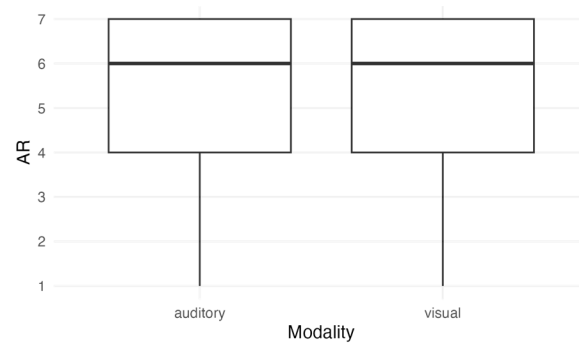


Figure 4. Distribution of acceptability ratings per perception modality.

Next, Figure 5 allows us to assess whether perceptual modality interacts with agreement, as predicted in H1b. For constructions with *ver*, acceptability ratings for singular and plural agreement are almost indistinguishable, with highly overlapping distributions and no clear difference in central tendency. This suggests that visual perception does not strongly bias speakers' evaluations toward either agreement option. In contrast, constructions with *oír* show a more pronounced differentiation: singular agreement tends to receive higher ratings than plural agreement, in line with the hypothesis that auditory perception favors a unitary construal of the perceived event. Although this asymmetry between visual and auditory perception aligns with the predictions underlying H1b, the observed differences are modest and show substantial overlap. It therefore remains an open question whether the apparent modality-agreement pattern for *oír* reaches statistical significance, a point that requires confirmation through inferential analysis (see Section 5.3.2).

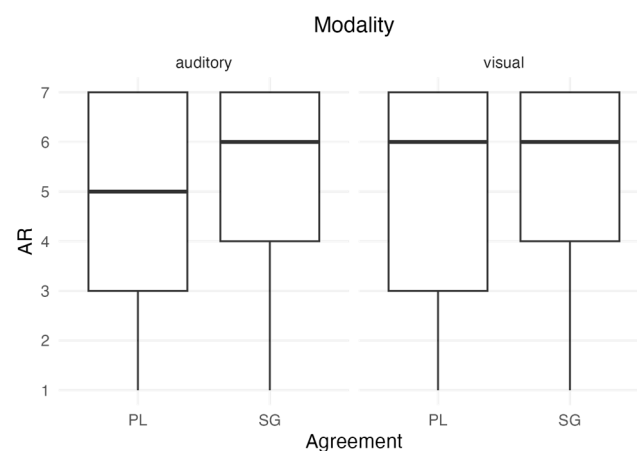


Figure 5. Distribution of acceptability ratings per perception modality and agreement pattern.

Figure 6 illustrates the differences in acceptability ratings assigned to sentences with human and inanimate plural referents. The preliminary inspection of the data distribution points towards a dissociation between human and inanimate NP₂s. Overall, participants rate stimuli containing human referents as more natural and acceptable than those with inanimate referents, a trend that seems to go in line with H2a.

Finally, Figure 7 explores the role of animacy in shaping agreement preferences, as proposed in H2b. For human referents, singular agreement is rated more favorably than plural agreement, with a higher median and a distribution skewed toward the upper end of the scale. In contrast, for inanimate referents, the pattern is reversed: plural agreement tends to receive higher acceptability ratings than singular agreement. This pattern runs counter to the prediction that higher conceptual salience characteristic of animate entities would promote plural agreement. Taken together, these findings do not support prediction

H2b and do not show a uniform animacy-based salience effect favoring one or the other agreement pattern.

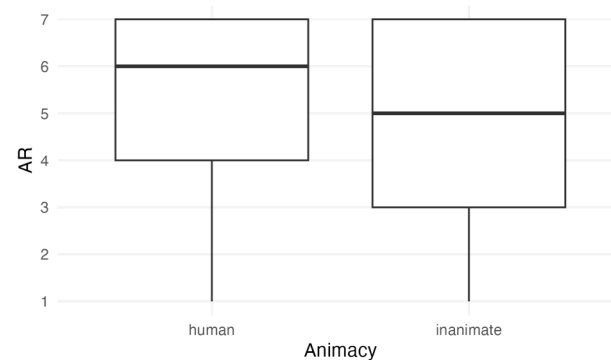


Figure 6. Distribution of acceptability ratings per human or inanimate NP₂s.

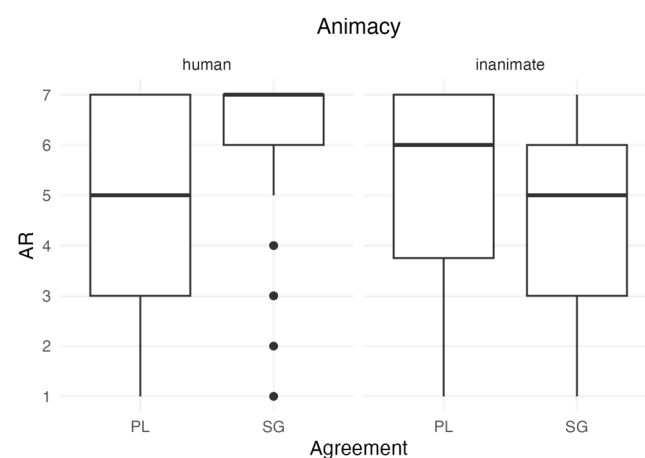


Figure 7. Distribution of acceptability ratings per animacy and agreement pattern.

After this preliminary inspection of the distribution of acceptability ratings, the next step is to statistically test the previously mentioned hypotheses and possible observed interactions. To this end, we examine the effects of the independent variables introduced in Section 5.1 (Modality, Agreement, and Animacy) on the dependent variable, namely the acceptability ratings (AR), as well as their possible interactions. These analyses directly test the predictions formulated in Section 5.2, concerning the expected Modality effect (H1a), the interaction between Modality and Agreement (H1b), the Animacy effect (H2a), and the interaction between Animacy and Agreement (H2b). To evaluate the significance and relative contribution of these predictors, we fitted a Cumulative Link Mixed Model (CLMM). The details of this model specification and the main results are described below.

5.3.2. Multivariate Analysis CLMM

As the dependent variable, namely the acceptability ratings, was measured on a Likert scale, it was treated as ordinal data. Following methodological recommendations in the literature (e.g., Jamieson, 2004; Grilli & Rampichini, 2011; Christensen & Brockhoff, 2013), we fitted a Cumulative Link Mixed Model using the *clmm()* function from the *ordinal* package (Christensen, 2022) in R (version 4.2.0: R Core Team, 2022). The fixed-effects structure contained the categorical predictors Modality (visual vs. auditory), Agreement (singular vs. plural), Animacy (human vs. inanimate), as well as their two- and three-way interactions. We also included Sentence Length (in words) as a control variable, given that longer sentences tend to receive lower ratings than shorter sentences (e.g., Featherston, 2005). A maximal random-effects structure was specified (Barr et al., 2013),

with random intercepts for items and participants and random by-participant slopes for all predictors excluding the control variable to avoid anticonservative inference. For categorical predictors, deviation-coded contrasts were applied (−0.5, 0.5), whereas the continuous control variable was centered around its mean. After fitting the model, we conducted pairwise comparisons using the *emmeans* package (Lenth, 2022) to examine the significant interactions.

Table 2 presents the fixed-effects estimates from the CLMM; the full output, including random effects, is provided in Appendix A.2.

Table 2. Output of CLMM (fixed effects only), showing estimates on the log-odds scale.

Fixed Effects ¹	Estimate	SE	z-Value
Modality	−0.03	0.33	−0.08
Agreement	0.33	0.33	0.98
Animacy	−1.07	0.30	−3.61 ***
SentenceLength	0.17	0.18	0.98
Modality * Agreement	−0.40	0.57	−0.69
Modality * Animacy	−2.15	0.58	−3.72 ***
Agreement * Animacy	−2.45	0.63	−3.87 ***
Modality * Agreement * Animacy	2.61	1.15	2.26 *

¹ Number of observations: 4000; number of groups: participant 100, item 40. Formula in R: `clmm (AR~Modality * Agreement * Animacy + SentenceLength + (1|item) + (1 + Modality * Agreement * Animacy | participant))`.
* $p < 0.05$, *** $p < 0.001$.

First, as was observed in the exploratory analysis (Section 5.3.1), neither Modality nor Agreement show a significant main effect, suggesting that, taken in isolation, perceptual modality (*ver* vs. *oír*) and agreement type (singular vs. plural) do not systematically influence acceptability judgments. However, the model did reveal a significant main effect of Animacy: overall, sentences with human NP₂s were about 2.9 times more likely to receive higher ratings than those with inanimate NP₂s (based on the inverse of the odds ratio of 0.34 calculated as $e^{-1.07}$; $z = -3.61$, $p < 0.001$). Concretely, this means that perception–verb sentences containing more salient, namely human, referents were systematically rated as sounding more acceptable than those containing inanimate referents, which confirms Hypothesis 2a.

However, contrary to our expectations (formulated as an outcome of the corpus study in Section 4) and despite the visual trend observed in Figure 5, there was no evidence for an interaction between Modality and Agreement when averaged across both levels of Animacy. This means that, once human and inanimate referents are considered together, speakers do not systematically prefer plural agreement with visual verbs or singular agreement with auditory verbs in their acceptability ratings. In other words, the modality-driven preferences observed in the corpus do not surface as a general evaluative bias when speakers rate the sentences in isolation. As such, Hypothesis 1b is not supported at the global level of analysis, even though more fine-grained patterns emerge once animacy is taken into account.

Indeed, there were significant two-way interactions of Animacy with Modality ($p < 0.001$) and Animacy with Agreement ($p < 0.001$), and a significant three-way interaction ($p < 0.05$), indicating that Animacy modulates the relationship between the other two predictors. As can be seen from Figure 8, which illustrates the distribution of Likert-scale acceptability ratings across levels of the three predictor variables, the effects of both Modality and Agreement reversed depending on Animacy. Sentences with *ver* ‘to see’

were generally rated as more acceptable than those with *oír* ‘to hear’ if they contained human NP₂s, whereas in the case of inanimate NP₂s, *oír* scored higher than *ver*. Similarly, compared to plural agreement, sentences with singular agreement received higher ratings with human NP₂s but not with inanimate NP₂s. These observations were supported by omnibus tests conducted separately for each Animacy level: significant main effects of Modality (human: $F(1, \infty) = 5.83, p = 0.016$; inanimate: $F(1, \infty) = 6.23, p = 0.013$) and Agreement (human: $F(1, \infty) = 10.15, p = 0.001$; inanimate: $F(1, \infty) = 4.31, p = 0.038$) were observed for both types of referents. Additionally, omnibus tests showed that for human NP₂s, there was a significant two-way interaction between Modality and Agreement ($F(1, \infty) = 4.42, p = 0.036$), which was not the case for inanimate NP₂s ($F(1, \infty) = 1.25, p = 0.264$).

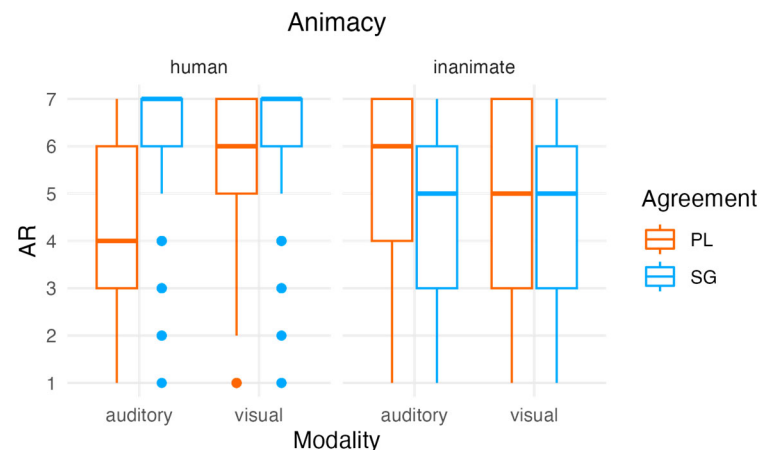


Figure 8. Acceptability ratings (y -axis: 7-point Likert scale) plotted as a function of Animacy (facets: human vs. inanimate), Modality (x -axis: auditory vs. visual), and Agreement (orange for plural, blue for singular). The boxplots display the median and the interquartile range for each condition.

These patterns suggest that the perceptual construal evoked by visual and auditory verbs is not fixed but is modulated by the nature of the perceived stimulus. Human referents appear to align more naturally with visual perception, which privileges individuated and potentially agentive entities, whereas inanimate referents show a closer alignment with auditory perception, which tends to construe the perceived situation at a more global, event-based level. Simultaneously, speakers seem to prefer singular agreement with human referents and plural agreement with inanimate referents. Crucially, the latter pattern runs counter to Hypothesis 2b, which predicted that higher conceptual salience, particularly for human figures, would favor plural agreement, while lower salience would favor singular agreement. Instead, the highest acceptability ratings are obtained for human–singular configurations in both perceptual modalities.

One possible interpretation is that this reversal is, at least partly, driven by the presence of Differential Object Marking (DOM). In all sentences with human referents in the singular condition, the NP₂s were explicitly marked as direct objects, whereas the remaining stimuli lacked DOM. This explicit morphosyntactic marking may enhance the prominence and identifiability of the human referent as the DO of the perception verb, aligning more naturally with singular agreement and thereby boosting acceptability. From this perspective, the higher ratings for human–singular sentences may also reflect the interaction between salience and grammatical encoding: explicit object marking reinforces the construal of the human NP₂ as one salient, individuated participant, reducing the pressure for plural agreement. This could explain why these sentences are rated highest overall, despite running counter to the original animacy-based prediction.

Finally, pairwise comparisons based on estimated marginal means from the CLMM (Table 3) demonstrated that for both levels of Animacy, Agreement significantly affected

acceptability ratings in the auditory but not in the visual modality. For human NP₂s, sentences with *oír* in singular agreement were about 11 times more likely to receive higher ratings than those with *oír* in plural agreement (based on the inverse of the odds ratio of 0.09 calculated as $e^{-2.40}$; $z = -3.75$, $p < 0.001$). Inanimate NP₂s displayed the opposite pattern, albeit a weaker one (as shown *supra* by the absence of a significant interaction in the omnibus test): *oír* in plural agreement was about 4 times more likely to belong to a higher Likert category compared to singular agreement (based on the odds ratio of 3.86 calculated as $e^{1.35}$; $z = 2.22$, $p < 0.05$).

Table 3. Pairwise comparisons of Agreement across levels of Animacy and Modality ¹.

Agreement (PL—SG)					
Animacy	Modality	Estimate	SE	z-Ratio	p-Value
human	auditory	−2.40	0.64	−3.75	<0.001
human	visual	−0.70	0.63	−1.12	0.265
inanimate	auditory	1.35	0.61	2.22	0.026
inanimate	visual	0.44	0.58	0.76	0.447

¹ based on estimated marginal means from the cumulative link mixed model. Estimates are on the log-odds scale; positive estimate values indicate higher acceptability ratings for plural relative to singular agreement.

Simultaneously, it can be said that for both levels of Animacy, Modality affected ratings in the plural but not in the singular agreement (Table 4). While human–singular sentences—possibly due to the presence of DOM—scored at ceiling regardless of perception modality, human–plural sentences were about 7 times more likely to receive higher ratings with *ver* than with *oír* (based on the inverse of the odds ratio of 0.15 calculated as $e^{-1.90}$; $z = -3.16$, $p < 0.01$). Notably, this result goes in line with (a part of) H1b. Sentences with inanimate referents, however, did not follow the same trend, as participants preferred plural marking with *oír* rather than with *ver*, the former sentences being about 5 times more likely to belong to a higher Likert category compared to the latter (based on the odds ratio of 4.76 calculated as $e^{1.56}$; $z = 2.61$, $p < 0.01$).

Table 4. Pairwise comparisons of Modality across levels of Animacy and Agreement ¹.

Modality (<i>oír</i> — <i>ver</i>)					
Animacy	Agreement	Estimate	SE	z-Ratio	p-Value
human	PL	−1.90	0.60	−3.16	0.002
human	SG	−0.20	0.59	−0.33	0.739
inanimate	PL	1.56	0.60	2.61	0.009
inanimate	SG	0.65	0.61	1.07	0.287

¹ based on estimated marginal means from the cumulative link mixed model. Estimates are on the log-odds scale; positive estimate values indicate higher acceptability ratings for auditory relative to visual modality.

What implications do these results have for the research hypotheses formulated in Section 5.2? The analysis confirms the central role of Animacy, with sentences containing human NP₂s receiving higher acceptability ratings overall than those with inanimate NP₂s (H2a). However, the direction of the Animacy × Agreement effect runs counter to Hypothesis 2b, which predicted higher ratings for plural agreement with human referents. Instead, singular agreement is preferred with human NP₂s, whereas plural agreement tends to be favored with inanimate NP₂s. This reversal suggests that the animacy-based predictions in H2b might be modulated by an additional grammatical factor not explicitly incorporated into the original hypotheses, namely, DOM. The presence of DOM in human–singular

configurations appears to yield higher acceptability for singular marking as compared to plural marking, where DOM is absent. Remarkably, ratings for human–plural configurations follow the pattern predicted by H1b: sentences in visual modality scored higher than those in auditory modality. For inanimate NP₂s, which are not marked by DOM in any of the conditions, plural agreement appears as generally more acceptable than singular agreement. Moreover, counter to Hypothesis 1b, for inanimate referents *oír* aligns with plural marking better than *ver*.

6. Discussion and Conclusions

Overall, it is fair to conclude that the results of the acceptability-rating task offer a more nuanced picture of agreement variability in perception–verb constructions than the corpus analysis suggested. While the corpus study clearly showed an effect of perceptual modality on agreement configurations—with *ver* favoring plural agreement and *oír* favoring singular agreement—the experimental data did not fully replicate this pattern. No global interaction between Modality and Agreement emerged, indicating that when speakers are asked to rate the naturalness of these constructions in isolation, they do not systematically prefer one agreement type over the other depending on perceptual modality. This discrepancy between corpus and experimental evidence might reflect a task effect: corpus data capture production, where speakers make spontaneous linguistic choices shaped by discourse context and communicative goals, whereas the acceptability task taps into processing and metalinguistic evaluation, where sentences are assessed without a discourse frame and thus without the pragmatic pressures that favor certain configurations.

Despite the absence of an overall visual modality dominance effect or an interaction of Modality with Agreement, the analysis confirmed that both the singular and plural variants are cognitively active and equally represented in speakers' internal grammar. The balanced acceptability ratings for both agreement patterns suggest that speakers treat them as coexisting constructions, each linguistically valid and licensed by the grammar of contemporary Spanish.

At the same time, the results clearly identified Animacy as a decisive factor in acceptability, modulating the relationship between Modality and Agreement. As expected, sentences with human NP₂s received higher ratings overall than those with inanimate NP₂s. Yet the direction of this effect was the inverse of what was initially hypothesized: rather than promoting plural agreement, the presence of human referents favored singular marking, while inanimate referents were more compatible with plural agreement. This inversion can be explained by the role of Differential Object Marking (DOM). As DOM is almost categorically associated with human and definite objects, its presence reinforces the DO status of NP₂ and thus enhances the acceptability of singular agreement.

The results showed another clear asymmetry between human and inanimate NP₂s. For human referents with plural agreement (not marked by DOM), acceptability ratings pattern as predicted by H1b, with sentences involving visual perception receiving higher scores than those involving auditory perception. In contrast, for inanimate NP₂s (also not marked by DOM) plural agreement is overall rated as more acceptable than singular agreement, regardless of perceptual modality. Notably, and contrary to H1b, in this inanimate domain plural agreement showed a slight preference for *oír* over *ver*. At a more abstract level, this divergence suggests that the interaction between perceptual modality and agreement is not uniform across referent types but is mediated by the conceptual status of NP₂.

Overall, these findings demonstrate that speakers can modulate their choice between alternating constructions depending on both referential properties (e.g., animacy and the presence of DOM) and perceptual modality (visual vs. auditory), although the latter plays a lesser role than anticipated.

From a broader theoretical perspective, these results lend further support to the view that agreement phenomena reflect the interface between grammatical form and conceptual construal. The fact that both agreement patterns are accepted by native speakers indicates that multiple constructions coexist within the grammar of Spanish. At the same time, the divergence between corpus and experimental outcomes underscores the importance of combining production-based and rating-based methodologies to capture the full range of grammatical variation.

For future research, it will be crucial to complement these acceptability data with production experiments that elicit spontaneous speech. Such studies would clarify how speakers actually construct and select agreement patterns in real—but controlled—communicative situations, and how contextual or discourse factors influence their choices. Furthermore, new experimental paradigms could be designed to further test the role of salience in contexts (a) where DOM cannot interfere, thus only with inanimate NP₂s, for instance, (b) by introducing adjectival modifiers that contribute specific perceptual-conceptual properties (for instance, indicating brightness, movement, dynamicity, etc.), or (c) by systematically testing animate referents across both agreement types and across conditions with and without DOM, including configurations that run counter to the grammatical norm, such as *se oyen/ven jugar a los niños* or *se oye/ve jugar los niños*. These controlled manipulations would make it possible to test whether the link between salience and agreement persists independently of the grammatical marking associated with DOM.

Author Contributions: Conceptualization, R.E. and M.B.; methodology, R.E. and M.B.; software, R.E. and M.B.; validation, R.E. and M.B.; formal analysis, R.E. and M.B.; investigation, R.E. and M.B.; resources, R.E. and M.B.; data curation, R.E. and M.B.; writing—original draft preparation, R.E. and M.B.; writing—review and editing, R.E. and M.B.; visualization, R.E. and M.B.; supervision, R.E.; project administration, R.E. and M.B.; funding acquisition, R.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Special Research Fund of Ghent University, grant number BOF19/GOA/013.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the Faculty of Psychology and Educational Sciences at Ghent University (protocol code 2021/30, 13 April 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All materials, data, and analysis scripts are publicly available on OSF (<https://osf.io/m4r6f/>, accessed on 19 December 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

1	first person
2	second person
3	third person
AR	acceptability rating
CDE	Corpus del Español
CREA	Corpus de Referencia del Español Actual
DO	direct object
DOM	differential object marking

Inf	infinitive
NP	noun phrase
NP2	(subordinate) second noun phrase
PL	plural number
SG	singular number
PV	perception verb
VP	verb phrase

Appendix A

Appendix A.1

Table A1. Extralinguistic variables characterizing the 100 participants who were included in the analysis (based on their responses to the sociodemographic questionnaire).

Variable		Levels	Number Participants
Gender		female	38
		male	62
		other	-
Age	(median split)	≤26 y.o.	51
		>26 y.o.	49
Education		secondary school	4
		high school	18
		university	65
		vocational training	12
		other	1
Mobility	(lived in another province/abroad for >1 year)	yes	41
		no	59
Active L2	(high proficiency, used daily)	yes	53
		no	47
Many L2s	(min. two L2s, any proficiency, any frequency of use)	yes	47
		no	53
Reading habits	(books per year)	no books	8
		1–2 books	21
		3–5 books	26
		6–12 books	26
		13–25 books	13
		26–50 books	6
		>50 books	-
Social media use	(hours per day)	<15 min.	10
		15–30 min.	15
		30–60 min.	23
		1–2 h	21
		2–3 h	14
		3–4 h	10
		>4 h	7

Appendix A.2

Table A2. Full output of the cumulative link mixed model (including random effects and their correlation), showing estimates on the log-odds scale.

Fixed Effects ¹	Estimate	SE	z-Value	p-Value
Modality	−0.03	0.33	−0.08	0.936
Agreement	0.33	0.33	0.98	0.328

Table A2. Cont.

Fixed Effects ¹	Estimate	SE	z-Value	p-Value
Animacy	−1.07	0.30	−3.61	<0.001
SentenceLength	0.17	0.18	0.98	0.325
Modality * Agreement	−0.40	0.57	−0.69	0.490
Modality * Animacy	−2.15	0.58	−3.72	<0.001
Agreement * Animacy	−2.45	0.63	−3.87	<0.001
Modality * Agreement * Animacy	2.61	1.15	2.26	0.024
Random effects participant	Var	SD	Corr	
(Intercept)	2.86	1.69		
Modality (visual)	1.46	1.21	−0.35	
Agreement (singular)	2.99	1.73	−0.61	0.45
Animacy (inanimate)	2.26	1.50	−0.05	0.62
Modality (visual) * Agreement (singular)	0.48	0.69	0.38	−0.93
Modality (visual) * Animacy (inanimate)	2.16	1.47	0.01	−0.61
Agreement (singular) * Animacy (inanimate)	5.58	2.36	0.25	−0.59
Modality (v.) * Agreement (sg.) * Animacy (inan.)	2.17	1.47	0.20	0.29
Random effects item	Var	SD		
(Intercept)	0.76	0.87		

¹ Number of observations: 4000, number of groups: participant 100, item 40. Formula in R: clmm (AR~Modality * Agreement * Animacy + SentenceLength + (1|item) + (1 + Modality * Agreement * Animacy|participant)).

Notes

- ¹ In the absence of an explicit citation, the examples provided have been created by the authors for illustrative purposes.
- ² Particularly noteworthy, and relevant for this study, is that the DO is clearly syntactically marked by Differential Object Marking (DOM) through the preposition *a*. This construction may illustrate how formal cues of objecthood can interact with the agreement system, challenging the traditional boundary between subject and object control in Spanish.
- ³ Psycholinguistic models of agreement production differ in their assumptions about processing flow. In feedforward models, agreement is generated in a linear, top-down sequence, that is, from subject selection to verb inflection, without subsequent monitoring, which explains attraction effects when distractors intervene. In contrast, feedback models assume continuous interaction between grammatical and conceptual levels, allowing for revision of the verb form when semantic interference or attraction is detected (cf. Bock & Miller, 1991). This second model seems to be more compatible with the empirical data, especially in cases where semantic or even pragmatic factors come into play, as is often the case in the types of non-standard agreement discussed in this paper.
- ⁴ A chi-square test confirms that this difference is statistically significant ($\chi^2 = 9.26$, $p < 0.001$), with Cramer's *V* indicating a medium effect size (= 0.22).
- ⁵ As González-Vilbazo et al. (2013) point out, such tasks provide a more accurate view of speakers' internal competence or their I-language.
- ⁶ Their sociodemographic characteristics are presented in Table A1 included in Appendix A.1.
- ⁷ In the corpus data, the infinitive construction appears only marginally with the voluntary perception verbs *mirar* 'look' and *escuchar* 'listen'. For this reason, only the involuntary ones were included in the experimental materials.

References

- Acuña-Fariña, J. C., Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, 143, 108–128. [CrossRef]
- Avendaño, C. S. (2007). "Para que la gente se enteren": La concordancia *ad sensum* en español oral. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 33(2), 205–226.
- Baltais, M., Van Hulle, S., & Hartsuiker, R. J. (2026). *Corpus-based productivity informs acceptability ratings: Evidence from Spanish inchoatives*. Ghent University. Manuscript submitted for publication.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. [CrossRef] [PubMed]
- Bello, A. (1964). *Gramática de la lengua castellana destinada al uso de los americanos* (A. Alonso, Ed.). Losada.

- Bock, K., & Miller, C. (1991). Broken Agreement. *Cognitive Psychology*, 23, 45–93. [CrossRef]
- Bolinger, D. (1968). *Aspects of language*. Harcourt, Brace & World.
- Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Kraemer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69–94). Royal Netherlands Academy of Arts and Sciences.
- Christensen, R. H. B. (2022). *Ordinal—Regression models for ordinal data*. R package version 2022.11-16. Available online: <https://CRAN.R-project.org/package=ordinal> (accessed on 19 December 2025).
- Christensen, R. H. B., & Brockhoff, P. B. (2013). Analysis of sensory ratings data with cumulative link models. *Journal de La Société Française de Statistique & Revue de Statistique Appliquée*, 154(3), 58–79.
- Claes, J. (2015). Competing constructions: The pluralization of presentational haver in Dominican Spanish. *Cognitive Linguistics*, 26(1), 1–30. [CrossRef]
- Corbett, G. (2003). Agreement: Terms and boundaries. In W. E. Griffin (Ed.), *The role of agreement in natural language: TLS 5 proceeding* (pp. 109–122). Texas Linguistic Forum.
- Corbett, G. (2006). *Agreement*. Cambridge University Press.
- Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.
- Del Rey Quesada, S. (2022). Accusativus cum Infinitivo y otras construcciones de infinitivo latinizante: Caracterización sintáctica y uso en la literatura erasmiana doctrinal del siglo XVI. *Zeitschrift für romanische Philologie*, 138(2), 483–505. [CrossRef]
- Duranti, A. (1997). *Linguistic anthropology*. Cambridge University Press.
- Enghels, R. (2007). *Les modalités de perception visuelle et auditive: Différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français*. Max Niemeyer.
- Enghels, R. (2019). Linguistic reflections of object perception versus event perception. Agreement ad sensum in the pronominal infinitive construction in Spanish. *Syntaxe et Sémantique*, 20, 107–123. [CrossRef]
- Evans, N., & Wilkins, D. (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language*, 76(3), 546–592. [CrossRef]
- Featherston, S. (2005). The decathlon model of empirical syntax. In S. Kepser, & M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives* (pp. 187–208). De Gruyter.
- Felser, C. (1999). *Verbal complement clauses: A minimalist study of direct perception constructions*. John Benjamins.
- Gili Gaya, S. (1980). *Curso superior de sintaxis española*. Vox.
- Givón, T. (2001). *Syntax: An introduction*. John Benjamins.
- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.
- González-Vilbazo, K., Bartlett, L., Downey, S., Ebert, S., Heil, J., Hoot, B., Koronkiewicz, B., & Ramos, S. (2013). Methodological considerations in code-switching research. *Studies in Hispanic and Lusophone Linguistics*, 6, 119–138. [CrossRef]
- Gómez Torrego, L. (1992). *La impersonalidad gramatical: Descripción y norma*. Arco/Libros.
- Granvik, A., Hatakka, V., Silvennoinen, O., Erkkilä, R., & Mäntylä, E. (2025). Beyond corpus data—Complementary and alternative methods in cognitive linguistics. *Review of Cognitive Linguistics*, 23(2), 327–344. [CrossRef]
- Grilli, L., & Rampichini, C. (2011). Multilevel models for ordinal data. In R. S. Kenett, & S. Salini (Eds.), *Modern analysis of customer surveys: With applications using R* (pp. 391–411). John Wiley & Sons.
- Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject–verb agreement. *Journal of Memory and Language*, 48(4), 760–778. [CrossRef]
- Hilpert, M. (2014). *Construction grammar and its application to English*. Edinburgh University Press.
- Ibarretxe-Antuñano, I. (1999). *Polysemy and metaphor in perception verbs: A cross-linguistic study* [Ph.D. thesis, University of Edinburgh].
- Jackendoff, R. (1983). *Semantics and cognition*. MIT Press.
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1217–1218. [CrossRef] [PubMed]
- Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume I: Theoretical prerequisites*. Stanford University Press.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.
- Lenth, R. V. (2022). *Emmeans: Estimated marginal means, aka least-squares means*. R package version 1.8.1-1. Available online: <https://CRAN.R-project.org/package=emmeans> (accessed on 19 December 2025).
- Maldonado, R. (1999). *A media voz: Problemas conceptuales del clítico se (en español)*. Universidad Nacional Autónoma de México.
- Mallinson, G., & Blake, B. J. (1981). *Language typology: Cross-linguistic studies in syntax*. North-Holland.
- Mare, M., & Pato, E. (2018). *Parecen que lo olvidan...* Hyper-agreement in non-standard Spanish. *Borealis: An International Journal of Hispanic Linguistics*, 7, 71–95.
- Martínez, J. A. (1999). La concordancia. In I. Bosque, & V. Demonte (Eds.), *Gramática descriptiva de la Lengua Española* (pp. 2695–2786). Espasa Calpe.
- Melis, C. (2022). Alignment changes with Spanish experiential verbs. In E. Dahl (Ed.), *Alignment and alignment change in the Indo-European family* (pp. 246–276). Oxford University Press.

- Mendikoetxea, A. (1999). Construcciones con *se*: Medias, pasivas e impersonales. In I. Bosque, & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 1631–1722). Espasa Calpe.
- Miller, G., & Johnson-Laird, P. (1976). *Language and perception*. Cambridge University Press.
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27. [CrossRef]
- RAE & ASALE. (2009). *Nueva gramática de la lengua española*. Espasa.
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 19 December 2025).
- Rodríguez Espiñeira, M. J. (2000). Percepción directa e indirecta en español. Diferencias semánticas y formales. *Verba*, 27, 33–85.
- Roegiest, E., & Enghels, R. (2008). La Reducción Oracional En La Construcción Factitiva Española. In H. J. Döhla, R. Montero Muñoz, & F. Báez de Aguilar (Eds.), *Lenguas en diálogo: El Iberorromance y su diversidad lingüística y literaria: Ensayos en homenaje a georg bossong* (pp. 289–312). Iberoamericana.
- Sánchez, M. E., Jaichenco, V., & Sevilla, Y. (2014). Errores de concordancia sujeto-verbo en la producción de oraciones en español: El papel de la distancia lineal y de los modificadores. *PSIENCIA. Revista Latinoamericana de Ciencia Psicológica*, 6(2), 55–63. [CrossRef]
- Sánchez, M. E., Sevilla, Y. A., & Jaichenco, V. I. (2013). Interferencias en la producción de la concordancia sujeto-verbo en el español. Un estudio sobre el rol de los factores semánticos y morfofonológicos. *Revista Argentina de Ciencias del Comportamiento*, 3, 15–23. [CrossRef]
- Sánchez López, C. (2002). Las construcciones con *se*: Estado de la cuestión. In C. Sánchez López (Ed.), *Las construcciones con se* (pp. 13–163). Visor Libros.
- Schütze, C. T. (2016). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Language Science Press.
- Soto, C. J., & John, O. P. (2017). The next big five inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117–143. [CrossRef]
- Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press.
- Talmy, L. (2000). *Toward a cognitive semantics. Volume I: Concept structuring systems*. MIT Press.
- Torrego Salcedo, E. (1999). El complemento directo preposicional. In I. Bosque, & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 1779–1805). Espasa Calpe.
- Viberg, Å. (1984). The verbs of perception: A typological study. In B. Butterworth, B. Comrie, & Ö. Dahl (Eds.), *Explanations for language universals* (pp. 123–162). De Gruyter.
- Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61(3), 261–298. [CrossRef]
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.