

# An exploration of neuter determiner 'lo' and 'lo que' constructions in Spanish

Andrea Levinstein Rodriguez

University of Calgary

## Abstract

*This paper presents an exploratory overview of the syntactic properties of 'lo', a polysemic definite determiner that is standardly characterized as having a 'neuter' grammatical gender. I argue that 'lo' is better characterized as a referential pronoun lacking  $\varphi$  features, which refers to non-individuated entities. In addition to this referential 'lo', and following Zulaica Hernandez (2018), I argue that there is a non-referential, expletive 'lo', which relates a possible world to a predicate which may contain morphologically expressed  $\varphi$  features. However, the distribution and referentiality of 'lo' in 'lo que' constructions (prepositional clauses where 'lo' can be external or internal to the CP; see Plann, 1980) indicate that whether 'lo' is referential or expletive is not a matter of syntactic position, contrary to Zulaica Hernandez (2018). Referential 'lo' stands in contrast with most structurally defective pronouns, such as impersonal pronouns and expletives, in that it is a definite pronoun with a specific referent. This goes against the standard assumption that a nominal phrase that has properties associated with 'higher' layers of structure, such as definiteness (associated with DP), must therefore also have 'lower' layers like NumP and nP. Although this paper does not present a solution, it does point out that standard theory cannot account for it and argues for the need to continue developing our understanding of nominal structure.*

Keywords: Spanish, neuter 'lo', neuter pronouns, nominal structure, syntax

## 1. Introduction

Although Spanish has a two-gender grammatical system, there are a number of linguistic objects that have been traditionally classified as neuter, i.e., neither masculine nor feminine. These are: definite determiner *lo*, pronoun *ello*<sup>1</sup> and demonstratives *eso*, *esto* and *aquello* (proximal, medial and distal, respectively). These pronominal-like objects refer to, broadly speaking, ‘non-nominals’ – events, propositions, properties, etc. For example, *lo* in (1) refers to an utterance.

- 1) *Lo que me dijiste me dolió*<sup>2</sup>  
 LO that ACC.1SG said.2SG ACC.1SG hurt.3SG  
 ‘What you said to me hurt me’

*Lo*, in particular, has been the subject of much debate and attention in the Spanish literature for its highly polysemic nature and broad distribution. Although the semantic and syntactic properties of *lo* are by no means under-researched, the majority of this research has centred *lo* in a few specific contexts, while others have been largely ignored.

This paper investigates the semantic and syntactic properties of *lo* by centring what Susan Plann (1980) called *lo que* constructions, a term that encompasses two different types of relative clauses. Specifically, it explores how the *lo* in these constructions presents a challenge to previous typologies of *lo*. At the same time, it investigates how so-called neuter determiners in general, but particularly *lo*, present a challenge to existing conceptions of nominal structure.

The rest of this section introduces *lo que* constructions as they were originally defined by Plann (1980). Section 2 acts as a literature review by exploring the question that has been at the center of *lo* literature, that of its categorization. Section 3 introduces Stark & Pomino’s (2010) historical account of the semantic features of *lo*, and specifically its relationship with  $\varphi$  features, in the context of feature geometry. Section 4 provides an analysis and critique of Zulaica Hernandez (2018), a semantic analysis which differentiates two major types of *lo* linguistic objects based on their referentiality. Section 5 explores why the semantic properties of *lo*, as described in the previous sections, make troublesome predictions on the nature of its nominal structure. Section 6 is the conclusion.

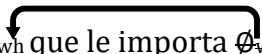
### *Lo que* constructions

Plann (1980) characterized *lo* as a definite article that can only select null ( $\emptyset$ ) neuter nouns. She identified two relative clause constructions, both with the surface string *lo que* at the edge of the relative clause but with different underlying structures. In the bare *lo que* construction, this article (henceforth ‘antecedent *lo*’) is part of the relative clause’s

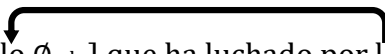
<sup>1</sup> The use of *ello* is considered archaic nowadays, having been replaced by *eso* and by a null pro-form (Zulaica Hernandez, 2018).

<sup>2</sup> Unless otherwise specified, examples are my own.

antecedent (2a). As diagrammed in (2b), this antecedent is base-generated in the matrix clause, where it selects a null noun modified by the relative clause.

- 2) a. *Luchó por lo<sub>i</sub>  $\emptyset_i$  que le importa*  
 Fought.1SG for LO that ACC.3SG matter.SG  
 ‘She fought for what matters to her’
- b. [PP por [NP lo  $\emptyset$  [CP  $\emptyset_{wh}$  que le importa  $\emptyset_{wh}$  ] ] ]
- 

In the *P lo que* construction, the article (henceforth ‘wh-internal *lo*’) selects a null wh-word<sup>3</sup> within the relative clause, which modifies an antecedent with matching (lack of)  $\varphi$  features (3a). As diagrammed in (3b), *lo* is base-generated within the embedded clause and is pied-piped with the rest of the extended nominal projection to Spec, CP.

- 3) a. ... *Ha conseguido aquello<sub>i</sub> por lo<sub>i</sub> que ha luchado*  
 have.3SG achieved DISTAL.N<sup>4</sup> for LO that have.3SG fought  
 ‘She has achieved that **which** she has fought for.’
- (Retrieved from *Corpus del Español* (Davies, 2004))
- b. [NP aquello [CP [PP por lo  $\emptyset_{wh}$  ] que ha luchado por lo  $\emptyset_{wh}$  ] ] ]
- 

Plann (1980) was using a Transformational Grammar approach, which posited syntactic movement as a series of transformations between underlying and surface structure. For this reason, the focus of her research is on demonstrating that the ‘antecedent *lo*’ construction has the structure in (4a) and the ‘wh-internal *lo*’ construction has the structure in (4b), as well as on identifying the transformations that result in the surface order. The properties of *lo* in-and-of-itself did not receive much attention.

- 4) a. [NP lo  $\emptyset$  [CP que... ] ]
- b. [CP [PP P [NP lo  $\emptyset_{wh}$  ] ] que... ]

Nonetheless, by identifying *lo* as an article, Plann (1980) participates in the central debate surrounding *lo*: the categorization question. The following section gives an overview of this debate and of the specific *lo* constructions at the heart of it.

<sup>3</sup> Although nowadays the standard assumption is that wh-phrases are DPs, Plann (1980) assumes that they are nouns which may take an article as Specifier. This distinction is not relevant to the present paper. Suffice to say that *lo*, as well as the Prepositional Phrase, are part of the extended projection of the wh-phrase.

<sup>4</sup> As I show in Section 3, ‘neuter’ is not a descriptively adequate label to describe pronominals that do not express  $\varphi$  features. Nonetheless, demonstratives of this type will be glossed as N (‘neuter’) for simplicity’s sake.

## 2. The category of *lo*<sup>5</sup>

The syntactic literature on *lo* centers around the question of how to best categorize *lo* and, relatedly, whether *lo* is a single polysemic linguistic object or multiple objects (each belonging in a different category). This section presents a brief overview of some of the most influential attempts to answer these questions. Section 2.1 presents the ‘main’ categorization question, i.e., whether *lo* is an article or a pronoun. Section 2.2 presents approaches that explore or contest the possibility of *lo* being anything other than a determiner in a given context.

### 2.1. Article or pronoun?

The main overarching question of this debate is whether *lo* is an article or a pronoun. At first glance, this might be considered a non-issue, since a linguistic object defined as *a definite article that selects an inherently null noun* is not meaningfully or functionally different from a linguistic object defined as *a personal pronoun*. However, the bulk of this debate took place before or around the introduction of the DP Hypothesis (Abney, 1987). Therefore, the issue underlying this debate is the structural importance of *lo*, i.e., whether it is the head of the nominal domain or its specifier.

The prominent grammarian Andres Bello is credited as the first person to attempt to categorize *lo* (Luján, 2004). He claimed that *lo* was the ‘weak’ (i.e., clitic) version of the ‘neuter’ pronoun *ello*, which should be used when the pronoun is followed by a modifier (Bello, 1847, as cited in Luján, 2004). By contrast, other early analyses of *lo* categorized it as an article, which selected either a null noun or a modifier that has been ‘nominalized’ (see Contreras, 1973, section 4). This ‘article analysis’ seems to have been the standard assumption, the one made by researchers for whom the broad categorization of *lo* was not a main concern (e.g., Contreras, 1973; Plann, 1980).

Recent analyses of *lo* (Hamalainen, 2004; Stark & Pomino, 2010; Zulaica Hernandez, 2018) present the ‘article vs. pronoun’ debate as an active and ongoing polemic. They each classify *lo* as a pronoun: Hamalainen (2004) following Bosque & Moreno (1990); Stark & Pomino (2010) and Zulaica Hernandez (2018) for independent semantic reasons<sup>6</sup>. However, I have not encountered any arguments in favour of the ‘article’ categorization that are dated any later than the 1980s. This suggests that a consensus has been reached in favour of the ‘pronoun’ categorization. Nonetheless, authoritative dictionary sources (e.g., Real Academia Española, 2021; Diccionario del Español de México, 2021) categorize *lo* as an article in most

<sup>5</sup> This section is unique in that most of the sources referenced here were written in Spanish. When relevant, translations and glosses are my own.

<sup>6</sup> Stark & Pomino (2010) claim to side-step the debate, but operate under the assumption that the pronominal system of Latin (and subsequently, Spanish) is based on its semantic feature geometry, and analyse *lo* as part of this paradigm. Zulaica Hernandez follows Roberts (2003), according to whom both DPs with definite articles and pronouns are the same in that they both require a unique referent that is ‘weakly familiar’ (implicitly available in the common ground). The difference between both is a matter of relative salience – the pronoun must refer to the most salient possible referent.

contexts, which might give the impression that the ‘article’ side of the debate is more widespread than it actually is.

## 2.2. One *lo* or multiple *lo*?

The other side of the categorization debate concerns whether the *lo* found in any specific context should be characterized as a determiner or as something else, like an adverb or a degree word. This debate is centred around pairs such as (5-6). In (5), the adjective *simpático* ‘friendly’ shows default gender and number morphology, as is typical for adjectives selected by *lo* (see Section 3). In (6), the adjective *simpáticas* shows number and gender concord with the predicate *estas chicas* ‘these girls’. The *lo* in sentences like (6), which has an ‘intensifier’ interpretation, is sometimes characterized as an adverb or quantifier. However, authors like Contreras (1973), Gutiérrez-Rexarch (1999) and Bosque & Moreno (1990) have attempted to fold it into the nominal paradigm<sup>7</sup>.

- 5) *Me sorprende lo simpático / \*simpáticas de estas chicas*  
 ACC.1SG surprise.3SG LO friendly.M.SG friendly.F.PL of MED.F.PL girls  
 ‘I am surprised by the friendliness of these girls.’
- 6) *Me sorprende lo simpático / simpáticas que son estas chicas*  
 ACC.1SG surprise.3SG LO friendly.M.SG friendly.F.PL that be.3PL MED.F.PL girls  
 ‘I am surprised by how friendly these girls are.’

(Contreras, 1973, pp. 20-21)

Contreras’s (1973) analysis differentiates between the ‘non-anaphoric *lo*’ in the sentences above and the ‘anaphoric *lo*’ that acts as a pronominal object and pro-predicate. She claims that the latter is a pronoun, while the former is a definite article that ‘fuses’ with a null pronoun. The difference between (5) and (6) is explained as the result of timing differences in the transformation of each sentence. Gutiérrez-Rexach (1999), following Kayne’s (1994) anti-symmetry framework, argues that the ‘degree’ interpretation comes from a null degree operator located in Spec, DP. This operator is specified for the same  $\phi$  features as the predicate, triggering adjective agreement.

Bosque & Moreno (1990) present the most influential analysis of the categorization of *lo*, in which they argue for a unified analysis. They characterize *lo* as a variable pronoun whose range and denotation are valued by its right-branching modifiers. These modifiers are what trigger the adverb-like interpretation of (6), or the quantifier-like interpretation in (9) below. Bosque and Moreno place themselves in opposition to the ‘article’ analysis that presents these modifiers as being nominalized. They argue that, if this was the case, *lo* would have a wider distribution than it does<sup>8</sup>.

<sup>7</sup> The properties of *lo* in these structures is beyond the scope of this paper, but it will be briefly addressed in Section 4.

<sup>8</sup> Specifically, *lo* can only be modified by prepositional phrases headed by the preposition *de* ‘of, about’. If *lo* could nominalize prepositional phrases, there is no principled reason why it would make a distinction between this phrase and others.

Bosque & Moreno (1990) identify three *lo* types, differentiated by their denotations. The Individuating type denotes a set, where *lo* refers to everything to which the set applies. For example, in (1), reproduced as (7) below, *lo* refers to an utterance because the relative clause restricts its potential referents to utterances, as only utterances can be ‘said’. This category also contains ‘nominalized’ adjectives such as (8), where the potential referents are anything that could be described using the adjective, including objects. Individuating *lo* is used as the ‘elsewhere’ category by authors who adopt it, e.g., Hamalainen (2004).

- 7) *Lo que me dijiste me dolió*  
 LO that ACC.1SG said.2SG ACC.1SG hurt.3SG  
 ‘What you said to me hurt me.’

- 8) *Me encanta lo auténtico*  
 ACC.1SG love. SG LO authentic  
 ‘I love that which is authentic.’

(Hamalainen, 2004, p. 194)

The Qualitative type denotes the maximal expression of a quality or circumstance. This is the type in (6) above, where *lo* denotes the maximal amount of the property of ‘friendliness’. The final type is the Quantitative type, e.g. (9), where the modifier of *lo* values an interpretation that roughly translates to ‘the proper amount’.

- 9) *No duermo lo suficiente*  
 Not sleep.1SG LO sufficient  
 ‘I don’t sleep enough.’

(Bosque & Moreno, 1990, p. 32)

Although it purports to be a unified account, this typology centres around what can be termed ‘*lo* + modifier’ constructions. It does not account for those cases where *lo* does not require a modifier to receive an interpretation. In section 4, we will bring the focus back to *lo que* constructions and briefly introduce other *lo* types that have been ignored in most of the literature, showing how the narrow focus of Bosque & Moreno (1990) prevent them from achieving a truly unified account of *lo*. Before that, though, we must look at the most distinctive characteristic of *lo*: its complicated relationship with  $\varphi$  features.

### 3. The $\varphi$ features of *lo*

One of the defining characteristics of *lo* is non-concord. *Lo* can only select adjectives and modifiers with default  $\varphi$  feature morphology (10a) or with no  $\varphi$  features (10b). This has led to a traditional characterization of *lo* and other  $\varphi$ -rejecting pronouns<sup>9</sup> as ‘neuter’, i.e., as having a grammatical gender distinct from ‘masculine’ and ‘feminine’.

<sup>9</sup> Archaic ‘strong’ pronoun *ello* and demonstratives *eso*, *esto*, *aquello*.

- 10) a. *Lo rojo / \*roja / \*rojos / \*rojas*  
 LO red red.SG.F red.PL.M red.PL.F  
 'That which is red.'
- b. *Lo que pasó*  
 LO that happened.3SG  
 'That which happened.'

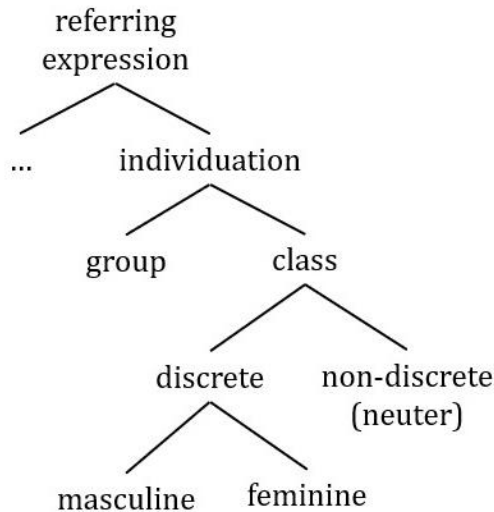
There are a few issues with this characterization, the most pressing one being that there are no neuter nouns in Spanish. The grammatical gender of pronouns with inanimate referents is usually determined by the  $\phi$  features of their referent (Hualde et al., 2012), which begs the question of where *lo* could be getting the neuter feature from<sup>10</sup>. Moreover, there is no principled reason for neuter gender to exclude plural morphology in Spanish, further suggesting that *lo* is not grammatically neuter, but devoid of  $\phi$  features.

Stark & Pomino (2010) claim that not only is 'neuter' an incorrect characterization of *lo*, but that the semantic features that have been used in the literature to describe it (e.g., [+abstract], [+proposition], [-animate]) fail to capture the semantic characteristics of the potential referents of *lo*. Instead, they characterize *lo* as a marker of semantic non-individuation. The semantic feature [individuation] refers to the property of being a discrete entity or set. In other words, *lo* cannot refer to individuals, only to 'undifferentiated categories' and things that cannot be individuated, like propositions. Crucially, semantic individuation is distinct from syntactic individuation, i.e., the implementation of mass-count distinctions. Mass nouns in Spanish are specified for grammatical gender and cannot serve as referents for *lo* any more than count nouns can. Mass nouns, while being uncountable, are not truly non-individuated, as they refer to a set (Pomino & Stark, 2009).

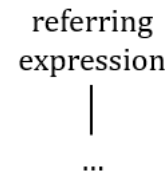
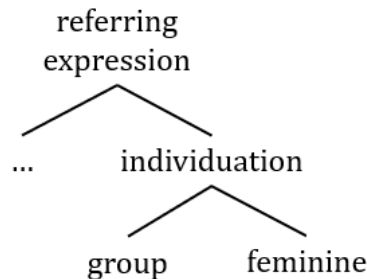
According to Stark & Pomino (2010), non-individuated pronouns are the result of the consolidation of various semantic categories that were morphologically expressed in Latin, Spanish's parent language. They claim that in Latin, grammatical gender is dependent on the expression of the feature [discrete], which encodes mass/count distinctions. The Latin neuter was associated with non-discreteness, as illustrated in (11)<sup>11</sup>.

<sup>10</sup> If we characterize *lo* as an article that selects a null noun, we could say that the null noun is neuter. However, to postulate a null neuter noun when no overt null nouns exist seems like an *ad-hoc* solution.

<sup>11</sup> Stark and Pomino replace the feature [animate] in Harley & Ritter (2002) with the feature [discrete]. The authors motivate this choice by claiming that the gender features [masculine] and [feminine] could not have been dependent on [animate], since Latin had plenty of inanimate nouns that were lexically specified for masculine and feminine gender. This choice does not affect their representation of the Spanish pronominal system or of *lo* as a marker of non-individuation; that node of the tree, regardless of its contents, was reduced into the individuation/non-individuation distinction.

11) Feature geometry of Latin 3<sup>rd</sup> person pronouns

Over time, semantic features like [class] and [discrete] disappeared from the pronominal system, and gender features became directly associated with the mother node [individuation] (12a). Since non-individuating referring expressions do not have the individuation feature (12b), they cannot have number [group] and gender [feminine] features, triggering the default morphological form *lo*.

12) a. Feature geometry of Spanish 3<sup>rd</sup> person pronouns      b. Feature geometry of *lo*

Stark & Pomino's (2010) proposal provides a straightforward distinction between what can be a potential referent of *lo* (individuals) and what cannot (non-individuals). This distinction provides a principled account for why *lo* lacks  $\varphi$  features without needing to invoke an otherwise unattested neuter grammatical gender. However, it does not address those cases where *lo* seems to be modified by an adjective that expresses number or gender morphology, or where it seems to refer to an individual. These cases are explored in the following section.

4. The referential properties of *lo*

Zulaica Hernandez (2018) presents a comprehensive semantic analysis of the 'neuter' pronominal system, i.e., *lo*, demonstratives *eso*, *esto* and *aquello*, and a null pro-form which

he claims is in complementary distribution with *lo*. He identifies two types of *lo*: referential and denotational. Referential *lo* is a pronoun that can only refer to non-individuals, in line with Stark & Pomino (2010). Denotational *lo* is not a true pronoun in that it has no referent. Rather, it is an expletive mostly associated with pseudo-cleft constructions.

The first two sections introduce referential *lo* (4.1) and denotational *lo* (4.2), broadly classifying ‘antecedent *lo*’ and ‘wh-internal *lo*’ in this typology. Section 4.3 shows how Zulaica Hernandez (2018) fails to account for the distribution of *lo que* constructions, as these constructions seem to be able to accommodate both referential and denotational *lo*.

#### 4.1. Referential *lo*

Referential pronouns refer to the most salient antecedent available the context of the utterance. This antecedent does not necessarily need to be in the discourse, but only implicitly available in the common ground (Roberts, 2003). Zulaica Hernandez identifies referential pronoun *lo* as appearing in the following contexts: as an accusative pronoun<sup>12</sup> (13), a relative pronoun (14), and a pro-predicate (15). These pronominals can only refer to non-individuals. For example, in (13), the referent is a proposition; in (14), it is an event; and in (15), it is a property.

- 13) *Juan llegará tarde, ya te lo había dicho*  
 Juan arrive.FUT.3SG late, already ACC.2SG LO had told.3SG  
 ‘Juan will arrive late, I had already told you **so**.’

(Modified from Zulaica Hernandez, 2018, p. 21)

- 14) *Juan nunca llegó, lo cual me sorprendió*  
 Juan never arrived.3SG LO which ACC.1SG surprised.3SG  
 ‘Juan never arrived, **which** surprised me.’

- 15) *Juan es muy descuidado, pero yo no lo soy*  
 Juan be.3SG very careless but I not LO be.1SG  
 ‘Juan is very careless, but I am not \_\_.’

While *lo que* constructions are not specifically named, ‘wh-internal *lo*’ can be considered as part of the ‘relative pronoun’ category<sup>13</sup>. For example, the referent in (16) is an outcome of some sort, presumably an action or change-of-state. The situation with ‘antecedent *lo*’ is more complicated. The introduction to the paper asserts that the *lo* in ‘free relative’ constructions<sup>14</sup> is denotational. However, the body of the text and all relevant examples refer only to pseudo-cleft constructions (see section 4.2). This leaves the statute of ‘antecedent *lo*’ in non-copular constructions somewhat ambiguous, as I will address in section 4.3.

<sup>12</sup> Not to be confused with 3SG.M accusative pronoun *lo*, which refers to masculine nouns.

<sup>13</sup> ‘Relative pronoun’ is a broad term for the wh-phrase in the context of relative clauses (Brucart, 2016).

<sup>14</sup> Relative clauses whose antecedent is either null or a simple determiner, a category which includes ‘antecedent *lo*’ constructions. Not to be confused with the ‘free relatives’ in English that have a null complementizer and wh-phrase.

- 16) ... *Ha conseguido aquello por lo que ha luchado* ...  
 have.3SG achieved DISTAL,N for LO that have.3 SG fought  
 ‘She has achieved that **which** she has fought for.’

(Retrieved from *Corpus del Español* (Davies, 2004))

#### 4.2. Denotational *lo*

Zulaica Hernandez (2018) highlights *lo*’s ability to ‘co-refer’ with an individuated predicate when it is the subject of a pseudo-cleft sentence such as (17). Contrary to the predictions in Stark & Pomino (2010), the DP *una silla* ‘a chair’ refers to a concrete individuated object with lexically specified  $\varphi$ -features. Zulaica Hernandez claims that this is not a case of true reference, as the *lo* in these sentences is a denotational pronoun. This denotational pronoun is not a pronoun *per se*, but a function that maps a possible world to the predicate argument of the copula. In example (17) below, *lo* maps the predicate *una silla* to the verb ‘to need’.

- 17) *Lo que necesito es una silla*  
 LO that need.1SG be.3SG a.F.SG chair  
 ‘What I need is a chair.’

The author argues that most of the constructions commonly studied in the literature are this denotational *lo*, especially the ones in Bosque & Moreno’s (1990) ‘individuating *lo*’ type<sup>15</sup>. While Zulaica Hernandez only illustrates this point using pseudo-cleft constructions, it is easy to see how this analysis can be extended to sentences like (8), reproduced below in (18). In these sentences, the ‘potential referent’ of *lo* might include individuated objects (e.g., specific cultural artifacts). The fact that these constructions are best thought of as having a ‘potential’ rather than an ‘actual’ referent suggests that the *lo* in these sentences might indeed be a denotational pronoun<sup>16</sup>. Denotational *lo* might also account for the presence of number and gender concord in ‘qualitative *lo*’ sentences such as (6), reproduced in (19). *Lo* in these cases might be mapping the predicate *estas chicas* ‘these girls’ to the possible world of ‘(being) friendly’, triggering concord.

- 18) *Me sorprende lo simpáticas que son estas chicas*  
 ACC.1SG surprise.3SG LO friendly.F.PL that be.3PL MED.F.PL girls  
 ‘I am surprised by how friendly these girls are.’
- 19) *Me encanta lo auténtico*  
 ACC.1SG love.SG LO authentic  
 ‘I love that which is authentic.’

<sup>15</sup> If we accept the premise of Stark & Pomino (2010), that *lo* is a non-individuation marker, then ‘individuating *lo*’ becomes an oxymoron.

<sup>16</sup> Zulaica Hernandez (2018) does not elaborate on how the function mapping would work outside of the context of pseudo-cleft constructions. While that question is beyond the scope of this paper, I speculate that phrases like *lo auténtico* might have an underlying small clause structure equating the predicate property with the *lo* subject.

Since ‘denotational *lo*’ can exist outside of pseudo-cleft constructions, this begs the question: Can ‘wh-internal *lo*’ be denotational? I address this issue in the following section.

### 4.3. Referential ‘antecedent *lo*’ and denotational ‘wh-internal *lo*’

Zulaica Hernandez (2018) draws a sharp distinction between the environments in which *lo* is denotational and the ones in which it is referential. Specifically, he claims that ‘referential *lo*’, being a clitic pronoun, can never be in subject position. This seems to indicate that ‘antecedent *lo*’, which can appear in subject position, is a denotational pronoun, while ‘wh-internal *lo*’, which cannot, is a referential pronoun. However, this proposed distribution falls apart when we examine the *lo que* constructions in more detail.

Let us first examine the case of referential ‘antecedent *lo*’, which, I argue, is found in sentences like (20a). The *lo* DP *lo que me dijiste* ‘what you said to me’ seems to have more nominal-like behaviour than the pseudo-cleft *lo*. The entire modified DP can be replaced by a demonstrative pronoun without the overall meaning of the sentence being affected (20b). Compare to the pseudo-cleft, where the relative clause cannot be replaced without changing the meaning of the sentence (21b). The meaning of (17), reproduced in (21a), is not equivalent to (21b) in the same way that (20a-b) are equivalent to each other.

- 20) a. ***Lo que me dijiste me dolió***  
 LO that ACC.1SG said.2SG ACC.1SG hurt.3SG  
 ‘What you said to me hurt me.’
- b. ***Eso me dolió***  
 PROX.NEUT ACC.1SG hurt.3SG  
 ‘That hurt me.’
- 21) a. ***Lo que necesito es una silla***  
 LO that need.1SG be.3SG a.F.SG chair  
 ‘What I need is a chair.’
- b. ***Eso es una silla***  
 PROX.NEUT be.3SG a.F.SG chair  
 ‘That is a chair.’

Moreover, ‘denotational *lo*’ needs to map to an available predicate in the discourse to receive an interpretation. In (20a), there is no viable candidate that *lo* could map to. This is in contrast with denotational *lo* in (21a), which maps to *la silla* ‘the chair’. Since (20a) is a grammatical and felicitous sentence, we must assume that *lo* is receiving an interpretation from somewhere. Since there is nothing available in the discourse, it must be retrieving its meaning from the common ground, aka. referencing. Therefore, I maintain that the

referential ‘antecedent *lo*’ that appears in (20a) is an entity distinct from Zulaica Hernandez’s denotational ‘antecedent *lo*’<sup>17</sup>.

Let us now turn to the case of denotational ‘wh-internal *lo*’. Zulaica Hernandez classifies the *lo* inside ‘relative pronouns’ as referential, using *lo cual* as his example. A relative clause ‘headed’ by *lo cual* cannot be the subject of a pseudo-cleft construction, as we can see when we try to turn (14) above into one (22). However, prepositional relative clauses with ‘wh-internal *lo*’ can be the subject of a pseudo-cleft, as shown in (23).’

- 22) \**Lo cual me sorprendió fue que Juan nunca llegó.*  
 LO which ACC.1SG surprised.3SG was that Juan never arrived.3SG  
 ‘Juan never arrived, which surprised me.’
- 23) *De lo que estábamos hablando era de la fiesta*  
 About LO that were.1PL talking was.3SG about the.F.SG party  
 ‘What we were talking about was about the party.’

One might be tempted to declare that ‘wh-internal *lo*’ is not (part of) a referential ‘relative pronoun’, but a denotational pronoun. However, this relative clause can only be sentence-initial in the context of a pseudo-cleft. In (24), paraphrased from (20a), the sentence is only grammatical if *lo* has a linguistic antecedent in the discourse. This suggests that, like ‘antecedent *lo*’, ‘wh-internal *lo*’ might have both a referential and a denotational form. This is a problem for Zulaica Hernandez (2018), who presents both *lo* types as being in complementary distribution.

- 24) \*(*Eso de lo que estábamos hablando me dolió*  
 about LO that were.1PL talking ACC.1SG hurt.3SG  
 ‘(That) which we were talking about hurt me.’

To the best of my knowledge, there is no principled reason why an expletive cannot have the same surface position as the ‘contentful’ version of that form<sup>18</sup>, so this issue does not pose an unsurmountable challenge to Zulaica Hernandez’s (2018) proposal. The issue is that distribution was acting as the primary diagnostic between referential and denotational *lo*. Since I have shown this diagnostic to be unreliable, we are left with no reliable way to differentiate between these two types of pronouns<sup>19</sup>.

<sup>17</sup> As to why this *lo* can be in subject position, a possible answer is that the clitic attaches to the relative clause, or another is that *lo* is not actually clitic in this specific context. Determining which is the case is a morpho-phonological question well beyond the scope of this paper.

<sup>18</sup> E.g., expletive *it* in *It is raining* vs. referential *it* in *It is eating my tomatoes*.

<sup>19</sup> Another issue with this proposal is that Zulaica Hernandez (2018) defines referential pronouns as only needing an implicit referent in the common ground. As (24) shows, ‘wh-internal *lo*’ needs an overt linguistic antecedent. In fact, out of the types of referential *lo* that Zulaica Hernandez identifies, only ‘neuter’ object *lo* can have an implicit referent. For example, it references a proposition in (i) as a response to an event that brings said proposition to mind. This issue is not addressed by the author.

(i) *Te lo dije*  
 ACC.2SG LO told.1SG  
 ‘I told you so’

Overall, Zulaica Hernandez (2018) strengthens the semantic analysis of Stark & Pomino (2010) by covering a gap in their proposal, i.e., cases where the non-individuation pronoun *lo* seems to refer to individuals. However, he does not provide clear guidelines for how to identify a referential vs. denotational pronoun outside of a few sample constructions. While his account does provide a broader analysis of the referential properties of more types of *lo*, compared to previous accounts, it nonetheless falls short because of its failure to account for environments where *lo* may be either referential or denotational.

## 5. The nominal structure of referential *lo*<sup>20</sup>

The semantic characterization of referential pronoun *lo* as [-individuation] provides a clear explanation for both its inability to co-refer with a nominal and its lack of  $\varphi$  features. At the same time, it leaves the syntactician with a conundrum – how do we represent the nominal structure of something that has definiteness, but no  $\varphi$  features?

Section 5.1 explains why the non-individuation of *lo* presents a challenge for the standard generative understanding of nominal structure, including the potential solutions that could be devised under this framework. Section 5.2 explores an alternative characterization of *lo* using the Universal Spine Hypothesis (USH) framework.

### 5.1. The *lo* conundrum

Nominal structure is standardly assumed to contain at least three layers: DP, NumP and NP. A *nP* layer between NP and NumP is also quite common. Syntactic features are introduced through the structure-building operation Merge, in which a feature or lexical item joins the previously formed phrase and creates a new phrase in which it is the head (Chomsky, 1995; 2001). NP is the base layer, associated with the semantic content of a nominal, as well as lexically specified features such as animacy. The features associated with *nP* vary depending on the source, but it is commonly associated with categorization (turning a bundle of features into a noun, verb, or adjective) and grammatical gender (Kramer, 2016). NumP is associated with the expression of number and countability. DP is associated with person and definiteness. Additionally, authors like Ritter & Wiltschko (2019) propose a nominal speech act structure which relates the discourse participants (speaker, addressee, and any other possible referents) to the common ground (GroundP).

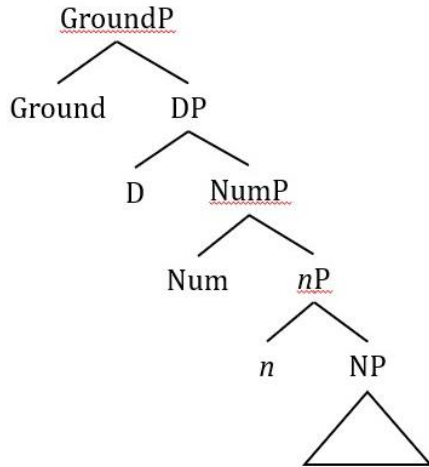
Since syntactic structure is built from the bottom up, with lower categories being selected by, and merging with, higher categories, this creates the implicational hierarchy depicted in (25). If a nominal has features associated with a category higher in the tree, it is predicted to have at least some of the features associated with lower categories. That is, a nominal specified for definiteness (D) will have features associated with grammatical number (Num) and gender (*n/N*). Conversely, a nominal that is unspecified for number/countability is

---

<sup>20</sup> Since denotational *lo* is not a true pronoun, but an expletive with no semantic content of its own, it seems safe to assume that it has little nominal structure, if any. Therefore, the rest of this paper focuses only on referential *lo*.

assumed to lack NumP. Therefore, it cannot have person or definiteness features associated with DP.

## 25) Nominal structure

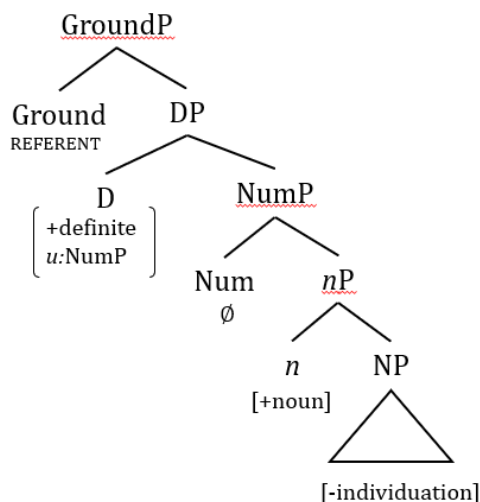
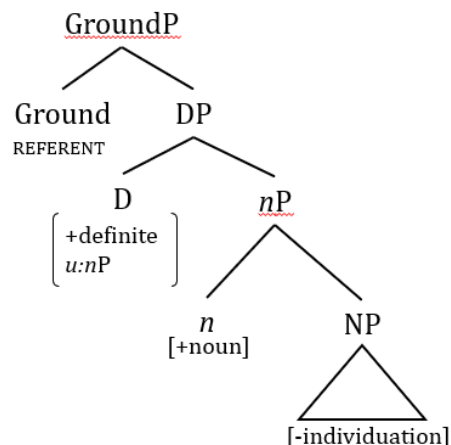


*Lo* challenges this conception of nominal structure because it does not follow the implicational hierarchy. The defining characteristic of *lo* is non-individuation. That is, it is not specified for grammatical number (singular/plural) and countability (mass/count), the syntactic features realized at Num<sup>21</sup>. Since *lo* does not appear to have a Num layer, it is predicted to be unspecified for definiteness (no DP) and referentiality (no GroundP), and yet it is positively valued for both.

There are two possible solutions to this dilemma, both of which would present a drastic departure from standard theory. The first one is to posit the existence of a Num layer that instantiates neither number nor countability, but only exists to preserve the structure. In other words, an expletive feature (26a). This feature does not have a functional role other than ‘existing between *n* and *D*’, has no semantic content by definition, and cannot be independently motivated. Therefore, this solution would be unacceptably *ad-hoc*. The second possible solution is to posit the existence of a Determiner, positively valued for definiteness, that selects a *nP*, rather than a NumP (26b). While more plausible than the previous solution, introducing a Determiner that can ‘skip layers’ of nominal structure, as it were, predicts the existence of all sorts of syntactic patterns that, to the best of my knowledge, are not attested in nominal structure.

<sup>21</sup> At this point, we must clarify the difference between default specification and non-specification, since they are morphologically indistinguishable. When a nominal is default specified, it has an unmarked interpretation. For example, a noun negatively marked for [plural] is interpreted as singular. When a feature is non-specified, there is no available interpretation for that feature. For example, a mass noun, which is unspecified for number, is neither singular nor plural. In these cases, we assume that the syntactic head associated with the unspecified feature is not present in the structure (Wiltschko, 2008).

## 26) a. 'Expletive Num' proposal

b. 'nP-selecting D' proposal<sup>22</sup>

It seems that the nominal structure of *lo* cannot be represented without a drastic, *ad-hoc* departure from standard theory. Therefore, it is worth considering whether a drastic but principled departure from standard theory might be able to provide a clearer understanding of nominal structure and the relationship between pronouns and  $\varphi$  features. In the following section, I explore the nominal structure of *lo* using the Universal Spine Hypothesis framework.

## 5.2. *Lo* and the Universal Spine Hypothesis

According to Universal Spine Hypothesis (USH), the syntax consists of a series of layers associated with specific syntactic functions, called the universal spine (Wiltschko, 2014). These layers are inherent to the structure and present in any utterance; lexical items associate with different points in the spine to gain functional properties. What makes the USH a promising alternative for the syntactic analysis of *lo* is that, since the structure is independent of the lexical content, the absence of content associated with a given layer does not automatically indicate the absence of that layer.

McDonald et al. (2022), working in this framework, propose that pronominals in languages like Japanese are not true pronouns, in the sense that they do not instantiate  $\varphi$ -features. Instead, pronominals in these languages express sociolinguistic traits of the speaker and the addressee, e.g., social status and conceptual gender, as well as of the relationship between the participants. McDonald et al. call these pronominals 'paranouns'. They propose that paranouns merge in the interactional layers of the spine, specifically in the grounding layer.

The grounding layer is composed of two phrases, which together represent the common ground. One represents the speaker's knowledge ( $\text{Ground}_{\text{Spk}}$ ) and the other represents the

<sup>22</sup> The features presented in these trees are for illustration purposes only, and not intended to be a complete representation of all the features associated with pronominal structure. In particular, I am unfamiliar with which features would be implemented at GroundP and have used REFERENT as an *ad-hoc* placeholder.

speaker's assumptions about the addressee's knowledge ( $\text{Ground}_{\text{Adr}}$ ). The Specifier of these phrases identifies the holders of these grounds, i.e., the speaker and the addressee. Paranouns merge at either of these Specifier positions to provide additional information about the speaker or addressee.

*Lo*, like paranouns, does not express  $\varphi$ -features but, unlike paranouns, it does not express sociolinguistic features either. It only expresses definiteness and the semantic content of the proposition/event/etc. that it references. This seems to suggest that *lo* is a pronoun, but pronouns are intrinsically linked with the expression of  $\varphi$ -features. McDonald et al. (2022) state that paranouns have content “*beyond* that of nouns, and beyond what is required for the representation of the discourse referent”<sup>23</sup> (p. 6). This is in contrast with pronouns, which stand in *for* the noun. Since *lo* stands in *for* referents that are not nouns, it might be a paranoun in a completely different way than the socio-linguistically oriented paranouns of Japanese.

Specifically, I suggest that *lo* might attach to the spine as the head of one of the grounding layers, directly providing reference without instantiating  $\varphi$  features. While this has not been attested yet, McDonald et al. (2022) speculate in their conclusion that pronominals that express properties related to the speech act situation, such as definiteness and familiarity, might contain informational structure. Therefore, there is no principled reason that prevents a pronominal from associating with the spine in that position. Unfortunately, the USH and nominal speech act structure are quite new theories. At this moment, these theories have not been developed enough to allow me to present a more detailed prediction of how *lo* might interact with the universal spine.

## 6. Conclusion

In this paper, I show that what is traditionally called the ‘neuter’ determiner *lo* is better characterized as a referential pronoun lacking  $\varphi$  features, which refers to non-individuated entities. However, unlike ‘structurally defective’ pronouns such as impersonal pronouns and expletives, *lo* is a definite pronoun with a specific referent. In addition to this referential *lo*, and following Zulaica Hernandez (2018), I argue that there is a non-referential, expletive *lo*, which relates a possible world to a predicate which may contain morphologically expressed  $\varphi$  features. However, the distribution and referentiality of *lo* in *lo que* constructions indicate that whether *lo* is referential or expletive is not a matter of syntactic position, contrary to Zulaica Hernandez (2018).

Moreover, I show that neither the standard generative approach to nominal structure nor Wiltschko's (2014) Universal Spine Hypothesis are currently able to describe the nominal structure of *lo*. However, the USH presents a more promising avenue for future research because the absence of features associated with Num, and indeed any  $\varphi$  features, does not represent a drastic departure from the present theory. I find it quite likely that a future incarnation of this theory, one that has explored how properties like definiteness integrate

---

<sup>23</sup> Emphasis is their own.

into the universal spine, will contain the tools needed to explain the syntactic and semantic composition of *lo*.

For now, this paper does not present solutions to the issues it presents beyond brief speculation. Or, it presents a wide range of potential avenues for future research. The biggest unanswered question in this paper pertains to the difference between ‘antecedent *lo*’ and ‘wh-internal *lo*’. As we saw in Section 4, the former can reference something implicitly available in the common ground, while the latter requires an antecedent overtly present in the discourse. Research on this question would have to include other referential *lo* types identified by Zulaica Hernandez (2018) such as ‘neuter’ object *lo* and predication *lo*.

Additionally, this proposal would have to extend to the rest of the ‘neuter’ pronominal system identified by Zulaica Hernandez (2018), i.e., demonstratives *eso*, *esto* and *aquello* as well as the ‘neuter’ null *pro*. Other avenues of research include looking into whether non-individuation pronouns/pronouns that reject  $\varphi$  features are attested in other languages and whether these languages might provide new insights into the Spanish system.

## References

- Abney, S. P. (1987). *The English noun phrase in its sentential aspect*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Bello, Andres. (1970). *Gramática de la Lengua Castellana* (R. Cuervo, Ed.). Sopena (Original work published 1847).
- Bosque, I., & Moreno, J. C. (1990). Las construcciones con *lo* y la denotación del neutro. *Lingüística*, 2, 5-50.
- Brucart, J. M. (2016). Oraciones de relativo. In J. Gutiérrez-Rexach (Ed.), *Enciclopedia de Lingüística Hispánica* (pp. 722-736). Routledge.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2001). Derivation by Phase. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 1-52). MIT Press.
- Contreras, H. (1973). Spanish non-anaphoric *lo*. *Linguistics*, 11(111), 5-30.
- Davies, M. (2004). *El uso del Corpus del Español y otros corpus para investigar la variación actual y los cambios históricos*. Tokyo: University Sophia.
- El Colegio de Mexico. (2021). *Lo*. In *Diccionario del Español de Mexico (DEM)*. Retrieved April 04, 2021, from <https://dem.colmex.mx/Ver/lo>.
- Gutiérrez-Rexach, J. (1999). The structure and interpretation of Spanish degree neuter constructions. *Lingua*, 109, 35-63.
- Hamalainen, T. (2004). *La dimensión referencial y atributiva de las expresiones determinadas e indeterminadas: Estudio sobre los artículos del español*. Doctoral Dissertation, University of Helsinki.
- Harley, H., & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, 78, 482-526.
- Hualde, J., Olarrea, A., & O'Rourke, E. (Eds.). (2012). *The handbook of Hispanic linguistics*. Wiley-Blackwell.
- Kayne, R. (1994). *The antisymmetry of syntax*. MIT press.

- Kramer, R. (2016). The location of gender features in the syntax. *Language & Linguistic Compass*, 10, 661-67.
- Luján, M. (2004). Determiners as pronouns. *Romance*, 47, 129-148.
- McDonald, B., Ritter, E. & Wiltschko, M. (2022). Pronouns and Paranouns: A new pronominal typology. [Manuscript to appear]. *Canadian Journal of Linguistics/Revue canadienne de linguistique*.
- Plann, S. (1980). *Relative clauses in Spanish without overt antecedents and related constructions*. University of California Press.
- Pomino, N. & Stark, E. (2009). Losing the neuter: The case of Spanish demonstratives. *Probus*, 21, 217-247.
- Real Academia Española. (2021). El. In *Diccionario de la lengua española* (23<sup>rd</sup> ed.). Retrieved April 04, 2021, from <https://dle.rae.es/el>.
- Ritter, E. (2005). Ambiguity, Generality and the Formalization of Person and Number Features. [Conference handout]. *Les geometries de traits: de la phonologie à la syntaxe*, Université de Paris.
- Ritter, E. & Wiltschko, M. (2019) Nominal speech act structure: Evidence from the structural deficiency of impersonal pronouns. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 64(4), 709-729.
- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, 26(3), 287-350.
- Stark, E., & Pomino, N. (2010). How the Latin neuter pronominal forms became markers of non-individuation in Spanish. In Stathi, K., Gehweiler, E., & König, E. (Eds.) *Grammaticalization: Current views and issues* (pp. 273-294). John Benjamins Publishing Company.
- Wiltschko, M. (2008). The syntax of non-inflectional plural marking. *Natural Language & Linguistic Theory*, 26(3), 639-694.
- Wiltschko, M. (2014). *The universal structure of categories: Towards a formal typology*. Cambridge University Press.
- Zulaica Hernandez, I. (2018). A denotation-driven reanalysis of the Spanish neuter pronominal system. *Lingua*, 213, 91-116.

## Contact Information

Andrea Levinstein Rodriguez

[andrea.levinsteinrod@ucalgary.ca](mailto:andrea.levinsteinrod@ucalgary.ca) | [a.levinstein93@gmail.com](mailto:a.levinstein93@gmail.com)

School of Languages, Linguistics, Literatures, and Cultures

University of Calgary

Craigie Hall D310

2500 University Dr. N.W.

Calgary, AB, T2N 1N4

Canada