

# Revista Española de Lingüística

*Órgano de la Sociedad Española de Lingüística*

***RSEL***

**52/1**

enero-junio 2022



REVISTA ESPAÑOLA DE LINGÜÍSTICA

(RSEL)

52/1

Edita

*SEL*

<https://doi.org/10.31810/RSEL.52.1>

# REVISTA ESPAÑOLA DE LINGÜÍSTICA (RSEL)

ISSN: 0210-1874 • eISSN: 2254-8769

Depósito Legal: M-24.769-1971

DIRECTORA: M. Victoria Escandell-Vidal (Universidad Complutense de Madrid).

SECRETARIOS: Luis Unceta Gómez (Universidad Autónoma de Madrid) / Marianela Fernández Trinidad (Universidad Complutense de Madrid).

CONSEJO DE REDACCIÓN: Ángeles Carrasco Gutiérrez (Universidad de Castilla-La Mancha), Eladio Duque Gómez (Universidad Complutense de Madrid), M.<sup>a</sup> Ángeles Gallego (Consejo Superior de Investigaciones Científicas), Aitor García Moreno (Consejo Superior de Investigaciones Científicas), Eveling Garzón Fontalvo (Universidad de Salamanca), Irene Gil Laforga (Universidad Complutense de Madrid), Ramón González Ruiz (Universidad de Navarra), Victoria Marrero Aguiar (Universidad Nacional de Educación a Distancia), Emilia Ruiz Yamuza (Universidad de Sevilla), Esperanza Torrego Salcedo (Universidad Autónoma de Madrid).

CONSEJO ASESOR: Olga Batiukova (Universidad Autónoma de Madrid), José Luis Blas Arroyo (Universidad Jaime I), Nicole Delbecque (Universidad de Lovaina), María Dolores Jiménez López (Universidad Rovira i Virgili), Rafael Marín Gálvez (Centre National de la Recherche Scientifique, Lille), Hella Olbertz (Universidad de Ámsterdam), Ana Isabel Ojea (Universidad de Oviedo), Begoña Sanromán Vilas (Universidad de Helsinki), Juan Pedro Sánchez Méndez (Universidad de Neuchâtel), Eusebia Tarriño (Universidad de Salamanca), Luis Unceta Gómez (Universidad Autónoma de Madrid), Jesús de la Villa (Universidad Autónoma de Madrid).

RESPONSABLE DE RESEÑAS: Ventura Salazar (Universidad de Jaén).

Los textos enviados para su publicación han de presentarse a través de la web: <http://revista.sel.edu.es/index.php/revista/about/submissions>. Deben ser originales e inéditos y ajustarse a las normas que aparecen en la página web de la *Revista Española de Lingüística*. Todos los artículos son sometidos al dictamen de al menos dos evaluadores designados por el Consejo de Redacción, mediante informes de carácter confidencial. Los derechos de publicación y difusión, bajo cualquier forma, son propiedad de la *RSEL*. Todo texto publicado en la revista obliga a sus autores a no cederlo a terceros, sin autorización previa de la revista. Todos los números de la revista se pueden encontrar en abierto en la web <<http://revista.sel.edu.es/>>.

REDACCIÓN: Sociedad Española de Lingüística, Centro de Ciencias Humanas y Sociales del CSIC, c/ Albasanz, 26-28, 28037 Madrid.

CORREO ELECTRÓNICO Y CORRESPONDENCIA: [RSEL.secretaria@gmail.com](mailto:RSEL.secretaria@gmail.com)



DISEÑO Y COMPOSICIÓN: Carmen Chíncoa & Carlos Curia ([produccionRSEL@gmail.com](mailto:produccionRSEL@gmail.com)).

SERVICIOS DE INFORMACIÓN: Los contenidos de la *RSEL* son recogidos sistemáticamente en *Bibliographie Linguistique/Linguistic Bibliography*, *CINDOC-Base de datos Sumarios ISOC*, *Dialnet*, *DOAJ*, *Dulcinea*, *CIRC*, *Latindex 2.0*, *ERIH PLUS*, *DICE*, *CABELLS*.

# ÍNDICE 52/1 (2022)

## ARTÍCULOS

- A triangulation study on gender agreement in Spanish  
by native Basque speakers* ..... 7  
*Una triangulación sobre la concordancia de género en el español  
de hablantes nativos en vasco*  
GORKA BASTERRETXEA SANTISO
- Hacia una caracterización sintáctica del género del sustantivo en español* ..... 39  
*Towards a syntactic characterisation of nominal gender in Spanish*  
ANTONIO FÁBREGAS
- A través de la subordinada:* ..... 97  
*hiperconcordancia en contextos de dativo experimentante en español*  
*Through the subordinate clause:*  
*Hyper-agreement in dative experiencer contexts in Spanish*  
IRENE FERNÁNDEZ-SERRANO
- Análisis y clasificación de las estructuras comparativas de desigualdad en español* ..... 135  
*Analysis and classification of Spanish inequality. Comparative structures*  
LAURA VELA-PLO

## NOTAS

- La palabra para 'hija' en indoeuropeo: un análisis comparativo* ..... 167  
*The word for 'daughter' in Indo-European: A comparative analysis*  
ESTEBAN NGOMO FERNÁNDEZ

- RESEÑAS ..... 185



## ARTÍCULOS





## A TRIANGULATION STUDY ON GENDER AGREEMENT IN SPANISH BY NATIVE BASQUE SPEAKERS

Gorka BASTERRETxea Santiso<sup>1</sup>  
*Georgetown University*

### *Abstract*

The Spanish spoken in the Basque Country contains some features due to its contact with Basque, one of those being the production of non-standard gender agreement (N-SGA) (Etxebarria Arostegui, 2007, 2008; Fernández Ulloa, 1997, 2006; Urrutia Cárdenas, 2006), since, unlike in Spanish (RAE, 2016), there is no such category in Basque (Zubiri & Zubiri, 2012). This characteristic is typically attributed to elder Basque native speakers without a high level of education (Fernández Ulloa, 1997), but this affirmation is not based on empirical studies. To investigate this gap, this study triangulates data from three sources: conversations, a grammaticality test, and a metalinguistic questionnaire taking education level and participants' gender as independent variables. Twelve young adult (between 21 and 29 years) native speakers of Basque have been recruited. The results suggest that the scope of the population that produces N-SGA is broader than what scholars (Fernández Ulloa, 1997) believed: participants produced N-SGA in speech, despite their grammar knowledge being evident. Finally, these participants believe that N-SGA forms part of the variety of Spanish in the Basque Country and their own idiolect.

*Keywords:* Basque Country; Spanish; Basque; gender agreement; language contact.

## UNA TRIANGULACIÓN SOBRE LA CONCORDANCIA DE GÉNERO EN EL ESPAÑOL DE HABLANTES NATIVOS EN VASCO

### *Resumen*

El español en el País Vasco contiene características particulares debido a su contacto con el vasco, como la concordancia no estándar de género (Etxebarria Arostegui, 2007,

---

1. gb790@georgetown.edu.  <https://orcid.org/0000-0002-9275-1949>

2008; Fernández Ulloa, 1997, 2006; Urrutia Cárdenas, 2006), ya que, a diferencia del español (RAE, 2016), dicha categoría no existe en el vasco (Zubiri y Zubiri, 2012). Esta característica es atribuida a hablantes nativos mayores del vasco sin un alto nivel de educación (Fernández Ulloa, 1997), pero no está basada en estudios empíricos. Por eso, se realizó una triangulación: conversaciones, un test de gramaticalidad, y un cuestionario metalingüístico, tomando el nivel de educación y el género de los participantes como variables independientes. Se reclutó a doce hablantes adultos jóvenes (entre 21 y 29 años) nativos del vasco. Los resultados indican que la población que produce concordancias no estándar de género es más amplia que lo que los académicos (Fernández Ulloa, 1997) creían: los participantes produjeron en el habla casos de concordancia no estándar, a pesar de ser evidente su conocimiento de la gramática. Finalmente, estos participantes identifican la concordancia no estándar como parte del español en el País Vasco y de su idiolecto.

*Palabras clave:* País Vasco; español; vasco; concordancia de género; contacto de lenguas.

RECIBIDO: 10/08/2021

APROBADO: 13/01/2022

## 1. INTRODUCTION

Spanish and Basque are two languages that have been in constant contact and have mutually influenced each other linguistically over time, and they continue to do so (Camus Bergarehe and Gómez Seibane, 2012b; Echenique Elizondo, 1997, 2004; Escobar, 2010; Etxebarria, 2004; Ezeizabarrena, 2009; Klee and Lynch, 2009; Montrul, 2013; Moreno-Fernández, 2020; Urrutia Cárdenas, 2006). This means that the Spanish spoken in the Basque Country<sup>2</sup> contains some particular and distinct characteristics that have arisen through its contact with Basque (Echenique Elizondo, 1997; Etxebarria-Arostegui, 2007; Moreno-Fernández, 2020). An example that illustrates this phenomenon is the production of non-standard gender agreement (N-SGA) in Spanish by Basque speakers (Etxebarria Arostegui, 2007, 2008; Fernández Ulloa, 1997, 2006; Urrutia Cárdenas, 2006). This is so because while Spanish realizes grammatical gender (RAE, 2016), Basque does not, meaning that the same form is used for masculine and feminine (Euskaltzaindia, 1991, 2002; Laka, 1996; Trask, 2003; Zubiri and Zubiri, 2012). Particularly, and despite the fact that the majority of works do not mention it (e.g., Etxebarria, 2004; Escobar,

---

2. Basque Country refers here to the three historical territories of Bizkaia, Gipuzkoa and Araba, leaving apart other territories in which Basque is also spoken, such as Nafarroa and the French Basque Country.

2010; Klee and Lynch, 2009; Moreno-Fernández, 2020; Moreno Fernández and Otero Roth, 2016; Urrutia Cárdenas, 1995), this feature is typically attributed to elder *euskaldun zaharrak*<sup>3</sup> who do not have a high level of education (Fernández Ulloa, 1997), and is based on generalizations instead of empirical investigations. However, after some personal observations of the Spanish that is produced in the Basque Country, it could be the case that the particular feature that has been introduced here does not exclusively form part of the speech of the population mentioned by the previous publications but might rather be a general and prominent characteristic of the Spanish spoken in the Basque Country, regardless of age and education. In fact, N-SGA production is believed to be common in language contact situations (e.g., Spanish in contact with Palenquero, Lipski, 2015; Tepehuán, Torres Sánchez, 2021; Papiamentu, Valdés Kroff et al, 2019; and English, Licerias et al., 2008) where the L1 does not produce grammatical gender (Cruz Rico et al., 2021; Husein, 2021), and in this case, in places where Spanish is the minority language (e.g., Anderson, 1999). This does not mean that those who may produce N-SGA have language learning difficulties (Anderson, 1999), or that they do not know about standard gender agreement (GA) rules in Spanish, but simply that they would not use it in speech. These are the reasons why it is necessary to conduct a study that analyzes GA in this variety of Spanish. The present paper proposes a triangulation study to determine whether there are any differences between (i) what young adult native speakers of Basque who have received formal education on Spanish grammar actually do in speech and (ii) what they think they would say in an imaginary conversation regarding GA, together with (iii) if they can identify this linguistic feature as part of the Spanish spoken in the Basque Country.

## 2. LANGUAGE CONTACT

The phenomenon being analyzed here is a case of language contact. Within this framework, the production of N-SGA in Spanish by Basque speakers could be considered a «linguistic interference», a term originally introduced by Sandfeld (1938) and later categorized by Weinreich (1966, p. 1) as «instances of deviation from the norms of either language which occur in the speech of bilinguals as a

---

3. This is the typical terminology used to refer to Basque native speakers, while *euskaldun berriak* refers to the people whose first language is not Basque (Fernández Ulloa, 1997, 2006; Klee and Lynch, 2009; Urrutia Cárdenas, 2006). *Euskaldun* means speaker of Basque (Etxebarria, 2004; Laka, 1996), *zaharra* means 'old', and *berria* 'new'.

result of their familiarity with more than one language». The linguistic interference between two languages is extended to every part of a language, among which syntax, morphology, phonology, and vocabulary are included (Thomason, 2001; Weinreich, 1966). Following the introduction of the original term, other nouns have been introduced to replace the ungrammatical connotation that the term interference contains (López Morales, 1993), one of them being «transference» (Clyne, 1967). As Gutiérrez and Silva Corvalán (1993) explain, one of the possible types of transferences is defined as «negative transference», meaning that the absence of a grammatical phenomenon in a language is transferred to the other language. An example that illustrates this case, and following Fernández Ulloa (2006), is the exact phenomenon being investigated in this paper, since, as it was introduced above, Basque does not produce grammatical gender as Spanish does. Transference is understood sometimes (e.g. Silva-Corvalán, 1994) as something common in bilinguals who use two linguistic systems. Another difference between «linguistic interference» and «transference» is that the first implies that language variation is produced due to a language's contact with another one, while the latter implies the consumption of a linguistic change or at least a linguistic change in progress (Ramírez Cruz, 2009). Moreover, there is even a third term: «linguistic convergence», which leaves behind the notion of ungrammaticality, and it can be understood as a «calque» or «pattern transfer» (Matras, 2009), or the merging of structures between languages (Moreno-Fernández, 2020). One of the features of convergence is the so-called «perceived similarity», according to which «a convergence process based on perceived similarity can be compatible with copying from A to B» (Palacios and Pfänder, 2014, p. 222). However, some features can be lost through language contact, e.g., gender in some dialects of Yiddish (Matras, 2009), which might be similar to what is being investigated here in the Spanish produced in the Basque Country.

### 3. GENDER AGREEMENT IN THE SPANISH IN THE BASQUE COUNTRY

The variety of Spanish spoken in the Basque Country has been widely investigated (e.g. Calleja Azpiazu, 2004; Camus Bergareche, 2011; Camus Bergareche and Gómez Seibane, 2010, 2012a; Echaide, 1968; Elordieta, 2003; Gómez Seibane, 2012; Landa, 1993; Landa and Franco, 1996; Oñederra, 2002, 2004; Paasch-Kaiser, 2015) but this in no way means that the description of this dialect is complete. To give an example, a characteristic that needs further investigation is the production of

N-SGA in this dialect (Badiola and Sande, 2018), as in *la tierra era más tranquilo*, while the expected standard agreement would be *tranquila*; ‘the earth [fem.] was calmer [fem.]’ (Fernández Ulloa, 1997, p. 202). This characteristic has been previously mentioned (e.g., Etxebarria Arostegui, 2007, 2008; Fernández Ulloa, 1997, 2006; Urrutia Cárdenas, 2006) but not many studies have been conducted on this region from a sociolinguistic point of view. The only experimental studies that have analyzed this phenomenon in the contact between Basque and Spanish are Parafita Couto et al. (2015) and Badiola and Sande (2018), in this case, in mixed (Basque and Spanish) DPs and not only in the Spanish that is produced in the Basque Country. Apart from these works, some other investigators have studied it from the perspective of language acquisition (e.g., Barreña, 1997; Larrañaga and Guijarro-Fuentes, 2012) or agreement in L3 English by Basque-Spanish bilinguals (e.g., Imaz Agirre and García Mayo, 2013). However, to date, no GA production study (e.g., spontaneous speech) has been carried in the variety of Spanish in the Basque Country, thus, we still do not know if the general population speaking this variety produces N-SGA since, according to Fernández Ulloa (1997), it is only a characteristic of elder *euskaldun zaharrak* without a high level of education (not based on an empirical study). This, in any case, is not a recent phenomenon as Etxague Burgos (2012) and Gómez Seibane (2008) also mentioned this feature as one of the characteristics that defined the written Spanish produced particularly in the province of Bizkaia during the 15<sup>th</sup> and 16<sup>th</sup> centuries, despite making it clear that it is not exclusive to this province since some examples were found in the Spanish produced in Nafarroa in the 18<sup>th</sup> century. However, we do not have any information on this regard in the current written Spanish in the Basque Country: do native Basque speakers produce N-SGA when writing in Spanish?

The production of N-SGA might be caused by the fact that while Spanish produces grammatical gender (RAE, 2016), Basque does not<sup>4</sup> (Euskaltzaindia, 1991, 2002; Laka, 1996; Trask, 2003; Zubiri and Zubiri, 2012), except for the familiar second person singular pronoun *hi*.<sup>5</sup> Regarding Spanish, traditionally two genders are described: masculine and feminine. The majority of nouns in Spanish have phonological characteristics that indicate their gender: the majority

---

4. Basque does not officially contain a grammatical gender, but some publications mention that some of the oral modern varieties of Basque make gender distinctions in some words (Euskaltzaindia, 1991, 2002; Laka, 1996; Trask, 2003; Zubiri and Zubiri, 2012). Euskaltzaindia (2002) and Laka (1996) characterize it as a marginal phenomenon.

5. This is the only element in Basque that distinguishes agreement morphemes depending on if the addressee is male or female (Alberdi, 1995; Euskaltzaindia, 1991, 2002; Hualde, 2003; Laka, 1996).

(99,9%) of nouns ending in *-o* are masculine, while the majority (96,3%) ending in *-a* are feminine (Goebel-Mahrle and Shin, 2020; Roca, 1989; Teschner and Russell, 1984). These are nouns with canonical endings but there are also nouns with non-canonical endings: nouns that do not end with their expected gender vowel or those that end with consonants (Cruz Rico, 2021; Goebel-Mahrle and Shin, 2020; Roca, 1989). Nevertheless, it is also notable that it is common to have «difficulties» with the acquisition of gender on the L2 (Anderson, 1999; Cruz Rico et al, 2021; Husein, 2021), regardless of whether the L1 produces grammatical gender or not (Bruhn de Garavito and White, 2002).

It has been noted that Fernández Ulloa (1997) attributes this N-SGA production to elder *euskaldun zaharrak* with a low education level despite this being a generalization that has not been confirmed with an empirical study. Thus, this confirms the need for further investigation. After some personal observations for being part of the population who speaks this variety of Spanish, it could be the case that this phenomenon is present among many speakers of Spanish in the Basque Country. This makes it particularly interesting to investigate this phenomenon among young adult native speakers of Basque who do not usually speak Spanish in their daily lives but have been educated on Spanish grammar during primary and secondary school. This is the perfect population to determine if they have the knowledge about GA rules in Spanish and whether they produce it in speech or not.

According to what has been described above and to analyze the production of GA in this variety of Spanish, the following research questions have been formulated:

1. To what extent do Basque native speakers who use Spanish infrequently and are formally educated produce N-SGA in speech production?
2. To what extent does the same group produce N-SGA when asked what they would say in a conversation (multiple-choice test)?
3. Does the population in question believe that the production of N-SGA is central to the Spanish spoken in the Basque Country?

Based on previous literature, the following hypotheses are considered:

1. Speakers will produce a considerable number of instances of N-SGA in Spanish compared to the instances that follow standard Spanish GA rules.

2. Participants will follow standard GA rules since they are not in a spontaneous conversation, and they are given time to think their answer.
3. They will not think that N-SGA is central to the Spanish spoken in the Basque Country. Instead, accent or intonation and the use of Basque words will be the principal characteristics that they will mention (Ciriza, 2009).

The methodology that has been employed in this paper will be described below. Then, the results will be analyzed to provide answers for the research questions. This will lead to the final part of the paper in which the discussion will be developed together with the conclusions and main implications of the investigation.

#### 4. METHODOLOGY

A total of 12 participants (six females and six males) were recruited in late 2019 and early 2020 following the judgment sampling method (Hoffman, 2014; Schilling, 2013; Silva-Corvalán and Enrique-Arias, 2017). These are young adult (age range: 21–29; age median: 24.91) native speakers of Basque who have learned Spanish in the elementary and high school in Model D.<sup>6</sup> With the aim of having participants who are native speakers of Basque and do not usually use Spanish in their daily lives (self-reported percentage of use of Basque and Spanish included in Table 1), these participants were recruited in Ondarroa (Bizkaia), a town where 98.8% of the young population mainly uses Basque (Albizu Lizaso and Arana Arexolaleiba, 2018). Regarding participants' level of education, six of them had completed (or were finishing) an undergraduate or a professional training degree (lower education level for analysis), the other six had a Master's or a PhD (higher education level). The information related to the linguistic background of participants was gathered through a language background questionnaire that was specifically prepared for the present study and that they completed at the end of their participation.

---

6. Since 1982 Basque and Spanish are obligatory subjects in the Basque Country and its public education is divided into three bilingual systems: A, B and D (there is no letter C in Basque). In Model A, students learn all the subjects in Spanish and Basque is taught as a subject with two to four hours per week. In Model B, 50% of the subjects are taught in Basque and the other 50% of the subjects are taught in Spanish. Finally, in Model D, Basque is the language of instruction and Spanish is taught as a subject with two to four hours of instruction per week (Cenoz and Perales, 2007; Cid Abasolo, 2009; Etxebarria-Arostegui, 2007; Hezkuntza Saila, 2020).

Participants	Most-used language	Family language		Friends language		Work/university language	
		Basque	Spanish	Basque	Spanish	Basque	Spanish
a_FH <sup>7</sup>	Basque	100%	-	100%	-	40%	60%
b_FL	Basque	90%	10%	90%	10%	70%	30%
c_FH	Basque	100%	-	80%	20%	100%	-
d_FH	Basque	100%	-	100%	-	75%	25%
e_FL	Basque	100%	-	100%	-	90%	10%
f_FL	Basque	80%	20%	100%	-	70%	30%
g_ML	Basque	100%	-	90%	10%	90%	10%
h_MH	Basque	100%	-	100%	-	85%	15%
i_ML	Basque	95%	5%	100%	-	50%	50%
j_ML	Basque	80%	20%	100%	-	90%	10%
k_MH	Basque	100%	-	100%	-	100%	-
l_MH	Basque	100%	-	95%	5%	90%	10%

Table 1. Self-reported percentage of use of Basque and Spanish

In order to be able to answer the research questions introduced above, a quasi-experimental triangulation study has been designed. Triangulation is understood here as the use of three different methods so as to develop an ample perspective on GA (not gender assignment) in the variety of Spanish produced in the Basque Country and thus obtain a more robust understanding of the analyzed linguistic phenomenon (Jaffe, 2014). Consequently, the data gathered for this study is threefold: spontaneous conversations in pairs, a multiple-choice test, and a metalinguistic questionnaire. The order in which participants completed each part was such that they were not aware of the linguistic aspect being studied: first the conversation, then the multiple-choice test and, finally, the questionnaire.

The spontaneous paired conversations with no researcher present (Schilling, 2013) consisted of a completely free 45-minute conversations between two participants of the same sex that were close friends or close family members so as to collect the most natural data possible. Some scholars (e.g., Hoffman, 2014; Schilling, 2013;

7. In order to identify each of the participants in this study, the low letter makes reference to the random letter that they were assigned to anonymize their information, the second letter makes reference to their self-reported gender (F = feminine, M = masculine), and the second letter is their level of education (H = higher, L = lower).



Silva-Corvalán and Enrique-Arias, 2017) comment on the ideal duration of the recording being between an hour and a half to two hours but it is believed that GA is a frequent phenomenon, thus, 45 minutes should be sufficient, also considering that after the first 10 minutes «the level of engagement should be high and the equipment will be forgotten or peripheral» (Hoffman, 2014: 35). With the aim of avoiding the Observer's Paradox (Labov, 1972), the investigator was not present in the conversations (only two participants at a time) and they were recorded with a cellphone with the hope that at some point participants would not feel its presence. The coding of GA was operationalized through accuracy measurements –accurate understood here as instances of standard agreement rules– of every instance of GA that arose in each conversation. For precision, each recording was listened to three times with the coding process repeated. The total number of cases that follow standard Spanish rules was counted. Then, ratios of GA instances were calculated following the suggestion by Guy (2014) and similar to previous studies (e.g., Anderson, 1999). This is due to the conversations not having any prerequisites or guidelines: the amount of GA instances that each speaker produced was variable (similar to Anderson, 1999).

Regarding the multiple-choice test, it was composed of 96 sentences with a blank space that participants had to fill in by choosing one of the two given options: there was a correct possibility and an incorrect one. From the 96 sentences of the test, only 24 are of interest to the current research (included in Appendix A) since the other 72 were distractors (related to features that form part of the analyzed variety of Spanish: the use of conditional where the subjunctive would be expected, absence of article before a noun, and *leísmo*) so that participants would not identify the feature being investigated. In these 24 sentences, three variables were controlled: nouns such as *el nuevo* 'the new', pronouns at the end of verbs such as in *recogerlo* 'gather it', and adjectives such as *complicado* 'complicated'. Each of these three variables had four sentences for each gender: two sentences in which the antecedent (always inanimate nouns) ended in the canonical vowel for its gender –for the feminine, *-a*, and *-o* for the masculine– and two sentences in which the antecedent had a non-canonical ending since these words are frequent in Spanish as well (Anderson, 1999). It is believed that canonical nouns give sufficient clues to determine gender due to their phonological characteristics, different from the non-canonical nouns that do not end with their expected vowel (Cruz Rico et al., 2021). The antecedents with which participants had to make the agreement are nouns that are among the 500 most frequent words in Spanish according to the

list of frequencies in the *Corpus de Referencia del Español Actual* (CREA, 2008) by the RAE. For consistency purposes, all the antecedents were accompanied by a determiner and words such as *agua* ‘water’ and *alma* ‘soul’ were excluded; despite being high frequency words according to the aforementioned dictionary, they are feminine nouns that in their singular form are accompanied by a masculine determiner for starting with an unstressed /a/ (Clegg, 2011). Also, for consistency, the antecedent and the blank space where the participants had to make their choice were distanced by a minimum of three and a maximum of five words (it is believed that the distance between nouns and adjectives increases the production of N-SGA, e.g., Goebel-Mahrle and Shin, 2020; Lipski, 2015; Ramírez Cruz, 2009). The results were coded by accuracy measurements, calculating the total number of instances that follow standard Spanish rules on GA.

In the metalinguistic interview, participants answered both general questions about the Spanish spoken in the Basque Country (e.g., characteristics that form part of it, what they think about them) as well as specific questions about GA (e.g., they were asked if they believed that N-SGA is a characteristic of the variety of Spanish spoken in the Basque Country and whether they think they produce it or not). This is the qualitative part of the paper from which general trends have been extracted.

In order to gauge GA –the dependent variable– two categorical independent variables have been considered: gender (female or male) and highest education level (undergraduate/professional training or Master/PhD)<sup>8</sup>. Participants’ gender has been included in previous studies to characterize language varieties (e.g., Falcón Ccenta et al., 2012). Level of education has also been mentioned (e.g., Palacios, 2006) as a key factor for N-SGA production in different language contact situations. According to the standards by Silva-Corvalán and Enrique-Arias (2017), at least five participants are necessary per external variable, and this has been respected for the two social independent variables in this research. Standard multiple linear regressions with random effects have been conducted to analyze the data from the conversation and the multiple-choice test by including both social and linguistic predictor. For this, the package *lme4* has been employed (Bates and Maechler, 2010). Finally, no test has been used for the third research question since it will be analyzed qualitatively.

---

8. In the results section, those who have an undergraduate or a professional training will be referred to as «lower education», and those with a Master or a PhD will be grouped under «higher education».

### 5. RESULTS

The number of N-SGA production rates per participant in both spontaneous conversations as well as the multiple-choice test are presented in Table 2. These results are illustrated by the plots in Figure 1.

Participants	Conversation		Multiple-Choice Test	
	Number	Percentage	Number	Percentage
a_FH	3/18	5,4%	3/24	12,5%
b_FL	0/4	0%	3/24	12,5%
c_FH	4/12	33,33%	4/24	16,66%
d_FH	1/5	20%	0/24	0%
e_FL	4/13	30,76%	1/24	4,16%
f_FL	0/7	0%	12/24	50%
g_ML	1/6	16,66%	11/24	45,83%
h_MH	1/14	7,14%	4/24	16,66%
i_ML	2/9	22,22%	10/24	41,66%
j_ML	2/10	20%	5/24	20,83%
k_MH	3/7	42,85%	4/24	16,66%
l_MH	3/12	25%	1/24	4,16%

Table 2. N-SGA production by speaker and experimental condition

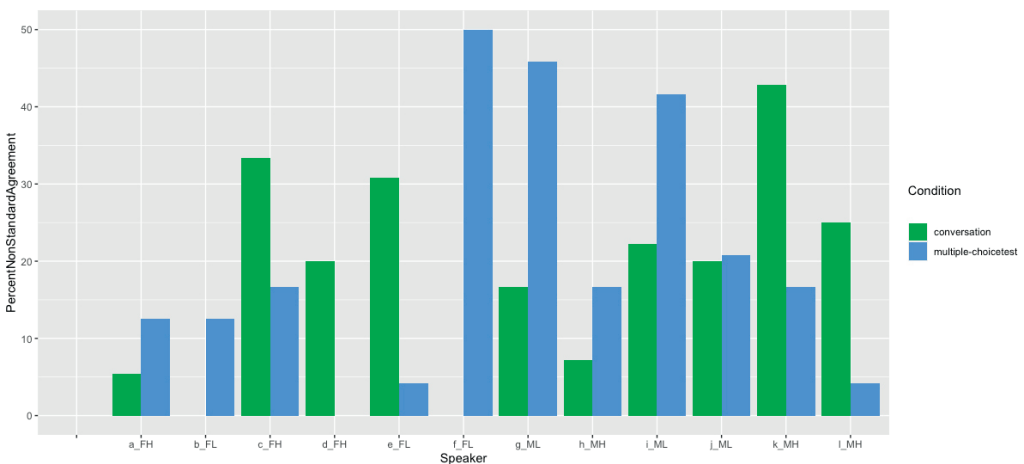


Figure 1. N-SGA production by speaker and experimental condition

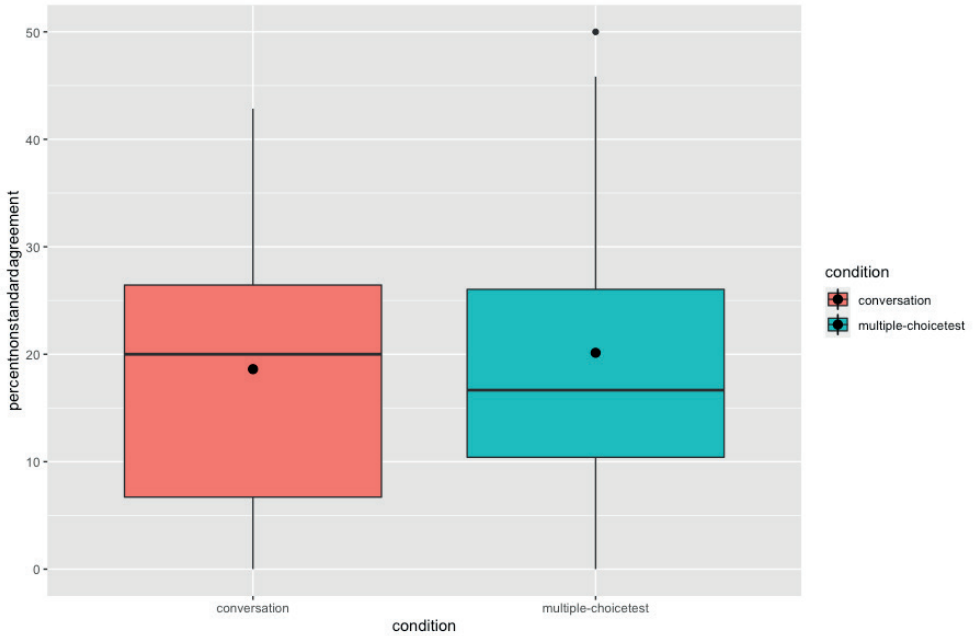


Figure 2. N-SGA Production by Experimental Condition

In Figure 2, the overall mean scores of N-SGA productions in each of the two conditions are presented. Specifically, there are more instances of N-SGA in Spanish in this study in the data gathered through the multiple-choice test (mean = 20.13, SD = 16.75,  $N = 12$ ), when compared to the rates obtained in the results gathered through the conversations that participants maintained in pairs among them (mean = 18.61, SD = 13.52,  $N = 12$ ). This might be due to the fact that the number of GA instances that each participant produced in the conversation was variable since the conversations were free and, thus, did not follow any specific guideline as in the multiple-choice test.

The results of the first research question show that, overall, participants produced many more instances in which they followed the standard rules for GA in Spanish in their speech, regardless of participants' gender and their level of education. This can be visualized in Figure 3 below since the scores of the conversations make reference to the ratios of those cases in which the standard GA rules were followed. This figure also indicates that, interestingly, those with a higher level of education had more instances of N-SGA compared to those participants with a lower level of education. With regards to participants' gender, Figure 3

demonstrates that female participants in the lower education group produced more cases of GA following the standard Spanish rules compared to male participants of both levels of education.

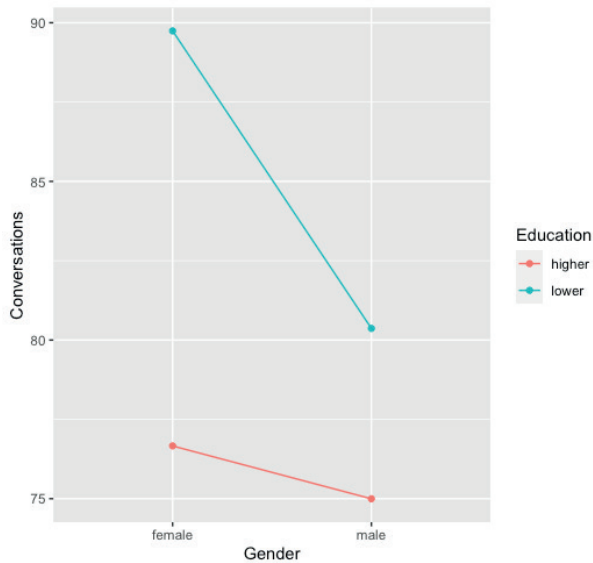


Figure 3. Participants' gender and education interaction on standard agreement scores in conversations

Regarding the results of the second research question, Figure 4 visualizes that in the test too, the number of instances –from a total of 24 points– in which participants chose the option that followed the standard Spanish rules for GA are much higher than the N-SGA cases. The same figure also indicates that overall, those with a higher education level chose more instances of what standard grammar rules expect for gender, in front of those with a lower level of education, regardless of participants' gender. Indeed, the females with a higher level of education are the ones with the highest scores. However, females with a lower education are quite close to the males with a higher education. Finally, the males who belong to the lower education level have the lowest number of cases of standard GA and they appear to be quite distanced from the rest of the participants. Nevertheless, the multiple linear regression will delimit if these differences and variables are significant.

The results of the multiple linear regression with random effects that has been conducted including three independent variables (participants' gender, their level of education, and type of experiment) reveal in Table 3 that none of the introduced

predictors has been found to be significant (participants' gender  $p = 0.216$ ; education  $p = 0.395$ ; type of experiment  $p = 0.807$ ). Considering the adjusted  $R^2 = 0.024$ , this model explains only 2.4% of the variation, which is very low. The result of this model is  $F(3,20) = 0.815$ ,  $p = 0.5$ , which together with the  $t$ -values (participants' gender  $t = 1.276$ ; education  $t = -0.869$ ; type of experiment  $t = 0.247$ ), indicate that the null hypothesis is accepted. Consequently, the multiple linear regression with random effects shows that variances are equal in this case.

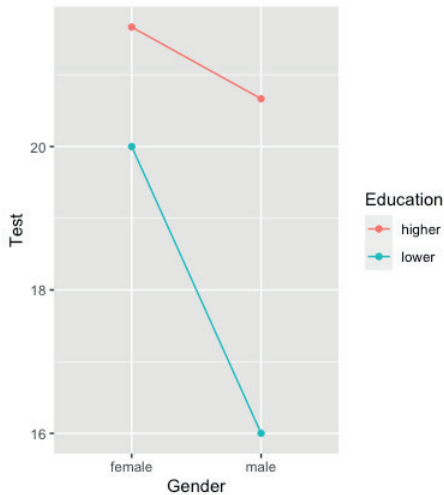


Figure 4. Participants' gender and education interaction on standard agreement scores in tests

	Estimate	Std. error	t-value	p-value
Intercept	22.714	10.670	2.129	0.0459
Participants' Gender	7.863	6.16	1.276	0.2164
Education	-5.355	6.16	-0.869	0.395
Experiment Type	1.522	6.16	0.247	0.807

Table 3. N-SGA by participants' gender, level of education and type of experiment

Since the multiple linear regression conducted in this study has shown a low effect of participants' gender, their highest level of education, and type of experiment conducted when explaining the rates of N-SGA, purely linguistic factors need to be considered now. Starting with the instances of N-SGA found in the pair

conversations, all cases of N-SGA are long distance agreement cases, as in (1)<sup>9</sup>, (2), (3), and (4), and no N-SGA has been found within the NP as in *el pared* ‘the wall [masc.]’ when the standard agreement would be *la pared* ‘the wall [fem.]’.

- (1) *Es* *super pequeño* (referring to *la casa*) [*pequeña*]  
 be.PRS.3SG very small-M (referring to *the house-F*) [small-F]  
 ‘It’s very small’.
- (2) *La mayoría que no sabemos hacerla* (referring to *maskilla*) [*hacerla*]  
 the majority that not know.PRS.1PL do-it-M (referring to *mask-F*) [do-it-F]  
 ‘The majority of us who don’t know to make it’.
- (3) *Es* *negro* (referring to *la camiseta*) [*negra*]  
 be.PRS.3SG black-M (referring to the *T-shirt-F*) [black-F]  
 ‘It’s black’.
- (4) *Está* *todo el día pegado a la televisión* (referring to *la madre*)  
 [*pegada*]  
 be.PRS.3SG all the day stuck-M to the television (referring to the mother-F)  
 [*pegada-F*]  
 ‘She’s watching TV the whole day’.

In addition, the antecedent which the elements of the sentence had to agree with were all inanimate, as in (1), (2), (3), (5), and (6), except for two cases, as shown in (4).

- (5) *Un montón de caseríos estaban quedándose viejas* [*viejos*]  
 a lot of farms-M be.IMP.3PL stay. CONT.REFL old-F [old-M]  
 ‘A lot of farms were becoming old’.
- (6) *La cara te despertabas lleno de picaduras* [*llena*]  
 the face-F you wake-up.IMP.2SG full-M of bites [full-F]  
 ‘You woke up with the face full of bites’.

The majority of the N-SGA instances found in the conversations are singular in terms of number as in (1), (2), (3), (4), and (6), except for a few of them (7 out of 24), as in (5). Finally, these results would suggest that young adult native speakers of Basque follow the masculine default strategy when producing GA in Spanish as in the majority of cases (19 out of 24), they produced masculine GA when the

9. For space limitations, one example has been taken from each conversation. The word that represents the N-SGA case is underlined. The option that follows Spanish rules is presented between brackets.

standard rules indicated feminine agreement. This was the case even if the antecedent's ending was canonical (*-a* for feminine words, RAE, 2016): 14 out of 19 cases as in (1), (2), (3), and (6).

	Estimate	Std. error	<i>t</i> -value	<i>p</i> -value
Intercept	0.833	0.076	10.855	<2e-16
Antecedentfem_noncan	-0.208	0.108	-1.919	0.056
Antecedentmasc_can	0.166	0.108	1.535	0.125
Antecedentmasc_noncan	0.166	0.108	1.535	0.125
Agreeing_elementnoun	-0.125	0.108	-1.151	0.250
Agreeing_elementpro	0.083	0.108	0.768	0.443
Antecedentfem_noncan * agreeing_elementnoun	0.291	0.153	1.900	0.058
Antecedentmasc_can * agreeing_elementnoun	-0.083	0.153	-0.543	0.587
Antecedentmasc_noncan * agreeing_elementnoun	-0.166	0.153	-1.085	0.278
Antecedentfem_noncan * agreeing_elementpro	0.041	0.153	0.271	0.786
Antecedentmasc_can * agreeing_elementpro	-0.125	0.153	-0.814	0.416
Antecedentmasc_noncan * agreeing_elementpro	-0.375	0.153	-2.442	0.015

Table 4. Interaction between gender and canonicity of the antecedent and agreeing element upon N-SGA in the multiple-choice test

Then, by analyzing the linguistic predictors (gender and canonicity of the antecedent and type of agreeing element [noun, pronoun, or adjective]) in the N-SGA instances found in the multiple-choice test, a multiple linear regression has been conducted to analyze the impact of their interaction in the dependent variable. The results in Table 4 suggest that there is a significant interaction only between the antecedent elements that are masculine and that end with their non-canonical vowel [every ending except *-o*, RAE, 2016] (Masculine non-canonical antecedents agreeing with pronouns  $p = 0.015$ ) as in sentence (7) and (8) in Appendix A. This suggests that there is a higher tendency to produce N-SGA when the antecedent is masculine and non-canonical and the agreeing element is a pronoun attached to a noun. Considering the adjusted  $R^2 = 0.061$ , this model explains only 6.1% of the variation, which is very low. The result of this model is  $F_{(3,11)} = 2.702$ ,  $p = 0.002$ , which indicates that the null hypothesis is rejected. Consequently, the multiple linear regression with random effects shows that variances are not equal.

Finally, the qualitative data of the third research question suggest that, on the one hand, «accent» is the most repeated characteristic by the participants (though none of them mentioned «intonation») when they are asked to name some of the



characteristics that form part of the Spanish spoken in the Basque Country; it is actually mentioned by all the participants except two of them. Nevertheless, only two of the participants mentioned the «use of Basque words». Other characteristics that were mentioned were, for example, *pronunciar y enfatizar la ‘s’* ‘to pronounce and emphasize the «s»’, *cambiar la estructura de la frase* ‘to change the structure of the sentence’, and *palabras al final de la frase como ‘pues’* ‘words like «so» at the end of the sentence’. On the other hand, the most interesting result is related to the linguistic feature being analyzed in this paper. When they were asked to name some of the characteristics of the variety of Spanish spoken in the Basque Country, only six of them named it. Nevertheless, when they were asked directly if they believe that N-SGA is a characteristic of the aforementioned variety, 10 of them agreed. Indeed, some of them declared that it is *uno de los errores principales* ‘one of the principal mistakes’, *suele pasar a menudo* ‘it happens often’, and *es una característica del español hablado por gente cuya lengua materna es el euskera, ya que en euskera los nombres no tienen género* ‘it’s a characteristic of the Spanish spoken by people whose first language is Basque, since in Basque nouns don’t have gender’. Moreover, 11 of them revealed that they believe they commit N-SGA (for instance, one participant declared that *me suele entrar duda de qué es lo correcto* ‘I doubt what the correct form is’) and only two of them feel as though this is an important mistake for communication. The others think things like *pienso que se puede comunicar bien a pesar de esos errores* ‘I think that you can communicate appropriately despite those mistakes’, and *no causa problemas de comunicación, aunque a veces produce cierta gracia en la otra persona y de vez en cuando pausa la conversación* ‘it doesn’t cause communication problems, although sometimes it’s kind of funny for the other person and it sometimes stops the conversation’. That might also be the reason why when they were asked if they feel they need to improve any aspect of their own Spanish, only four of the total participants declared that GA might be something they need to improve. However, none of the participants remembers if this is a characteristic that Spanish grammar teachers in primary and secondary school insisted on correcting.

## 6. DISCUSSION

The aim of the present study was to analyze N-SGA by young adult Basque native speakers who do not usually speak Spanish in their daily lives but who have received classes on Spanish grammar. In terms of the first research question,

it was predicted that speakers would commit a considerable number of N-SGA in Spanish. However, the results of the first research question that have been presented above suggest that the initial hypothesis cannot be completely confirmed since participants had greater ratios of instances in which they followed standard Spanish grammar rules than those in which they committed N-SGA (similar to e.g., Torres Sánchez, 2021). One of the reasons that may explain this result is that fewer and uneven instances (similar to Anderson, 1999) of GA were found in the pair conversations, contrary to what was expected before conducting the study. This is due to the initial decision that aimed to achieve as natural conversations as possible, bearing in mind that this way the context is less controlled (Schilling, 2013): it is believed that this objective was achieved due to the personal and deep conversations participants had with each other. It is also interesting that despite the fact that the conducted analysis showed that the differences between the independent variables (participants' gender, education level, and type of experiment) were not statistical, further research is necessary to justify why those who have a higher education level had higher ratios of N-SGA when logic would predict the opposite.<sup>10</sup> Although some studies defend that the level of education has no effect in N-SGA (e.g., Balam et al., 2021), the result in this study contradicts previous studies that mention that those with lower education levels are the ones that produce more N-SGA (e.g., Díaz Barajas and Orozco, 2019; Palacios, 2006; Torres Sánchez, 2021). Again, the previously mentioned fact about the different amount of speech produced by each participant in the conversations might be the best explanation.

Regarding the purely linguistic factors that were included in the analysis of the conversations, one of the facts that has been tested with the conversation data is that the distance between agreeing elements is a factor that probably favors the production of N-SGA in the Spanish spoken in the Basque Country, as it is the case in other language contact situations (e.g., Goebel-Mahrle and Shin, 2020; Lipski, 2015; Ramírez Cruz, 2009). In addition, the antecedent's animacy has also been shown to be a good indicator for N-SGA as the majority were inanimate nouns: the origin of the gender of inanimate nouns is random when compared to that of animate nouns as it is later acquired both by L1 and L2 speakers (Andersen,

---

10. Apart from logic, this was an unexpected result also based on other previous studies (e.g., Etxebarria-Arostegui, 2007, 2008; Paasch-Kaiser, 2015; Urrutia Cárdenas, 1991, 1995) that indicate that, for example, in the same variety of Spanish analyzed here, the use of the conditional in *if*-clauses, and not the subjunctive –as expected by standard grammars– is more common among those with a lower level of education and culture.

1984; Cuza and Pérez-Tattam, 2016; Fernández-García, 1999; Hernández-Pina, 1984). Then, previous studies also found more N-SGA with inanimate nouns (Fernández-García, 1999; Finnemann, 1992). Another important factor to explain the production of N-SGA in this study is the use of the masculine default gender, an important variable in other previous studies (e.g., Balam et al., 2021). However, it was interesting to note that, contrary to other studies analyzing N-SGA in other language contact situations (e.g., Alarcón, 2011; Bianchi, 2012; Boers et al., 2020; Goebel-Mahrle and Shin, 2020; Ramírez Cruz, 2009), the majority of the antecedents with which speakers produced N-SGA ended with their corresponding canonical vowel.

With reference to the linguistic factors that may explain the production of N-SGA in the multiple-choice test, results suggest that there is no significant interaction between the variables introduced regarding the antecedent and the agreeing elements, except when the antecedent is a non-canonical masculine noun and the agreeing element is a pronoun attached to a noun. Apparently, more N-SGA is produced with sentences including this condition.

With respect to the second research question, the hypothesis that participants would follow GA rules in the test due to it not being a spontaneous conversation is only partially confirmed. The previous section showed that the results of the grammatically expected answers were very high; the majority of participants had very few instances of N-SGA. Moreover, the results showed that the difference between variables was not statistical. However, in this case, unlike the results of the previous research question, those with a higher education level had higher ratios of expected agreement. This confirms the results found for N-SGA in other language contact situations (e.g., Díaz-Barajas and Orozco, 2019; Palacios, 2006, 2021). These results imply that the participants have a knowledge of what the Spanish grammar rules dictate for GA.

In general, it is in a way understandable not to have obtained statistically significant differences and results from the analyzed data since this is the first empirical research on this subject matter on this variety of Spanish. In fact, as Larson-Hall (2016, p. 144) explains by citing Cohen (1988), «effect sizes are likely to be small when one is undertaking research in an area that has been little studied, as researchers may not yet know what kinds of variables they need to control for». Another reason, as defended by Drager (2018) may be the limited number of participants.

With respect to the third research question, the hypothesis has been confirmed only in part. In accordance with previous studies (e.g., Ciriza, 2009), «accent» was the most frequently mentioned characteristic. However, following Ciriza's (2009) results, and Thomason's (2001, p. 10) claim that «the most common specific type of influence [between languages] is the borrowing of words», it was hypothesized that participants would mention the use of Basque words as part of the Spanish spoken in the Basque Country: the results show that this part of the hypothesis is not confirmed. This might be explained by the fact that these participants do not regularly speak Spanish in their daily lives, and consequently, they might not be very aware of introducing Basque words when speaking in Spanish. Regardless of this fact, the results in this study suggest that the population under study shows a tendency to believe that the production of N-SGA is a possibility that is present among them and, therefore, not only among elder *euskaldun zaharrak* with a low education level, contrary to Fernández Ulloa's (1997) affirmation.

These results, in any case, are sufficient to prove the original aims of this study: (i) to contradict the previously existing view of this linguistic feature as one characteristic only of elder *euskaldun zaharrak* without a high level of education (Fernández Ulloa, 1997) (despite participants' considerable knowledge of what Spanish grammar rules demand in speech), (ii) to amplify its attribution to include a broader and more accurate set of speakers, and (iii) to support the idea that N-SGA production is common among many bilinguals (Anderson, 1999; Bruhn de Garavito and White, 2002; Cruz Rico et al., 2021; Husein, 2021). This is an example of the linguistic convergence between Basque and Spanish among speakers with this profile (Matras, 2009; Moreno-Fernández, 2020), suggesting that the production of N-SGA in the Spanish spoken in the Basque Country may contradict the results of those scholars who have analyzed N-SGA in language contact situations. Overall, the results collected through the three methods demonstrate a tendency to believe that these bilinguals, despite having a high knowledge of standard Spanish GA rules, they still produce some N-SGA, and that they are quite aware of the presence of this linguistic feature in the variety of Spanish in the Basque Country in general, and in their idiolect in particular. However, more research is needed to discuss whether this linguistic phenomenon is a case of linguistic transfer or interference, as suggested by Palacios (2021) and Ramírez Cruz (2009).

## 7. CONCLUSION

The objective of the present study was to start a line of investigation on the N-SGA production by speakers of Spanish in the Basque Country from a socio-linguistic point of view (and not in mixed DPs). This study serves to demonstrate that young adult native speakers of Basque who do not typically speak Spanish and have different education degrees produce some instances of N-SGA, contrary to what scholars might have believed until now. Hence, this feature is not exclusive of elder *euskaldun zaharrak* with a low education level. The data in this paper shows that the population in question has a tendency to produce some instances of N-SGA in their speech (although apparently quite a small amount), that they know the rules of Spanish regarding GA, and that they are aware of the N-SGA production despite they do not believe that it is an important impediment to communication. Then, this feature might be a case of linguistic transfer and a community-wide established linguistic variety of the Spanish spoken in the Basque Country, although more studies should be conducted to confirm this hypothesis.

In future research, it would be convenient to gather more participants and to collect oral data through sociolinguistic interviews since, as encouraged by scholars (e.g., Palacios, 2021), natural data should be used to analyze language contact phenomena. Apart from the number of participants, future research also needs to include more diverse population in the Basque Country: speakers with different language backgrounds and education, as well as citizens from different generations, to name a few. The language in which participants received their education should be considered. Finally, by combining production and language attitudes in the future will help us obtain a more robust understanding of the linguistic phenomena arising from language contact situations. In any case, this paper should be taken as a positive and solid starting point for the analysis of GA in the spontaneous Spanish spoken in the Basque Country.

### ACKNOWLEDGEMENTS

First, I would like to thank all the participants of this study. Also, thank you to Víctor Fernández-Mallat, Lara Bryfonski, and Annie Ornelles for their valuable comments. Finally, thank you to the anonymous reviewers and any other person who has helped to develop this project in one way or another.

## APPENDIX A: GENDER AGREEMENT ITEMS IN THE MULTIPLE-CHOICE TEST

No me gusta el gobierno actual. Espero que \_\_\_\_\_ sea mejor.  
el nuevo  
la nueva  
No sé

No me ha gustado este año. Espero cambiar para \_\_\_\_\_.  
el próximo  
la próxima  
No sé

Este papel no sirve. Pásame \_\_\_\_\_ de allí.  
el blanco  
la blanca  
No sé

Hay que actualizar ese valor. Necesitamos que sea \_\_\_\_\_.  
el moderno  
la moderna  
No sé

Hay que cuidar del mundo. No se puede \_\_\_\_\_.  
abandonarlo  
abandonarla  
No sé

El tiempo es crucial. Siempre hay que \_\_\_\_\_ en cuenta.  
tenerlo  
tenerla  
No sé

No me gusta ese nombre. Creo que me gustaría \_\_\_\_\_.  
cambiarlo  
cambiarla  
No sé

Ese sector se hunde. Hay que \_\_\_\_\_.

reestructurarlo

reestructurarla

No sé

No quepo por ese paso. Para mí es \_\_\_\_\_.

estrecho

estrecha

No sé

Llevamos mucho tiempo con el proyecto. Igual es muy \_\_\_\_\_.

largo

larga

No sé

No puedo más con este mes. Espero que el siguiente sea \_\_\_\_\_.

distinto

distinta

No sé

Ese es el lugar. Creo que ya estaba \_\_\_\_\_ desde ayer.

reservado

reservada

No sé

No te gusta esta música. Seguro que prefieres \_\_\_\_\_.

el antiguo

la antigua

No sé

Esa empresa no es recomendable. Necesitamos \_\_\_\_\_.

el apropiado

la apropiada

No sé

Escucha esa voz. Tienes que decirme si es \_\_\_\_\_.

el adecuado

la adecuada

No sé

Me encanta esa capital. Pero no tanto como \_\_\_\_\_.

el primero

la primera

No sé

La vida es así. Es importante \_\_\_\_\_ a diario.

disfrutarlo

disfrutarla

No sé

La casa tiene muchos años. Debo \_\_\_\_\_.

renovarlo

renovarla

No sé

Hay que afrontar la realidad. Pero puede ser difícil \_\_\_\_\_.

superarlo

superarla

No sé

No se puede pasar por esa calle. Primero tienen que \_\_\_\_\_.

renovarlo

renovarla

No sé

No me gusta la política. Me parece muy \_\_\_\_\_.

aburrido

aburrida

No sé



Me encanta esa ciudad al sur. Es mi \_\_\_\_\_.

favorito

favorita

No sé

Hoy en la noche no veo nada. Está más \_\_\_\_\_ de lo normal.

oscuro

oscura

No sé

Me horrorizó la situación de ayer. Me pareció \_\_\_\_\_.

incómodo

incómoda

No sé

## REFERENCES

- Alarcón, I. V. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition*, 14(3), 332-350. DOI: <https://doi.org/10.1017/S1366728910000222>
- Alberdi, J. (1995). The development of the Basque system of terms of address and the allocutive conjugation. In J. I. Hualde, J. A. Lakarra & R. L. Trask (Eds.), *Towards a history of the Basque language* (pp. 275-294). Amsterdam: John Benjamins.
- Albizu Lizaso, E., & Arana Arexolaleiba, G. (2018). Bermeo eta Ondarroa: Hizkuntzaren garapenean eragiten duten faktoreak. *BAT Soziolinguistika Aldizkaria*, 107(2), 133-164.
- Andersen, R. (1984). What's gender good for, anyway? In R. Andersen (Ed.), *Second languages: A cross-linguistic perspective* (pp. 77-99). Boston: Newbury House.
- Anderson, R. T. (1999). Loss of gender agreement in L1 attrition: Preliminary results. *Bilingual Research Journal*, 23(4), 389-408. DOI: <https://doi.org/10.1080/15235882.1999.10162742>
- Badiola, L., & Sande, A. (2018). Gender assignment in Basque/Spanish mixed determiner phrases: A study of simultaneous bilinguals. In L. López (Ed.), *Code-switching – Experimental answers to theoretical questions. In honor of Kay González-Vilbazo* (pp. 15-38). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/ihll.19.02bad>
- Balam, O., Lakshmanan, U., & Parafita Couto, M. C. (2021). Gender assignment strategies among simultaneous Spanish/English bilingual children from Miami, Florida. *Studies in Hispanic and Lusophone Linguistics*, 14(2), 241-280. DOI: <https://doi.org/10.1515/shll-2021-2045>

- Barreña, A. (1997). Desarrollo diferenciado de sistemas gramaticales en un niño vasco-español bilingüe. In A. T. Pérez-Leroux & W. R. Grass (Eds.), *Contemporary perspectives on the acquisition of Spanish. Vol. 1: Developing grammars* (pp. 55-74). Sommerville (MA.) Cascadilla.
- Bates, D., & Maechler, M. (2010). Package 'lme4'. Reference manual for the package, available at: <https://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Bianchi, G. (2012). Gender in Italian–German bilinguals: A comparison with German L2 learners of Italian. *Bilingualism: Language and Cognition*, 16(3), 538–557. DOI: <https://doi.org/10.1017/S1366728911000745>
- Boers, I., Sterken, B., van Osch, B., Parafita Couto, M. C., Grijzenhout, J., Tat, D. (2020). Gender in unilingual and mixed speech of Spanish heritage speakers in The Netherlands. *Languages*, 5(4), 2–35. DOI: <https://doi.org/10.3390/languages5040068>
- Bruhn de Garavito, J., & White, L. (2002). The second language acquisition of Spanish DPs: The status of grammatical features. In A. T. Pérez-Leroux & J. Muñoz Licerias (Eds.), *The acquisition of Spanish morphosyntax* (pp. 153-178). Amsterdam: Kluwer Academic Publishers. DOI: [https://doi.org/10.1007/978-94-010-0291-2\\_6](https://doi.org/10.1007/978-94-010-0291-2_6)
- Calleja Azpiazu, N. (2004). Alineamiento fonético de acentos tonales en el castellano de Vitoria. *Estudios de Fonética Experimental*, 13, 39–63.
- Camus Bergareche, B. (2011). El castellano de San Sebastián: desarrollo y caracterización. *Oihenart*, 26, 59–101.
- Camus Bergareche, B., & Gómez Seibane, S. (2010). Basque and Spanish in a 19<sup>th</sup> century city: San Sebastián. *Ianua: Revista Philologica Romanica*, 10, 223–239.
- Camus Bergareche, B., & Gómez Seibane, S. (Eds.). (2012a). *El castellano del País Vasco*. Leioa: Euskal Herriko Unibertsitatea/Universidad del País Vasco.
- Camus Bergareche, B., & Gómez Seibane, S. (2012b). Introducción: El castellano del País Vasco. In B. Camus Bergareche & S. Gómez Seibane (Eds.), *El castellano del País Vasco* (pp. 1-18). Leioa: Euskal Herriko Unibertsitatea/Univeridad del País Vasco.
- Cenoz, J., & Perales, J. (2007). Las comunidades vascohablantes. In M. T. Turell (Ed.), *El plurilingüismo en España* (pp. 133-151). Edicions a Petició.
- Cid Abasolo, K. (2009). Bilingüismo, monolingüismo y sesquilingüismo en la Comunidad Autónoma Vasca. *Revista de Filología Románica*, 26, 229–242.
- Ciriza, M. (2009). *Dialect divergence and identity in Basque Spanish* (Publication No. 3391910) [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest Dissertations & Theses Global.
- Clegg, J. H. (2011). A frequency-based analysis of norms for Spanish noun gender. *Hispania*, 94(2), 303–319. DOI: <http://dx.doi.org/10.1133/hpn.2011.0045>
- Clyne, M. (1967). *Transference and triggering*. The Hague: Marinus Nijhoff.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum Associates.
- Cruz Rico, H., Mauder, E., Barceló-Coblijn, L., & Parafita Couto, M. C. (2021). Niños bilingües español-neerlandés en Ámsterdam y en Mallorca: evidencia de la importancia del input y del contexto en la adquisición del género gramatical en español. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas (RNAEL)*, 15(30), 14–38. DOI: <https://doi.org/10.26378/rnlael1530440>

- Cuza, A., & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition*, *19*(1), 50-68. DOI: <https://doi.org/10.1017/S1366728914000893>
- Díaz-Barajas, K., & Orozco, L. (2019). Dinámica de la concordancia de género gramatical en una variedad de español de contacto: el español purépecha. *Lengua y Migración*, *11*(2), 35-62.
- Drager, K. (2018). *Experimental research methods in sociolinguistics*. London: Bloomsbury.
- Echaide, A. M. (1968). *Castellano y vasco en el habla de Orio*. Pamplona: Institución Príncipe de Viana.
- Echenique Elizondo, M. T. (1997). Castellano y lengua vasca en contacto. ¿Hubo una lengua criolla a fines del siglo XIX en Bilbao? *Analecta Malacitana*, *20*(1), 59-71.
- Echenique Elizondo, M. T. (2004). La lengua vasca en la historia lingüística hispánica. In R. Cano Aguilar (Coord.), *Historia de la lengua española* (pp. 59-80). Barcelona: Ariel.
- Elordieta, G. (2003). The Spanish intonation of speakers of a Basque pitch-accent dialect. *Catalan Journal of Linguistics*, *2*, 67-95. DOI: <https://doi.org/10.5565/rev/catjl.44>
- Escobar, A. M. (2010). Variación lingüística en español. In J. I. Hualde, A. Olarrea, A. M. Escobar & C. E. Travis (Eds.), *Introducción a la lingüística hispánica* (2nd ed.) (pp. 391-444). Cambridge: Cambridge University Press.
- Etxague Burgos, J. (2012). El transfondo lingüístico y cultural del habla vizcaína. In B. Camus Bergareche & S. Gómez Seibane (Eds.), *El castellano del País Vasco* (pp. 21-44). Leioa: Euskal Herriko Unibertsitatea/Universidad del País Vasco.
- Etxebarria, M. (2004). Español y euskera en contacto: Resultados lingüísticos. *Revista Internacional de Lingüística Iberoamericana*, *4*, 131-145.
- Etxebarria-Arostegui, M. (2007). Sociolinguistics of Spanish in the Basque Country and Navarre. *International Journal of the Sociology of Language*, *184*, 37-58. DOI: <https://doi.org/10.1515/IJSL.2007.013>
- Etxebarria Arostegui, M. (2008). Sociolingüística de las lenguas en contacto. *Oihenart*, *23*, 77-96.
- Euskaltzaindia. (1991). *Euskal gramatika: Lehen urratsak-I*. Bilbo: Euskaltzaindia.
- Euskaltzaindia. (2002). *Euskal gramatika laburra: Perpaus bakuna*. Bilbo: Euskaltzaindia.
- Ezeizabarrena, M. J. (2009). Development in language mixing: Early Basque-Spanish bilingualism. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 57-89). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/lald.50.04eze>
- Falcón Centa, P., Chumbile Vásquez, R., & Canturín Narrea, J. (2012). Características morfosintácticas del castellano amazónico en hablantes indígenas bilingües en Lima. *Escritura y Pensamiento*, *15*(30), 79-112. <https://revistasinvestigacion.unmsm.edu.pe/index.php/letras/article/view/7993>
- Fernández-García, M. (1999). Patterns of gender agreement in the speech of second language learners. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics: Papers from the 2<sup>nd</sup> Hispanic Linguistics Symposium* (pp. 3-15). Sommerville (MA.): Cascadilla Press.

- Fernández Ulloa, T. (1997). Lenguas en contacto: caracterización del castellano del País Vasco y actitudes hacia la lengua. In J. M. Oro & J. Varela (Eds.), *Actas del I Congreso Internacional «Adquisición y aprendizaje de lenguas segundas y sus literaturas»* (pp. 199-214). Universidad de Santiago de Compostela, Campus de Lugo.
- Fernández Ulloa, T. (2006). Influencias morfosintácticas de la lengua vasca en el castellano actual. *Oihenart*, 21, 73-99.
- Finnemann, M. D. (1992). Learning agreement in the noun phrase: The strategies of three first-year Spanish students. *International Review of Applied Linguistics in Teaching*, 30(2), 121-136.
- Goebel-Mahrle, T., & Shin, N. L. (2020). A corpus study of child heritage speakers' Spanish gender agreement. *International Journal of Bilingualism*, 24(5-6), 1088-1104. DOI: <https://doi.org/10.1177/1367006920935510>
- Gómez Seibane, S. (2008). *Concordancias de género en documentación vizcaína. Siglos XV-XVI* [Unpublished doctoral dissertation]. Universidad de Deusto.
- Gómez Seibane, S. (2012). Contacto de lenguas y orden de palabras: OV/VO en el español del País Vasco. *LEA*, XXXIV(1), 5-25.
- Gutiérrez, M. J., & Silva Corvalán, C. (1993). Clíticos del español en una situación de contacto. *Revista Española de Lingüística*, 23 (2), 207-220. DOI: <https://doi.org/10.31810/RSEL.23.2>
- Guy, G. R. (2014). Words and numbers: Statistical analysis in Sociolinguistics. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 194-210). Oxford: Wiley Blackwell.
- Hernández-Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Siglo XXI.
- Hezkuntza Saila. (2020). *Hezkuntza-sistemari buruzko estatistikak*. Gasteiz: Eusko Jaurlaritza (Basque Government), Hezkuntza Saila. DOI: <https://www.euskadi.eus/informazioa/ikastetxeen-helibidetegia-kontzeptuak-eta-definizioak/web01-a2hestat/eu/>
- Hoffman, M. (2014). Sociolinguistic interviews. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 25-41). Oxford: Wiley Blackwell.
- Hualde, J. I. (2003). Finite forms. In J. I. Hualde & J. Ortiz de Urbina (Eds.), *A grammar of Basque* (pp. 205-245). Berlin: Mouton de Gruyter.
- Husein, A. (2021). Acquisition of agreement structures by Ghanaian Arabic learners. *Al-Adab Journal*, 1(137), 23-46. DOI: <https://doi.org/10.31973/aj.v1i1137.915>
- Imaz Agirre, A., & García Mayo, M. P. (2013). Gender agreement in L3 English by Basque/Spanish bilinguals. *Linguistic Approaches to Bilingualism*, 3(4), 415-447. DOI: <https://doi.org/10.1075/lab.3.4.02ima>
- Jaffe, A. (2014). Anthropological analysis in Sociolinguistics. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 213-229). Oxford: Wiley Blackwell.
- Klee, C. A., & Lynch, A. (2009). Contacto del castellano con el vasco, el catalán y el gallego en España. In C. A. Klee & A. Lynch (Eds.), *El español en contacto con otras lenguas* (pp. 27-77). Washington (DC): Georgetown University Press.

- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Laka, I. (1996). *A brief grammar of Euskara, the Basque language*. Leioa: Euskal Herriko Unibertsitatea.
- Landa, A. (1993). Los objetos nulos determinados del español del País Vasco. *Lingüística*, 5, 131-146.
- Landa, A., & Franco, J. (1996). Two issues on null objects in Basque Spanish: morphological decoding and grammatical permeability. In K. Zagana (Ed.), *Grammatical theory and Romance languages* (pp. 158-168). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cilt.133.13lan>
- Larrañaga, P., & Guijarro-Fuentes, P. (2012). The linguistic competence of early Basque-Spanish bilingual children and a Spanish monolingual child. *International Journal of Bilingualism*, 17(5), 577-601. DOI: <https://doi.org/10.1177/1367006911435704>
- Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R*. London: Routledge.
- Liceras, J. M., Fuertes, R. F., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua*, 118(6), 827-851. DOI: <https://doi.org/10.1016/j.lingua.2007.05.006>
- Lipski, J. M. (2015). From 'more' to 'less': Spanish, Palenquero (Afro-Colombian creole) and gender agreement. *Language, Cognition and Neuroscience*, 30(9), 1144-1155. DOI: <https://doi.org/10.1080/23273798.2014.975727>
- López Morales, H. (1993). *Sociolingüística*. Madrid: Gredos.
- Matras, Y. (2009). *Language contact*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511809873>
- Montrul, S. (2013). *El bilingüismo en el mundo hispanohablante*. Oxford: Wiley-Blackwell. DOI: <https://doi.org/10.5944/rdh.20.2013.12907>
- Moreno-Fernández, F. (2020). *Variedades de la lengua española*. London: Routledge.
- Moreno Fernández, F., & Otero Roth, J. (2016). *Atlas de la lengua española en el mundo* (3<sup>rd</sup> ed.). Barcelona: Fundación Telefónica & Ariel.
- Oñederra, L. M. (2002). La cabeza es para andar: sobre el castellano en zona vasca. In M. T. Echenique Elizondo & J. Sánchez Méndez (Eds.), *Actas del V Congreso Internacional de Historia de la Lengua* (pp. 263-273). Madrid: Gredos.
- Oñederra, L. M. (2004). El español en contacto con otras lenguas: español-vasco. In R. Cano-Aguilar (Ed.), *Historia de la lengua española* (pp. 1103-1115). Barcelona: Ariel.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85-110. DOI: <https://doi.org/10.1017/S0267190510000115>
- Paasch-Kaiser, C. (2015). *El castellano de Getxo*. Berlin: De Gruyter. DOI: <https://doi.org/10.1515/9783110366518>
- Palacios, A. (2006). Variedades del español hablado en América: una aproximación educativa. In E. de Miguel Aparicio & M. C. Buitrago Gómez (Eds.), *Las lenguas españolas: un enfoque filológico* (pp. 175-196). Madrid: Ministerio de Educación y Ciencia. DOI: <https://doi.org/10.4312/vh.15.2.171-172>

- Palacios, A. (2021). Sobre el contacto y los contactos: Algunas reflexiones a partir del análisis de los sistemas pronominales átonos de zonas de contacto lingüístico. In A. Palacios & M. Sanchez Paraíso (Eds.), *Dinámicas lingüísticas de las situaciones de contacto* (pp. 47-76). Berlin: De Gruyter. DOI: <https://doi.org/10.1515/9783110701364-003>
- Palacios, A., & Pfänder, S. (2014). Similarity effects in language contact: Taking the speakers' perceptions of congruence seriously. In J. Besters Dilger, C. Dermarkar, S. Pfänder & A. Rabus (Eds.), *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity* (pp. 219-238). Berlin: De Gruyter.
- Parafita Couto, M. C., Munarriz, A., Epelde, I., Deuchar, M., & Oyharçabal, B. (2015). Gender conflict resolution in Spanish-Basque mixed DPs. *Bilingualism: Language and Cognition*, 18(2), 304-323. DOI: <https://doi.org/10.1017/S136672891400011X>
- Ramírez Cruz, H. (2009). La inconcordancia de género y número en el contacto de lenguas. *Forma y Función*, 22(2), 165-195.
- Real Academia Española. (2008). *Corpus de Referencia del Español Actual (CREA) – Listado de frecuencias*. [http://corpus.rae.es/frec/1000\\_formas.TXT](http://corpus.rae.es/frec/1000_formas.TXT)
- Real Academia Española. (2016). *Nueva gramática de la lengua española: Manual*. Madrid: Asociación de Academias de la Lengua Española y Real Academia Española.
- Roca, I. M. (1989). The organization of grammatical gender. *Transactions of the Philological Society*, 87(1), 1-32. DOI: <https://doi.org/10.1111/j.1467-968X.1989.tb00617.x>
- Sandfeld, K. (1938). Les interférences linguistiques. *Rapports du 4<sup>o</sup> Congrès International des Linguistes*, 60-65.
- Schilling, N. (2013). *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511980541>
- Silva-Corvalán, C. (1994). *Language contact and change. Spanish in Los Angeles*. Oxford: Clarendon Press.
- Silva-Corvalán, C., & Enrique-Arias, A. (2017). *Sociolingüística y pragmática del español*. Washington (DC): Georgetown University Press.
- Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics*, 1, 115-132.
- Thomason, S. G. (2001). *Language contact: An introduction*. Washington (DC): Georgetown University Press.
- Trask, R. L. (2003). The Noun phrase: nouns, determiners and modifiers; pronouns and names. In J. I. Hualde & J. Ortiz de Urbina (Eds.), *A Grammar of Basque* (pp. 113-170). Berlin: Mouton de Gruyter.
- Torres Sánchez, N. (2021). Análisis de la concordancia de género y número en el español de bilingües o'dam (tepehuano del sureste) y español. In E. Blestel & A. Palacios (Eds.), *Variedades del español en contacto con otras lenguas* (pp. 57-66). Bern: Peter Lang. DOI: <https://doi.org/10.3726/b17748>
- Urrutia Cárdenas, H. (1991). Variedades lingüísticas, competencia gramatical de transición y enseñanza en una situación de bilingüismo. In H. López Morales (Ed.), *La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre Aportes de la lingüística a la enseñanza del español como lengua materna* (pp. 212-228).

- Urrutia Cárdenas, H. (1995). Morphosyntactic features in the Spanish of the Basque Country. In C. Silva-Corvalán (Ed.), *Spanish in four continents: Studies in language contact and bilingualism* (pp. 243-259). Washington (DC): Georgetown University Press.
- Urrutia Cárdenas, H. (2006). El bilingüismo en la CAV (Comunidad Autónoma Vasca): aspectos lingüísticos y educativos. *Oihenart*, 21, 481-520.
- Valdés Kroff, J. R., Rooijackers, F., & Parafita Couto, M. C. (2019). Spanish grammatical gender interference in Papiamentu. *Languages*, 4(4), 1-12. DOI: <https://doi.org/10.3390/languages4040078>
- Weinreich, U. (1966). *Languages in contact. Findings and problems*. Berlin: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110802177>
- Zubiri, I., & Zubiri, E. (2012). *Euskal gramatika osoa*. Gorliz: IkasBook, SL.

Department of Spanish and Portuguese  
Georgetown University  
3700 O St. NW  
ICC 400  
Washington, DC 20057





## HACIA UNA CARACTERIZACIÓN SINTÁCTICA DEL GÉNERO DEL SUSTANTIVO EN ESPAÑOL

Antonio FÁBREGAS<sup>1</sup>  
Universidad de Tromsø

### *Resumen*<sup>2</sup>

Este artículo presenta una propuesta sobre cómo integrar sintácticamente el género en español. La propuesta es que la mezcla de propiedades derivativas y flexivas que exhibe el género español se debe a que los exponentes relevantes (*-o* y *-a*) materializan una secuencia de núcleos que se dividen en dos áreas diferenciadas, una referida a la clasificación de entidades (clasificadores nominales) y otra referida a la definición de rasgos *phi* (el género en sentido estricto). Una vez divididas estas dos áreas, es posible dar cuenta de las nueve clases de sustantivos que el español diferencia por su correlación entre género y exponente de una forma directa que no implica disolver la distinción entre los procesos de formación de palabras y los procesos que introducen las marcas gramaticales de una categoría.

*Palabras clave:* género; clasificadores; concordancia; alternancias de género; nanosintaxis.

## TOWARDS A SYNTACTIC CHARACTERISATION OF NOMINAL GENDER IN SPANISH

### *Abstract*

This article presents a syntactic proposal about the internal structure of gender in Spanish. We argue that the mixture of derivational and inflectional properties exhibited by Spanish gender is explained by the fact that the relevant exponents (*-o* and *-a*) materialize a sequence of heads that are divided into two differentiated areas, one

---

1. antonio.fabregas@uit.no.  <https://orcid.org/0000-0001-9907-5878>.

2. Agradezco a Julio Borrego Nieto, Álvaro Recio Diego, Jaume Mateu, Raquel González Rodríguez y dos revisores anónimos sus comentarios y observaciones a versiones previas de este trabajo. Todos los errores que perseveren en el texto actual son mi exclusiva responsabilidad

referring to the classification of entities (nominal classifiers) and another referring to the definition of *phi* features (gender proper). Once these two areas are differentiated, it is possible to account for the nine classes of nouns that Spanish differentiates by their gender and morphological marking behaviour in a direct way that does not require dissolving the distinction between word formation processes and the processes that introduce the grammatical markers of a category.

*Keywords:* gender; classifiers; concordance; gender alternations; nanosyntax.

RECIBIDO: 04/06/2021

APROBADO: 16/12/2021

## 1. PROBLEMAS DE GÉNERO

El género es uno de los rasgos gramaticales de análisis más controvertido y polémico en las lenguas naturales (Hockett, 1958; Fodor, 1959; Corbett, 1991; Comrie, 1999; Dahl, 2000; Hellinger y Bussman, 2001-2003; Picallo, 2008; Acquaviva, 2009; Kramer, 2015; Audring, 2016). La controversia radica en tres aspectos fundamentales. El primero es que el género, al contrario que otros accidentes gramaticales de sustantivos o verbos, no parece condicionar otras propiedades de la misma lengua. No todas las lenguas tienen marcado de género (cf. 1), sin aparentes consecuencias globales para su estructura gramatical; por ejemplo, el español tiene género en los sustantivos mientras que el inglés no. Esto contrasta marcadamente con la presencia o ausencia de rasgos de número o tiempo en la lengua, donde se han identificado distintas consecuencias gramaticales (cf. Borer, 2005; Ramchand y Svenonius, 2006).

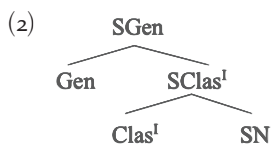
- (1) a. la niña ~ el niño  
b. the child

En segundo lugar, el género se comporta idiosincrásicamente en términos de cómo se asigna un valor a cada sustantivo o incluso de cuántos valores se reconocen dentro de cada lengua. Descriptivamente, este aspecto altamente impredecible se refiere a qué factores emplean las lenguas para determinar la asignación de género a los sustantivos –con divisiones clásicas entre rasgos formales o semánticos; cf. Corbett, 1991, 2007, 2008–, qué particiones hace el sistema de género en cada lengua. En tercer lugar, desde la perspectiva de la concordancia en rasgos formales, son frecuentes los casos de discordancia total o parcial (cf. González Alonso y otros, 2021) o los casos en que el valor de género de un sustantivo no coincide con los

rasgos que fuerza en su concordancia con adjetivos y determinantes (cf. Fábregas y Pérez, 2008; Collins y Postal, 2012). En definitiva, todos estos problemas aluden a la misma cuestión: qué estatuto gramatical cabe dársele al género dentro de un análisis explicativo de cómo se integra con el resto de rasgos gramaticales de una lengua.

La ausencia de una respuesta clara a esta pregunta ha inclinado la balanza hacia los estudios que concluyen que el género es un rasgo puramente morfológico –en el sentido de «no analizable mediante reglas gramaticales»– cuya presencia en una lengua es un mero capricho de su componente léxico, con apenas más consecuencias que forzar concordancia en algunos casos. El objetivo de este trabajo es el de contribuir a integrar el género dentro de la sintaxis de las lenguas a través de una propuesta puramente sintáctica de cómo es la estructura sintáctica interna del género en español y cómo se reflejan estas estructuras a través de sus exponentes. Si bien, al centrarnos en una sola lengua, no vamos a explorar qué consecuencias gramaticales tiene el hecho de que una lengua marque género o no, esperamos con esta contribución proporcionar algunas bases para un análisis estructural y explicativo de la marca de género en español.

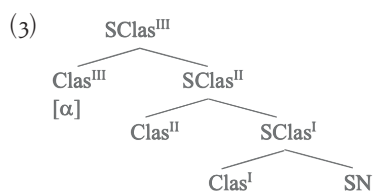
La estructura de este artículo es la siguiente: dado que nos centramos en el español, en la próxima sección presentaremos los principales problemas internos a esta lengua referidos al análisis del género en los sustantivos (§2). En la sección 3 presentaremos nuestras suposiciones teóricas sobre la naturaleza del género, su expresión gramatical y el procedimiento de inserción de vocabulario para materializar las estructuras sintácticas. La sección 4 mostrará el núcleo de nuestra propuesta sintáctica, que es la que se expone en (2).



Argumentaremos que el problema de encontrar un lugar para el género en la gramática de una lengua se disuelve cuando uno distingue mediante estructuras sintácticas diferenciadas el «género» entendido como un clasificador nominal responsable de dividir los sustantivos de una lengua y el género entendido como un rasgo funcional que contribuye a identificar la referencia de una entidad dentro de la estructura funcional del sintagma nominal. Dicho de otra manera: en español hay dos posiciones distintas –al menos– para lo que tradicionalmente se ha llamado «género»; si las dos nociones se han confundido es porque las posiciones

son adyacentes en la estructura sintáctica y se materializan sintéticamente con el mismo conjunto de exponentes morfológicos. Propondremos que ambas posiciones corresponden a distintas áreas sintácticas en la estructura funcional de un nombre (Wiltschko, 2014) tal y como se representa en su forma más básica para (2), que corresponde a un sustantivo como *perro* o *libro*: las proyecciones que llamamos «Clas(ificador)» corresponden al género como clasificador nominal, dentro del área donde el sustantivo se define como predicado y donde el clasificador contribuye a definir un conjunto de propiedades descriptivas que definen la clase de entidad denotada por él. Por el contrario el ‘género’ en su sentido funcional pertenece al área donde se introducen los rasgos formales que intervienen en la concordancia y contribuyen a denotar a un ejemplar de una clase en lugar de a la clase completa de entidades. El género tradicionalmente considerado, desde esta perspectiva, se puede entender como el exponente morfológico que da contenido léxico externalizable al material que está en estas dos regiones, el clasificador y el género gramatical como rasgo nominal.

De igual manera, el área del género como clasificador nominal es ella misma compleja, y tiene la estructura de (3), donde seguimos a Taraldsen (2010) en su propuesta de que los clasificadores nominales en lenguas de género léxico deben ser descompuestos en secuencias de núcleos sintácticos. Propondremos que los datos del español quedan explicados de forma sistemática si se admite la existencia de tres clasificadores nominales, cuya diferencia depende exclusivamente del tamaño del constituyente sintáctico que materializan y de la información formal que codifican.



En la sección 5 nos concentraremos en la materialización de las estructuras de (2) y (3) para derivar los patrones de marcación de género del español (cf. tablas 1 y 2 *infra*). Mostraremos que las aparentes discordancias entre género y marca morfológica pueden analizarse a partir de la interacción entre el material lexicalizado por cada uno de los exponentes y la elección de un valor particular para el clasificador. Finalmente, la sección 6 presenta las implicaciones y predicciones más amplias del análisis, y concluye este trabajo.

## 2. PROBLEMAS DE GÉNERO EN ESPAÑOL

Para el caso del español, existen varias cuestiones polémicas que los investigadores han tratado de dilucidar tanto desde perspectivas internamente morfológicas como desde perspectivas más orientadas hacia la sintaxis. Estos problemas son la manifestación específica en español de las preguntas teóricas que se acaban de exponer y pueden resumirse en tres clases, que se presentan a continuación.

### 2.1. *El problema del estatuto gramatical del género*

Se ha propuesto repetidamente que el género en español parece desempeñar un doble papel, incluso si nos restringimos a su uso con sustantivos (Ambadiang, 1994; Murillo, 1999; Lliteras, 2008; Serrano Dolader, 2010; Mendívil, 2020). En los pares de (4) la diferencia de género es predecible semánticamente, y no altera la denotación del sustantivo, por lo que ha sido considerada flexiva, bajo una definición específica de flexión donde se incluyen en ella los procesos que alteran las propiedades formales de una palabra pero no su información léxica.

(4) niño ~ niña, perro ~ perra, abuelo ~ abuela, prisionero ~ prisionera...

En contraste, en pares como (5) la alternancia de género va asociada a un cambio no sistemático en la denotación de la palabra –por ejemplo «árbol» vs. «fruta del árbol»–. Esta diferencia de contenido es suficiente para tratar al género como un rasgo léxico, y por lo tanto parte de la morfología derivativa, en este segundo conjunto de ejemplos.

(5) manzano ~ manzana, cesto ~ cesta, cargo ~ carga, río ~ ría, madero ~ madera...

Por tanto, la cuestión es si el género debe entenderse como un rasgo flexivo, uno derivativo o una mezcla de ambos; alternativamente, la pregunta puede plantearse desde una perspectiva más general: ¿es el género un rasgo funcional, sin consecuencias para la clase de concepto que debe expresar una palabra, o es un rasgo léxico que nos ayuda a definir el contenido de un sustantivo pero hace poco o nada en su estructura funcional? En muchos casos, como Serrano Dolader (2010), la respuesta parece ser mixta: el género debe ser ambas cosas. El problema es que en tal caso estamos forzando la disolución de la división entre propiedades léxicas o no predecibles de una estructura y las propiedades gramaticales, regidas por reglas, de la misma estructura, algo que no consideramos un resultado deseable.

## 2.2. *El problema de la correspondencia entre morfema y género*

Dejando a un lado el problema de cómo se predice qué género se asigna a un sustantivo, en español el principal problema que suscita la identificación del género es que no siempre hay correspondencia formal entre la terminación y el valor de género que tiene un sustantivo. La inmensa mayoría de los sustantivos acabados en *-o* son masculinos (6a), y los que acaban en *-a* son femeninos (6b), pero existen excepciones en las dos direcciones (7).

- (6) a. cosa, mata, rosa, puerta, ventana, lluvia...  
b. cazo, martillo, libro, tomo, plato, agobio...
- (7) a. día, problema, planeta, idioma, cura, marica, artista, colega, mapa, Papa, pijama, tranvía, tanga, emblema, idiota, sibarita, cosmopolita...  
b. mano, radio, foto, moto, disco, seo, nao, libido, dinamo, polio  
c. la testigo, la médico, la abogado, la soprano

Desde Harris (1991), esto ha llevado a una distinción entre las nociones de «marca de palabra» y «género», que tiene un considerable apoyo tipológico: los morfemas visibles se pueden considerar marcas de palabra o clasificadores nominales y no marcas de género (cf. Kramer, 2015), por lo que casos de (7) se tratan como excepciones léxicas en las que la marca de palabra no refleja transparentemente el género de la forma.

Con todo, los casos de cada una de las subclases mostradas en (7) son asimétricos desde varias perspectivas, la principal de las cuales es que la cantidad de sustantivos de (7a) es considerablemente mayor que los de (7b), que se reducen generalmente a menos de una docena de sustantivos. Sin embargo, la clase se incrementa algo en las variedades, cada vez menos extendidas, en que el femenino de ciertos sustantivos animados en *-o* conserva la marca de palabra del sustantivo (7c).

Los casos de (7a) y los casos de (7b-c) no son equivalentes tampoco en términos morfológicos. Como veremos, la marca de palabra en los casos de *-a* masculina (7a) es estable cuando el sustantivo se combina con el diminutivo *-it-* (*el problem-it-a*)<sup>3</sup>,

---

3. Nótese que estamos hablando de un diminutivo *-it-* y no del comportamiento de estos sustantivos con los morfemas apreciativos en general. El motivo es que no todos los morfemas apreciativos tienen el mismo comportamiento; así, *-azo* o *-ucho* imponen su propia marca de palabra regular en todos los casos (*problemazo*, *el artistucho*). El diminutivo *-it-* no impone su propia marca de palabra, incluso al unirse a adverbios con marca de palabra compleja (*lejos > lejitos*); posiblemente la causa última de esta asimetría es que *-it-* tiene el comportamiento de un especificador que no proyecta estructura por encima del sustantivo léxico con el que se combina, mientras que otros morfemas apreciativos se comportan como núcleos que toman al SN como complemento y proyectan sus propios clasificadores en la estructura (Fábregas, 2013).

incluso en casos animados (*el cur-it-a*, *el colegu-it-a*), mientras que la vocal *-o* en (7b) es más inestable en esos mismos contextos (*la man-it-a*) y decididamente suele reemplazarse en los casos animados (*la testigu-it-a*). Esto nos llevará a proponer asimetrías estructurales entre las configuraciones que subyacen a (7a) y a (7b-c).

La propuesta de diferenciar marcas de palabras y género ha sido un primer paso para explicar los casos de (7). Sin embargo, no se ha integrado estructuralmente de forma suficiente –al surgir en contextos lexicalistas, no existen propuestas sintácticamente explícitas que discutan qué estatus tiene cada una de las dos nociones ni cómo se combinan entre ellas–, ni tampoco ha permitido hasta ahora establecer generalizaciones que vayan más allá del terreno de las excepciones léxicas para dar cuenta de estos casos.

Para complicar las cosas más aún, el español tiene dos clases más de sustantivos que tradicionalmente se distinguen por la aparente marca de palabra que contienen: aquellos acabados en *-e* átona (8) y aquellos acabados en consonante (9) o vocales distintas de *-a*, *-o* o *-e* átonas (10), que se consideran de marca de palabra cero. Existen términos tanto masculinos como femeninos que pertenecen a estas dos clases.

- |      |   |            |
|------|---|------------|
| (8)  | a. noche, clase, nieve, tarde, liebre...        | femeninos  |
|      | b. coche, cable, diente, baile, cliente...      | masculinos |
| (9)  | a. ración, moral, nuez, avestruz, sal, mujer... | femeninos  |
|      | b. papel, lápiz, reloj, túnel, camión, arroz... | masculinos |
| (10) | a. grey, mamá, ley, hurí, tribu...              | femeninos  |
|      | b. rey, sofá, café, alhelí, tabú...             | masculinos |

Una propiedad sistemática de estos sustantivos es que, en el contexto del diminutivo en *-it-*, se combinan con las marcas de palabras prototípicas de masculino y femenino (*noch-ecit-a*, *papel-it-o*, *rey-ecit-o*...). Si fuera cierto que estos sustantivos tienen *-e* o *-ø* como marca de palabra, esta sería inestable en estos sustantivos, al igual que la clase de *mano* ~ *manita*.

### 2.3. *El problema de las alternancias morfológicas en correlación con el género*

El tercer problema es que, pese a la aparente idiosincrasia y arbitrariedad de las terminaciones de género en español, en los sustantivos variables en género hay una tendencia muy marcada a obtener una regularidad máxima en femenino.

No todos los sustantivos alternan entre los dos géneros, y cuando lo hacen la morfología usada para expresar estas alternancias no siempre funciona de la misma manera. En cuanto a la correlación entre género y marca de palabra cuando el

sustantivo puede producir concordancia en masculino o femenino, se documentan varios subcasos: alternancia entre *-o* y *-a*, donde la segunda es femenina (11a), alternancia entre *-e* y *-a*, donde la segunda también es femenina (11b), alternancia entre ausencia de marca y *-a*, de nuevo con la misma consideración de género para la segunda (11c), y *-a* común para masculino y femenino (11d).

- (11) a. gato ~ gata  
 b. presidente ~ presidenta  
 c. juez ~ jueza  
 d. colega ~ colega

De forma interesante, en estos cuatro pares, si bien el masculino difiere en su marca de sustantivo, se produce un sincretismo en el femenino, donde en los cuatro casos se emplea *-a* para este valor, que además siempre se interpreta composicionalmente. Este sincretismo se clasifica como «horizontal» (Caha y Pancheva, 2012) porque implica dos dimensiones gramaticales: la marca de palabra y el valor de género que fuerza en la concordancia.

	Masculino	Femenino
gato ~ gata	-o	-a
presidente ~ presidenta	-e	
juez ~ jueza	-ø	
colega ~ colega	-a	

Tabla 1. Sincretismo en femenino

El problema de las alternancias de género está complicado por la existencia de otros casos –algunos de los casos tradicionalmente considerados de sustantivos comunes en cuanto al género–, donde al igual que en la clase (11d) el género se expresa sin alteración morfológica de la palabra, donde el femenino es idéntico al masculino. Esto incluye sustantivos acabados en *-o* (12a), *-e* (12b) y consonantes (12c).

- (12) a. el testigo ~ la testigo  
 b. el cónyuge ~ la cónyuge  
 c. el fiscal ~ la fiscal

De forma interesante, los pares de (12a-c) tienen una tendencia creciente a convertirse en pares alternantes, y por tanto a asimilarse a los casos de (11). En esta propiedad también se puede encontrar otra asimetría entre las palabras masculinas



acabadas en *-a* y las femeninas acabadas en *-o* en (7): mientras que pares como (12a) tienden a regularizarse (*el testigo ~ la testiga*), pares como los de (11d), en cambio, no se regularizan a formas marcadas como masculinas (*\*el colega ~ la colega*).<sup>4</sup>

Todas las propiedades y contrastes que se acaban de notar dan lugar a nueve grupos de sustantivos en español, dependiendo de los factores que hemos discutido hasta ahora: si alternan entre masculino y femenino, si la marca de palabra que tienen refleja transparentemente el valor de género o no, si el femenino se interpreta composicionalmente o influye en el significado idiomático del sustantivo y si la marca de palabra se preserva o no con el diminutivo *-it-*. Nos proponemos en este trabajo analizar los nueve grupos, que se resumen en la siguiente tabla.

Grupo	Alternancia de género	Marca de palabra regular o no	Género composicional o idiomático	Preservación de la marca de palabra
a) niño - niña	Sí	Regular	Composicional	Sí
b) colega, idiota	Sí	No ( <i>-a</i> en masculino)	Composicional	Sí
c) planeta, día	No	No ( <i>-a</i> en masculino)	Idiomático	Sí
d) mesa, gorra	No	Regular	Idiomático	Sí
e) techo, gorro	No	Regular	Idiomático	Sí
f) jefe - jefa / el jefe - la jefe	Sí	No ( <i>-e</i> en masculino, posiblemente en femenino)	Composicional	No ( <i>jef-ecit-o</i> )
g) juez - jueza / el juez - la juez	Sí	No ( <i>-ø</i> en masculino, posiblemente en femenino)	Composicional	No ( <i>juec-ecit-o</i> )
h) coche, papel, hurí	No	No ( <i>-e o -ø</i> en masculino o femenino)	Idiomático	No ( <i>coch-ecit-o, papel-it-o</i> )
i) mano, foto, polio	No	No ( <i>-o</i> femenina)	Idiomático	Generalmente no ( <i>man-it-a, testigu-it-a</i> ), pero sometida a variación ( <i>la man-it-o, la fot-it-o</i> )

Tabla 2. Nueve grupos de sustantivos por su marca de género en español

4. Un revisor anónimo, al que agradecemos la observación, apunta la existencia de casos como *azafato* y *modisto*, donde la forma masculina ha sido creada históricamente después de la femenina. Desde la perspectiva sincrónica que adoptamos, y sin negar el interés del estudio de qué reordenamiento se produjo históricamente con estos términos, consideramos que los hablantes que usan productivamente estas formas masculinas tratan *azafato* y *modisto* como nombres del primer grupo de sustantivos, igual que *niño*.

Nótese de entrada que la alternancia de género coincide con la composicionalidad en su lectura. Pero antes de pasar al análisis, haremos explícitas nuestras suposiciones teóricas en la siguiente sección.

### 3. BASES TEÓRICAS: GÉNERO, CLASE NOMINAL Y MATERIALIZACIÓN DE SINTAGMA

Esta sección se dedica a presentar las bases teóricas que asumiremos en este trabajo: comenzaremos presentando qué clase de papel asumimos para el género dentro de la estructura nominal (§3.1), para luego centrarnos en la distinción entre clase nominal y género (§3.2), presentar la teoría sintáctica y de rasgos que asumimos (§3.3) y detallar el procedimiento de inserción léxica que empleamos en el estudio (§3.4).

#### 3.1. *La naturaleza del género entre las propiedades gramaticales del sustantivo*

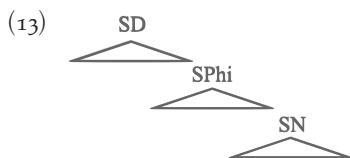
El género es un procedimiento que algunas lenguas emplean para establecer divisiones dentro del dominio nominal. Las divisiones pueden ser de muy distinta naturaleza y atender a distintos factores conceptuales no siempre predecibles: entes animados y no animados, machos y hembras, ríos y montañas, objetos comestibles, plantas, fuego y otras cosas peligrosas, y, en definitiva, una gran variedad de conceptos que con todo probabilidad no deben estar representados directamente en la estructura sintáctica de un sintagma nominal y deben ser tratados como categorías semánticas conceptuales (Aikhenvald, 2000; Gil, 2011).

Sean como sean esas categorías conceptuales, toda lengua con género emplea el género para hacer divisiones dentro de un dominio semántico definido por las propiedades que denota el sustantivo. Déchaine (2018) toma esta idea de la partición como el centro de su teoría y propone que el género no es más que un introductor de una función de partición que divide las raíces dentro de los contextos nominales en distintas clases. Déchaine (2018: 17) define partición de la siguiente manera (mi propia traducción).

Una función de partición aplicada a un conjunto  $X$  es el conjunto de los subconjuntos no vacíos de  $X$  tal que cada elemento  $x$  perteneciente a  $X$  está exactamente en uno de esos subconjuntos. La función de partición es exhaustiva –cada elemento  $x$  del conjunto  $X$  está clasificado– y no intersectiva –cada elemento  $x$  del conjunto  $X$  es asignado a solo un subconjunto–.

De esta manera, el género en una lengua que lo posee no hace más que tomar el conjunto potencialmente abierto de sustantivos léxicos y lo divide dentro de subconjuntos que, una vez agrupados, definen clases nominales. Un conjunto de seis miembros puede ser particionado de distintas maneras –por ejemplo definiendo seis subconjuntos unarios, o tres subconjuntos de dos miembros cada uno–, lo cual implica que cada lengua puede definir arbitrariamente cuántas clases de género tendrá.

Déchaine (2018: 18) destaca que esta función de partición no es un rasgo especial del género, sino que es compartido con todas las proyecciones funcionales que dominan a un elemento léxico. Cada proyección funcional introduce una partición basada en el contenido gramatical o semántico que contienen –por ejemplo, el número establece particiones basadas en la idea de pluralidad, singularidad, dualidad, etc–. Así, el género no es especial y tampoco esperamos que tenga una localización única dentro de la estructura funcional del sustantivo. Supongamos, como ilustración, una estructura funcional nominal con tres áreas, como hace Wiltschko (2014).



En estas tres áreas se desempeñan funciones distintas: el área llamada del sintagma nominal o SN es el área que se encarga de clasificar entidades por las propiedades léxicas y descriptivas del sustantivo; es el área en que se ha localizado el género en las lenguas indoeuropeas, que luego se han tomado como modelo arbitrario para intentar identificar la misma propiedad en otras lenguas. El área que representamos en (13) como SPhi, y que Wiltschko (2014) asocia más generalmente con los rasgos funcionales nominales –rasgos *phi*, donde está mínimamente el género, y tal vez la persona o, como en nuestro caso, el género gramatical– da una perspectiva a las clases consideradas, por ejemplo en términos de las agrupaciones que se forman a partir de ellas, o restringe su referencia solo a los miembros hembra del grupo. Dicha perspectiva se toma como punto de partida para definir la posible referencialidad de la estructura, algo a lo que se dedica el área del sintagma determinante (SD), que ancla la descripción proporcionada por el SN al contexto referencial.

La función de partición que se ha llamado tradicionalmente «género» puede aparecer en cualquiera de estas tres áreas a través de las lenguas. Así podemos tener lenguas donde la partición de género tiene lugar en el nivel del sintagma determinante

(Luraghi, 2011), donde el género interactúa con la definitud y topicalidad, como en Plains Cree, donde los sustantivos tienen género variable dependiendo de su estatus informativo y su referencialidad. En árabe (Acquaviva, 2009, 2018), el género se enlaza con el número y la contabilidad, de forma que el género se emplea para hacer particiones que afectan a la naturaleza de los grupos y cantidades que se construyen sobre la denotación del nombre. En español, en cambio, el género claramente es un fenómeno que está en la órbita del sustantivo como elemento léxico, estableciendo diferencias entre clases de entidades, lo cual lleva a autores como Picallo (2008) o Kramer (2015) a proponer que es una propiedad del dominio del SN.

Desde esta perspectiva, el género no es más que el nombre que se ha dado tradicionalmente a una función de partición en las lenguas clásicas más estudiadas –latín, griego, y otras lenguas indoeuropeas–, para dar nombre a dicha función cuando se sitúa en las capas más bajas de la estructura del sintagma nominal, concretamente en el área del SN, pero en tanto que función de partición no es distinta de otras proyecciones funcionales que igualmente dividen a las entidades del conjunto en subclases.

Lo que hace especial al género en español, como ya se ha notado en muchos estudios, es que esa partición divide las entidades del conjunto de una manera que puede tener implicaciones tanto para el contenido conceptual del sustantivo como para su comportamiento gramatical, como notan Ambadiang (1994), Serrano-Dolader (2010) o Mendivil (2020) para el español, con el resultado de que la clasificación del género oscila entre lo derivativo y lo flexivo. Dedicemos a este problema la siguiente subsección.

### 3.2. *Clasificador nominal y género*

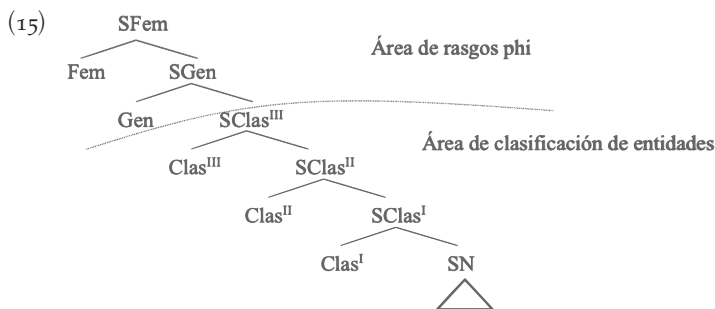
Pazó (1989) y Harris (1991) discuten en detalle los casos en que la marca morfológica de un sustantivo en español no refleja de forma transparente la concordancia de género que ese sustantivo impone a los adjetivos y determinativos –*el dí-a, la man-o*–. Para resolver el problema, ambos proponen una misma solución: la de diferenciar el género en tanto que clasificador nominal –la llamada «marca de palabra»– y el género en tanto que rasgo sintácticamente activo que entra en procesos de concordancia. Así, los morfemas átonos finales que aparecen en los sustantivos –y también adjetivos o algunos adverbios– del español no corresponderían a la marca de género, porque no se asocian siempre con el masculino o femenino, sino que serían marcas de palabra que actúan como clasificadores nominales. Esta teoría encaja en principio con la

propuesta de Déchaine (2018), y adoptaremos aquí sus rasgos esenciales: el género, como accidente gramatical tradicional, no es una noción primitiva, sino que debe ser descomponible al menos en dos partes, su papel léxico –el valor como clasificador nominal– y su papel funcional –el rasgo que produce la concordancia con otras formas dentro del sintagma nominal–.

La distinción entre clasificador y género ha tenido éxito en los estudios tipológicos, donde se ha observado repetidamente que en los casos en que hay discordancia entre la marca del sustantivo y la flexión, la concordancia siempre copia el género de la forma, nunca su clasificador:

- (14) a. la man-o izquierd-a
- b. \*l-o man-o izquierd-o

Nuestra propuesta es la de diferenciar estructuralmente entre estas dos nociones, la marca de palabra –clasificador nominal– y la información de género. En nuestra propuesta cada una de estas dos nociones está representada en la estructura mediante dos secuencias de núcleos sintácticos que se ordenan de forma rígida entre ellas, como sigue: la secuencia del género por encima de la secuencia del clasificador nominal, siendo ambas, como veremos, descomponibles en más de un núcleo.



La región de los SClas en (15) se descompondrá aquí en tres núcleos (Clas<sup>I</sup>, Clas<sup>II</sup>, Clas<sup>III</sup>). Pertenecen los tres al área de (13) donde se desempeña la función de clasificar las entidades a través de sus propiedades denotativas, en colaboración con SN, que a su vez puede ser descompuesto en distintos núcleos sintácticos en los que no entraremos en este trabajo, pero donde suponemos que se sitúan los morfemas apreciativos y los nominalizadores. El clasificador nominal, pues, tendrá efectos en la clasificación conceptual –no sintáctica– del sustantivo, y establecerá relaciones de legitimación léxica de las raíces contenidas en el sustantivo, como veremos a continuación.

La región que se inicia con SGen en (15), y que a su vez descompondremos en dos capas (SFem, SGen), pertenece al área donde se incluyen los rasgos *phi* de la estructura nominal y se define la perspectiva con la que se presenta la clase de entidades definida en el área inferior. Es la presencia de los núcleos del área de Gen y su contenido de rasgos lo que motiva la concordancia de adjetivos y determinantes.

Por tanto, el español a través del género establece una doble partición en el dominio nominal:

- a. A través de los clasificadores nominales, el género establece particiones conceptuales no composicionales que manipulan la información denotativa de la raíz.
- b. A través del género como rasgo *phi*, el género establece una partición secundaria sobre el concepto expresado por la suma del clasificador y la raíz, que se interpreta siempre composicionalmente como las hembras dentro de una clase de entidades.

Desde este punto de vista, la marca de género en español puede verse como el elemento que, en el dominio nominal, enlaza el área de clasificación –donde se define el contenido del sustantivo léxico– con el área funcional más alta –los rasgos nominales que entran en relaciones de concordancia con otros elementos–<sup>5</sup>.

Con respecto a la combinación de los clasificadores nominales con las distintas raíces que se categorizan como sustantivos, nuestras asunciones son las siguientes. Partimos de la teoría de que las raíces son elementos sintácticos acategoriales (Marantz, 1997, 2000; Embick, 2000) que carecen por completo de rasgos sintácticos o semánticos por sí mismos (Arad, 2005; Borer, 2013), y se reducen a actuar como meros índices fonológicos que permiten la selección de exponentes morfológicos

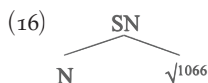
---

5. Doy por hecho, por tanto, que los adjetivos también concuerdan con las proyecciones del área de rasgos *phi* y no con los clasificadores, como se verá después. El orden entre las dos áreas, al ser rígido, explica que la concordancia no copie nunca el clasificador nominal, sino el valor de género: incluso si el clasificador contiene un rasgo interpretable para establecer concordancia, la información que posee el género en el área de rasgos *phi* está por encima de ella y por lo tanto será la que sea copiada por los determinantes y otros elementos concordantes.

(i) [SGen Gen<sub>[a]</sub>][SClas Clas<sub>[b]</sub>][SN]]

Los modificadores y determinantes se introducen por encima de SGen –Cinque (2010)–. Desde esa posición, cuando el adjetivo o determinante busca un rasgo de género para satisfacer su concordancia, encuentra el valor más próximo en Gen, que está estructuralmente ante el elemento concordante y Clas. Esto es una manifestación del principio de Minimidad Relativizada (Rizzi, 1990), y explica directamente por qué la concordancia en español es con el género y no con la clase nominal.

individuales, como *gat-*, *perr-*, *cas-*. Al igual que Borer (2013), representamos el índice fonológico como un cardinal superescrito a la raíz, que será interpretable solo cuando se busquen los exponentes dentro del repertorio léxico al momento de materializar la estructura sintáctica. Para convertir la raíz en un sustantivo, es necesario formar una estructura sintáctica en la que sobre la raíz se proyecte la estructura funcional del sustantivo, mínimamente N como categorizador.



La ausencia de categoría gramatical en la raíz hace que no podamos codificar en la raíz rasgos no interpretables que pidan un clasificador nominal concreto. En lugar de esto, seguiremos a Harley y Noyer (2000) y a Acquaviva (2009), quien –al analizar el género en italiano– propone que la relación que se da entre la raíz y el clasificador es una de legitimación en el contexto del clasificador. Es decir: la raíz carece de información que pida la presencia de estructura nominal, pero sí está asociada léxicamente con una entrada donde se restringe su distribución superficial solo a contextos en que aparece dominada por un clasificador con un valor determinado.

Así, en aras de la concreción, un exponente como *plátan-* estaría asociado a una entrada como la de (17), donde se indica que este morfema solo está legitimado como materialización del SN si sobre N está el clasificador al que llamaremos I.

(17) *plátan-* Contexto de legitimación: [+Clas<sup>I</sup>]

Un exponente como *problem-*, por su parte, quedará legitimado mediante otro clasificador, al que llamaremos en este trabajo ClasII.

(18) *problem-* Contexto de legitimación: [+Clas<sup>II</sup>]

A la vez, la combinación del clasificador y el exponente usado para el nombre léxico formará una construcción idiomática con significado conceptual no predecible; como ya adelantaban Marantz (1995) y Acquaviva (2009), todos los nombres léxicos forman construcciones idiomáticas con sus clasificadores, de manera que los valores del clasificador codefinen el concepto que expresa el nombre léxico.

Como se verá más adelante, suponemos que todos los exponentes nominales deben restringir su distribución de la misma manera –esto es, no existen exponentes nominales del español que no tengan una restricción para el clasificador–, tal y como se espera de la función de partición, que debe ser exhaustiva. Esto quiere decir que

trataremos los nombres variables (*gato ~ gata*) como elementos que en realidad contienen siempre el mismo clasificador, y donde la diferencia que se percibe en la superficie en la marca de palabra deriva enteramente del valor del género como rasgo *phi*, no del clasificador.

### 3.3. *Suposiciones sobre la concordancia*

Nuestra propuesta parte de la idea básica de que el género tradicionalmente considerado debe descomponerse en dos áreas que pertenecen a distintas regiones de la estructura funcional del sustantivo: un clasificador nominal y un género propiamente dicho. El género en este sentido restrictivo produce concordancia con adjetivos y determinativos. Por ello, seamos explícitos acerca de nuestras suposiciones sobre la concordancia.

Asumimos que la concordancia es una relación sintáctica, no morfológica, que se produce a distancia –no necesita movimiento explícito (cf. Zaenen, Maling y Thráinsson, 1986)– entre dos elementos, la sonda y la meta. La sonda es el elemento que busca una meta para satisfacer su concordancia. Suponemos, con Béjar y Rezac (2009), que la relación de concordancia es bidireccional: la sonda en principio está por encima de la meta y busca en su dominio estructural un elemento que contenga el rasgo que trata de satisfacer; si no lo encuentra, prueba a buscar en el dominio más alto la meta que le dé valor.

El modelo habitual hasta mediados de los 2000 para explicar la concordancia era el llamado «modelo de la desactivación de la bomba». Su idea central es que la concordancia surge de la asimetría entre rasgos interpretables / valorados y rasgos no interpretables / no valorados (Chomsky, 2000). El elemento concordante –la sonda– contiene rasgos no interpretables para la semántica y la fonología por carecer de valor, y busca cotejarlos con una meta que contiene la versión valorada e interpretable de esos rasgos. Si esos rasgos no interpretables no se valoran durante el curso de la construcción de la estructura sintáctica, el resultado es la agramaticalidad de la estructura resultante, como si esos rasgos hubieran sido el detonador de una bomba que debía ser desactivada.

No obstante, la condición de los rasgos fue revisada en la bibliografía posterior (Pesetsky y Torrego, 2007), y como resultado de esto también la teoría sobre la concordancia (Preminger, 2011, 2014). Preminger (2014), que define el modelo de concordancia que seguiremos aquí, disocia por completo la interpretabilidad de la valoración permitiendo que una derivación sea convergente sin valor, mientras



sea interpretable. Si una sonda con un rasgo interpretable pero sin valor (19a) no encuentra en la derivación sintáctica una meta con un valor para el rasgo, este toma un valor por defecto antes de la transferencia a las interficies (19b).

- (19) a. [iF: \_\_\_\_\_]  
 b. [iF: valor DEF]

Por ejemplo, esta es la manera en que Preminger (2014) explica la emergencia de la concordancia neutra cuando el adjetivo toma como sujeto de predicación una cláusula u otro elemento no nominal: ya que la cláusula carece de información de género –en nuestra perspectiva, porque no proyecta una estructura funcional nominal–, el género del adjetivo no puede valorarse. Esto no produce agramaticidad, sino que hace que se introduzca el valor por defecto para ese rasgo. (20) ilustra esta valoración por defecto en género neutro para el caso del noruego bokmål, que diferencia masculino y femenino (20a) frente a neutro (20b) con sus terminaciones morfológicas, manifestado con el exponente *-t*.

- (20) a. Bokene           dine    er   god-e  
       libro.f-los       tuyos   son   buen-f.pl  
       ‘Tus libros son buenos’  
       a. At du skriver er god-t.  
       que tú escribes es buen-neut  
       ‘Que escribas es bueno’

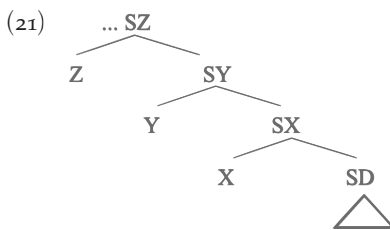
### 3.4. *Procedimiento de lexicalización*

La última pieza que emplearemos en nuestro análisis del género en español es el algoritmo de materialización que asumimos en este trabajo, que es el que se ha desarrollado en nanosintaxis (Starke, 2009, 2011, 2018; Caha, 2009; Abels y Muriungi, 2008; Fábregas, 2007, 2016, 2020; Taraldsen, 2010; Vanden Wyngaerd, 2018; Baunaz y Lander, 2017, 2018). La nanosintaxis es un modelo de análisis sintáctico que se construye sobre tres principios:

- (i) No hay un componente morfológico independiente. La sintaxis es la responsable de construir todas las estructuras gramaticales. Esto incluye las supuestas matrices de rasgos: en lugar de proponer matrices complejas de rasgos, la nanosintaxis construye las matrices de rasgos mediante la combinación de núcleos sintácticos independientes.

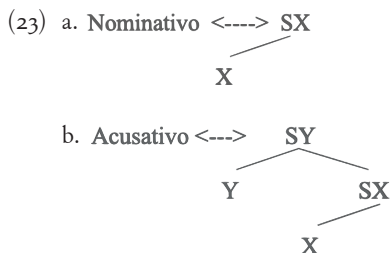
- (ii) Las estructuras sintácticas son materializadas por exponentes de vocabulario que corresponden a estructuras sintácticas (Principio de la Materialización de Sintagma)
- (iii) La materialización de las estructuras sintácticas por medio de exponentes no puede dejar ningún rasgo sintáctico sin identificar (Principio de Lexicalización Exhaustiva).

Ilustremos estos principios con los exponentes de caso, con el razonamiento tomado de Caha (2009). Caha (2009) propone que el caso no ha de verse como una propiedad léxica o morfológica, sino más bien como el resultado de la combinación sintáctica de un conjunto de núcleos que siguen una jerarquía bien establecida, como la de (21), que dominan directamente a los determinantes (cf. Fillmore, 1968). Se rechazan así estructuras léxicas en las que los rasgos se combinan de forma compleja en el interior de un solo núcleo sintáctico indescomponible, como las de (22).



- (22) a. [X]  
 b. [Y, X]  
 c. [Z, Y, X]

Los casos tradicionales (nominativo, acusativo, genitivo...) se obtienen incrementando monotónicamente la secuencia de núcleos sintácticos. En la suposición de que el caso menos complejo sintácticamente es el nominativo (Blake, 2000), un exponente que manifiesta caso nominativo sería un exponente que materializa sólo el SX (23a). El acusativo –suponiendo que este es el siguiente caso en complejidad, como de hecho propone Blake (2000) y confirma Caha (2009)– se asociaría a un exponente que materializa el sintagma SY y el sintagma SX al mismo tiempo, es decir, tiene la entrada de (23b). El dativo sería un exponente que materializa SZ con SY y SX, y los casos más complejos de la jerarquía de Blake (2000) implicarían todos la construcción de capas funcionales adicionales que se suman a las anteriores.



Esto tiene dos consecuencias: (i) los exponentes materializan no núcleos, sino configuraciones sintácticas que corresponden a sintagmas sintácticamente contruidos, porque la sintaxis es la encargada de unir los rasgos gramaticales para formar estructuras complejas (21), ya que el léxico no puede ser generativo y (ii) la entrada de un exponente se asocia a una estructura sintagmática, como las de (23). Este segundo principio es enunciado como el Principio de la Materialización de Sintagma o *Phrasal Spell Out*.

(24) Principio de Materialización de Sintagma (PMS)

La entrada de un exponente asocia al exponente no a un conjunto de rasgos contenidos en una terminal sintáctica, sino a una estructura sintagmática donde los rasgos aparecen organizados dentro de una configuración.

La propuesta de que los exponentes se asocian a estructuras sintácticamente complejas permite explotar como parte del mecanismo de materialización las propiedades del constituyente complejo que cada exponente identifica. Concretamente, la nanosintaxis parte de la hipótesis antilexicalista de que la inserción del vocabulario que se emplea para materializar una forma está siempre subordinada a las propiedades de la estructura sintáctica que materializa. Así, los exponentes no pueden eliminar, ignorar o sobrescribir la información sintáctica de la estructura, como sí sucede por ejemplo en la Morfología Distribuida (Noyer 1997), donde los rasgos sintácticos pueden ser ignorados por los exponentes. En nanosintaxis no hay operaciones morfológicas, por lo que el vocabulario debe respetar la información de la sintaxis y no más. La nanosintaxis formula este principio de la siguiente manera (Fábregas, 2007).

(25) Principio de Lexicalización Exhaustiva (PLE)

Todo rasgo sintáctico debe ser identificado por la inserción de un exponente.

Veamos entonces cómo funciona la materialización exacta de una estructura sintáctica. Supongamos el caso del ruso, comenzando por un caso que distinga los tres casos relevantes en nuestra discusión, que serán nominativo, acusativo y dativo.

- (26) a. gero-i  
 héroe-nom  
 b. gero-ja  
 héroe-ac  
 c. gero-ju  
 héroe-dat

La idea es que para esta declinación, el ruso tiene (al menos) tres exponentes, cuyas entradas son las siguientes:

- (27) a. -i <----> SX  
 X
- b. -ja <----> SY  
 Y SX  
 X
- c. -ju <----> SZ  
 Z SY  
 Y SX  
 X

Supongamos que la sintaxis produce la estructura de (28), para que sea lexicalizada.

- (28)
- 
- ```

  graph TD
    SY --- Y
    SY --- SX
    SX --- X
    SX --- SD
    style SD fill:none,stroke:none
  
```

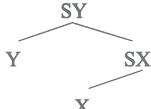
Esta es la estructura del caso acusativo, que se compone de dos capas de rasgos de caso. El PLE en (25) exige que cada parte esté identificada por un exponente. Supongamos que SD ya ha sido lexicalizado por el exponente correspondiente a la base nominal, *gero-*. La sintaxis examina la estructura configuracional de las entradas

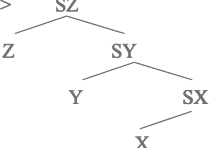
léxicas y elige la entrada de (27b). La entrada de (27a) está eliminada por el PLE (25), porque no identifica el rasgo Y. La entrada de (27b), en cambio, encaja perfectamente con la configuración sintáctica.

Veamos ahora un caso de sincretismo, como el del sustantivo correspondiente a ‘ventana’.

- (29) a. okn-o  
           ventana-nom  
       b. okn-o  
           ventana-ac  
       c. okn-u  
           ventana-dat

El mismo exponente se emplea en el acusativo y en el nominativo. Por tanto, debemos tener una entrada léxica que sea compatible tanto con el acusativo (SY + SX) como con el nominativo (SX). Dado el Principio de Lexicalización Exhaustiva, la entrada léxica de -o debe ser la de (30a), donde -o puede emplearse en acusativo; la de -u debe ser la de (30b).

- (30) a. -o <--->
- 
- ```

graph TD
  SY --> Y
  SY --> SX
  SX --> X
  
```
- b. -u <--->
- 
- ```

graph TD
  SZ --> Z
  SZ --> SY
  SY --> Y
  SY --> SX
  SX --> X
  
```

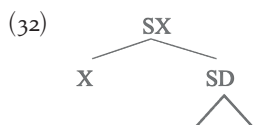
Así, cuando el sustantivo aparece en caso acusativo, -o es un encaje perfecto y obtenemos okno; cuando la sintaxis genera el caso dativo, con SZ, -u desbanca a -o por el PLE (25), ya que (30a) no identifica SZ. Pero, ¿qué sucede para que se pueda emplear -o también en nominativo? En este punto es donde entra en juego el Principio del Superconjunto (PS), que se deriva directamente de la Materialización de Sintagma y del Principio de Lexicalización Exhaustiva:

- (31) Principio del superconjunto (Caha, 2009: 67)

Un exponente léxico materializa un constituyente sintáctico si la configuración asociada al exponente es un superconjunto propio o no del constituyente sintáctico.

Un nudo sintáctico de la entrada léxica coincide con un nudo sintáctico de la estructura si y solo si los dos nudos, y todos los nudos que contengan, son idénticos.

De forma menos técnica, esto quiere decir que un exponente tiene que identificar todos los rasgos de la estructura en la que se inserta, pero pueden ignorarse rasgos que están en la entrada léxica del exponente, siempre y cuando esos rasgos no interfieran con los que se materializan en la sintaxis. En términos prácticos esto quiere decir que para introducir *-o* en una configuración de nominativo, como (32), deben darse dos condiciones: la estructura sintáctica del nominativo debe estar representada en la configuración a la que se asocia *-o*, y esa representación del nominativo no puede contener dentro de ella material que no esté presente en la estructura sintáctica.



En el supuesto de que SD ya ha sido materializado por la base *okn-*, si comparamos (32) con la entrada léxica de (30a) vemos que se cumple el principio del superconjunto: el SX que representa el nominativo está contenido en (30a), y SX no contiene nudos adicionales que no estén presentes en el árbol de (32).

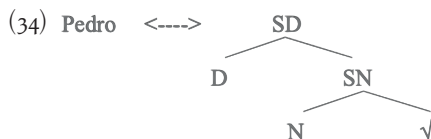
Podría pensarse que el exponente de (30b), correspondiente a un dativo, también cumple el principio del superconjunto. Es cierto. La razón de que no se introduzca es la Condición de Panini o *Elsewhere condition* (Kiparsky, 1973): dadas dos entradas léxicas igualmente legítimas para materializar una estructura, la que gana es aquella que está más cerca de la representación sintáctica. Para usar (30a) en (32) solo hay que ignorar una capa estructural, la de SY, mientras que para usar (30b) sería necesario ignorar dos capas estructurales. Enunciamos la condición de Panini como sigue:

(33) Condición de Panini

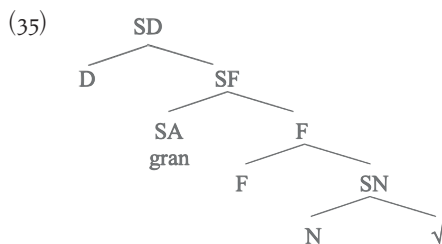
Si más de un exponente puede materializar la misma configuración sintáctica por el principio del superconjunto, se escoge el exponente con el menor número de material superfluo (contado en capas estructurales).

El PS (31) tiene una consecuencia empírica inmediata. La coincidencia entre exponente y representación sintáctica se aplica no solo a un nudo, sino a todo el constituyente sintáctico que se lexicaliza. Esto tiene la consecuencia de que, si la sintaxis introduce elementos adicionales dentro del material que debería materializar el exponente, ese exponente solo podrá lexicalizar la parte de la estructura que esté

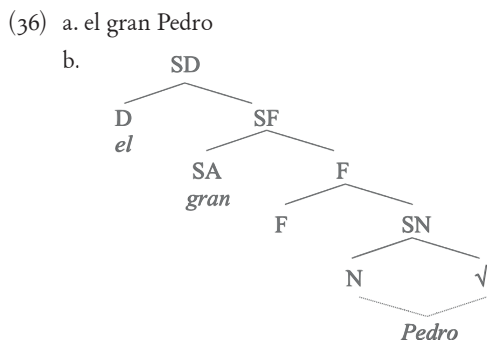
por debajo del material adicional. Veamos un ejemplo práctico. Supongamos que una lengua tiene un exponente como el de (34) que materializa no sólo un SN sino también la capa del SD (lo cual sería una versión del análisis de Longobardi, 1994 para los nombres propios). Imaginemos que ese exponente es por ejemplo *Pedro*.



Supongamos que la sintaxis produce una estructura con un modificador adjetival que se interpone entre SD y SN, como en (35).



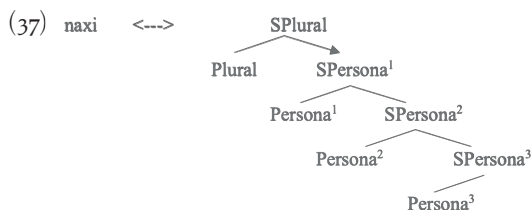
Por el PLE, (34) no puede lexicalizar toda la estructura, porque no incluye en su entrada el SF ni el SA que aparece en (35). El PS busca entonces el elemento más grande que sí respete la condición, y ese elemento es SN: en el nudo SN se satisface la condición de que el constituyente es idéntico en la entrada léxica y en la representación sintáctica. Así pues, *Pedro* materializa el SN, y SD tiene que ser materializado por otro exponente distinto.



Dicho de otra forma: cuando se introduce el adjetivo, se rompe en sintaxis el constituyente que materializa el exponente *Pedro*, y este exponente debe introducirse en el subconstituyente que coincida con la parte baja de su entrada léxica. Por tanto, una forma de comprobar si un exponente materializa un constituyente más grande es tratar de interponer material adicional en el interior de ese constituyente. Si el exponente materializa el sintagma mayor –en este caso, SD– interponer material bajo él fuerza al exponente a reducirse a identificar solo el material de la parte baja de la estructura, y otros exponentes deberán materializar la parte alta. Este corolario al PS de (33), que fuerza al exponente a materializar solo la parte baja de su entrada léxica cuando aparecen elementos en el árbol que no identifica léxicamente, se conoce como la Condición del Ancla (Abels y Muriungi 2008).

En este trabajo, el nudo más bajo de la entrada léxica de adjetivo o sustantivo léxico es siempre la raíz sobre la que se construye, por lo que la Condición del Ancla implica que un sustantivo debe identificar necesariamente su raíz. Como veremos en este trabajo, las predicciones del PS en lo que toca al material interpuesto en la estructura sintáctica nos permitirán determinar que no existen desinencias *-e* o *-ø* en español (§5.4), sino que en esos casos el exponente del nombre léxico también identifica las desinencias.

El PS se aplica a entradas léxicas, y permite dar cuenta de casos de sincretismo vertical, que se aplica a una sola dimensión gramatical, como la manifestación del caso en el ejemplo ruso que se vio en (29). La nanosintaxis también permite que una entrada léxica remita a otra entrada léxica, permitiendo así que el principio del superconjunto se aplique en dos lugares distintos de la representación del exponente, y dando cuenta así de sincretismos llamados horizontales donde se combinan dos o más dimensiones gramaticales, como por ejemplo los sincretismos que surgen al combinar el número con la persona. Veamos el ejemplo que proporciona Vanden Wyngaerd (2018: 285-287) para entender cómo se da cuenta de estos sincretismos horizontales en la nanosintaxis. (37) es la entrada simplificada del pronombre *naxi* de la lengua atabasca *slave*, que se emplea en el plural tanto para la primera como para la segunda persona –es decir, es sincrético entre *nosotros* y *vosotros*–.





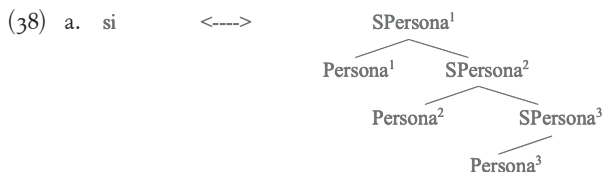
Como se ve en la tabla 3, tenemos un caso de sincretismo horizontal en que se cruzan los rasgos de número y de persona: el *slave* diferencia el singular del plural y también, en singular, la primera y la segunda persona. Estamos, por tanto, hablando de sincretismos multidimensionales.

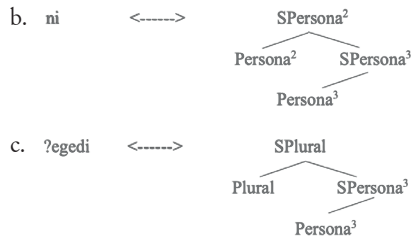
|          | Singular | Plural |
|----------|----------|--------|
| 1persona | si       | naxi   |
| 2persona | ni       |        |
| 3persona | ?edi     | ?egedi |

Tabla 3. Los pronombres personales del slave

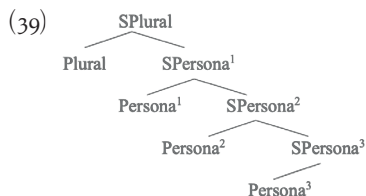
Nótese la flecha que sale del nudo más alto, *SPlural*, en (37). Esta flecha es técnicamente un ‘puntero’ (*pointer*; Starke, 2011, 2013, Caha y Pancheva, 2012). El puntero es, dentro de la entrada léxica de un exponente, el elemento que relaciona la entrada de ese exponente con las piezas léxicas que materializan los rasgos de otra dimensión gramatical. En nuestro ejemplo, el puntero relaciona el exponente para el plural con los exponentes que expresan los rasgos de persona. De esta manera, (37) expresa la generalización de que en slave el plural de la primera y de la segunda persona se materializan con el mismo exponente, porque *naxi* reemplaza a *si* y *ni* en contextos plurales.

Estos punteros solo pueden aparecer entre núcleos que expresan dimensiones gramaticales distintas, como en este caso el número y la persona, porque lo que hacen es expresar que la pieza léxica que expresa el rasgo de encima del puntero remite a la pieza léxica que expresa los rasgos por debajo del puntero. El efecto práctico que tiene este puntero es permitir que el principio del superconjunto se aplique al complemento de la flecha, es decir, *SPersona<sup>1</sup>*, para la primera persona. Veamos esto paso a paso. Supongamos, para empezar, que los pronombres correspondientes a *yo*, *tú* y *ellos* tienen las siguientes entradas léxicas (simplificados con respecto a Vanden Wyngaerd, 2018):



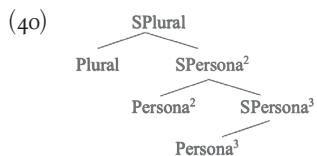


Veamos cómo los principios descritos previamente derivan las formas correctas. Supongamos que la sintaxis genera un árbol como el de (39) –que sintácticamente expresa la primera persona plural–.



Si examinamos las entradas de (37) y (38), vemos que (37) es la única opción. Ni (39a) ni (39b) expresan los rasgos de número plural, mientras que (39c) carece de rasgos de Persona<sup>1</sup> o Persona<sup>2</sup>, por lo que quedan todas eliminadas por el PLE de (25).

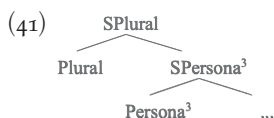
De forma crucial, el puntero en (37) permite que *naxi* identifique el rasgo plural pero que se reduzca, mediante el Superconjunto de (35), el constituyente que se expresa debajo del puntero. Supongamos que la estructura sintáctica es ahora (40).



De nuevo, las entradas de (38) están eliminadas por el PLE: o no identifican todas las capas de persona, o no identifican el rasgo plural. (37) es la única entrada que identifica tanto plural como el rasgo de persona asociada a la segunda persona. El puntero permite que se aplique el PS en el área de persona; en esencia, (37) nos dice en este contexto que *naxi* reemplaza a la entrada de (38b) cuando aparece el rasgo de plural.

En cambio, si la sintaxis produce una forma plural de tercera persona, (37) no se emplea, sino que vence (38c) por la condición de Panini de (33): aunque tanto (37)

como (38c) pueden identificar todos los rasgos, (38c) carece de material superfluo, y (37) tiene dos capas de tal material, las que corresponden a la primera y segunda personas. Por ello *naxi* no vence a *?egedi* en la estructura de (41).



Hemos visto en la tabla 2 que el español muestra un sincretismo vertical en el femenino: en los sustantivos alternantes, sean de la clase que sean, la *-a* expresa el femenino. Veremos que esto puede darse cuenta, evitando la homofonía entre tipos de *-a*, si se propone que la entrada para ese exponente contiene un puntero que relaciona el femenino como rasgo *phi* con la dimensión de los clasificadores nominales.

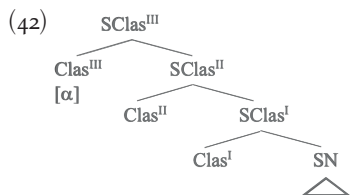
Terminamos aquí la discusión general de las bases teóricas de nuestro análisis; en la sección siguiente, vamos a presentar y justificar nuestra propuesta acerca de la estructura sintáctica del género en español.

#### 4. LAS PROPIEDADES DERIVATIVAS Y LAS PROPIEDADES FLEXIVAS: LA ESTRUCTURA SINTÁCTICA DEL GÉNERO

Pasemos ahora a hablar del análisis que proponemos para el género en español, que dividiremos en dos partes, siguiendo los dos niveles que se admiten en la nanosintaxis: en esta sección discutiremos la estructura sintáctica que subyace a la estructura del género, mientras que en la siguiente sección nos ocuparemos de las relaciones entre esa estructura y los exponentes particulares. Esta sección tiene tres partes: la sección 4.1 se ocupa de mostrar la estructura propuesta para el área de los clasificadores nominales; la sección 4.2 se concentra en el área de género como rasgo funcional. En la sección 4.3 discutiremos la relación sintáctica entre las dos áreas.

##### 4.1. *Región de clasificadores nominales*

Nuestra propuesta es que la región de los clasificadores nominales en español se divide en tres núcleos –veremos más abajo que en algunas variedades puede ser necesario contar con un cuarto núcleo–. Siguiendo los principios de la nanosintaxis, proponemos que esos núcleos se combinan unos con otros, en una secuencia funcional rígida, como en (42).



Estos núcleos son los que legitiman a ciertos sustantivos dentro del contexto sintáctico, tal y como se vio en la sección 3.2 (Acquaviva, 2009). El SN debe entenderse aquí como la base de un nombre léxico, antes de combinarse con el morfema que se ha identificado como la marca de palabra.

Veamos cuáles son las tres clases de clasificador nominal que estamos proponiendo. Los siguientes ejemplos muestran un ejemplo de cada sustantivo en esa capa, y a continuación justificamos empíricamente la división.

- (43) a. perr-o  
 b. problem-a, coleg-a  
 c. mes-a

La primera capa del clasificador nominal es la que contienen los sustantivos del tipo de *perr-o*. Estos sustantivos se caracterizan por cuatro propiedades significativas que los diferencian de las otras tres clases.

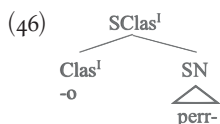
- i) Estos sustantivos son variables en género, y marcan regularmente con *-o* el masculino y con *-a* el femenino (*perro* ~ *perra*, *niño* ~ *niña*).
- ii) El femenino en estos sustantivos se identifica con el género femenino y da lugar a una interpretación composicional. El femenino *perr-a* indica la misma clase de entidad que el masculino *perr-o*, pero esta vez en una versión que alude al género biológico asociado a las hembras de la clase de entidades. En los sustantivos que poseen este clasificador nominal la versión femenina depende de la presencia de un rasgo *phi* de género, no del clasificador en sí, y esto es lo que hace que en ellos el género se interprete como una propiedad flexiva que no afecta al contenido conceptual del sustantivo.
- iii) La construcción de un nombre genérico, que alude a toda la clase de entidades, con estos sustantivos implica siempre la eliminación del femenino. Como es bien sabido, en (44a) es posible interpretar el sustantivo como denotador de la clase total a la que se refiere el nombre léxico (Borik y Espinal, 2012, 2015, 2020), mientras que en (44b) solo se alude a las entidades hembra dentro de la clase de entidades.

- (44) a. El perro es un mamífero.  
b. #La perra es un mamífero.

iv) La combinación de estos sustantivos con el diminutivo *-it-* no implica cambio en el morfema final del sustantivo.

- (45) a. perr-it-o  
b. perr-it-a

Proponemos que el núcleo Clase<sup>I</sup>, el más básico, es el que se asocia a la *-o* como marca de palabra.



Los sustantivos de Clase<sup>II</sup>, contruidos sobre los sustantivos de Clase<sup>I</sup>, tienen las siguientes propiedades:

i) Son sustantivos cuya vocal final es siempre *-a* y no existe variación en la vocal si el género es variable. Esto incluye los casos de derivados en *-ist-* tanto como otros sustantivos comunes animados acabados en *-a* para masculino y para femenino, junto a la clase relativamente abundante de sustantivos masculinos acabados en *-a*, cuyo comportamiento –como veremos– es paralelo a estos.

- (47) a. el artista ~ la artista  
b. el colega ~ la colega

(48) el problema, el idioma, el planeta, el cura, el marica, el tema...

ii) En estos sustantivos, cuando hay movilidad de género, el femenino se interpreta de forma composicional, como en los sustantivos de Clase<sup>I</sup> –los individuos hembra de la clase de entidades designada por el nombre léxico–, pese a que no se refleje directamente en su morfología superficial.

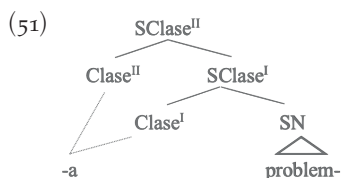
iii) Al igual que los sustantivos de Clase<sup>I</sup>, la versión genérica singular de los sustantivos de la Clase<sup>II</sup> requiere el uso del masculino, cuando hay versión femenina.

- (49) a. El artista es un creador.  
b. #La artista es una creadora.

iv) Al igual que los sustantivos de la Clase<sup>I</sup>, la combinación con el diminutivo *-ito* no altera la marca de palabra del sustantivo, incluso cuando la *-a* corresponde al género masculino.

- (50) a. *problem-it-a* (\**problem-it-o*)  
 b. *el colegu-it-a* (\**el colegu-it-o*)

Nuestra propuesta es simplemente que los sustantivos que deben ser legitimados bajo Clase<sup>II</sup> proyectan sobre la capa del SN un área de clasificador nominal compleja, donde aparecen los núcleos Clase<sup>I</sup> y Clase<sup>II</sup>. La diferencia es que el exponente *-a* es el que materializa, por materialización de sintagma (28), el conjunto Clase<sup>I</sup>-Clase<sup>II</sup>.



Los sustantivos de Clase<sup>III</sup> son los nombres invariables en género femenino marcados con *-a*. Estos se caracterizan por las siguientes propiedades.

- i) En ellos no hay variación de género (*mes-a*, *cart-a*, *puert-a*), pero la *-a* marca transparentemente el género femenino. Casos de aparente variación de género como *cas-a* / *cas-o* o *banc-a* / *banc-o* no son realmente tales, sino que se deben a la existencia de sustantivos léxicos que pueden ser legitimados bajo clasificadores III o clasificadores I.
- ii) El femenino en estos sustantivos no tiene propiedades composicionales en su semántica, incluso si el sustantivo puede interpretarse como animado: no hay interpretación alguna que se asocie con la lectura flexiva de «miembros de sexo femenino», como por ejemplo en *la persona*, o *la víctima*. Esta falta de composicionalidad se ve aún más clara en los sustantivos no animados que pertenecen a esta clase, o en los casos en que se establecen oposiciones superficiales como *manzano* ~ *manzana*, donde se ha dicho que el género tiene propiedades derivativas porque afecta al contenido léxico de la base.
- iii) La versión genérica de estos sustantivos, animados o no, es femenina, y el femenino no excluye la interpretación de clase. En esto estos sustantivos contrastan con los pertenecientes a las otras dos clases.

- (52) a. La víctima es quien más sufre en un crimen.  
b. La manzana es un alimento saludable.

iv) No obstante, al igual que los sustantivos de las otras dos clases, la combinación con el diminutivo *-it-* mantiene sistemáticamente la vocal final.

- (53) a. la victim-it-a  
b. la manzan-it-a

La clase III está asociada sistemáticamente al género femenino. Nuestra propuesta es que la clase III se construye sobre la clase II proyectando una tercera capa. Frente a las otras dos capas, esta capa contiene un rasgo valorado pero no interpretable de género femenino –llamémoslo a partir de ahora [ $\alpha$ ] para evitar la confusión entre el término formal femenino y la interpretación de género biológico hembra–.

- (54) Clase<sup>III</sup>  
[uGen:  $\alpha$ ]

Como veremos en la sección 4.3, este rasgo es el que será copiado por el género en el área de rasgos nominales. Lo que nos importa en este punto es que el rasgo alfa, cuando el sustantivo tiene el clasificador III, está introducido dentro de la región de los clasificadores nominales, que pertenece al área en la que se define el nombre léxico como un conjunto de propiedades que denotan una clase de objetos –es el área de la clasificación de entidades dentro del dominio nominal–. En esta área, por tanto, la presencia del rasgo alfa influye en la categorización conceptual del sustantivo. Por este motivo, el rasgo no recibe una interpretación composicional de ‘miembros femeninos de la clase de entidades denotada por el sustantivo léxico’, sino que da lugar a significados idiomáticos que se imbrican con la raíz y el sustantivo léxico para dar, virtualmente, cualquier tipo de interpretación conceptual. Podemos resumir este principio como en (55).

- (55) El material del área de clasificación nominal influye en el concepto expresado por el nombre léxico.

Así surge lo que se ha entendido como la aportación léxica o derivacional del género: cuando el rasgo alfa está presente en el clasificador nominal produce construcciones idiomáticas que dan contenido conceptual a la raíz, como ya notó Acquaviva (2009) sobre principios teóricos diferentes pero que convergen en esta explicación. Tomemos como ilustración el par *manzano* ~ *manzana*. La

idea principal es que el exponente *manzan-*, en su entrada léxica, especifica que puede legitimarse en bajo dos clasificadores distintos, frente por ejemplo al caso de *problem-* en (18), que solo admite un clasificador:

(56) *manzan-* Contexto de legitimación:  $[\{S\text{Clas}^I / S\text{Clas}^{III}\}]$

Esto permite que se formen dos nombres léxicos distintos, *manzan-o*<sup>I</sup> y *manzan-a*<sup>III</sup>; cada uno de estos dos nombres léxicos son, en esencia, construcciones idiomáticas cuya semántica conceptual está almacenada en el léxico enciclopédico:

(57)  $[[\text{manzan}]\text{o}]$  ‘árbol que produce frutos carnosos, típicamente rojos y redondos’  
 $[[\text{manzan}]\text{a}]$  ‘fruto comestible carnosos, típicamente rojo y redondo’

Dado que estamos hablando de una clasificación conceptual, son innumerables las lecturas idiomáticas que se pueden listar. Sabemos que en español se dan varios casos de la de (57), junto a otras interpretaciones donde parece hablarse de diferencias de tamaño o colectividad (*banco* ~ *banca*, *cesto* ~ *cesta*, *río* ~ *ría*), junto a muchas otras opciones que no son sistematizables en términos composicionales o sintácticos porque dependen de las entradas conceptuales idiomáticas que se asocian a cada nombre léxico: *gorro* ~ *gorra*, *ventana* ~ *ventano*, *peso* ~ *pesa*, etc.

Desde el punto de vista de la materialización, proponemos que la estructura del clasificador III, que contiene al II y al I, se lexicaliza también con *-a*.

(58)

```

graph TD
    SClaseIII[SClaseIII] --- ClaseIII[ClaseIII [α]]
    SClaseIII --- SClaseII[SClaseII]
    ClaseIII --- a[-a]
    ClaseIII --- ClaseI1[ClaseI]
    SClaseII --- ClaseII[ClaseII]
    SClaseII --- SClaseI[SClaseI]
    SClaseI --- ClaseI2[ClaseI]
    SClaseI --- SN[SN mes-]
  
```

Esto implica que, aunque haya tres niveles sintácticos para el clasificador, solo proponemos dos exponentes que materializan esta región, y que se exponen en (59), donde aún no hemos representado qué área del género funcional expresan.

(59) a.  $-\text{o} \leftrightarrow$

```

graph TD
    SClaseI[SClaseI] --- ClaseI[ClaseI]
  
```

b.  $-\text{a} \leftrightarrow$

```

graph TD
    SClaseIII[SClaseIII] --- ClaseIII[ClaseIII]
    SClaseIII --- SClaseII[SClaseII]
    ClaseIII --- a[-a]
    ClaseIII --- ClaseI1[ClaseI]
    SClaseII --- ClaseII[ClaseII]
    SClaseII --- SClaseI[SClaseI]
    SClaseI --- ClaseI2[ClaseI]
  
```



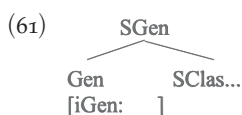
Mediante el PLE (25), cuando aparecen las proyecciones de Clase<sup>II</sup> o Clase<sup>III</sup>, (59a) no puede emplearse –porque dejaría esa parte de la estructura sin identificar–. Esto implica que *-o* es reemplazado por *-a* tan pronto como aparezca Clase<sup>II</sup>. Por el Principio de Panini (33), cuando aparece Clase<sup>II</sup>, se emplea el exponente de (59b) aunque Clase<sup>III</sup> no esté presente en la estructura, porque (59b) es el exponente de los dos que tiene menos material sin identificar, pero identifica toda la estructura proporcionada por la sintaxis.

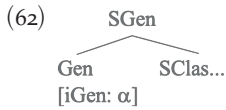
Antes de pasar a nuestra explicación del género, hagamos un par de observaciones generales. Nótese que no estamos proponiendo que exista un clasificador nominal correspondiente a *-e* (*coche*) o a *-ø* (*papel*). Como veremos en la próxima sección (§5.3), trataremos estos sustantivos como casos de los otros tres clasificadores donde el exponente del sustantivo materializa también las capas del clasificador nominal, algo que se hace evidente cuando esos sustantivos se combinan con el diminutivo *-it-*. La principal prueba es que estos sustantivos, en tales casos, deben materializar una marca de palabra *-o* o *-a*, que son las representadas arriba en (59). Esto es una manifestación del principio general que se presentó en la sección 3.4, que la interposición de material adicional fuerza a que se aplique el principio del superconjunto, como en (36). Esto indica que cuando no aparece el diminutivo es el exponente el que cubre estas áreas, y cuando aparece el diminutivo no puede materializarlas.

- (60) a. el president-it-o (\*president-it-e)  
 b. el papel-it-o (\*papel-it-(e))

#### 4.3. Capa de género: solo hay un valor de género para el español

Nuestra propuesta es que el género en español contiene solo un valor, el que corresponde a alfa, que se ha clasificado normalmente como femenino. Por tanto, el masculino será tratado aquí como la ausencia de valor de género. Un masculino es, por tanto, una estructura como la representada en (61), en la que el género interpretable pasa a las interficies sin valor alguno. El femenino implica una estructura como la de (62), donde Gen está valorado con el rasgo alfa, que como veremos puede proceder de dos fuentes.





Así pues, tomamos el masculino en realidad como la ausencia de género en español. Veamos tres pruebas para esta afirmación.

Comenzando por la prueba más directa, el hecho es que los adjetivos adoptan la versión tradicionalmente llamada «masculina» en los contextos en que se predicen de un sujeto que no es nominal, y por tanto carece de rasgos de género. Algunas gramáticas han tratado este género como una forma de neutro (por ejemplo, RAE y ASALE, 2009: §13.5n), que casualmente coincide en absolutamente todos los adjetivos del español con la terminación del masculino para el mismo adjetivo. Desde nuestra perspectiva, sin embargo, esta equivalencia no es una casualidad, sino que se debe a que en ambos casos tenemos la misma situación, es decir, ausencia de valor de género que pueda asignarse al SGen del adjetivo.

- (63) a. Marta es {buena / \*bueno}.  
 b. Que vengas es {bueno / \*buena}.  
 c. Venir a tiempo es {bueno / \*buena}.

Un segundo argumento procede de los sustantivos genéricos en singular, como se adelantó en la subsección anterior. Borik y Espinal (2012, 2015) analizan casos como los de (64) proponiendo que en ellos el sustantivo denota una clase porque los rasgos nominales –rasgos *phi*– se encuentran ausentes de la estructura.

- (64) El gato es un mamífero.

Su propuesta parte de la idea general de Gehrke y McNally (2015), que propone una división entre la lectura de clase de entidades y eventos y su lectura de ejemplar. Dentro de esta propuesta (cf. también Ramchand, 2018), el SN define una clase de entidades, no ejemplares particulares de esa clase –como vimos arriba, los clasificadores nominales manipulan la información conceptual de los nombres léxicos y de las clasificaciones de entidades a las que dan lugar–. Para denotar individuos particulares, ejemplares de la clase, es necesaria la presencia de rasgos nominales sobre la estructura del SN. De esta manera, Borik y Espinal (2018) proponen que en (64) el sustantivo carece de información de número, lo cual permite que se refiera a una clase completa formada por una pluralidad de entidades y no a un

único individuo de la clase. El sustantivo, por tanto, no estaría en singular, sino que carecería de número.

Nuestra propuesta es extender esta hipótesis –que las versiones genéricas de un sustantivo en singular carecen de número– a todos los rasgos nominales o rasgos *phi*: estos sustantivos también carecen de la proyección SGen. Un sustantivo genérico en singular es una proyección del área del SN directamente dominada por el área del SD sin el área intermedia de rasgos *phi*.

Una prueba independiente de que esta idea tiene sentido nos la proporciona el noruego. En noruego hay un contraste claro entre la interpretación genérica y la de ejemplar de los sustantivos, que se ilustra en (65).

- (65) a. Pizza-en er god.  
 pizza-la es buen-m/f  
 ‘La pizza (=este ejemplar de pizza) está buena’
- b. Pizza-ene er gode.  
 pizza-las son buen-pl  
 ‘Las pizzas (=estos ejemplares de pizza) están buenas’
- c. Pizza er godt.  
 pizza es buen-neut  
 ‘La pizza (=la clase de objetos) está buena’

En noruego hay género neutro además del género masculino y femenino, que en la mayoría de los adjetivos se expresan con la misma concordancia. Dejando a un lado las propiedades del determinante –en noruego, como en otras lenguas germánicas, la lectura genérica se obtiene sin determinante explícito–, lo que vemos en (65c) es que cuando se emplea el sustantivo en su valor de clase el adjetivo toma género neutro, independientemente de que el sustantivo sea masculino o femenino en contextos no genéricos. No es posible proporcionar valores de número (65b) o de género (65a) en esta interpretación. Los datos se siguen de forma natural si la lectura genérica singular implica la falta de todos los rasgos *phi*, tanto número como género, y por tanto el adjetivo debe tomar valor por defecto, que en noruego es el que tradicionalmente llamamos neutro.

Nuestra propuesta es que lo mismo sucede en español, solo que ahí no hemos llamado neutro a la forma del adjetivo. Por tanto, la estructura interna del sujeto en (66) –*el gato*, como nombre de clase– es como la que sigue, donde faltan todas las proyecciones del área de los rasgos *phi*.



- (70) a. est-e, est-a, est-o  
 b. es-e, es-a, es-o  
 c. aquel, aquell-a, aquell-o  
 d. él, ell-a, ell-o  
 e. el, l-a, l-o

Somos conscientes de que esta propiedad no queda explicada del todo en nuestra teoría: parece que la estructura de los pronombres admite un segundo valor para el género además de alfa, llamémoslo beta, que explique qué sucede con la versión aparentemente masculina. Dejaremos esto para investigación posterior, aunque desde nuestra perspectiva es probable que lo que suceda aquí es que la estructura del SD en el pronombre permita introducir una segunda proyección de género interpretable, como nota Déchaine (2018) para otras lenguas. En este trabajo nos concentraremos más bien en la propiedad de que la forma en *-o* es la que corresponde al neutro, es decir, ausencia de valor, y se emplea en los contextos en que el pronombre se refiere a entidades no nominales que por tanto carecen necesariamente de género.

- (71) Juan llegó tarde. Esto (=llegar tarde) molestó a Pedro.

Si el tradicionalmente llamado ‘masculino’ en adjetivos y sustantivos, marcado con *-o*, es en realidad la manifestación de un género sin valor, esta correlación se explica de forma natural. De lo contrario, estaríamos ante otra casualidad léxica que, por algún capricho, invierte el valor de *-o* de masculino a neutro entre los sustantivos y los pronombres.

#### 4.3. *El valor de género y la interacción con la marca de palabra*

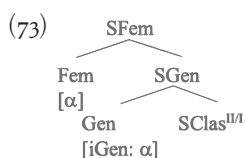
Examinemos ahora el género femenino y la relación con el rasgo alfa. En nuestra propuesta, el género femenino –entendido en el sentido estricto de que es un rasgo del dominio de los rasgos nominales que se emplea en la concordancia– se obtiene cuando el SGen tiene el rasgo alfa. Como hemos visto, una primera forma en que el rasgo alfa puede aparecer en SGen es mediante concordancia con el rasgo que introduce Clas<sup>III</sup>.

- (72)
- 
- ```

  graph TD
    SGen --> Gen
    SGen --> SClasIII
    Gen --- GenFeat["[iGen: α]"]
    SClasIII --> ClasIII
    SClasIII --> Ellipsis["..."]
    ClasIII --- ClasFeat["[α]"]
  
```

En estos casos, alfa está presente en el área de clasificación del nombre léxico, por lo que construye una estructura idiomática con el nombre y no tiene un valor composicional.

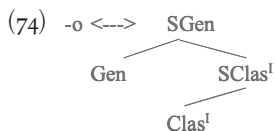
Existe una segunda forma en que el rasgo alfa puede aparecer en Gen: cuando se introduce directamente en el área de flexión nominal, como un rasgo de Gen; en estos casos, no hay Clas<sup>III</sup>. Siguiendo los principios nanosintácticos, proponemos que este núcleo recibe valor del rasgo alfa de una proyección superior a él, a la que llamaremos SFem.



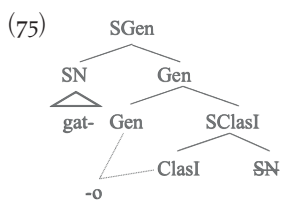
En este contexto sintáctico, el núcleo Gen primero trata de buscar un valor en su dominio; no lo encuentra; siguiendo a Béjar y Rezac (2009), entonces prueba a buscar su meta en la estructura más alta, donde encuentra alfa en Fem. Este valor se copia en Gen, donde es interpretable. El resultado sigue siendo el mismo: el valor de género que se ha denominado tradicionalmente como femenino termina en Gen, donde podrá forzar concordancia femenina con adjetivos y determinantes.

No obstante, en esta configuración el rasgo alfa no está presente nunca en el área de los clasificadores nominales, y solo está presente en el área de los rasgos *phi*. En tales casos, el rasgo alfa no puede formar una estructura idiomática con el nombre léxico, y por tanto no afecta a la conceptualización del sustantivo. Su significado se construye, entonces, sobre la denotación del nombre léxico, composicionalmente seleccionando solo los individuos que son hembras dentro de la clase de entidades. Es en estos casos cuando el género tiene un comportamiento flexivo, sencillamente porque el rasgo alfa aparece únicamente en la estructura funcional del sustantivo y no en las capas de clasificadores nominales. Por tanto, esperamos que sea en estos casos donde el género femenino tiene interpretación composicional, y el rasgo alfa tendrá su origen en una configuración como la de (73) en todas las clases de sustantivos de la tabla 2 donde el femenino es composicional.

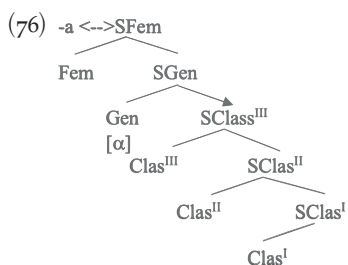
Antes de pasar al siguiente apartado, donde derivaremos cada clase morfológica de sustantivos, veamos cómo las entradas léxicas propuestas para los clasificadores nominales incorporan la materialización del género. La entrada de (74) es la que proponemos para el exponente *-o*; proponemos que su entrada léxica incluye también la materialización de SGen cuando no contiene alfa.



Así, *gat-o* se materializaría de la siguiente manera. Suponemos movimiento del nombre léxico al especificador de SGen para obtener el orden de morfemas.



Al mismo tiempo, revisaremos la entrada de *-a* como sigue:



Como puede verse, en esta entrada estamos introduciendo un puntero entre el área de la información de rasgos nominales y el área de clasificación. Técnicamente, esto permite al exponente *-a* expresar el género femenino con las tres clases de sustantivos, independientemente de si contienen Clas<sup>III</sup>, Clas<sup>II</sup> o Clas<sup>I</sup>, ya que el puntero permite que el principio del superconjunto se aplique en la estructura intermedia (recuérdese la discusión de 41). El dato empírico que justifica esta decisión es el siguiente hecho: la versión femenina de los sustantivos de las tres clases notadas es invariablemente expresada mediante el mismo exponente, *-a*, con independencia de si el sustantivo es variable o no en género.

- (77) a. la *gat-a*  
 b. la *coleg-a*  
 c. la *cam-a*

Es decir, el sincretismo horizontal de la tabla 1, en el que todas las formas femeninas se tratan con el exponente que define la clase III, independientemente

de si el sustantivo pertenece a la clase I o a la clase II, será explicada mediante la entrada de (76). Como en los casos examinados en la sección 3.4, el puntero es el que permite capturar la idea de que la presencia de alfa en Gen disuelve la distinción entre esas tres clases.

Esto querrá decir, como veremos en la sección 5.4, que la forma en que analizaremos los sustantivos femeninos en *-o* (*la mano*) implicará una entrada léxica para el exponente que bloquee la presencia de *-a*. Examinaremos esos casos más en detalle en esa sección.

Antes de mostrar la forma específica en que estas dos entradas léxicas dan cuenta de todos los patrones de morfología de género en las clases de sustantivos del español, queremos dirigir la atención del lector hacia el hecho de que en ningún caso hemos tenido que proponer homofonía. Nuestra propuesta es que el género en español usa solo dos exponentes: *-o*, con la entrada de (74), que cubre parte del área de género y parte del área de clasificador nominal, y *-a* con la entrada de (76), que cubre todos los casos en que hay un valor alfa (femenino) en Gen. En otras palabras: estamos unificando todos los casos de *-a* como parte de una misma entrada léxica, y no estamos admitiendo entradas léxicas para las supuestas desinencias *-e* o *-ø*, ni para los casos especiales donde la *-o* da lugar a género femenino invariable. Sí tendremos que proponer que en una variedad del español hay un caso de homofonía que implica a *-o*, en forma de un cuarto clasificador relacionado con las vocales de enlace, pero esa entrada léxica no está en la gramática de otros muchos hablantes.

Veamos en la siguiente sección de qué manera este grupo mínimo de entradas léxicas permiten dar cuenta de las clases de sustantivos en español que se definen por su marca de género.

## 5. LA MARCACIÓN DE GÉNERO: LAS PROPIEDADES DE MATERIALIZACIÓN DEL GÉNERO

Esta sección muestra cómo se pueden derivar las clases morfológicas de sustantivo en español a partir de la estructura y los dos exponentes que hemos presentado. Los grupos relevantes que deben ser examinados, y que fueron introducidos en la tabla 2, son los siguientes, ordenados por las similitudes estructurales y exponenciales que se discutirán en las páginas siguientes:

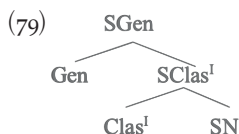
- (78) a. Grupo primero: Sustantivos variables en género (*-o* / *-a*) de femenino composicional (*mono* ~ *mona*)



- b. Grupo segundo: Sustantivos con dos valores de género, pero invariables (-a) de femenino composicional (*el ~ la colega*)
- c. Grupo tercero: Sustantivos de género masculino invariable en -a (*el planeta*)
- d. Grupo cuarto: Sustantivos de género femenino invariable en -a (*la mesa*) de femenino no composicional.
- e. Grupo quinto: Sustantivos de género masculino invariable en -o (*el techo*) de masculino no composicional.
- f. Grupo sexto: Sustantivos variables en género (-e / -a) de femenino composicional (*jefe ~ jefa*), usados a veces como invariables (*el ~ la jefe*)
- g. Grupo séptimo: Sustantivos variables en género (-ø / -a) de femenino composicional (*juez ~ jueza*), usados a veces como invariables (*el ~ la juez*)
- h. Grupo octavo: Sustantivos masculinos o femeninos no acabados en -o o en -a (*coche, papel*)
- i. Grupo noveno: Sustantivos femeninos invariables en -o (*la mano*).

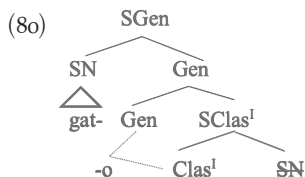
5.1. Grupo primero: marca de palabra -o de género variable y mutable a -a

La primera clase de sustantivos es la más frecuente entre los nombres animados, referidos tanto a personas -*amigo ~ amiga, vecino ~ vecina, tío ~ tía*- como a animales -*perro ~ perra, zorro ~ zorra, cerdo ~ cerda*-. La estructura de la forma tradicionalmente llamada masculina en -o es la de (79).

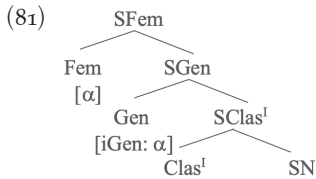


Estos sustantivos están legitimados bajo el clasificador I, que carece de información de género. En el nivel de rasgos *phi*, el masculino introduce género sin valor, lo que resulta en la inserción del valor por defecto para el género.

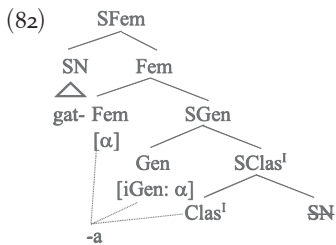
De entre las dos entradas léxicas para las desinencias, la que corresponde a -o (74) es la que se emplea: es la coincidencia perfecta, sin excluir ningún núcleo, de la estructura proyectada sobre el SN.



La versión femenina de estos sustantivos introducen el rasgo alfa en el nivel de los rasgos *phi*, donde su interpretación es composicional y no afecta a la categorización semántica.



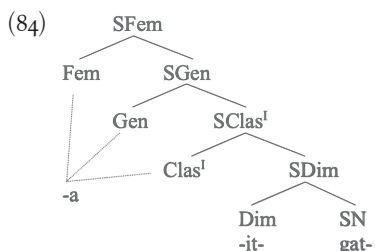
En la materialización, la entrada para *-o* en (74) no puede emplearse por el PLE (25). La entrada de *-a* en (76), en cambio, encaja. El puntero en esa entrada permite que se aplique el PS (31) a los clasificadores: en este caso, el clasificador I es el único presente en la estructura. Por la condición del ancla, *-a* lexicaliza el conjunto entre Fem y Clas¹.



Una propiedad muy relevante de esta clase es que la presencia del diminutivo *-it-* no altera la desinencia. En nuestra propuesta, esto se sigue del hecho de que el exponente usado para el SN no lexicaliza ninguno de los núcleos de clasificador nominal o de género. De esta manera, el diminutivo puede proyectarse dentro del área de clasificación nominal, inmediatamente sobre SN, sin alterar el constituyente que lexicaliza el exponente nominal o las desinencias, como se ve en (83).<sup>6</sup>

(83) gat-a ~ gat-it-a

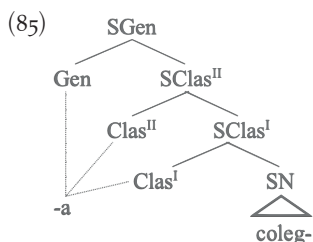
6. Como observa un revisor anónimo al que agradecemos el comentario, esta propuesta fuerza a un tratamiento irregular de combinaciones como *la testigo* para aquellos hablantes que no regularizan la marca *-a* en estos casos. Nuestras opciones solo son dos: o bien para esos hablantes estos sustantivos se asimilan al noveno grupo (*mano*) con la única salvedad de que admiten la alternancia de género o bien esos hablantes analizan la *-o* final como parte de la raíz. Las dos teorías pueden coexistir y aplicarse a distintos hablantes: quienes tienen la primera solución diminutivizarán como *la testigu-it-a* y quienes analicen la segunda solución optarán por *la testigu-it-o*.



5.2. Grupos segundo, tercero, cuarto y quinto de sustantivos

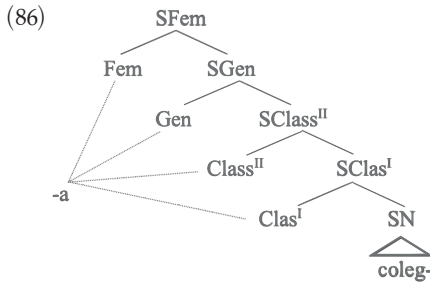
Veamos ahora el segundo y tercer grupo de sustantivos, ambos caracterizados por pertenecer a la clase II de clasificadores nominales. Aquí es relevante distinguir dos casos, el de los sustantivos con dos géneros (*el colega ~ la colega*) y el de los sustantivos de género masculino en *-a*. Comencemos por el primero.

El grupo segundo de sustantivos es, como se espera de la alternancia sistemática en género, animado. La estructura sintáctica de su masculino es como la del primer grupo, pero ahora aparece el clasificador II junto al clasificador I. Este clasificador carece también de rasgo alfa, por lo que el masculino se obtiene por defecto cuando no se introduce alfa en el área de los rasgos *phi*.



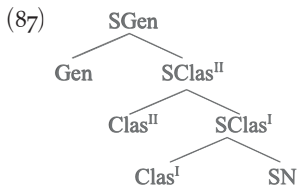
En cuanto a la materialización, la entrada de *-o* (74) no puede emplearse, porque ese exponente no identifica Clas<sup>II</sup> y su inserción violaría el PLE. La entrada de *-a* en (76) sí puede emplearse, aplicando doblemente el superconjunto, por encima y por debajo del puntero. Por debajo del puntero, *-a* identifica Clas<sup>II</sup> y Clas<sup>I</sup>, respetando la condición del ancla; por encima del puntero, *-a* se reduce al núcleo más bajo también, Gen, ya que Fem no está presente.

La versión femenina es paralela también al primer grupo de sustantivos: el rasgo alfa procede de Fem, introducido en el área de los rasgos *phi* y por tanto interpretado composicionalmente.



En estos casos, *-a* se emplea como exponente: en el área por encima del puntero hay correspondencia perfecta con el material que identifica este exponente, y en el área de debajo se aplica el superconjunto igual que en la versión masculina. Por tanto, la estructura sintáctica de los sustantivos de este grupo es mínimamente diferente a la del primer grupo: la presencia de *Clas<sup>II</sup>* bloquea la inserción de *-o* y fuerza la inserción de *-a* por el PLE.

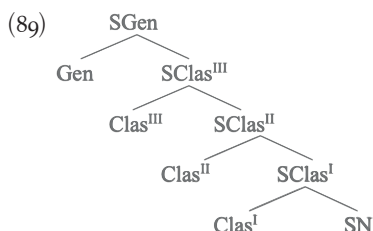
El tercer grupo de sustantivos son nombres como *cura*, *problema*, *idioma*, *planeta* o *día*, donde la *-a* marca género masculino invariable. En tales casos, la estructura sintáctica es como en (87). El núcleo *Fem* nunca se introduce en estos casos porque sería no interpretable –no podría delimitar un conjunto de hembras dentro de la clase–, y por tanto no hay rasgo alfa en ninguna posición de la estructura.



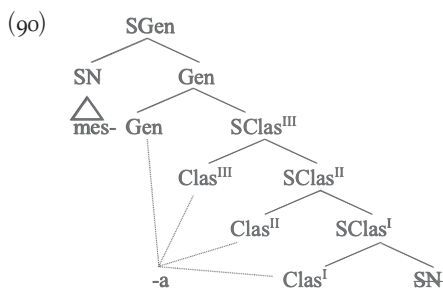
Al igual que en el primer grupo, estos sustantivos nunca alteran la marca de palabra en el diminutivo, aunque sean sustantivos masculinos. Como en el primer grupo, esta propiedad se sigue de que el exponente que materializa el sustantivo léxico no materializa ninguna parte del área del género o del clasificador nominal. De esta manera, el material interpuesto entre *SN* y estas áreas no influye en la elección de exponentes.

- (88) a. el artist-a ~ el artist-it-a  
 b. el problem-a ~ el problem-it-a

Pasemos ahora al cuarto grupo de sustantivos, aquellos que son femeninos en *-a*, pero donde el género no se interpreta composicionalmente como referido a las hembras de la clase. Esta clase incluye sustantivos tanto animados *-persona, víctima-* como no animados *-casa, mesa, ventana, puerta-*. Su estructura sintáctica es como en (94): la presencia de *Clas<sup>III</sup>* introduce el rasgo alfa, que es copiado por *Gen*.



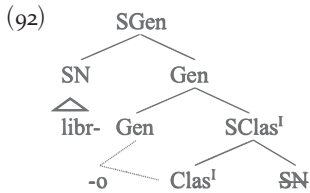
En cuanto al exponente usado, el PLE impide que se use *-o* y fuerza la inserción de *-a*. Este exponente no aplica el PS por debajo del puntero, ya que contiene las tres capas de clasificador nominal; por encima del puntero, en cambio, sí lo aplica: SFem no aparece proyectado, y su lexicalización se reduce al núcleo *Gen* con el rasgo alfa.



Como de nuevo el exponente nominal queda restringido al material contenido en SN, la presencia del diminutivo *-it-* no altera la desinencia usada.

(91) la cas-it-a

El quinto grupo de sustantivos, los masculinos invariables en *-o*, se sigue trivialmente de la entrada del exponente *-o*. Su estructura sintáctica es la de (97), al igual que la versión masculina de los sustantivos del grupo primero, y por los mismos motivos el exponente *-o* es el introducido en estos casos, de nuevo sin que el diminutivo afecte a los exponentes.



### 5.3. Los grupos séptimo y octavo son subyacentemente de las clases I-III

Pasemos ahora a los grupos sexto, séptimo y octavo, todos ellos formados por sustantivos que no emplean *-o* o *-a* en su forma morfológica. La propiedad fundamental de estos sustantivos es que, frente a los anteriores grupos, la combinación de ellos con el diminutivo *-it-* altera el morfema final empleado al menos en el masculino:

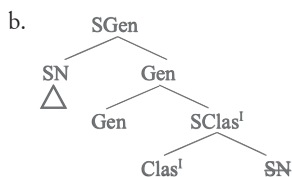
- (93) a. el presidente ~ la presidenta  
 b. el president-it-o ~ la president-it-a
- (94) a. el juez ~ la juez-a  
 b. el juec-ecit-o ~ la juec-ecit-a
- (95) a. el coche  
 b. el coch-ecit-o
- (96) a. la canción  
 b. la cancion-cit-a

Los grupos de (93) y (94) varían según los hablantes entre la manifestación de (93a)-(94a), donde el femenino está marcado con *-a* sin el diminutivo, y la manifestación de (97), donde no hay marca *-a* en el femenino, pero sí se presenta en combinación con el diminutivo.

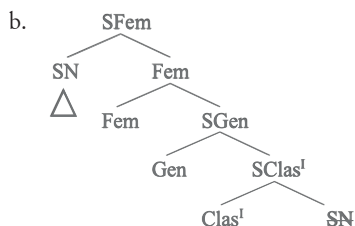
- (97) a. la presidente ~ la president-it-a  
 b. la juez ~ la juec-ecit-a

Estos patrones en combinación con el diminutivo son justamente los que esperamos si los sustantivos de los grupos sexto y séptimo son subyacentemente de la clase I y lo único que sucede con ellos es que el exponente que materializa el nombre léxico identifica también los núcleos de clasificador nominal y de género. La estructura de *el jefe / el juez* y *la jefa / la jueza* sería la siguiente:

(98) a. el jefe

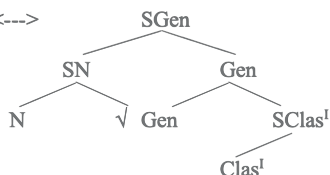


(99) a. la jefa



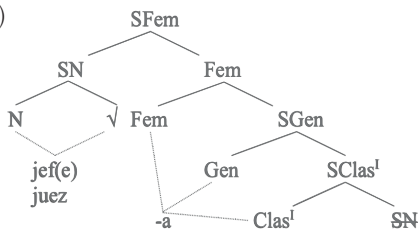
Sintácticamente no hay diferencia con los sustantivos del grupo primero, pero sí la hay en el exponente usado: en estos casos, el exponente *jefe / jefa* materializan un sintagma que incluye al SGen, pero no al SFem.

(100) jefe, juez <--->



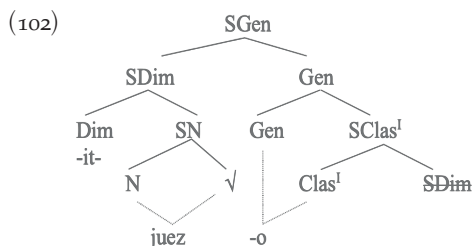
En la forma femenina, el SFem está presente. Por el PLE, la entrada de (100) no puede usarse, porque no identifica Fem ni el rasgo alfa de Gen. Así, cuando aparece el femenino para materializar el conjunto formado por Fem-Gen-Clas¹ se utiliza la entrada de *-a* (76), como en los sustantivos del primer grupo. Por el PS (31), sin embargo, el exponente nominal se reduce para identificar el SN, que es el elemento de la estructura que contiene la raíz.

(101)



Mínimamente, la única diferencia entre las variedades donde la forma femenina añade *-a* y las que no es que, en las primeras, el constituyente máximo que lexicaliza el exponente para el sustantivo incluye SFem, lo cual hace que el PLE esté respetado cuando se introduce el exponente en la versión femenina.

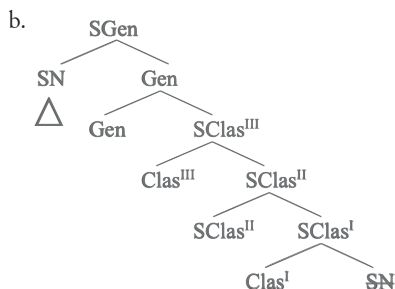
Veamos ahora cómo la presencia del diminutivo *-it-* fuerza que cambie la desinencia. Tomemos la versión masculina como ilustración:



La presencia de Dim entre el SN y el SGen rompe el constituyente sintáctico que expresan estos exponentes en (100), por lo que el exponente tiene que reducirse por el PS y limitarse a materializar SN. Esto deja los núcleos de género y clasificador nominal sin identificar, lo cual fuerza la inserción de una desinencia *-o* (cuando aparece Clas<sup>I</sup>) o *-a* (en caso de que apareciera Clas<sup>III</sup>, Clas<sup>II</sup> o Fem en la estructura).

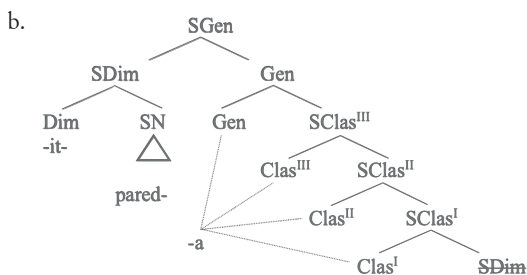
Las formas invariables del grupo octavo son esencialmente iguales en el proceso de materialización: son casos en que el exponente nominal materializa también los núcleos de género; por el mismo motivo, el diminutivo fuerza la inserción de las marcas de palabra regulares, ya que rompe el constituyente sintáctico de la entrada. Tomemos como ilustración el sustantivo femenino *pared*, que subyacemente es un sustantivo de la clase III, con femenino no composicional.

(103) a. la pared





(104) a. la paredita



Los sustantivos que terminan en vocales tónicas o en *-i / -u* se comportan esencialmente de la misma manera *-papá ~ pap(a)-it-o*,<sup>7</sup> *café ~ cafe-tit-o*, *tribu ~ tribu-cit-a-*, por lo que proponemos que en ellas la vocal final también es parte no segmentable del exponente que materializa también el SN, como en los casos anteriores.

Por tanto, estamos proponiendo que los sustantivos masculinos o femeninos que no terminan en *-o* o *-a* átonas no forman clases distintas de sustantivos en el nivel sintáctico, sino simplemente que, en tanto que exponentes, materializan las proyecciones de clasificador y género, y subyacentemente pertenecen a las otras clases que se materializan con *-o* o *-a*, como muestra el comportamiento cuando se combinan con *-it-*.

Cabe preguntarse entonces qué diferencia a los sustantivos de la clase sexta y los de la séptima para que unos tengan una vocal final *-e* y otros terminen en consonantes u otras vocales. Preliminarmente, nos gustaría sugerir que la diferencia es puramente fonotáctica: la *-e* final es epentética y se introduce en fonológicamente para evitar situaciones en las que la consonante final no es un segmento legítimo a final de palabra o existe otra condición fonológica que se satisface construyendo una segunda sílaba.

Resulta llamativo, en este sentido, que los sustantivos terminados en *-e* sean casi siempre casos en que la consonante anterior no es /r, s, d, l, n, q/, que son las únicas consonantes legítimas en posición final (*adobe, noche, bloque, llave, nieve, torre, jefe, esfinge*, y los nominalizadores *-nte* o *-aje*, por ejemplo), o donde la consonante final es correcta pero forma una secuencia compleja con otras consonantes (*dirigible, arrastre, cisne, alcalde, conde, alambre, cofre*, o sufijos

7. Como anota un revisor anónimo, se da variación entre *papito, papaño* e incluso *papacito*, donde la cuestión está en la preservación o no de la vocal final que tratamos como parte del exponente nominal. Nuestra propuesta de nuevo fuerza a la conclusión de que esta alternancia está guiada por principios fonológicos.

como *-stre*) que requieren un núcleo vocálico. En otros casos, la inserción de *-e* puede justificarse para producir una palabra prosódica de dos sílabas (*frase, clase, cese, pose*). Solo existe un número relativamente reducido de casos de sustantivos acabados en *-e* cuya consonante final habría sido un segmento posible a final de palabra y donde no hay otros requisitos prosódicos evidentes que justifiquen la inserción de una vocal epentética (*ábside, cariátide, curare, índole, sístole*, o sufijos como *-oide*), generalmente préstamos en los que sugeriríamos que la *-e* final debe verse como parte del exponente del nombre léxico, no como un morfema segmentable o como un elemento epentético. No obstante, consideramos necesario elaborar un análisis fonológico detallado para apoyar esta propuesta, algo que dejaremos para investigación posterior.

#### 5.4. Grupo noveno: palabras en *-o* femeninas

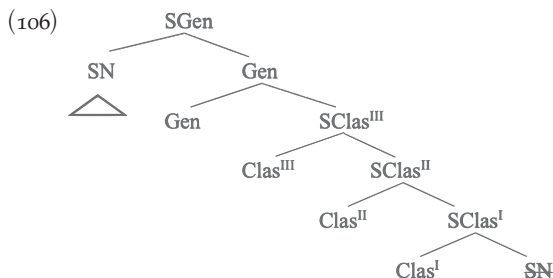
Pasemos al noveno y último grupo de sustantivos, el más complejo por la variación a la que están sometidos: los sustantivos femeninos en *-o*. Este grupo es muy reducido, y en él se suelen citar apenas nueve casos, de los que *mano* es realmente el principal ejemplo. En este grupo se encuentran unos pocos sustantivos que proceden originalmente de acortamientos (*moto, foto, radio, polio*), junto a dos voces de poco uso (*nao, seo*) y un par de préstamos de otras lenguas (*soprano, libido, dinamo*). A estos cabe añadir los casos femeninos de masculinos en *-o* para aquellos hablantes que emplean la forma en *-o* aún en esos casos: *la testigo, la piloto, la médico*, o incluso *la arquitecto*, que en la actualidad se van asimilando cada vez más al grupo primero (*la testiga, la médica...*). Nuestra propuesta para estos sustantivos depende de su comportamiento con el diminutivo, que de nuevo es variable a través de distintos hablantes.

##### 5.4.1. En las variedades del tipo «manita»

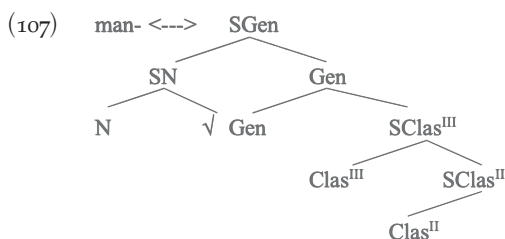
Generalmente, el español europeo y buena parte del americano restablecen la marca de palabra *-a* en el diminutivo de *mano*:

(105) la man-it-a

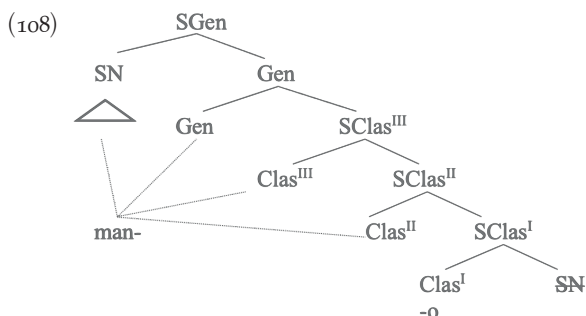
Comencemos presentando las variedades que tienen este comportamiento. Para ellas proponemos que subyacemente *man-o* es un sustantivo con el clasificador III.



Lo que hace especial a este sustantivo está en el nivel del exponente: *man-*, al igual que *jefe* o *juez*, materializa parte de las proyecciones de género y clasificador nominal, por lo que la inserción de Dim rompe el constituyente sintáctico y hace que surja la marca de palabra *-a* como en *jef-ecit-a*. Al contrario de *jefe*, sin embargo, esta entrada léxica no cubre el núcleo *Clas<sup>I</sup>*, que queda disponible para que *-o* se inserte en su lugar:



Esto resulta en que el núcleo *Clas<sup>I</sup>* debe ser identificado por otro exponente, debido al PLE (25). Por la Condición de Panini (33), el material que debe ignorarse en la entrada de *-o* es menor que el que requeriría la entrada de *-a*, lo cual da lugar a la siguiente materialización:

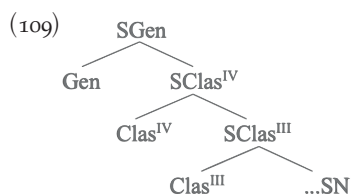


La propuesta, por tanto, exige que el sustantivo femenino en *-o* adquiera la marca *-a* con *-it-*. Esto sucede mayoritariamente en los sustantivos animados del grupo (*la sopran-it-a, la architect-it-a, la mediqu-it-a...*), que tratamos de la misma manera –solo que en este caso, al ser de género variable y con femenino composicional, proponemos que los sustantivos correspondientes son subyacentemente del primer grupo–. La prueba del diminutivo *-it-* es de difícil aplicación con algunas de las palabras del grupo por motivos fonológicos (*seo, nao*) o semánticos (*libido, polio, dinamo*), lo cual no nos da el suficiente margen para evaluar la propuesta aplicada a ellos.

En cambio, con los sustantivos que proceden de acortamientos la situación es, para muchos hablantes, más variable: *una fotita* compite con *una fotito*, más natural para muchos hablantes peninsulares, que también se decantan decididamente por *mi motito* frente a *mi motita* o la *radiíto* frente a *la radiíta*. En estos casos, no se repone la marca de palabra regular, lo cual implica que para estos hablantes, en su variedad, la entrada léxica de *moto, foto* o *radio* no puede tener la estructura interna de *mano* en (107), ya que el diminutivo no altera la forma.

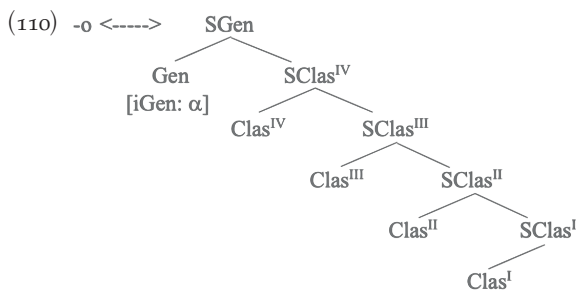
#### 5.4.2. Variedades del tipo «manito»

Para los hablantes que prefieren estas versiones con *-o* tras el diminutivo, parece necesario complicar el análisis levemente proponiendo un clasificador nominal adicional –llamémoslo Clas<sup>IV</sup>–, que se introduce para un conjunto pequeño de sustantivos sobre el clasificador III, y por tanto donde el género copia el rasgo alfa de este clasificador. Esta clase IV estaría restringida a un conjunto muy pequeño de voces, la palabra *mano* para algunos hablantes, y otras fruto del acortamiento de otra palabra más larga donde la *-o* era originalmente una vocal de enlace.



Paralelamente a esto, para estos hablantes, es necesario postular una entrada léxica adicional, que materializa los escasos sustantivos del grupo IV –que creemos que en español peninsular son solo *foto, radio* y *moto*, para aquellos hablantes que sistemáticamente mantienen la marca de palabra *-o* en el diminutivo femenino,

a los que se añade *mano* en otras zonas geográficas, y *soprano* para los hablantes que diminutivizan como *sopran-it-o-*.



Esta *-o* sería el único caso de homofonía que requeriríamos en el análisis para los hablantes que tienen el comportamiento que se acaba de describir, al ser un segundo exponente *-o* distinto de (74) que se emplea en géneros femeninos con sustantivos de la clase IV.<sup>8</sup> Bajo esta premisa, el diminutivo no interfiere en la materialización de la desinencia, porque el exponente para el SN no materializa ninguna parte de la estructura de clasificadores nominales o de género.

## 6. CONCLUSIONES

En este trabajo hemos propuesto que el género tradicional en español debe verse como la materialización de dos áreas a su vez complejas que pertenecen a distintos dominios de la estructura nominal: los clasificadores nominales y el género como rasgo *phi*. La aparente ambigüedad del género entre la derivación y la flexión se debe a que el mismo exponente, la llamada «marca de palabra» materializa núcleos

8. La cuestión que subyace a este problema es que, en el fondo, los hablantes pueden introducir capas adicionales de los clasificadores nominales sin que esto afecte al comportamiento funcional del sustantivo, ya que esto puede hacerse sin afectar a la definición del género en el área de los rasgos *phi*. De hecho, recientes movimientos sociales entre los hablantes están extendiendo en varios grupos el uso de *-e* como morfema adicional de género con valor no marcado (cf. Gil, 2020 para una revisión crítica). El hecho de que este morfema *-e* se emplee también en la concordancia con adjetivos y sustantivos, convirtiendo a los adjetivos de dos terminaciones en sustantivos de tres (*unas alumnas buenas, unos alumnos buenos, unas alumnas buenas*) sugiere con fuerza que se trata de un clasificador adicional que carece de rasgo alfa y posiblemente se introduzca por debajo del clasificador I, que para esos hablantes lleva un rasgo adicional de género marcado; nótese que el diminutivo no parece impedir la presencia de este morfema *-e* de género inclusivo (*amigues ~ amiguites*), frente a la *-e* epentética que se ha analizado en la sección 5.3. El análisis de estos casos se deja para análisis posterior.

de las dos áreas. Asimismo, hemos argumentado que el masculino en los sustantivos debe considerarse como un neutro, es decir, la ausencia de valor de género.

Hemos propuesto solo dos exponentes, *-o* y *-a*, con entradas sintácticamente complejas que, mediante la aplicación del Principio del Superconjunto, dan cuenta de 8 grupos de sustantivos que difieren en el clasificador que toman y en el tamaño del exponente nominal. El noveno grupo de sustantivos, *mano-foto*, no requiere más exponentes en las variedades que regularizan al femenino en contextos de *-it-* (*manita*, *fotita*), mientras que en las variedades donde se preserva (*manito*, *fotito*) requieren introducir un cuarto clasificador nominal cuyo exponente es una segunda *-o*, homófona con el exponente asociado al masculino y posiblemente relacionada con la *-o-* que aparece como vocal de enlace en algunos compuestos.

Este trabajo admite distintas extensiones. La más evidente sería extender el análisis a los pronombres, donde existen tres clases de género. Debido a que los pronombres implican la materialización de estructura funcional alta, dentro del área del determinante, cabe plantearse la posibilidad de que este valor adicional de género marcado como realmente masculino sea el efecto de una partición dentro del área D que añade un rasgo beta  $[\beta]$  a la estructura nominal, que solo puede tener  $[\alpha]$  o la ausencia de rasgo, pudiendo así definir una división a tres elementos.

Otra extensión directa sería al análisis de la marca de género en los adjetivos, donde parecen encontrarse al menos cuatro clases relevantes que pueden analizarse en paralelo a algunos de los grupos de sustantivos discutidos aquí:

- (111) a. guapo ~ guapa (cf. niño ~ niña)
- b. hipócrita ~ hipócrita (cf. el artista ~ la artista)
- c. español ~ española (cf. juez ~ jueza)
- d. pobre (cf. el estudiante ~ la estudiante)

Nuestro trabajo no ha analizado de qué manera se pueden dar sistematicidades entre los clasificadores empleados en español y los tipos de raíces o sufijos nominales. Por tanto, igualmente, cabría investigar como extensión de este estudio qué clase de sistematicidades conceptuales existen en la combinación entre clasificadores nominales y raíces, o entre los nominalizadores que introducen su propia clase nominal y la estructura interna de las bases que nominalizan. Naturalmente, lo que he propuesto en este trabajo solo rasca la superficie de la estructura sintáctica del género de los sustantivos en español, y quedan muchos otros aspectos por explorar; esperamos, no obstante, haber al menos hecho una propuesta lo bastante convincente

y explícita para funcionar como punto de partida de otras investigaciones que hagan posible la integración total del género en la estructura sintáctica.

## REFERENCIAS BIBLIOGRÁFICAS

- Abels, K. y Muriungi, P. K. (2008). The focus particle in Kĩitharaka: Syntax and semantics. *Lingua*, 118, 687-731.
- Acquaviva, P. (2009). Roots and lexicality in distributed morphology. *York Essex Morphology Meeting*, 2, 1-21.
- Acquaviva, P. (2018). Categorization as noun construction. En É. Mathieu, M. Dali y G. Zareikar (Eds.). *Gender and noun classification* (pp. 44-63). Oxford: Oxford University Press.
- Aikhenvald, A. (2000). *Classifiers: A typology of noun classification devices*. Oxford: Oxford University Press.
- Ambadiang, T. (1994). *La morfología flexiva*. Madrid: Taurus.
- Arad, M. (2005). *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.
- Audring, J. (2016). Gender. En *Oxford Bibliographies*. <https://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0066.xml>
- Baunaz, L y Lander, E. (2018). Nanosyntax: The basics. En L. Baunaz, L. Haegeman, K. De Clercq y E. Lander (Eds.). *Exploring nanosyntax* (pp. 3-56). Oxford: Oxford University Press.
- Baunaz, L. y Lander, E. (2017). Syncretisms with nominal complementizers. *Studia Linguistica*, 52, 1-34. DOI: <https://doi.org/10.1111/stul.12076>
- Béjar, S. y Rezac, M. (2009). Cyclic agree. *Linguistic Inquiry*, 40, 35-73.
- Blake, B. J. (2000). *Case*. Cambridge: Cambridge University Press.
- Borer, H. (2005). *In name only*. Oxford: Oxford University Press.
- Borer, H. (2013). *Taking form*. Oxford: Oxford University Press.
- Borik, O. y Espinal, M. T. (2012). On definite kinds. *Recherches linguistiques de Vincennes*, 41, 123-145.
- Borik, O. y Espinal, M. T. (2015). Reference to kinds and to other generic expressions in Spanish: definiteness and number. *The Linguistic Review*, 32(2), 167-225.
- Borik, O. y Espinal, M. T. (2020). Numberless kinds: evidence from Russian, *Catalan Journal of Linguistics*, 19, 231-260.
- Caha, P. (2009). *The nanosyntax of case*. Tesis doctoral inédita. CASTL: Universidad de Tromsø.
- Caha, P. y Pancheva, M. (2012). Contiguity beyond linearity: Modeling cross-dimensional syncretisms. Conferencia presentada en *Workshop on the representation and selection of exponents*, Universidad de Tromsø. 7 de junio.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (2000). Minimalist inquiries: the framework. En M. Roger, D. Michaels y J. Uriagereka (Eds.). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89-155). Cambridge (Mass.): MIT Press.
- Cinque, G. (2010). *The syntax of adjectives*. Cambridge (Mass.): MIT Press.
- Collins, C. y Postal, P. (2012). *Imposters*. Cambridge (Mass.): MIT Press.

- Comrie, B. (1999). Grammatical gender systems: a linguist's assessment. *Journal of Psycholinguistic Research*, 28, 457-466.
- Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. (2007). Gender and noun classes. En T. Shopen (Ed.). *Language typology and syntactic description* (pp. 241-279). Cambridge: Cambridge Univ. Press.
- Corbett, G. (2008). Systems of gender assignment. En Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie (Eds.). *The World Atlas of Language Structures Online*. Múnich: Max Planck Digital Library.
- Dahl, Ö. (2000). Animacy and the notion of semantic gender. En B. Unterbeck (Ed.). *Gender in grammar and cognition, part 1: Approaches to gender* (pp. 99-116). Berlín: Mouton de Gruyter.
- Déchaine, R. M. (2018). Partitioning the nominal domain. En É. Mathieu, M. Dali y G. Zareikar (Eds.). *Gender and noun classification* (pp. 17-40). Oxford: Oxford University Press.
- Embick, D. (2000). Features, syntax and categories in the Latin perfect. *Linguistic Inquiry*, 31, 185-230.
- Fábregas, A. (2007). An exhaustive lexicalisation account of directional complements. *Nordlyd*, 34(2), 165-199.
- Fábregas, A. (2013). Diminutives as heads or specifiers. *Iberia*, 5, 1-44.
- Fábregas, A. (2016). *Las nominalizaciones*. Madrid: Visor.
- Fábregas, A. (2020). *Morphologically derived adjectives in Spanish*. Ámsterdam: John Benjamins.
- Fábregas, A. y Pérez, I. (2008). Gender agreement of adverbs in Spanish, *Journal of Portuguese Linguistics*, 7(2), 25-45.
- Fodor, I. (1959). The origin of grammatical gender. *Lingua*, 8, 186-214.
- Gehrke, B. y McNally, L. (2015). Distributional modification: The case of frequency adjectives. *Language*, 91(4), 837-870.
- Gil, D. (2011). Numeral classifiers. En M. S. Dryer y M. Haspelmath (Eds.). *The World Atlas of Language Structures Online*. Múnich: Max Planck Digital Library. <http://wals.info/feature/55A>
- Gil, J. M. (2020). Las paradojas excluyentes del «lenguaje inclusivo»: Sobre el uso planificado del morfema flexivo *-e*. *Revista Española de Lingüística*, 50(1), 65-84. DOI: <http://dx.doi.org/10.31810/RSEL.50.1.3>
- González Alonso, J., Cunnings, I., Fujita, H., Miller, D. y Rothman, J. (2021). Gender attraction in sentence comprehension, *Glossa*, 6(1), 1-27.
- Harley, H. y Noyer, R. (2000). Formal versus encyclopedic properties of vocabulary: evidence from nominalisations. En B. Peeters (Ed.). *The lexicon-encyclopedia interface* (pp. 349-374). Nueva York: Elsevier.
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22, 27-62.
- Hellinger, M. y Bussman, H. (2001-2003). *Gender across languages: The linguistic representation of women and men*. Ámsterdam: John Benjamins.
- Hockett, C. F. (1958). *A course in modern linguistics*. Nueva York: Macmillan.
- Kiparsky, P. (1973). 'Elsewhere' in phonology. En P. Kiparsky y S. A. Anderson (Eds.). *A festschrift for Morris Halle* (93-106). Nueva York: Holt, Rinehart and Winston.
- Kramer, R. (2015). *The morphosyntax of gender*. Oxford: Oxford University Press.



- Llitas, M. (2008). Del género derivativo al género flexivo. En *Gramma-Temas 3: España y Portugal en la tradición gramatical* (pp. 125-148). León: Universidad de León.
- Llitas, M. (2014). Los nombres ortónimos en tiempos de la fundación de la Real Academia Española. En M. L. Calero y otros (Eds.), *Métodos y resultados actuales en historiografía de la lingüística* (pp. 434-444). Muñiste: Nodus.
- Longobardi, G. (1994). Reference and proper names, *Linguistic Inquiry*, 25(4), 609-665.
- Luraghi, S. (2011). The origin of the Proto-Indo-European gender system: Typological considerations. *Folia Linguistica*, 45(2), 435-64.
- Marantz, A. (1995) *Cat as a phrasal idiom: consequences of late insertion in distributed morphology*. Manuscrito, MIT.
- Marantz, A. (1997). No escape from syntax. *University of Pennsylvania Working Papers in Linguistics*, 4, 201-25.
- Marantz, A. (2000). *Words*. Manuscrito, MIT.
- Mendivil, J. L. (2020). El masculino inclusivo en español. *Revista Española de lingüística*, 50(1), 35-64. DOI: <http://dx.doi.org/10.31810/RSEL.50.1.2>
- Murillo, J. E. (1999): La marcación del género en los sustantivos del español: entre la flexión y la derivación, *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 25(1), 181-192.
- Noyer, R. (1997). *Features, positions, and affixes in autonomous morphological structure*. Nueva York: Garland.
- Pazó, J. (1989). *Morfología léxica del español. La estructura de palabra en nombres y adjetivos*. Tesis doctoral inédita, UAM.
- Pesetsky, D. y Torrego, E. (2007). The syntax of valuation and the interpretability of features. En S. Karimi, V. Samian, y W. K. Wilkins (Eds.), *Phrasal and clausal architecture* (pp. 262-294). Ámsterdam: John Benjamins.
- Piccolo, C. (2008). Gender and number in Romance. *Lingue e linguaggio*, 7, 47-66.
- Preminger, O. (2011). *Agreement as a fallible operation*. Tesis doctoral. MIT.
- Preminger, O. (2014). *Agreement and its failures*. Cambridge (Mass.): MIT Press.
- RAE y ASALE (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Ramchand, G. (2018). *Situations in syntax*. Cambridge (Mass.): MIT Press.
- Ramchand, G. y Svenonius, P. (2006). Mapping a parochial lexicon onto a universal semantics. En T. Biberauer (Ed.). *The limits of syntactic variation* (pp. 219-245). Ámsterdam: John Benjamins.
- Rizzi, L. (1990). *Relativized minimality*. Cambridge (Mass.): MIT Press.
- Serrano Dolader, D. (2010). El género en los sustantivos: ¿flexión y/o derivación?. En J. F. Val, y M. Horno (Eds.). *La gramática del sentido: léxico y sintaxis en la encrucijada* (pp. 49-270). Zaragoza: Prensas Universitarias.
- Starke, M. (2009). Nanosyntax: A short primer to a new approach to language. *Nordlyd*, 36(1), 1- 6. <http://septentrio.uit.no/index.php/nordlyd/index>
- Starke, M. (2011). *Towards an elegant solution to language variation: Variation reduces to the size of lexically stored trees*. Manuscrito, CASTL. Disponible en <LingBuzz/001183>.
- Starke, M. (2013). *Auxiliaries and structural gaps: Current issues in nanosyntax*. Serie de conferencias, CRISSP, Hogeschool- Universiteit Brussel. 18, 20, 22 de marzo.

- Starke, M. (2018). A Note on Kim's Korean question particles seen as pronouns. En L. Baunaz, L. Haegeman, K. De Clercq y E. Lander (Eds.). *Exploring nanosyntax* (pp. 111-115). Oxford: Oxford University Press.
- Taraldsen, T. (2010). The nanosyntax of Nguni class prefixes and concords. *Lingua*, 120, 1522-1548.
- Vanden Wyngaerd, G. (2018). The feature structure of pronouns: A probe into multidimensional paradigm. En L. Baunaz, L. Haegeman, K. De Clercq y E. Lander (Eds.). *Exploring nanosyntax* (pp. 277-304). Oxford: Oxford University Press.
- Wiltschko, M. (2014). *The universal structure of categories*. Cambridge: Cambridge University Press.
- Zaenen, A., Maling, J. y Thráinsson, H. (1986). Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory*, 3, 441-483.

## A TRAVÉS DE LA SUBORDINADA: HIPERCONCORDANCIA EN CONTEXTOS DE DATIVO EXPERIMENTANTE EN ESPAÑOL

Irene FERNÁNDEZ-SERRANO<sup>1</sup>  
*Universitat Autònoma de Barcelona. Centre de Lingüística Teòrica*

### *Resumen*<sup>2</sup>

Este artículo ofrece una aproximación al fenómeno que denominamos «hiperconcordancia»: concordancia de número entre el verbo principal y un argumento de la oración subordinada en estructuras biclausales (*me gustan que los planes salgan bien*). A nivel descriptivo, se centra en estructuras con verbos psicológicos con experimentante dativo y muestra que la hiperconcordancia es posible con cláusulas subordinadas finitas y no finitas. A nivel teórico, se defiende la adopción de la «restricción de caso único» (Nevins, 2004) sobre la operación de Concordancia (Chomsky, 2000, 2001). El análisis propuesto combina esta restricción con la hipótesis de que persona y número constituyen sondas independientes (Rigau, 1991, entre otros) lo que permite dar cuenta del efecto de desbloqueo de la subordinada y de que la concordancia a distancia solamente pueda ser parcial.

*Palabras clave:* Concordancia; número; subordinación; variación sintáctica; dativo experimentante.

### THROUGH THE SUBORDINATE CLAUSE: HYPER-AGREEMENT IN DATIVE EXPERIENCER CONTEXTS IN SPANISH

### *Abstract*

This paper explores the phenomena we dub «hyper-agreement»: number agreement between the main verb and an argument of the subordinate clause in biclausal structures

---

1. irene.fernandez.serrano@uab.cat  <https://orcid.org/0000-0003-0343-0586e>.

2. Este trabajo se ha beneficiado de las ayudas concedidas por el MCIU (FPU16/01327), el MINECO (FFI2017-87140-C4-1-P, IPs: Á. Gallego y J. Mateu) y la AGAUR-Generalitat de Catalunya (2017SGR634, Grup de Lingüística Teòrica). Agradezco a Ángel Gallego y Ricardo Etxepare sus observaciones a una versión previa de este trabajo. También quiero dar las gracias a tres revisores anónimos, cuyos comentarios y sugerencias han sido de gran utilidad para mejorar distintos aspectos del artículo.

(*me gustan que los planes salgan bien*, lit. ‘me like<sub>3pl</sub> that plans work out’). At an empirical level, it focuses on structures with dative experiencers required by a psych-verb and shows that hyper-agreement is possible with both finite and non-finite subordinate clauses. In light of these data, it is argued that Agree (Chomsky, 2000, 2001) must not violate the Single Case Constraint (Nevins, 2004). The analysis put forward combines this constraint with the hypothesis that person and number are independent Probes (Rigau, 1991, among others). This proposal accounts both for the unlocking effect on the subordinate clause and for the fact that the long-distance agreement is restricted to partial agreement.

*Keywords:* Agreement; number; subordination; syntactic variation; dative experiencer.

RECIBIDO: 07/01/2021

APROBADO: 12/04/2021

## 1. INTRODUCCIÓN<sup>3</sup>

La existencia de patrones «inesperados» de concordancia ha suscitado gran interés en el estudio de las lenguas romances (v. p. ej. D’Alessandro y Pescarini, 2016) sin excepción del español. Algunos fenómenos conocidos son la falta de concordancia en las pasivas reflejas, las formas concordantes de *haber* existencial o de *hacer* impersonal. Ilustramos estos fenómenos con ejemplos tomados de Martínez, 1999, p. 2771:

- (1) *Se arregla* paraguas
- (2) A pesar del gran tráfico por las fiestas, no *hubieron* accidentes
- (3) *Hacen* ya dos horas que vino.

En los últimos años, muchos investigadores han dado cuenta de otros patrones poco conocidos o estudiados gracias a la creciente disponibilidad de datos lingüísticos en línea (De Benito y Estrada, 2016) y que facilitan el estudio de distintas variedades y perfiles sociolingüísticos. Algunos de ellos son, por ejemplo, la concordancia a través de complementos preposicionales (Gallego, 2016, 2019), el uso de *se* por *os* (Estrada y De Benito, 2016) o casos de concordancia en contextos de ascenso (Mare y Pato, 2018):

---

3. Las abreviaciones utilizadas en este trabajo son las siguientes: ACUS = acusativo, AE = argumento externo, AI = argumento interno, CdA = condición de actividad, D = determinante, DAT = dativo, def = defectivo, EXP = experimentante, n = (rasgo de) número, N = nombre, NGL = *Nueva Gramática de la Lengua Española*, NOM = nominativo, p = (rasgo de) persona, pl = plural, SC = sintagma complementante, SN = sintagma nominal, SP = sintagma preposicional, SV = sintagma verbal, V = verbo, v\* = verbo pequeño.

- (4) Dijo que se *hablaron* con las autoridades (tomado de Gallego, 2019, p. 96).
- (5) Feliz día de san Juan no *quemarse* con las brasas  
(tuit tomado de Estrada y De Benito, 2016, p. 105, énfasis mío).
- (6) Estas celebridades *parecen* que se llevan una gran diferencia de años  
(tomado de Mare y Pato, 2018, p. 74).

En (4) se observan casos de concordancia a larga distancia en los que el verbo concuerda con un SN dentro de un SP. En este trabajo se presentan datos<sup>4</sup> que muestran que la concordancia es capaz de atravesar una barrera clausal:

- (7) A mí me *chiflan* [oír *esas canciones*] (tomado de Martínez, 1999, p. 2773).
- (8) «La verdad que a uno no le *gustan* [que pasen estas cosas] [...]» (CORPES, «Selección argentina inicia preparación», en *Los Tiempos*. Cochabamba, Bolivia, 24-03-09. Disponible en <lostiempos.com>)<sup>5</sup>.

Utilizaremos la etiqueta «hiperconcordancia» para aludir a este fenómeno<sup>6</sup>. Nos centraremos en los casos de hiperconcordancia que aparecen en contextos con verbos que seleccionan un experimentante dativo como *chiflar* o *gustar*. El dato de (7), recogido por Martínez en la *Gramática Descriptiva de la Lengua Española* (Bosque y Demonte, 1999) demuestra que esta concordancia ha sido observada en este contexto sintáctico por otros investigadores, pero no conocemos un estudio pormenorizado de la misma. Asimismo, cabe destacar que la concordancia a través de cláusulas finitas ejemplificada en (8) no parece haber sido tratada en la bibliografía hasta ahora (exceptuando Fernández-Serrano, 2017).

El objetivo de este trabajo es doble. Por un lado, el de ofrecer una aproximación empírico-descriptiva del fenómeno. Para ello, se consideran datos extraídos de diferentes corpus digitales y de la red social Twitter. Por otro lado, el de sentar las

4. Se reproducen literalmente los ejemplos extraídos de corpus y de Twitter, por lo que pueden observarse faltas de ortografía, de puntuación o de otra índole. Los corchetes y la cursiva siempre son nuestros.

5. Un revisor señala que en el ejemplo de (8) presenta la omisión del verbo copulativo (*la verdad [es] que*). Es relevante notar que, pese a ser un ejemplo de prensa, este dato proviene de una transcripción de una entrevista. Precisamente esta omisión, propia de la lengua oral, sugiere se ha reproducido de forma fiel el discurso del entrevistado.

6. Calcamos esta denominación del término «hiperascenso» (ingl. *hyper-raising*, Ura, 1994), que describe el fenómeno por el cual el sujeto concordado se eleva desde una cláusula subordinada a la oración matriz. Este fenómeno, documentado en distintas lenguas (Martins y Nunes, 2005; Farghal, 2020; entre otros) se ha observado recientemente en español (Mare y Pato, 2018), como ejemplificamos en (6). Pese a estar estrechamente relacionado con el fenómeno de la hiperconcordancia, no lo trataremos aquí por motivos de espacio.

bases para una propuesta teórica en el marco sonda-meta (Chomsky, 2000, 2001). Existen al menos dos retos principales: (i) cómo dar cuenta de la permeabilidad de estas cláusulas, y (ii) restringir la posibilidad de que un mismo SN pueda recibir más de un Caso<sup>7</sup> estructural como consecuencia de la operación de Concordancia (ingl. *Agree*) con más de una sonda. La aportación principal de este trabajo responde a la segunda cuestión.

Defenderemos la hipótesis de que la hiperconcordancia es el resultado de dos operaciones de concordancia: una completa (en el dominio subordinado) y una defectiva (a distancia). Prueba de esta operación defectiva es el hecho de que la hiperconcordancia solo puede manifestarse como concordancia de número. Este análisis es consistente con la idea de que el «Caso estructural» es la contrapartida de una relación sonda-meta no defectiva, pero contraviene la «condición de actividad» (Chomsky, 2000, 2001), ya que defiende que un SN que ha recibido Caso es visible para la concordancia (Nevins, 2004, Bošković, 2003, 2007).

El trabajo se organiza de la siguiente manera: la sección 2 está destinada a la presentación de los datos y la caracterización del fenómeno; la sección 3 ofrece un análisis de la concordancia de número; la sección 4 aborda la cuestión de la localidad y se esboza una propuesta. Finalmente, la sección 5 concluye.

## 2. DATOS DE HIPERCONCORDANCIA EN ESPAÑOL

En esta sección se caracteriza la hiperconcordancia en contextos de predicados psicológicos con experimentante dativo en español mediante datos de diferentes fuentes. Véase un par mínimo básico en (9), donde se ejemplifica este fenómeno en contextos de subordinada no finita y finita respectivamente:

- (9) a. (A María) le *gustan* [leer *novelas*].  
 b. (A María) le *gustan* [que *los niños* lean el libro].

En primer lugar, se introduce brevemente el contexto sintáctico (sección 2.1); a continuación se detallan los datos relevantes (sección 2.2 y sección 2.3); y, finalmente, se aborda la naturaleza dialectal de estos.

---

7. Con la mayúscula se diferencia 'Caso' como término gramatical de otras acepciones de la palabra que puedan aparecer en el texto. Igualmente, se utiliza 'Concordancia' como traducción del término *Agree* (Chomsky 2000, 2001) para distinguirlo de *agreement* 'concordancia'.

### 2.1. Contexto sintáctico: estructuras dativo-nominativas

En distintas lenguas romances se documentan vacilaciones en la concordancia (Rigau, 1991, 1993, 1997, 1999a, 1999b, 2005; Costa 2001; Manzini y Savoia, 2002; entre otros) en estructuras con un «sujeto derivado» (Rigau, 1990, p. 8), es decir, en las que el argumento que controla la concordancia del verbo se genera estructuralmente en una posición propia de los objetos<sup>8</sup>, como es el caso de las oraciones con verbos inacusativos o de las pasivas reflejas<sup>9</sup>. Este tipo de estructuras también se caracterizan por admitir argumentos proposicionales. Como se verá, la combinación de estos factores constituye un contexto donde la hiperconcordancia puede ser esperable.

La concordancia a larga distancia se ha detectado en diferentes lenguas en los llamados contextos de reestructuración, en los que el verbo principal y el verbo subordinado forman una unidad o perífrasis (Bhatt, 2005; Chung, 2004; entre otros). Esta posibilidad se ha observado en español en contextos de pasivas reflejas, con verbos auxiliares modales o aspectuales y otros predicados que pueden considerarse semiauxiliares (Hernanz y Rigau, 1984; Mendikoetxea, 1999; entre otros).

(10) \*Yo {pueden/quieren} vender todas las flores.

(11) a. Se {puede/pueden} vender todas las flores.

b. Se {consiguió/consiguieron} vender todas las flores.

Nótese que, de acuerdo con estas propuestas, los ejemplos de (11) no constituyen ejemplos de hiperconcordancia, ya que se analizan como estructuras monoclausaes. Sin embargo, si se pone en duda el estatus semiauxiliar de verbos como *conseguir* en (11b), puede considerarse que la concordancia está efectivamente cruzando una barrera clausal. Retomaremos esta cuestión en la sección 4.

Si se tiene en cuenta lo expuesto hasta ahora, es esperable encontrar otros ejemplos de hiperconcordancia en estructuras donde la cláusula de infinitivo es un sujeto derivado. Nos centramos a tal efecto en estructuras con ciertos predicados «psicológicos» o de «afección psíquica» (v. Marín, 2015 y ref. allí citadas) que requieren un AE dativo (experimentante) y un AI nominativo:

---

8. Asumimos que los sujetos se generan en una posición estructuralmente inferior al núcleo T(iempo), en la posición denominada de argumento externo (AE), mientras que los objetos (o los sujetos derivados) se generan como complemento de la proyección V en la posición de argumento interno (AI).

9. Algunos autores consideran que las pasivas reflejas son estructuras transitivas canónicas, en las que *se* es un sujeto nominativo, y atribuyen la concordancia con el AI a un fenómeno morfológico (Ormazabal y Romero, 2019 y ref. allí citadas). No adoptaremos esta visión aquí.

- (12) a. A María le {*gustan, encantan, molestan, preocupan*} las tormentas.  
 b. A María le {*gusta, encanta, molesta, preocupa*} [llegar a casa].  
 c. A María le {*gusta, encanta, molesta, preocupa*} [que haya tormenta].

Este tipo de predicados se consideran inacusativos (Belletti y Rizzi, 1988<sup>10</sup>; Acedo-Matellán y Mateu, 2016), entre otras razones, porque concuerdan con el AI<sup>11</sup>. Esto los diferencia de otros verbos que también se clasifican como psicológicos por motivos semánticos:

- (13) a. María {*ama, odia*} las tormentas.  
 b. María {*ama, odia*} [llegar a casa] / [que haya tormenta].

Por tanto, los contextos susceptibles de mostrar concordancia interclausal son fruto de la combinación de dos factores: que su AI pueda ser proposicional y que el AE no controle la concordancia. Como muestra (13), predicados como *amar* cumplen el primer requisito, pero no el segundo. Este punto es esencial, ya que parece indicar que la hiperconcordancia se manifiesta en contextos determinados por la sintaxis e independientemente de posibles factores semánticos. Compárese a este efecto el ejemplo de (9) con (14):

- (14) a. \*María *odian* [leer novelas].  
 b. \*Tú *aman* [que los niños lean].

No es tampoco casual que estos verbos no seleccionen subordinadas de indicativo, pues se ha demostrado que las cláusulas de infinitivo y subjuntivo son más transparentes (Picallo, 1985; Torrego y Uriagereka, 1992; entre otros; v. sección 4). Los datos que aquí se presentan sugieren que esta transparencia puede dar lugar a concordancia a larga distancia.

Antes de pasar a la descripción de los datos, cabe considerar, como nos sugiere un revisor anónimo, otros posibles contextos de concordancia interclausal:

10. En los ejemplos de (12) los verbos pertenecen a los tipos II y III en la clasificación de Belletti y Rizzi (1988). Como es bien conocido, los verbos de clase II como *molestar* también admiten un patrón transitivo en el que el experimentante es un objeto acusativo (*las tormentas la molestan*) (v. Mendivil, 2005). Solo tendremos en cuenta, por las razones expuestas, las estructuras con experimentante dativo independientemente de esta clasificación.

11. Si bien es cierto que algunas de sus propiedades no son prototípicamente inacusativas (v. Rigau, 1990; Pesetsky, 1995), para la discusión que aquí nos interesa utilizaremos este término de forma más general para referirnos a estructuras en las que no se asigna Caso acusativo.



- (15) Tú eres (uno) de los que {apoya/apoyan} esta propuesta  
 (16) Nosotros somos los que lo {afirmamos/?afirman}

En (15) y (16) la concordancia puede darse de forma inversa a lo que se ha mostrado anteriormente: entre el verbo subordinado y un argumento de la oración principal. Sin embargo, de nuevo, catalogar estos datos como ejemplos de hiperconcordancia depende del análisis que se les confiera. Si se defiende la hipótesis que el antecedente asciende a la oración matriz desde una posición interna a la subordinada relativa (Rivero, 1982, *apud* Brucart, 2015, p. 733), la concordancia podría haber tenido lugar de forma interna a la subordinada y la hiperconcordancia sería aparente. De no poder mantener este tipo de análisis, estos ejemplos constituirían, en efecto, una prueba de que la hiperconcordancia en español no es exclusiva de las estructuras referidas en esta sección.

No es posible ofrecer aquí una visión exhaustiva del fenómeno<sup>12</sup>. Por ello este trabajo se centra en las estructuras aludidas más arriba y deja la cuestión de las relativas, así como la del hiperascenso —v. (5)—, para futuras investigaciones.

## 2.2. Hiperconcordancia en cláusulas no finitas

Como ya apuntábamos en la introducción, no disponemos de una descripción detallada de las propiedades ni de la distribución de la hiperconcordancia en estructuras dativo-nominativas en español<sup>13</sup>. Solamente se han encontrado dos trabajos en los que se recogen ejemplos de esta índole. Uno de ellos es *Morfosintaxis del español coloquial* (Vigara Tauste, 2005), en el que se describen ejemplos de conversación espontánea de «carácter urbano», principalmente recogidos en Madrid (p. 35):

- (17) a. Luego además, tú que eres un... fonetista me imagino que *te chiflarán* escuchar los rodajes (Esgueva y Cantarero, 1985, IV, 80)

12. Estos datos sugeridos por el revisor ofrecen complicaciones adicionales: por un lado, en cuanto al tratamiento de las relativas semilibres, que no es unitario en la bibliografía (v. Brucart, 2015), por otro lado, en cuanto al tratamiento de la concordancia. En (15) el indefinido *uno* es el núcleo de una estructura partitiva (NGLE §33.90-p, §33.10h), contexto típico de concordancia *ad sensum*. En (16) la copulativa enfática de relativo fuerza la concordancia en 1ª persona del plural, aunque la 3ª persona parece no ser completamente agramatical, de ahí que indiquemos esta opción con un ? en el ejemplo, si bien «se recomienda evitar» (v. NGLE §40.11f). Por otro lado, con otros pronombres la 3ª persona en el verbo de la relativa es mucho más natural (*soy yo el que llamé*, NGLE §40.11d). Este contraste podría estar relacionado con la discordancia de persona del tipo *los estudiantes lo afirmamos* (v. nota 15). Por motivos de espacio no podemos aquí discutir todas estas cuestiones.

13. El término no aparece como tal en la bibliografía descriptiva del español, v. nota 5.

- b. Sólo me *gustan estudiar* algunas asignaturas y no tengo facilidad para estudiar.  
(tomados de Vigara Tauste, 2005, p. 220, énfasis en el original)

La autora señala que este tipo de discordancias<sup>14</sup> son muy usuales, pese a ser incorrectas<sup>15</sup>. Volveremos a esta cuestión en la sección 2.4.

El segundo trabajo es el capítulo sobre concordancia de la *Gramática descriptiva de la lengua española* (Bosque y Demonte, 1999). Reproducimos los datos ofrecidos por Martínez, 1999:

- (18) a. A mí me *chiflan* oír *esas canciones* [\*A mí me *chifla* oír*las*].<sup>16</sup>  
b. Me *gustan* más estudiar *otras asignaturas*.  
c. *Ya sé que te duelen tener que desmentir esos rumores*.  
(tomados de Martínez, 1999, p. 2773, énfasis en el original)

En este caso, el autor considera que los hablantes analizan de forma indebida la estructura como monoclausal, es decir, toman los dos verbos como un solo complejo perifrástico (v. sección 4).

Encontramos datos similares en corpus y en la red social Twitter<sup>17</sup>:

- (19) a. a mí me *gustan* [ver *estas obras*, todas estas obras en Valladolid] (CORPES, oral, España)  
b. —Si a la niña no le *gustan* [ver *esas cosas*], yo no le he de traer la cabeza que le he ofrecido —replicó Parra— (CORDE, *Amalia*, José Marmol, Argentina).

14. Un revisor anónimo considera que lo que aquí denominamos hiperconcordancia es una mera discordancia, entre el verbo (pluralizado) y la cláusula (en singular). Esta es una posibilidad lógica teniendo en cuenta que el español permite otras discordancias como la de persona con sujetos plurales (*los estudiantes nos esforzamos*). Es importante destacar que se ha defendido ampliamente en la bibliografía que esta discordancia es solo aparente, ya que responde a la presencia de un pronombre tácito con rasgos de 1ª persona del plural (Hurtado, 1985; entre otros. v. Höhn, 2015 y ref. allí citadas). Además, no nos parece casual la presencia de un argumento subordinado plural en los ejemplos que discutimos. Esto sugiere, como defendemos aquí, que existe una relación de concordancia entre el verbo principal y dicho argumento subordinado.

15. Estos ejemplos están marcados con \* en el original, sin embargo, todos los datos que reportados lo incluyen (p. ej. \*¿ *puedes cerrar la ventana?*, p. 68). La autora no explicita a qué se refiere con este símbolo, ni tampoco hace alusión a la noción de (a)gramaticalidad. Por ello, preferimos no reproducir el \* en los ejemplos.

16. Más adelante ponemos en duda la agramaticalidad de la secuencia entre corchetes, ya que existen datos que reflejan la concordancia entre el verbo matriz y el objeto subordinado pronominalizado —v. (23)—.

17. El origen de los tuits corresponde a la información que el usuario/a proporciona en su perfil público de Twitter.

- c. me *encantan* [hacer *dulces*] (PRESEEA, Valencia, España).
- d. ¿y de qué nacionalidad le *gustan* [ver *las películas*]? (PRESEEA, Monterrey, México).
- (20) a. Realmente ya no me *interesan* [mantener *vínculos*] (@wilseveno7, 28-12-20, Medellín, Colombia, tuit).
- b. Como me *entretienen* [hacer *estas cosas*] (@ManuelaSuarezJ, 22-03-14, Salazar, Argentina, tuit).
- c. Así me *gustan* a mi [ver *los pasos* de palio] (@barbara\_rusher1, 10-05-15, Sevilla, España, tuit).
- d. Me *encantarían* [tener *esos músculos*] (@GLandaburo, 29-01-17, Bahía Blanca, Argentina, tuit).

En estos ejemplos se observa que, aparte de la concordancia, estas estructuras no presentan un comportamiento distinto a estructuras biclausales de la misma índole. En (19d), vemos que es posible extraer desde la cláusula subordinada; en (20c), el experimentante aparece entre el verbo flexionado y el infinitivo<sup>18</sup>; y en (20d) la flexión del verbo principal en condicional (*encantaría*) sugiere que el presente no es un factor limitante<sup>19</sup>.

Si bien estos ejemplos contienen distintos predicados psicológicos, podría pensarse que puede existir alguna restricción léxico-semántica, más allá de las propiedades que indicamos anteriormente (v. sección 2.1). Por ejemplo, el predicado *gustar* parece ser el más proclive a permitir esta concordancia. Sin embargo, creemos que esto se debe no solo a un factor de frecuencia de uso<sup>20</sup>, sino que también responde al carácter coloquial del fenómeno (v. sección 2.4). En otras palabras, la hiperconcordancia es

18. El pronombre átono debe aparecer en posición proclítica obligatoriamente, por lo que es el SP dativo el que puede interponerse entre el verbo flexionado y el infinitivo. Nótese que esta posición del OI es obligatoria ante la aparición del foco antepuesto *así* (\**así a mí me gustan*) (v. NGLE §40.3c). Este comportamiento, propio de los sujetos, podría tomarse como una prueba más de que el OI de los verbos psicológicos comparte propiedades con los sujetos pese a no concordar con el verbo (Masullo, 1993; Fernández Soriano, 1999; López, 2007; entre otros).

19. Es necesario indagar en más detalle si el presente (o el aspecto imperfectivo) inciden de forma significativa en la aparición del fenómeno. Un revisor anónimo sugiere, dados los numerosos ejemplos con *ver*, que un verbo de percepción en la cláusula de infinitivo también podría ser un factor favorecedor.

20. Si consultamos la frecuencia normalizada de los verbos psicológicos del tipo que aquí nos interesa en el CORPES, podemos comprobar como el verbo *gustar* (frec. 306,57) es el más frecuente, seguido por *importar* (frec. 130,33) en una búsqueda por lema para incluir todas las formas verbales. En el *Corpus del español* (a partir de ahora CdE) solo aparece el número total de apariciones, en este caso 1.016.203 para *gustar* y solo 318.108 para *importar*. Es necesario notar que estos resultados no excluyen la acepción

poco esperable con verbos propios de un registro formal. La existencia de datos con predicados de naturaleza coloquial parece reforzar esta hipótesis<sup>21</sup>:

- (21) a. *Me chocan* [ver *las fotos* de hace dos años] (@emireth98, 07-05-19, Monclova, México, tuit).  
 b. *Me flipan* [ver *mapas antiguos*] (@ivanGuevara\_, 25-09-12, Málaga, España, tuit).  
 c. *Me matan* [hacer *las cosas* corriendo] (@MiguelJG29, 24-08-14, Jérez de la Frontera, España, tuit).  
 d. *Me molan* [leer *indirectas* por el Twitter] (@A\_blackjack, 09-10-13, Badajoz, España, tuit).  
 e. Saludos desde Argentina, me *copan* [ver *tus videos*] (@EmmaSoul23, 05-03-17, Argentina, tuit).

En este sentido también es esperable encontrar ejemplos con formas analíticas de predicados psicológicos (Di Tullio, 2015):

- (22) a. *Me dan* vergüenza [ver *las propagandas* del gobierno] (@Tato\_Martinez, 02-06-13, Concepción del Uruguay, Argentina, tuit).  
 b. Te haces mayor... ya no te *hacen* ilusión [tener *esas melenas*] (@borjavilarinho, 14-01-13, Pamplona, España, tuit).

Pasemos ahora a considerar qué propiedades pueden indicarnos cuál es el estatus de la subordinada aparentemente «transparente» y qué rasgos intervienen en la concordancia. Empecemos por esta última cuestión.

---

en la que *importar* es transitivo, lo que evidencia aún más la gran diferencia en frecuencia entre el uso de *gustar* y otros verbos psicológicos.

También es interesante comprobar qué sucede en una base de datos de Twitter como la del ASinEs; en ella de nuevo es *gustar* el verbo que más aparece seguido esta vez de *encantar*. Ofrecemos los datos de los paquetes de tuits de España y México como ejemplo:

(i)		España	México
	<i>gustar</i>	18.378	17.393
	<i>encantar</i>	10.369	9.078
	<i>importar</i>	2.793	3.698

21. En la versión preliminar del artículo habíamos escrito «que existan datos con predicados de naturaleza coloquial *parecen* indicar que [...]». Estamos de acuerdo con un revisor en que esto demuestra que este tipo de concordancia es gramatical (aunque incorrecta, por lo que se ha corregido este caso). En este ejemplo, además, la oración subordinada se sitúa en la posición del sujeto preverbal, por lo que estaríamos ante un dato que combina hiperconcordancia con hiperascenso.

Hasta ahora se han mostrado ejemplos en los que el argumento concordante es un sintagma nominal de tercera persona. Dado que los verbos psicológicos en español permiten la concordancia de persona, como ejemplifica (23a), es necesario comprobar si esta concordancia puede mantenerse a larga distancia, (23b).

- (23) a. Le gustamos nosotros.  
 b. \*Le *gustamos* [vernos].

Si el juicio sobre (23b)<sup>22</sup> es correcto, hay que descartar que esta agramaticalidad radique en la imposibilidad de concordar con un clítico acusativo. De hecho, Martínez, 1999 da cuenta de esta restricción, patente en el ejemplo que ofrecimos en (18a) y que recuperamos aquí:

- (24) A mí me chiflan oír *esas canciones* [\*A mí me chifla oír*las*].

Ahora bien, contamos con ejemplos que parecen contradecir esta observación. En (23) se observa que el verbo principal concuerda con el objeto de la subordinada pese a estar pronominalizado:

- (25) a. Igual posta me *divierten* [hacerlos], son fáciles y tienen visitas (@nicosanmarcos, 31-07-20, Llavallol, Argentina, tuit).  
 b. cmo m *gustan* [veros tan felicisssssss] (@JuanPeGarciaRoy, 25-02-13, Córdoba, España, tuit).

El dato de (25b) permite comprobar, además, que la intuición de (23b) parece ser correcta: el verbo principal flexiona en tercera persona del plural, pese a que el objeto es un clítico de segunda persona, lo que sugiere que la hiperconcordancia solamente afecta al número.

### 2.3. Hiperconcordancia en cláusulas finitas

La posibilidad de concordar con un argumento de una cláusula subordinada finita no se encuentra recogida en los ejemplos proporcionados por Martínez, 1999, ni en ninguna obra de la que seamos conocedores. Sin embargo, también hemos

---

22. Si bien la agramaticalidad de (23b) se basa en nuestros propios juicios, es patente que esta secuencia es mucho peor que la concordancia de número a larga distancia, incluso entre los hablantes que no aceptan ninguna de las dos (coinciden en este juicio distintos lingüistas hablantes de español europeo consultados). Aunque de manera informal, también hemos detectado casos de hiperconcordancia en la lengua oral, pero nunca del tipo que ejemplificamos en (23b).

podido recuperar ejemplos de corpus y en mayor medida de Twitter. Como se ha adelantado, estas evidencias corroboran la observación de que las cláusulas de subjuntivo son más permeables que las de indicativo.

Una diferencia notable entre los datos de subordinación no finita discutidos en la sección anterior y los que ahora nos ocupan es el estatus del sujeto subordinado. La ausencia de sujeto de infinitivo predice que la hiperconcordancia solo sea posible con el objeto de la subordinada, es decir, si el infinitivo es un verbo transitivo. Por el contrario, en contextos finitos, no se da tal restricción, ya que esperamos que el argumento más cercano (estructuralmente) al verbo principal controle la concordancia. Este será el AE en verbos transitivos, inergativos y copulativos y el AI en el caso de los inacusativos.

- (26) a. No me *importan* mucho [que *las personas* me traten mal] (@Katcitap, 18-05-14, Venezuela, tuit).  
 b. No me *sorprenderían* [que *ustedes dos* fueran novios] jajja (@solchu138, 08-01-17, México, tuit).  
 c. me *encantan* [que *los planes* salgan bien]<sup>23</sup> (@Aagarozein, 02-05-18, tuit. Madrid, España, tuit).  
 d. me *encantan* [que vuelvan *las tendencias*] (Chile, CdE. Elblogdelupi.com).

No parece ser relevante tampoco que el sujeto sea o no explícito, como vemos en los siguientes ejemplos extraídos de CORPES y del PRESEEA:

- (27) a. Me *gustan* [que *cierren* dramáticamente] (CORPES, Giroud, Pavel «Dejando atrás La edad de la peseta», en *Revista de cine cubano on line* <<http://www.cubacine.cult.cu/es/publicaciones>>, La Habana, 2006-12).  
 b. Suenan así porque así le *gustan* [que *suenen*] (CORPES, Brown, Nelson «Ave castisimas». Managua: Centro Nicaragüense de Escritores, 2007).  
 c. ¿y a ti cómo te *gustan* [que te traten normalmente?] (PRESEEA, Sevilla, España, 26-03-2016).

Los ejemplos en los que el sujeto subordinado no es de 3ª persona parecen indicar, de nuevo, que la concordancia solo afecta al número<sup>24</sup>:

23. Este ejemplo parece ser especialmente aceptado por los hablantes pues su búsqueda entrecomillada ofrece cuantiosos resultados tanto en Twitter como en Google (1670 resultados, búsqueda realizada el 04-01-2021).

24. La mayoría de ejemplos encontrados con sujeto explícito parecen responder a otro tipo de variación relacionada con el doblado de clíticos.

- (28) a. no me *gustan* [que *tengamos* las misma estatura] (@luana\_sclabi, 25-10-20, Santa Fe, México, tuit).  
 b. Nos *gustan* [que *seáis* un poquito originales] (@LorenaGarcia18, 12-05-15, Granada, España, tuit).

Para los casos con subordinadas no finitas mostramos evidencias de concordancia con el objeto pronominalizado —v. (23)—. Es difícil testar esta posibilidad en contextos no finitos, ya que la mayoría de ellos presentan ambigüedad entre la concordancia con el sujeto —elidido en (27)<sup>25</sup>— o con el objeto pronominalizado:

- (29) A mi me *gustan* [que *los* estén juntando] (@CamilaV03962430, 15-11-20, Santiago de Chile, Chile, tuit).

Es necesario buscar contextos en los que el verbo subordinado sea impersonal y que a la vez permita un objeto acusativo<sup>26</sup>:

- (30) a mi si me *gustan* [que se *las* hayan aprobado] (@TirsoSolis, 05-05-20, Panamá, tuit).

En definitiva, los datos que aquí se presentan sugieren que la hiperconcordancia en contextos dativo-nominativos no está restringida a estructuras con subordinada no finita, sino que también pueden darse con cláusulas finitas en subjuntivo.

---

(i) A ellos le molestan que nosotras opinemos de fútbol (@GNRftEminem, 08/08/13, Buenos Aires, Argentina, tuit)

En este caso se produce una discordancia entre el complemento indirecto plural y el clítico dativo singular (NGLE §35.2k), fenómeno documentado en distintas variedades y extendido a diversas construcciones (v. Bogard, 1992; Roca, 1993; Caicedo, 2011):

(ii) a. Luis nunca le da dinero a niños (tomado de Roca, 1993, p. 254).

b. Le pedimos a tres mujeres que hicieran la prueba facial (tomado de Caicedo, 2011, p. 19).

Es cierto que en el ejemplo con *molestar* (i) parece que los hablantes no solo utilizan la forma singular del dativo, sino que además transfieren el plural a la flexión verbal. Aunque es una cuestión interesante, la dejamos pendiente para futuras investigaciones.

25. Los ejemplos encontrados en los que el sujeto elidido no es de 3ª persona del plural son faltas de puntuación y no constituyen evidencia de que haya concordancia con el objeto pronominalizado:

(i) Es por si te gustan[,] que los veas dos veces (@melanie\_\_casado, 12-01-19, tuit).

26. El dato de (30) es particularmente interesante, puesto que la subordinada es una impersonal con *se* que muestra concordancia con el AI pronominalizado, opción normalmente considerada agramatical (*se aprobaron las medidas* / \**se las aprobaron*) (v. Sánchez López, 2002; Ordóñez y Treviño, 2016; entre otros).

#### 2.4. Sobre la naturaleza del fenómeno

Para finalizar la descripción del fenómeno es necesario hacer un breve apunte sobre su naturaleza dialectal. El variado origen geográfico de los datos que se han presentado en las secciones anteriores (secciones 2.2. y 2.3.) sugieren que la hiperconcordancia en contextos dativo-nominativos es una opción para hablantes de distintas áreas dialectales, tanto de la Península como de Hispanoamérica. Aunque no podemos descartar que el fenómeno pueda estar influido por la diatopía, parece más probable que responda a factores diastráticos y/o diafásicos.

Por un lado, es necesario notar que los ejemplos de corpus recogidos provienen de la lengua oral, como los extraídos del corpus PRESEEA, o la reproducen, como en (8) y (19b). Este hecho parece corroborar las observaciones de Vigarra Tauste (2005), quien también relaciona este fenómeno con la oralidad. En este sentido la autora recoge datos propios de la lengua coloquial, que toma como sinónimo de conversacional, y señala que este tipo de variante diafásica también puede presentar diferencias en cuanto al eje diastrático en el que se sitúe el hablante (p. 12).

El carácter no normativo o incorrecto del fenómeno que nos ocupa parece indiscutible (y así lo indican Vigarra Tauste, 2005 y Martínez, 1999), pero precisamente por ello es plausible que estos datos sean gramaticales para hablantes situados en una diastratía popular. Nótese que los datos escritos en plataformas como redes sociales reflejan un tipo de lengua mucho más cercana a la lengua oral y espontánea, sin presión normativa<sup>27</sup> (Estrada y De Benito, 2016 y ref. allí citadas). Por este motivo nos parece reseñable la gran cantidad de muestras de hiperconcordancia que aparecen en Twitter.

Algunos corpus ya cuentan con muestras de fuentes en línea y permiten una aproximación cuantitativa a este tipo de datos. Son ejemplo de ello el CdE, formado por fragmentos de blogs y páginas webs, y el corpus de tuits en español del ASinEs<sup>28</sup>. A modo de ejemplo, en la tabla 1, se compara en ambos corpus el número de datos con hiperconcordancia que aparecen en oraciones con *gustar*,

27. Por ejemplo, el CORPES (versión 0.93) supera los 300 millones de formas ortográficas extraídas de fuentes escritas (ficción y no ficción), mientras que las de origen oral no llegan a los 5 millones (<https://www.rae.es/noticia/el-corpes-xxi-se-actualiza-con-la-incorporacion-de-21-millones-de-nuevas-formas>). Por ello este recurso es de gran utilidad para comprobar si un fenómeno se ha extendido a registros más formales y/o niveles más cultos de la lengua (Buena Fuentes y Sánchez, 2020), pero aquellos como el aquí discutido difícilmente aparecen representados.

28. El CdE cuenta con una muestra de 2.000 millones de palabras. El corpus público de ASinEs ([asines.org/twitter](https://asines.org/twitter)) está constituido por unos 13 millones de tuits, pero no está disponible la información sobre el número de palabras.



*importar* y *encantar* (v. nota 22), según el tipo de subordinada. Los resultados son más numerosos en el CdE que cuenta con una muestra mayor, pero en ambos casos se documenta el fenómeno.

O. principal	CdE web/dialects		ASinEs	
	Cláusula subordinada		Cláusula subordinada	
	no finita	finita	no finita	finita
<i>importar</i>	27	52	7	14
<i>gustar</i>	453	169	59	94
<i>encantar</i>	167	117	33	42
total	647	338	99	150

Tabla 1. Casos de hiperconcordancia en el CdE y en el ASinEs

Es llamativo que el número de ocurrencias según la subordinada sea distinto: la hiperconcordancia con subordinadas finitas parece más frecuente en los tuits que en las muestras de blogs. Sin embargo, esta asimetría se disuelve si se contrasta la aparición del fenómeno con el contexto en el que puede aparecer. Es decir, además del predicado, que la oración subordinada contenga un argumento plural. Considérese la tabla 2.

O. principal	CdE web/dialects		ASinEs	
	Cláusula subordinada		Cláusula subordinada	
	no finita	finita	no finita	finita
<i>gustar</i>				
3ª singular	4667	5219	1083	2420
3ª plural	453 (8,8%)	169 (3,1%)	59 (5,2%)	94 (3,7%)
total	5120	5388	1142	2514

Tabla 2. Número de ocurrencias y porcentaje de hiperconcordancia con el verbo *gustar*

En la tabla 2 se detallan los resultados con el verbo *gustar*, que es el predicado que ha ofrecido más resultados. Se ha calculado el porcentaje de ocurrencias de hiperconcordancia (*gustar* en 3ª persona del plural) respecto al total de ocurrencias de *gustar* en contextos biclausales en los que en la oración subordinada cuenta con un argumento plural<sup>29</sup>. Se observa entonces que la incidencia de la hiperconcordancia

29. Para ello se ha llevado a cabo la búsqueda de *gustar* + que + V3pl o Dpl y *gustar* + INF + Npl o Dpl.

es mayor en contextos con cláusulas no finitas en ambos corpus y, además, el porcentaje en contextos finitos es similar.

Finalmente, cabe destacar otra ventaja que ofrece la utilización de Twitter para la investigación lingüística: el acceso a numerosas muestras del mismo hablante. Esto permite examinar si el hablante produce el fenómeno de forma reiterada. Son ejemplo de ello los datos de (31) pertenecientes a la cuenta de una usuaria venezolana que utiliza reiteradamente estructuras «hiperconcordadas» (en total aparecen 14 tuits con hiperconcordancia de esta usuaria<sup>30</sup>):

- (31) a. A mi me *gustan* [que me cuenten *las películas*].  
 b. Me *gustan* [hablar con *los gays*], no se porque, pero ellos son muy chevere.  
 c. En serio que les *encantan* [que los *traten* mal].

De esta forma se puede paliar parcialmente la presencia de «falsos positivos» en la cuantificación de datos en línea que suelen estar plagados de errores (v. De Benito y Estrada, 2016). En otras palabras, el uso reiterado de esta concordancia por parte de hablantes concretos sugiere que el fenómeno es parte de la gramática de dichos hablantes.

### 3. ANÁLISIS DE LA CONCORDANCIA DE NÚMERO A LARGA DISTANCIA

Una vez acotado el contexto sintáctico y la naturaleza dialectal del fenómeno, pasamos en esta sección a abordar el objetivo teórico del trabajo. En primer lugar, se presenta el marco teórico (sección 3.1.) y, a continuación, se ofrece una propuesta de análisis (secciones 3.2. y 3.3.).

#### 3.1. *Sintaxis del caso y concordancia: supuestos teóricos*

La siguiente discusión se sitúa en el marco sonda-meta propuesto en Chomsky, 2000, 2001 y ss. En este sistema la concordancia gramatical es fruto de una operación de valoración de rasgos denominada «Concordancia» (ingl. *Agree*). Esta operación se caracteriza por establecer una relación entre una «sonda» (ingl. *Probe*) y una «meta»

---

30. Como bien señalan De Benito y Estrada (2018, nota 4), los usuarios de Twitter no han accedido explícitamente a la explotación científica de sus textos, aunque estos son públicos y pueden encontrarse mediante una búsqueda de Google. Por ello, si que se ha indicado la cuenta de la que se han extraído los tuits en el resto del texto, pero en este caso, al aludir a una cuenta concreta, parece pertinente mantener anónima la identidad de la usuaria.

(ingl. *Goal*) a distancia. La sonda es un elemento que entra en la derivación con unos rasgos no valorados<sup>31</sup> y busca en su dominio una meta con la que *cotejar* tales rasgos.

- (32) i. sonda[x:\_] ... meta[x:1]                      Búsqueda  
           |\_\_\_\_\_|  
       ii. sonda[x:1]✓ ... meta[x:1]                      Concordancia

Como vemos en el esquema anterior, la sonda con un rasgo X no tiene valor. Por ello, busca un elemento con un rasgo X valorado (representado aquí mediante un número) con el que coteja su rasgo y obtiene el mismo valor<sup>32</sup>.

En este sistema, es el verbo el que concuerda con el sujeto (Chomsky, 2000, p. 24, Gallego 2008, p. 30), y no viceversa, ya que la sonda es un núcleo funcional, T(iempo)<sup>33</sup> (también denominado Flexión) que determina la flexión verbal mediante la Concordancia con un SN que tiene rasgos nominales o *ragos*  $\varphi$  (número, persona y género) interpretables. Véase este proceso ilustrado en (33)<sup>34</sup>:

- (33) Me gustas tú  
       i. [<sub>SC</sub> C [<sub>ST</sub> Me T <sub>$\varphi$ [p:][n:]</sub>] [<sub>SV</sub> gustas tú <sub>$\varphi$ [p:2][n:sg]</sub>]]                      Búsqueda  
           |\_\_\_\_\_|  
       ii. [<sub>SC</sub> C [<sub>ST</sub> Me T <sub>$\varphi$ [p:2][n:sg]</sub>] [<sub>SV</sub> gustas tú <sub>$\varphi$ [2][sg]</sub>]]                      Concordancia

En (33) T sondea su dominio (mando-c) y encuentra una meta con rasgos interpretables, en este caso el pronombre *tú* con el que puede concordar —cotejar sus rasgos de persona y número.

El segundo aspecto relevante de este sistema es que, como consecuencia de la operación de Concordancia, el SN (meta) recibe Caso estructural. En el esquema anterior, el paso de Concordancia, se completaría de la siguiente manera:

- (34) ii. [<sub>SC</sub> C [<sub>ST</sub> Me T <sub>$\varphi$ [p:2][n:sg]</sub>] [<sub>SV</sub> gustas tú <sub>$\varphi$ [2][sg]NOM</sub>]]                      Concordancia (+Caso)

31. En el sistema propuesto por Chomsky, los rasgos no interpretables requieren ser cotejados, en otras palabras, son rasgos no valorados. Mantendremos aquí este sistema (v. Pesetsky y Torrego, 2004 para una crítica y una propuesta alternativa), por lo que utilizaremos ambos términos indistintamente.

32. Utilizamos números romanos en los esquemas para indicar los pasos de la derivación.

33. En el sistema propuesto por Chomsky (2000, 2001), T hereda los rasgos del núcleo C para evitar conflictos en la transferencia de fases (v. Gallego, 2010; entre otros). No entraremos en esta cuestión técnica y mantendremos la idea, generalmente aceptada, de que T es la sonda que asigna NOM al sujeto.

34. Para mayor claridad expositiva, obviamos aquí el rol del núcleo *v\**. Contemplamos esta cuestión más adelante.

De forma análoga a la asignación de nominativo, la asignación de Caso acusativo tiene lugar a partir de una relación de concordancia. El núcleo asignador de acusativo es el núcleo  $v^*$  ('*v* pequeño')<sup>35</sup>.

- (35) i.  $[sv v^*_{\varphi\{p:\}[n:]} \quad [sv \text{ver-los}_{[3][p]}]]$  Búsqueda  
 ii.  $[sv v^*_{\varphi\{p:3\}[n:pl]} \quad [sv \text{ver-los}_{[3][p][ACUS]}]]$  Concordancia (+Caso)

La relación entre T y el AI que ejemplificamos más arriba —v. (32)-(33)—, es la misma que la que se establece entre  $v^*$  y el AI —v. (33)—, con la diferencia principal de que en español la concordancia con  $v^*$  no se manifiesta como flexión verbal.

Existen dos condiciones sobre Concordancia que son principales para la discusión de la hiperconcordancia. La primera es conocida como *condición de «actividad»* (ingl. *Activity Condition*), CdA a partir de ahora, que establece que Concordancia «desactiva» la sonda y la meta para subsecuentes operaciones sintácticas. De la CdA se deriva que una meta se vuelve invisible al haber recibido Caso y que, por tanto, no puede concordar con más de un verbo ni recibir más de un Caso estructural. La predicción es que el verbo principal no puede concordar con un argumento de la subordinada, ya que este ha sido «desactivado» por una previa relación de Concordancia con una sonda más cercana. Este es el caso de una estructura con una subordinada de infinitivo en un dialecto sin hiperconcordancia:

- (36) Me T gusta ...  $[v^*_{[sv \text{ leer libros}_{+ACUS}]}]$   
 i. \_\_\_\_\_ Concordancia (+Caso)  
 ii. \_\_\_\_\_? \*Concordancia

La segunda condición implica que solamente las sondas  $\varphi$ -completas, con rasgos de persona y número<sup>36</sup>, pueden asignar Caso, me referiré a ella como «condición de integridad». Esta condición establece la posibilidad de que existan sondas incompletas o defectivas. Precisamente, la existencia de este tipo de sondas explica que sea posible la concordancia entre T y el AI en contextos monoclausales de pasivas

35. La proyección *Sv* fue introducida por Larson, 1988 como una proyección intermedia entre SV y ST. Aunque inicialmente se postuló para dar cuenta de las estructuras ditransitivas, ha sido adoptada en el marco generativista para toda clase de oraciones (v. Legate, 2003). De acuerdo con la visión minimalista (Chomsky, 1995 y ss.)  $v^*$  tiene una doble función: proporciona la proyección en la cual se ensambla el AE y asigna acusativo al AI.

36. No discutiremos el rol del rasgo de género, el cual sí que es relevante para el estudio de la concordancia de participio en algunas lenguas romances (D'Alessandro y Roberts, 2010; entre otros). V. Carstens (2001) para una revisión del concepto de sonda defectiva a partir del estudio de lenguas con concordancia verbal de género.

y estructuras inacusativas, en las que incluimos los verbos psicológicos que estamos tratando aquí (v. nota 10). En estas estructuras, se asume que el núcleo *v* es defectivo y, por ello, no hay asignación de acusativo. En (37) ofrecemos una imagen más completa de la derivación de una estructura monoclausal con verbo psicológico (cf. (33)), en la que es T y no *v* el núcleo asignador de Caso.

$$(37) \text{ Me T } [v_{\text{def}} [_{\text{SV}} \text{ gustas tú}_{+\text{NOM}/\text{ACUS}}]]$$

|\_\_\_\_\_|

A continuación, veremos en detalle por qué la hiperconcordancia no se ajusta a las dos condiciones sobre Concordancia que hemos expuesto más arriba y relacionaremos nuestra discusión con el debate previo presente en la bibliografía a raíz de datos de concordancia a larga distancia en distintas lenguas.

### 3.2. Propuesta

En el apartado anterior hemos visto que un requisito para que se produzca una relación entre sonda y meta, siguiendo el sistema de Chomsky (2000, 2001 y ss.), es que estos elementos no hayan participado ya en la operación de Concordancia. Sin embargo, los datos de hiperconcordancia expuestos en la sección 2 parecen contradecir esta restricción.

Consideremos de nuevo algunos de los ejemplos expuestos anteriormente:

(38) [...] me gustan [veros]

(39) me *encantan* [que *los planes* salgan bien]

En (38) el clítico acusativo demuestra que el AI de la subordinada ha recibido Caso gracias a la relación de Concordancia con el núcleo *v*\* de la subordinada. Sin embargo, esto no impide que la sonda T de la oración matriz pueda detectarlo como posible meta resultando en la morfología del plural del verbo principal.

En (39) sucede algo similar. El verbo *salgan* de la subordinada concuerda en persona y número con el sujeto *los planes*, prueba de que se ha establecido una relación de Concordancia completa entre T subordinado y el sujeto. Según la CdA, al haber recibido Caso nominativo, *los planes* está «desactivado». No obstante, de nuevo, la sonda T matriz lo detecta y concuerda con este argumento en número<sup>37</sup>.

37. Un revisor pone en duda que no haya concordancia de persona con el argumento subordinado, cuando este es de 3ª persona, como es el caso de *los planes* en (37), pues considera que no tenemos pruebas

Esta observación no es exclusiva del español. La posibilidad de concordar a través de una cláusula subordinada con SNs con Caso ha sido ampliamente documentada en diversas lenguas como el hindi (Bhatt, 2005), el tsez (Polinsky y Potsdam, 2001), el vasco (Etxepare, 2006, 2012) y diversas lenguas algonquinas como el Blackfoot (Frantz, 1978). (40) es un ejemplo del vasco en el que el auxiliar concuerda con el objeto (con Caso absoluto) de la cláusula nominalizada<sup>38</sup>:

- (40) [Nobela erromantikoak irakurtzea] gustatzen zaizkio  
 novela románticas<sub>A</sub> leer<sub>-N-Det-A</sub> gustar<sub>-hab</sub> Aux<sub>(3plA-3sgD)</sub>  
 'Le gusta leer novelas románticas' (tomado de Etxepare, 2006. p. 303, trad. mía).

Ante estas evidencias empíricas, algunos autores han concluido que la CdA no regula la concordancia (Bhatt, 2005; Bošković, 2003, 2007; Etxepare, 2006, 2012).

Por otro lado, Nevins, 2004 plantea cómo derivar los efectos que tiene la CdA a partir de otras restricciones independientes de la gramática. Este autor propone que lo que regula que un SN no pueda participar en operaciones sintácticas no está determinado por su estatus activo/inactivo, sino por una restricción más básica, la de que un SN no pueda recibir más de un Caso estructural, lo que denomina *Single Case Constraint* 'restricción de caso único' (trad. mía). Aunque Nevins (2004) basa su argumentación en datos sobre el movimiento, cabe preguntarse si, dado lo expuesto más arriba, esta alternativa a la CdA es operativa para la concordancia.

Volvamos pues a nuestros datos. Anteriormente vimos que un núcleo subordinado *v\** concuerda y asigna Caso acusativo al AI de la subordinada. Si este SN se mantiene visible, la sonda T de la oración principal puede detectarlo como meta viable para cotejar sus rasgos.

- (41) Me T gusta ... [T<sub>def</sub> [v\* [sv leer libros]]  
 i. | | +ACUS  
 ii. | | +NOM

---

suficientes para descartar datos como el de (23b) (*nos gustamos vernos*). Incluso si esto último es así, es indiscutible que en (38) (v. también (28)) no hay concordancia de persona con el clítico *os* y nos parece justificado hipotetizar que lo mismo sucede en el resto de datos. En el caso de que la concordancia plena se documente en un futuro, tampoco podemos descartar que existan distintas variantes del fenómeno, de forma similar a las que se documentan con *haber* existencial. Como señalamos más tarde en la nota 45, es bien sabido que los hablantes que aceptan la pluralización no necesariamente aceptan la flexión de persona (*habemos*), que ha recibido análisis independientes.

38. Como indica el autor, las abreviaciones de la glosa de (40) corresponden a A (absolutivo), N (nominalizado), Det (determinante), aux (auxiliar) y hab (habitual).

La siguiente cuestión es cómo explicar que *los libros* no reciba dos Casos estructurales: acusativo de  $v^*$  y nominativo de T, violando la restricción de caso único. Creemos que la respuesta está en el hecho de que la hiperconcordancia solo sea posible en número. De acuerdo con la condición de integridad, por la cual el Caso estructural solo es asignado por sondas completas (persona y número), la sonda T principal no asigna nominativo al objeto de la subordinada porque solamente concuerda en número:

- (42) i. [ $v^*$  [<sub>SV</sub> leer libros]]  
           |\_\_\_\_\_ |+ACUS           conc. completa + asignación de Caso  
       ii. T ... [ $v^*$  [<sub>SV</sub> leer libros]]  
           |.....|                   conc. defectiva (número)

Una ventaja de este análisis es que permite situar el *locus* de la variación en T matriz. En otras palabras, la hiperconcordancia solo es posible en dialectos que cuenten con un T defectivo en su inventario de categorías funcionales (Borer, 1984, Chomsky, 2001)<sup>39</sup>. Esquematizamos este parámetro en (43), donde utilizamos la nomenclatura H/-H para referirnos a dialectos con y sin hiperconcordancia respectivamente:

- (43) a. dialecto H            $T_{[número]}$   
       b. dialecto -H        $T_{[persona][número]}$

En definitiva, esta hipótesis establece un parámetro en cuanto al tipo de T (completo o defectivo) disponible en cada dialecto y predice que la hiperconcordancia será posible solamente si no se viola la restricción de caso único, vigente para las dos variedades. Para mayor claridad, ofrecemos las respectivas derivaciones en (44)-(45):

- (44) Me gustan [leer libros]  
       i. [ $v^*$  [<sub>n:3|p:pl</sub>] [<sub>SV</sub> leer libros<sub>+ACUS</sub>]]           Concordancia, + Caso  
       ii.  $T_{[n:pl]}$  ... [ $v^*$  [<sub>SV</sub> leer libros<sub>+ACUS</sub>]]       Concordancia,\*Caso  
       (45) \*Me gustan [leer libros]  
       i. [ $v^*$  [<sub>n:3|p:pl</sub>] [<sub>SV</sub> leer libros<sub>+ACUS</sub>]]           Concordancia, + Caso  
       ii. \* $T_{[n:|p:]}$  ... [ $v^*$  [<sub>SV</sub> leer libros<sub>+ACUS+NOM</sub>]]   \*Concordancia, +Caso

39. La idea de que la variación depende de las piezas léxicas disponibles en el lexicon, en concreto de los núcleos funcionales, se conoce como la *Conjetura de Borer-Chomsky* (Baker, 2008, p. 156).

Aunque en (44)-(45) se muestran ejemplos con cláusulas subordinadas de infinitivo, este análisis también da cuenta de los contextos finitos. La única diferencia es que en los casos finitos la sonda defectiva impide que el sujeto de la subordinada reciba nominativo por duplicado.

Antes de comentar las posibles predicciones de este análisis, es necesario hacer un apunte sobre el estatus de las sondas defectivas. Según la teoría en la que enmarcamos nuestro trabajo, una sonda defectiva no puede asignar Caso, ni tampoco puede reflejar morfológicamente la relación de concordancia (defectiva) que establece con la meta. Es decir, en un contexto de *v*-defectivo como el de los verbos psicológicos, se establecen dos relaciones de concordancia, la de *v* con el AI y la de T con este mismo argumento, pero solamente la segunda tiene reflejo morfológico:

(46) Me T<sub>[persona][número]</sub> [v<sub>[número]</sub> [sv gustas tú]]

Consideramos que esto no es un problema, ya que la relación de concordancia con *v*, solo se manifiesta explícitamente en forma de marca de acusativo. Si *v* defectivo no asigna Caso, que valore un rasgo de número no debería afectar a la derivación final<sup>40</sup>.

### 3.3. Concordancia de número en español: extensiones del análisis

En el apartado anterior hemos propuesto que en español dos sondas distintas pueden concordar con un mismo SN con la condición de que solo una de ellas sea asignadora de Caso. En esta sección exploramos dos posibles predicciones de esta hipótesis.

En primer lugar, esperamos encontrar otras pruebas, aparte de la hiperconcordancia, que pongan de manifiesto que la concordancia con argumentos con Caso es una posibilidad de la gramática. Constituyen evidencia de ello algunos fenómenos de concordancia bien conocidos, como es la variante concordada de *haber* existencial (NGLE §41.6b; Fernández Soriano y Taboas, 1999; Rodríguez-Mondoñedo, 2007; entre otros):

- (47) a. Hubieron dos estudiantes en la fiesta  
 b. Los hubieron (tomado de Rodríguez Mondoñedo 2007, p. 3).

40. Richards (2012) pone en duda un sistema en el que haya rasgos valorados en una sonda defectiva. Como alternativa, propone la existencia de núcleos totalmente defectivos, sin rasgos (v. Legate, 2003 y ss. para una argumentación sobre la necesidad de mantener la proyección de *v* en todas las derivaciones).



El AI de *haber* recibe Caso acusativo en la variante concertada, como prueba la pronominalización de (47b). Esto sugiere que el AI concuerda primero con *v* y luego con T en número<sup>41</sup>.

Asimismo, como adelantamos en la sección 1, ciertas variedades admiten la concordancia en impersonales con *se* con argumentos introducidos por una preposición, ya sea la *a* de Marcado Diferencial de Objeto (Martín Zorraquino, 1979; Mendikoetxea, 1999; Sánchez López, 2002; Ordóñez y Treviño, 2016; entre otros<sup>42</sup>) o, en menor medida, otras preposiciones (Gallego, 2016, 2019)<sup>43</sup>:

- (48) a. Se *evacuaron* a más de 120.000 damnificados  
 b. Se *rescataron* a los alpinistas  
 (tomados de Ordóñez y Treviño 2016, pp. 236 y 252).
- (49) a. Dijo que se *hablaron* con las autoridades.  
 b. En Santiago anoche se *informaron* de cuatro homicidios  
 (tomados de Gallego 2019, p. 96).

De nuevo, no solo observamos que un SN al que ya se le ha asignado Caso es una meta, sino que, además, la concordancia solo puede afectar al número:

- (50) Se \*{*pensa-mos/-áis*} en {*nosotros/vosotros*} (tomado de Gallego 2019, p. 102).

Si bien es cierto que la imposibilidad de que el verbo admita flexión de persona es una característica propia de las estructuras *con se* (López, 2007), esto no constituye un contrargumento para nuestra hipótesis: el rasgo de persona de T no puede concordar con el AI, ya sea porque no lo posee (es defectivo) o porque ha concordado con un clítico de 3ª persona (*se*) (v. Mendikoetxea, 1999; D'Alessandro, 2007; Planells, 2017). Por otro lado, el rasgo no valorado de número puede cotejarse con el rasgo interpretable del AI, aunque este SN haya recibido Caso de otra sonda.

---

41. Es conocido que algunas variedades permiten la flexión de 1ª y 2ª persona del plural del verbo *haber*:

(i) Hemos/habéis veinte alumnos en la clase

Sin embargo, (i) se considera un fenómeno independiente de la pluralización a la que nos referimos en (47) (v. p. ej. Pato, 2016, p. 382). Agradecemos a un revisor anónimo una nota aclaratoria sobre esta cuestión.

42. Ordóñez y Treviño (2016, p. 238) apuntan que la observación de este fenómeno se remonta a Bello (1874).

43. Ormazabal y Romero (2019) defienden que estos datos no pueden explicarse sintácticamente, sino que constituyen un fenómeno post-sintáctico. No compartimos este análisis, ya que creemos que no predice que algunos patrones sean imposibles, como por ejemplo la concordancia a través de adjuntos preposicionales (\**se hablan mucho en los pasillos*) (Colomina y otros, 2020).

A la luz de los ejemplos de (48)-(49), no es inesperado encontrar datos de hiperconcordancia a través de preposición. En (51) recuperamos algunos ejemplos de Twitter:

- (51) a. A mi me *interesan* [conocer [SP a *mis compañeros*]] (@CyntuL, 28-04-2014, Córdoba, Argentina, tuit)  
 b. Me *cuestan* [creer [SP en todas *las palabras* que me dijo]] (@DaniiA\_Dj, 21-09-2016, Argentina, tuit)  
 c. Me *gustan* [reirme [SP de *las desgracias* de los demás]] (@mayradmaslow, 08-02-2013, México, tuit)

Otra posible evidencia se encuentra en la bibliografía sobre el español de los denominados «hablantes de herencia» (v. Aalberse, Backus y Muysken, 2019 y ref. allí citadas). Estos hablantes, parecen aceptar la concordancia de número con el dativo experimentante en contextos de verbos psicológicos (Dvorak y Kirschner, 1982; Montrul, 2016 y ref. allí citadas):

- (52) a. A él le *gusta* los deportes.  
 b. A ustedes les *gustan* este libro.  
 (Dvorak y Kirschner, 1982; tomados de Julià, 2015, p. 10).

Julià (2015, p. 10) apunta que, pese a que los hablantes de herencia producen cruces entre el español y el idioma al que están expuestos (en estos estudios, el inglés), no permiten la concordancia de persona con el dativo (v. Pascual Cabo, 2013), error que sí es propio de aprendices de español como lengua extranjera (Dvorak, 1983, p. 25). Aunque es necesario indagar más sobre esta cuestión, de nuevo, se atestiguan datos que parecen indicar que ciertos hablantes aceptan esta concordancia:

- (53) a. *A algunos les dan* miedo el feminismo (@nadiacastilla4, 02-02-18, La Plata, Argentina, tuit).  
 b. [...] *a muchos nos preocupan* el no encontrar una oportunidad para iniciarnos (CdE, tecoloco.com.gt, Guatemala).  
 c. *A muchos les parecen* que tanta libertad para el individuo equivale a una anarquía total  
 (CdE, Argentina, tomado de Mare y Pato, 2018, p. 92).

Este tipo de ejemplos parecen estar relacionados con un fenómeno de aparición del pronombre nominativo propio de la lengua coloquial (NGLE §40.31) y especialmente presente en el sur de España (Ruiz Sánchez, 2013):

- (54) a. *Yo me parece que esto no tiene solución*<sup>44</sup>.  
 b. *Yo ya no me gusta nada la política*.  
 c. *Yo a mí me parece que en Texas no hay marismas*.

Es crucial notar que el pronombre nominativo coexiste con el dativo, que debe expresarse obligatoriamente como mínimo mediante el clítico (*me*). Por este y otros motivos, *yo* se ha analizado como un marcador discursivo, probablemente gramaticalizado (Ruiz Sánchez, 2013). Este análisis es consistente con que la presencia del nominativo no influya en la concordancia verbal<sup>45</sup>.

En definitiva, de estas evidencias se desprende que, pese a que el grado de aceptabilidad de estos fenómenos puede ser variable y, probablemente, responda a algunos factores dialectales (o incluso de procesamiento) que necesitarían de una investigación más pormenorizada, existen ciertos principios de la gramática que no pueden violarse.

#### 4. HIPERCONCORDANCIA Y LOCALIDAD

Hasta ahora hemos obviado un aspecto nada trivial en el análisis de la hiperconcordancia: el de cómo derivar la permeabilidad o transparencia de los dominios que permiten tal relación a distancia. Es esta, precisamente, la cuestión que ha suscitado y sigue suscitando más debate en la bibliografía sobre concordancia a larga distancia (v. Schütze, 2020 para un resumen). Esto no es sorprendente si consideramos que, para empezar, tampoco existe un consenso sobre por qué ciertos dominios son opacos.

Se ha demostrado, por ejemplo, que en español las cláusulas de infinitivo y de subjuntivo son más transparentes que las subordinadas de indicativo (Picallo, 1985; Torrego y Uriagereka, 1992; Gallego y Uriagereka, 2011; Uriagereka, 2015) para ciertos aspectos sintácticos. Sin embargo, para la concordancia, ninguna de ellas lo es, a excepción del fenómeno que hemos presentado aquí:

- (55) a. [Que conozcas a tus compañeros] {es/\*son} {importante/\*importantes}  
 b. Se {sabe/\*saben} [que las clases se harán a distancia]

Podemos establecer la misma comparación con las cláusulas de infinitivo que en ciertas condiciones sí que son más permeables, por ejemplo, en contextos de

44. Agradecemos estos ejemplos a un revisor, quien indica que (54c) es el título de una canción del grupo *No me pises, que llevo chanclas*.

45. Aunque sería interesante comprobar si la aparición del pronombre nominativo propicia la concordancia con el experimentante dativo que se reporta en (53).

elevación para la extracción del sujeto —v. (56)— (Torrego, 1998, 2002), aunque tampoco suelen serlo para la concordancia —v. (57)—.

- (56) Los niños<sub>i</sub> parecen [tener fiebre t<sub>i</sub>]  
 (57) [Conocer a tus compañeros] {es/\*son} {importante/\*importantes}

Una excepción son los procesos de reestructuración (Rizzi, 1978; Hernanz y Rigau, 1984; Wurmbrand, 1998 y ss.) por los que una estructura biclausal deviene monoclausal mediante la formación de una perífrasis (v. Paradís, 2019 para una revisión). De hecho, esta es la hipótesis que plantea Martínez (1999, p. 2773) para los datos de hiperconcordancia a través de cláusulas de infinitivo, como se indicó en la sección 2.2.

Sin embargo, no parece haber evidencias claras de que los casos de hiperconcordancia discutidos aquí sean fruto de procesos de reestructuración (v. Fernández-Serrano, 2017). En primer lugar, porque los verbos psicológicos no han sido clasificados como verbos de reestructuración (Etxepare, 2012, p. 337) y, en segundo lugar, porque esta hipótesis no da cuenta de la concordancia a través de cláusulas finitas<sup>46</sup>.

Según la teoría de fases (Chomsky, 2000 y ss.; Gallego, 2010), la opacidad se deriva del hecho de que el material sintáctico se ha transferido a las interfaces. Es decir, si la cláusula subordinada es una fase, ninguno de los elementos de esa cláusula, excepto el núcleo de fase y su especificador<sup>47</sup>, estarán ya disponibles para el resto de operaciones sintácticas. Dicho de otra manera y aplicado al caso que nos interesa: T matriz no puede cotejar sus rasgos con el sujeto de la subordinada, porque este ya no está en la sintaxis. En (58) enmarcamos el dominio transferido para ilustrar este proceso:

- (58) Me gusta que los niños lean
- |     |   |                       |
|-----|---|-----------------------|
| i.  | [ <sub>SC</sub> C que [ <sub>ST</sub> los niños T [ <sub>SV</sub> lean]]]]                              | Transferencia         |
| ii. | [ <sub>SC</sub> C [ <sub>ST</sub> Me T <sub>{p3}[n.sg]} [<sub>SV</sub> gusta [<sub>SC</sub> C que</sub> | Búsqueda/Concordancia |
|     | _____?]   |                       |

46. Agradecemos a un revisor anónimo que nos señale la existencia de propuestas como la de den Dikken (2017). Este autor sugiere que en las estructuras de hiperascenso el sujeto que se eleva se genera en la oración matriz, mientras que la oración subordinada es en realidad un complemento predicativo que modifica al sujeto (v. Farghal, 2020, cap. 3 para una discusión). De adoptar este análisis para los contextos con cláusulas finitas, la hiperconcordancia sería solo aparente, porque, de forma análoga a la reestructuración, se estaría postulando una estructura monoclausal. Para poder evaluar la hipótesis de la predicación es necesario explorar en profundidad la conexión entre hiperconcordancia e hiperascenso, lo que, como ya se ha indicado, reservamos para futuros trabajos.

47. Según establece la condición de impenetrabilidad de fase (Chomsky, 2001).

Mientras que algunos autores defienden que, al no encontrar una meta viable, T valora sus rasgos por defecto (Béjar, 2003, p. 78; Preminger, 2014; Kalin, 2018; entre otros), resultando en la flexión de 3ª persona del singular; otros han argumentado que las cláusulas también tienen rasgos  $\varphi$  y que, por tanto, pueden participar en relaciones de concordancia (Picallo, 2002; Bošković, 2003, 2007; entre otros). Dejemos por un momento esta posibilidad a un lado y volvamos al tratamiento de la opacidad en el sistema de fases.

En la sección 3 establecimos que una sonda defectiva no puede asignar Caso. Según la teoría de fases, solamente los núcleos de fase pueden ser sondas, de forma que Caso, concordancia y puntos de transferencia dependen de este elemento funcional y más concretamente, de sus rasgos  $\varphi$ . De esta forma, Chomsky, 2000 propuso que la presencia de un núcleo defectivo, no solo conlleva consecuencias para el Caso y la concordancia, sino que también previene que su complemento se transfiera, como ilustramos en (59):

$$(59) \quad T \dots [_{SC} C_{def} \text{ que } [_{ST} \text{ los niños } T [_{SV} \text{ lean}]]]]$$

|\_\_\_\_\_|

Postular que el núcleo de la fase de la oración subordinada no es  $\varphi$ -completo, predice la transparencia de la cláusula, pero no da cuenta de la concordancia que este núcleo sí que establece con el sujeto subordinado<sup>48</sup>. De la misma manera, en los contextos no finitos, vimos que el núcleo de fase  $v^*$  de la subordinada, no puede ser defectivo porque asigna acusativo a su complemento. Las alternativas son postular una teoría distinta para la asignación de Caso y concordancia<sup>49</sup> o considerar la opción que hemos introducido más arriba: suponer que las cláusulas son metas  $\varphi$ .

Esta opción es a grandes rasgos lo que propone Etxepare (2012) para la hiperconcordancia en vasco —v. (40)—. Según su análisis, T concuerda en persona con la cláusula y en número con el SN dentro de dicha cláusula.

48. Mare y Pato (2018) mantienen este análisis, por el cual la sonda T subordinada es defectiva, para los datos de hiperconcordancia en contextos de verbos de ascenso —v. (5)—. Sin embargo, no dan cuenta del problema que indicamos aquí, ya que no ofrecen una discusión teórica pormenorizada sobre la relación entre Caso y concordancia, como ellos mismos indican (p. 93).

49. Gallego (2010) desarrolla esta opción, siguiendo el sistema de asignación de Caso propuesto por Pesetsky y Torrego (2004). Según este autor, las cláusulas de subjuntivo son dominios encabezados por una fase C defectiva.

$$(60) \text{sonda}[\text{persona}][\text{número}] \dots \varphi [ \dots \text{SN} \varphi ]$$

$$\quad \quad \quad \underbrace{\hspace{10em}}$$

$$\quad \quad \quad | \dots \dots \dots \dots \dots \dots |$$

Esta aproximación puede compararse con un efecto que Concordancia parece tener sobre algunos dominios desde los cuales es imposible extraer material. Se ha observado que solamente después de haber concordado, estos dominios permiten la extracción (Rackowski y Richards, 2005, entre otros). Este efecto de «desbloqueo» (trad. mía desde el ingl. *unlocking*, Davis y Branán, 2019) parece estar sucediendo también en vasco: la primera relación de concordancia, «desbloquea» el dominio, la cláusula, para que la segunda relación pueda establecerse (cf. Preminger, 2011).

Un detalle nada trivial es que en vasco este fenómeno aparece en contextos con cláusulas nominalizadas. Esta lengua, además, exhibe concordancia entre el verbo auxiliar y dichas cláusulas. Por tanto, es necesario encontrar evidencias de que las cláusulas subordinadas del castellano (al menos las de subjuntivo e infinitivo) tengan una naturaleza análoga.

La idea de que las cláusulas subordinadas del español son nominales está presente en la bibliografía (Picallo, 2002; Castillo, 2018; Gallego y Uriagereka, 2019), respaldada por datos como los del dequeísmo (Ruiz Alonso, 2020)<sup>50</sup>. Si asumimos esta idea, podríamos entonces postular que el parámetro que diferencia un dialecto con hiperconcordancia de uno que no la acepta, no es la presencia o ausencia de un rasgo de persona en T, sino que sus rasgos constituyan o no sondas independientes:

$$(61) \text{Dialecto H: sondas independientes}$$

$$\text{i. } T[\text{persona}][\text{número}] \quad \varphi_{\text{SC}} [\text{SN}]$$

$$\quad \quad \quad \underbrace{\hspace{10em}}$$

$$\text{ii. } T[\text{persona}][\text{número}] \quad \varphi_{\text{SC}} [\text{SN}]$$

$$\quad \quad \quad | \dots \dots \dots \dots \dots \dots |$$

$$(62) \text{Dialecto -H: sonda única}$$

$$\text{i. } T_{\varphi} \dots \varphi_{\text{SC}} [ \text{V SN} ]$$

$$\quad \quad \quad \underbrace{\hspace{10em}}$$

---

50. Esta hipótesis enlaza con una idea más general de que V solo puede seleccionar complementos nominales (Kiparsky y Kiparsky, 1970; Kayne, 2014; entre otros).

Si las sondas son independientes, cada una establece una relación de concordancia: la persona encuentra la cláusula, mientras que el número, como la cláusula ya ha sido «desbloqueada», se coteja con un SN incrustado. Por el contrario, si ambos rasgos sondean como un solo haz, la cláusula es la única meta disponible.

La idea de que los rasgos pueden operar como sondas independientes ha sido ampliamente defendida en la bibliografía (Rigau, 1991; Sigurðsson, 1996; Chomsky, 2000; Sigurðsson y Holmberg, 2008; Béjar y Rézac, 2003; entre otros), así como la jerarquía entre persona y número (Taraldsen, 1995; Anagnostopoulou, 2003). Asimismo, Richards (2004, p. 169) propone una parametrización en términos similares: en inglés los rasgos de la sonda T se cotejan por separado, mientras que en islandés lo hacen de forma conjunta. Nuestra intuición es que algo similar podría estar sucediendo a escala microparamétrica en español.

Cabe también preguntarse si el resto de contextos en los que hemos visto efectos de transparencia (sección 3.3.) pueden explicarse del mismo modo. Si bien algunos autores han postulado que los dativos pueden contar con una capa de 3ª persona (Richards, 2004; Rézac, 2008<sup>51</sup>), habría que confirmar esta hipótesis para el español, así como ver las posibles implicaciones para los casos de concordancia con un argumento acusativo.

Finalmente, si tenemos en cuenta la teoría de fases, es necesario explicar cómo el material de la oración subordinada permanece accesible para la sonda de la oración principal, pese a ser un dominio que debe ser transferido. Puede considerarse para ello la idea de que el material transferido es visible para el resto de la derivación mientras no sea modificado (Chomsky, 2016; Chomsky, Gallego y Ott, 2019), lo que permitiría establecer la relación a larga distancia de (60).

Con estos precedentes en mente, creemos que este esbozo de análisis puede ser un punto de partida prometedor para dar cuenta de los fenómenos de transparencia que hemos expuesto a lo largo de este artículo. Dejamos pues esta labor para trabajos futuros.

---

51. Este tipo de análisis se ha discutido a la luz de los denominados «sujetos caprichosos» (ingl. *quirky subjects*) del islandés. En esta lengua existen argumentos no nominativos que se comportan como sujetos en el resto de aspectos (v. por ejemplo Sigurðsson, 1996; Sigurðsson y Holmberg, 2008). Respecto al español, algunos autores han considerado que los experimentantes de los sujetos psicológicos también son sujetos caprichosos (Masullo, 1993; Fernández-Soriano, 1999; entre otros), no sin controversia (v. Mendívil, 2003).

## 5. CONCLUSIONES

Este trabajo ha ofrecido una primera aproximación descriptiva al fenómeno de la hiperconcordancia, en concreto, a la concordancia entre cláusulas en contextos con verbos psicológicos y experimentante dativo. La exploración de este fenómeno ha permitido indagar en el análisis de los rasgos que intervienen en la concordancia y en la asignación de Caso y ha suscitado preguntas no triviales sobre el estatus de la subordinación en español.

A nivel empírico se ha mostrado que esta concordancia responde a la ausencia de concordancia con un AE y a la presencia de una cláusula subordinada, no solo de infinitivo, como sugería la bibliografía, sino también de subjuntivo. También se ha demostrado que este tipo de concordancia afecta al número únicamente y se produce con el argumento plural estructuralmente superior de la cláusula subordinada (el objeto en las no finitas y el sujeto en las finitas).

Los datos reportados se han extraído principalmente de fuentes orales, presentes en corpus, y fuentes en línea. Estas últimas, especialmente la red social Twitter, se han mostrado particularmente útiles para detectar fenómenos como el de la hiperconcordancia, perteneciente al lenguaje coloquial y a diastratías populares.

Respecto a la vertiente teórica, hemos enmarcado el fenómeno en la bibliografía generativista que se ocupa de los fenómenos de concordancia a larga distancia en lenguas tipológicamente diversas. En este sentido, se han revisado algunas de las condiciones que regulan la relación entre Caso y concordancia en el sistema sonda-meta (Chomsky, 2000, 2001). Se ha defendido que los efectos derivados de la «condición de actividad», incompatible con la hiperconcordancia, pueden explicarse mediante una restricción independiente sobre la asignación de Caso, la «restricción de caso único» (Nevins, 2004). Este análisis explica por qué la hiperconcordancia solo puede ser parcial (en número): si la concordancia fuera completa, el SN subordinado recibiría dos Casos estructurales, uno del verbo subordinado y otro del verbo principal. Esta propuesta permite explicar otros datos de variación en los que se establece una relación de concordancia, siempre parcial, con un SN que ya ha recibido Caso. Como se ha puesto de manifiesto, la adopción del término hiperconcordancia requiere defender que las estructuras referidas no constituyen contextos monoclausales. En este trabajo se ha mantenido que, efectivamente, las estructuras estudiadas constituyen oraciones complejas. Para dar cuenta de la transparencia de las cláusulas en dialectos con hiperconcordancia, se ha considerado la hipótesis de que en estos dialectos los rasgos de persona y número constituyen sondas independientes (Richards, 2004) y por ello inducen un efecto de «desbloqueo». Esta hipótesis debe seguir siendo explorada,



especialmente a la luz de fenómenos asociados como el hiperascenso y de la posibilidad de que otro tipo de cláusulas como las relativas también puedan ser permeables.

Departamento de Filología Española  
Edificio B (Filosofía y letras), Campus UAB  
08193 Bellaterra (Barcelona)

## REFERENCIAS

### *Fuentes consultadas*

- ASinEs: Gallego, Á. (2014). *Atlas Sintáctico del Español*. <http://asines.org/twitter>.  
CdE: Davies, M. (2002). *Corpus del Español: 100 million words, 1200s- 1900s. Web / Dialects*. <http://www.corpusdelespanol.org>  
CORPES: Real Academia Española. *Corpus del Español del Siglo XXI*. <http://web.frl.es/CORPES/>  
CORDE: Real Academia Española. *Corpus Diacrónico del Español*. <http://corpus.rae.es/cordenet.html>  
PRESEEA (2014). *Corpus del Proyecto para el estudio sociolingüístico del español de España y de América*. Alcalá de Henares: Universidad de Alcalá. <http://preseea.linguas.net>

### *Referencias bibliográficas*

- Aalberse, B. y M. (eds.) (2019). *Heritage languages. A language contact approach*. Amsterdam: John Benjamins.  
Acedo-Matellán, V. y Mateu, J. (2015). Los verbos psicológicos: raíces especiales en estructuras corrientes. En R. Marín (Ed.). *Los predicados psicológicos* (pp. 81- 104). Madrid: Visor.  
Anagnostopoulou, E. (2003). *The syntax of ditransitives: Evidence from clitics*. Nueva York: Mouton de Gruyter.  
Baker, M. (2008). *The syntax of agreement and concord*. Nueva York: CUP.  
Béjar, S. (2003). *Phi-Syntax: A theory of agreement*. Tesis doctoral. University of Toronto. <https://twpl.library.utoronto.ca/index.php/twpl/article/view/6503>  
Béjar, S. y Rézac, M. (2003). Person licensing and the derivation of PCC effects. En A. T. Pérez-Leroux y Y. Roberge (Eds.). *Romance linguistics: Theory and acquisition* (pp. 49-62). Amsterdam: John Benjamins.  
Belletti, A. y Rizzi, L. (1988). Psych Verbs and Theta Theory. *Natural Language and Linguistic Theory*, 6, 291-352.  
Bello, A. (1874). *Gramática de la lengua castellana destinada al uso de los americanos*. París: Roger & Chernoviz.  
Bhatt, R. (2005). Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory*, 23, 757-807. DOI: <https://doi.org/10.1017/CBO9781139644037>

- Bogard, S. (1992). El estatus del clítico de complemento indirecto en español. En R. Barriga y J. García (Eds.). *Reflexiones lingüísticas y literarias: volumen I: Lingüística* (pp. 171-186). México, D.F: El Colegio de México.
- Borer, H. (1984). *Parametric syntax*. Dordrecht: Foris.
- Bošković, Ž. (2003). Agree, phases, and intervention effects. *Linguistic Analysis*, 33, 54-96.
- Bošković, Ž. (2007). On the locality and motivation of move and Agree: An even more minimal theory. *Linguistic Inquiry*, 38, 589-644. DOI: <https://doi.org/10.1162/ling.2007.38.4.589>
- Brucart, J. M. (1999) La estructura del sintagma nominal: las oraciones de relativo. En I. Bosque y V. Demonte (eds.) *Gramática descriptiva de la lengua española* (pp. 395-521). Madrid: Espasa.
- Brucart, J. M. (2015): Oraciones de relativo. En J. Gutiérrez-Rexach (Ed.). *Enciclopedia de Lingüística Hispánica* (pp. 722-736). Londres y Nueva York: Routledge.
- Caicedo, S. (2011). La discordancia entre el complemento indirecto y el pronombre átono le, un fenómeno en aumento. *Íkala. Revista de lenguaje y cultura*, 16(29), 17-43.
- Carstens, V. (2001). Multiple agreement and case deletion: against phi-(in)completeness. *Syntax*, 3, 147-163. DOI: <https://doi.org/10.1111/1467-9612.00042>
- Castillo, L.(2018). The case of the clauses. Evidence from ECM in Spanish. Seminario permanente SemLin, 24 de octubre de 2018, UPV/EHU, Vitoria.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: the framework. En R. Martin, D. Michaels y J. Uriagereka (Eds.). *Step by step* (pp. 89-155). Cambridge, Mass.: MIT Press.
- Chomsky, N. (2001). Derivation by phase. En Kenstowicz, M. (Ed.). *Ken Hale: A life in language* (pp. 1-52). Cambridge, MA: MIT Press.
- Chomsky, N. (2016). A dialogue with Noam Chomsky. Jornada Centro de Lingüística Teórica UAB, 6 de noviembre, Residència d'investigadors, Barcelona. [https://www.youtube.com/watch?v=4Ep\\_92woINQ](https://www.youtube.com/watch?v=4Ep_92woINQ)
- Chomsky, N., Gallego, Á. y Ott, D. (2019). Generative Grammar and the Faculty of language: Insights, questions and challenges. En Á. Gallego y D. Ott (Eds.). *Generative syntax. Questions, crossroads, and challenges, Catalan Journal of Linguistics* (número especial) (pp. 229-261). DOI: <http://doi.org/10.5565/rev/catj.l.288>
- Chung, S. 2004. Restructuring and Verb-Initial order in Chamorro. *Syntax*, 7, 199-233. DOI: <http://doi.org/10.1111/j.1368-0005.2004.00070.x>
- Colomina, P., Fernández Serrano, I., Gallego, Á. y Roca, F. (2020). Desplazamiento de caso en español: Del leísmo al SE no paradigmático. *XLIX Simposio internacional de la SEL*, 21-24 de enero, Universitat Rovira i Virgili, Tarragona.
- Costa, J. (2001). Postverbal subjects and agreement in unaccusative contexts in European Portuguese. *Linguistic Inquiry*, 18, 1-17.
- Cuervo, C. (2010). Some dative subjects are born, some are made. En C. Borgonovo y M. Echevarría y P. Prévost (Eds.). *Selected proceedings of the 12th Hispanic linguistics symposium* (pp. 26-37). Somerville, MA: Cascadilla Proceedings Project.
- D'Alessandro, R. y Pescarini, D. (2016). Agreement restrictions and agreement oddities. En S. Fischer y C. Gabriel (Eds.). *Manual of grammatical interfaces in Romance* (pp. 267-294). Berlín: de Gruyter.

- D'Alessandro, R. y Roberts, I. (2010). Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter. *Natural Language and Linguistic Theory* 28(1), 41-72. DOI: <https://doi.org/10.1007/s11049-009-9085-1>
- D'Alessandro, R. (2007). *Impersonal si constructions*. Ámsterdam: de Gruyter.
- Davis, C. y Branam, K. (2019). Agreement and unlocking at the edge. *Proceedings of the Linguistic Society of America*, 4 (16/1-15). DOI: <https://doi.org/10.3765/plsa.v4i1.4512>
- Davies, Mark. (2002). *Corpus del español: 100 million words, 1200s- 1900s. Web/ Dialects*. <http://www.corpusdelespanol.org>
- De Benito, C. y Estrada, A. (2016). Si Alvar levantara la cabeza: cómo tratar con informantes virtuales, *Revista Internacional de Lingüística iberoamericana* 2(28), pp. 7-17.
- De Benito, C., y Estrada, A. (2018). Aproximación metodológica al estudio de la variación lingüística en las interacciones digitales. *Revista Estudios del Discurso Digital (REDD)*, 1, 74-122. DOI: <https://doi.org/10.24197/redd.1.2018.74-122>
- Di Tullio, A. (2015). Variantes sintéticas y analíticas de los predicados psicológicos», en R. Marín (Ed.). *Los predicados psicológicos* (pp. 185-209). Madrid: Visor.
- Dvorak, T. (1983). Subject-object reversals in the use of *gustar* among New York Hispanics. En Elias-Olivares, L. (Ed.). *Spanish in the US setting: Beyond the Southwest* (pp. 21-36). Rosslyn, Va.: National Clearinghouse for Bilingual Education.
- Dvorak, T. y Kirschner, C. (1982). Mary likes fishes: Reverse psychological phenomena in New York Puerto Rican Spanish. *Bilingual Review / La revista bilingüe*, 59-65.
- Esgueva, M. y Cantero, M. (1985). *El habla de la ciudad de Madrid (Materiales para su estudio)*. Madrid: CSIC.
- Estrada, A. y De Benito, C. (2016). Variación en las redes sociales: datos twilectales. *Revista internacional de Lingüística Iberoamericana*, 28(2), 77-111.
- Etxepare, R. (2006). Number long distance agreement in (substandard) Basque. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International journal of Basque Linguistics and Philology*, 40(1-2), 303-350.
- Etxepare, R. (2012). *Agreement configurations in Basque: Lessons from distance*. Tesis de habilitación. CNRS-IKER, Bayona.
- Farghal, T. M. (2020). *Hyperraising in Jordanian Arabic*. Tesis doctoral. Michigan State University. <https://d.lib.msu.edu/etd/49549>
- Fernández-Serrano, I. (2017). *Long distance agreement in Spanish*. Tesis de maestría, UB-UAB. <http://diposit.ub.edu/dspace/handle/2445/116923>
- Fernández-Soriano, O. (1999). Two types of impersonal sentences in Spanish: Locative and dative subjects. *Syntax*, 22, 101-140. <DOI: <https://doi.org/10.1111/1467-9612.00017>
- Fernández-Soriano, O. y Táboas, S. (1999). Construcciones impersonales no reflejas. En I. Bosque y V. Demonte (Coords.). *Gramática descriptiva de la lengua española* (pp. 1723-1778). Madrid: Espasa Calpe.
- Frantz, D. G. (1978). Copying from complements in Blackfoot. En E. Cook y J. Kaye (Eds.). *Linguistic studies of native Canada* (pp. 89-110). Vancouver: University of British Columbia Press.
- Gallego, Á. (2008). Teoría del caso y sintaxis minimista, *Revista Española de Lingüística*, 3(82), 21-46.

- Gallego, Á. (2010). *Phase Theory*. Ámsterdam: John Benjamins.
- Gallego, Á. (2016). Three types of prepositions in Spanish SE sentences. Consequences for cross-dialectal studies. *Dialectologia*, 17, 51-70
- Gallego, Á. (2019). Long distance agreement in Spanish dialects. En L. Franco, M. Marchis y M. Reeve (Eds.). *Agreement, case and locality in the nominal and verbal domains* (pp. 85-108). Berlín: Language Science Press.
- Gallego, Á. y Uriagereka, J. (2011). Dos tipos de argumentos y la distinción indicativo/subjuntivo. *Cuadernos de la ALFAL*, 3, 188-199.
- Gallego, Á. y Uriagereka, J. (2019). *Off the chain*. Manuscrito.
- Hernanz, M. L. y Rigau, G. (1984). Auxiliarietat i reestructuració, *Els Marges*, 31, 29-51.
- Höhn, G. F. K. (2015). Unagreement is an illusion: Apparent person mismatches and nominal structure. *Natural Language & Linguistic Theory*, 34(2), 543-592. DOI: <https://doi.org/10.1007/s11049-015-9311-y>
- Hurtado, A. (1985). The unagreement hypothesis. En L. King y C. Maley (Eds.). *Selected Papers from the XIIIth Linguistic Symposium on Romance: Languages, Chapel Hill, N.C., 24-26 March 1983* (pp. 187-211). John Benjamins. DOI: <https://doi.org/10.1075/cilt.36.12hur>
- Julià, M. (2015). *Hacia una transitivización: los hablantes de herencia y los verbos psicológicos en español*. Tesis de maestría, Universitat Autònoma de Barcelona.
- Kayne, R. (2014). Why this isn't a complementizer. En P. Svenonius (Ed.). *Functional structure from top to toe* (pp. 188-231). Nueva York: Oxford University Press.
- Kalin, L. (2018). Opacity in agreement. En P.W. Smith, J. Mursell y K. Hartmann (Eds.). *Agree to agree: Agreement in the Minimalist Programme* (pp. 149-177). Berlín: Language Science Press.
- Kiparsky, P. y Kiparsky C. (1970). Fact. En M. Bierwisch y K. E. Heidolph (Eds.). *Progress in linguistics: A collection of papers* (pp. 143-173). La Haya: Mouton.
- Larson, R. (1988). On the double object construction. *Linguistic Inquiry*, 19, 335-391.
- Legate, J. (2003). Some interface properties of the phase. *Linguistic Inquiry*, 34, 506-516. DOI: <https://doi.org/10.1162/ling.2003.34.3.506>
- López, L. (2007). *Locality and the architecture of syntactic dependencies*. Nueva York: Palgrave Macmillan.
- Manzini, R. y Savoia, L. (2002). Parameters of subject inflection in Italian dialects. En P. Svenonius, (Ed.). *Subjects, expletives and the EPP* (pp. 157-119). Oxford: OUP
- Mare, M. y Pato, E. (2018). Parecen que lo olvidan: hyper-agreement in non-standard Spanish, *Borealis*, 7(1), 93-96. DOI: <https://doi.org/10.7557/1.7.1.4317>
- Marín, R. (ed.) (2015). *Los predicados psicológicos*. Madrid: Visor.
- Martínez, J. A. (1999). La concordancia. En I. Bosque y V. Demonte (Coords.). *Gramática descriptiva de la lengua española* (pp. 2695-2786). Madrid: Espasa Calpe.
- Martins, A. M. y Nunes, J. (2005). Raising issues in Brazilian and European Portuguese. *Journal of Portuguese Linguistics*, 4, 53-77.
- Masullo, P. (1993). Two types of quirky subjects: Spanish versus Icelandic. *Proceedings of NELS*, 23, 303-317.
- Mendikoetxea, A. (1999). Construcciones medias, pasivas e impersonales. En I. Bosque y V. Demonte (Coords.), *Gramática descriptiva de la lengua española* (pp. 1631-1722). Madrid: Espasa Calpe.

- Mendivil, J. L. (2003). La estructura ergativa de gustar y otros verbos de afeción psíquica en español. *Actas del V Congreso de Lingüística General: León 5-8 de marzo de 2002* (pp. 2003-2014). Madrid: Arco Libros.
- Mendivil, J. L. (2005). El comportamiento variable de *molestar*: *A Luisa le molesta que la molesten*. En G. Wotjaky otros (Eds.). *Entre semántica léxica, teoría del léxico y sintaxis* (pp. 261-272). Fráncfort: Peter Lang.
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: CUP.
- Nevins, A. (2004). Derivations without the activity condition. *MIT Working Papers in Linguistics*, 49, 283-306.
- Ordóñez, F. y Treviño, E. (2016). Agreement and D.O.M. with *se*: A comparative study of Mexican and peninsular Spanish. En M. Kato y F. Ordóñez (Eds.). *The morphosyntax of Portuguese and Spanish in Latin America* (pp. 236-258). Nueva York: Oxford University Press.
- Ormazabal, J. y Romero, J. (2019). The formal properties of non paradigmatic SE. *Borealis*, 8(1), 55-84. DOI: <https://doi.org/10.7557/1.8.1.4704>
- Paradis, A. (2019). *L'ascens dels clítics: reestructuració i control*. Tesis doctoral, Universitat Autònoma de Barcelona. <https://hdl.handle.net/10803/668058>
- Pascual Cabo, D. (2013). *Agreement reflexes of emerging optionality in heritage speaker Spanish*. Tesis doctoral, University of Florida. <https://eric.ed.gov/?id=ED558260>
- Pato, E. (2016). La pluralización de *haber* en español peninsular. En C. De Benito y O. de Toledo (Eds.). *En torno a haber* (pp. 357-392). Peter Lang. DOI: <https://doi.org/10.3726/978-3-653-06099-7>
- Pesetsky, D. (1995). *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Pesetsky, D. y Torrego, E. (2004). Tense, case, and the nature of syntactic categories. En J. Guéron y J. Lecarme (Eds.). *The syntax of time* (pp. 495-537). Cambridge, MA: MIT Press.
- Picallo, M. C. (1985). *Opaque domains*. Tesis doctoral sin publicar, CUNY.
- Picallo, M. C. (2002). Abstract agreement and clausal arguments. *Syntax*, 52, 116-147. DOI: <https://doi.org/10.1111/1467-9612.00049>
- Planells, S. (2017). Patrones de variación (ideolectal) en torno a las construcciones con SE no- paradigmático. *Revista Española de Lingüística*, 47(2), 29-47.
- Polinsky, M. y Potsdam, E. (2001). Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory*, 19(3), 583-646.
- Preminger, O. (2011). Asymmetries between person and number in syntax: a commentary on Baker's SCOPA. *Natural Language & Linguistic Theory*, 29, 917-37. DOI: <https://doi.org/10.1007/s11049-011-9155-z>
- Preminger, O. (2014). *Agreement and its failures*. Cambridge, MA: MIT Press.
- Rackowski, A. y Richards, N. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry*, 36, 565-599. DOI: <https://doi.org/10.1162/002438905774464368>
- RAE-ASALE (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Rezac, M. (2008). The syntax of eccentric agreement: the person case constraint and absolutive displacement in Basque. *Natural Language & Linguistic Theory*, 26, 61-106. DOI: <https://doi.org/10.1007/s11049-008-9032-6>

- Richards, M. (2004). *Object shift and scrambling in North and West Germanic: A case study in Symmetrical Syntax*. Tesis doctoral, University of Cambridge. <https://ling.auf.net/lingbuzz/000177>
- Richards, M. (2012). On feature inheritance, defective phases, and the movement-morphology connection. En Á. Gallego (Ed.). *Phases: Developing the framework* (pp. 195-232). Berlín: De Gruyter Mouton.
- Rigau, G. (1990). Les propietats d'agradar: estructura temàtica i comportament sintàctic. *Caplletra*, 8, 7-20.
- Rigau, G. (1991). On the functional properties of AGR. *Catalan Working Papers in Linguistics*, 1, 235-260.
- Rigau, G. (1993). El comportamiento sintáctico de los predicados existenciales en catalán. *Revista de lenguas y literaturas catalana, gallega y vasca*, 3, 33-53.
- Rigau, G. (1999a). Los predicados impersonales relativos en las lenguas románicas. *Revista Española de Lingüística*, 29(2), 317-355.
- Rigau, G. (1999b). Relativized impersonality: deontic sentences in Catalan, en E. Treviño y J. Lema (Eds.). *Semantic issues in Romance syntax* (pp. 193-229). Ámsterdam: John Benjamins.
- Rigau, G. (2005). Number agreement variation in Catalan dialects. En G. Cinque y R. Kayne (Eds.). *The Oxford handbook of comparative syntax* (pp. 775-805). Nueva York: Oxford University Press.
- Rivero, M. L. (1982). Las relativas restrictivas con *que*. *Nueva Revista de Filología Hispánica*, 31, 195-234.
- Rizzi, L. (1978). A restructuring rule in Italian syntax. En S. Keyser (Ed.). *Transformational studies in European languages* (pp. 113-158). Cambridge MA: MIT Press.
- Roca, F. (1993). Object clitics in Spanish and Catalan. *Catalan Working Papers in Linguistics*, 2, 245-280.
- Rodríguez-Mondoñedo, M. (2007). *The syntax of objects: Agree and differential object marking*. Tesis doctoral, University of Connecticut. <https://opencommons.uconn.edu/dissertations/AAI3276641/>
- Ruiz Alonso, C. (2020). The nominal nature of embedded clauses. Seminario del CLT 20 de noviembre. Universitat Autònoma de Barcelona, Barcelona.
- Ruiz Sánchez, C. (2013). *Yo a mí me parece*: la gramaticalización de yo como marcador de discurso en el español coloquial. En A. M. Carvalho y S. Beaudrie (Eds.). *Selected proceedings of the 6th workshop on Spanish sociolinguistics* (pp. 1-10). Somerville, MA: Cascadilla Proceedings Project.
- Sánchez López, C. (Ed.) (2002). *Las construcciones con se*. Madrid: Visor.
- Schütze, (2020). Against some approaches to long-distance agreement without AGREE. En B. Bjorkman y D. Currie Hall (Eds.). *Contrast and representation in syntax: A festschrift for Elizabeth Cowper*. Oxford: Oxford University Press.
- Sigurðsson, H. Á. (1996). Icelandic finite verb agreement. *Working Papers in Scandinavian Syntax*, 57, 1-46.
- Sigurðsson, H. Á. y Holmberg, A. (2008). Icelandic dative intervention: Person and number are separate probes. En R. D'Alessandro, S. Fischer y G. H. Hrafnbjargarson (Eds.). *Agreement restrictions* (pp. 251-280). Berlín: de Gruyter.

- Taraldsen, K. T. (1995). On agreement and nominative objects in Icelandic. En H. Haider, S. Olsen y S. Vikner (eds.), *Studies in comparative Germanic syntax* (pp. 307-327). Dordrecht: Kluwer.
- Torrego, E. (1996). Experiencers and raising verbs. En R. Freidin (Ed.), *Current issues in comparative grammar* (pp. 101-120). Dordrecht: Kluwer.
- Torrego, E. (2002). Arguments for a derivational approach to syntactic relations based on clitics. En S. Epstein y D. Seely (Eds.), *Derivation and explanation in the Minimalist Program* (pp. 249-268). Oxford: Blackwell.
- Torrego, E. y Uriagereka, J. (1992). Indicative dependents. Manuscrito.
- Ura, H. (1994). Varieties of raising and feature-based bare phrase structure theory, *MIT Occasional Papers in Linguistics*, 7.
- Uriagereka, J. (2015). Subordinación: indicativos y subjuntivos. En Á. Gallego (Ed.), *Perspectivas de sintaxis formal* (pp. 265-298). Madrid: Akal.
- Vigara Tauste, A. M. (2005). *Morfosintaxis del español coloquial* (2ª ed.). Barcelona: Gredos.
- Wurmbrand, S. (1998). Restructuring control. En H. Lee y H. Moisset (Eds.), *Proceedings of the 22nd annual Penn linguistics colloquium* (pp. 175-188). Filadelfia: University of Pennsylvania.





## ANÁLISIS Y CLASIFICACIÓN DE LAS ESTRUCTURAS COMPARATIVAS DE DESIGUALDAD EN ESPAÑOL

Laura VELA-PLO<sup>1</sup>

Universidad del País Vasco UPV/EHU

### *Resumen*<sup>2</sup>

Este trabajo examina en profundidad las características morfosintácticas de las comparativas de desigualdad en español teniendo en cuenta principalmente datos sincrónicos y recalca la necesidad de reformar algunas propuestas y clasificaciones previas de las comparativas en español. Concretamente, se muestra que la arquitectura de las comparativas está determinada por: (i) el tamaño del estándar de comparación (sintagmático u oracional) y (ii) el tipo de vínculo entre la coda y el correlato comparativo (coordinación o dependencia). Así, se definen por fin las restricciones de selección sintáctica y semántica de los marcadores *que* y *de* en las estructuras comparativas de desigualdad.

*Palabras clave:* comparativas; coordinación; subordinación; elipsis; español.

---

1. [laura.vela@ehu.eus](mailto:laura.vela@ehu.eus).  <https://orcid.org/0000-0002-3368-8548>.

2. Esta investigación ha sido parcialmente financiada por el Programa de Becas Predoctorales 2016-2020, Departamento de Educación, Política Lingüística y Cultura (Gobierno Vasco), el proyecto VASTRUD (PGC2018-096870-B-I00), financiado por el Ministerio de Ciencia, Innovación y Universidades (MiCIU) / Agencia Española de Investigación (AEI) y el Fondo Europeo de Desarrollo Regional (FEDER, UE), el HiTT / Grupo de Investigación en Lingüística Formal (IT769-13 & IT1396-19, Gobierno Vasco), el GIU18 / 221 de la UPV/Grupos de investigación EHU, el proyecto VALAL (MINECO FFI2014- 53675-P, Gobierno de España) y UFI11/14 (Universidad del País Vasco, UPV/EHU). Este texto se presentó como ponencia invitada en el L Simposio y IV Congreso Internacional de la Sociedad Española de Lingüística (SEL; abril 2022) tras recibir el premio anual para tesis doctorales de la SEL. Agradezco especialmente al jurado de la SEL el reconocimiento otorgado, así como a Elena Castroviejo y Myriam Uribe-Etxeberria, y a los miembros de *Hizkuntzalaritza Teorikorako Taldea* (HiTT) e *Hizki Elkarte*a en Vitoria-Gasteiz, el fructífero debate sobre los datos aquí recopilados.

## ANALYSIS AND CLASSIFICATION OF SPANISH INEQUALITY COMPARATIVE STRUCTURES

### *Abstract*

This paper offers an in-depth morphosyntactic and semantic characterization of Spanish inequality comparatives by taking into account mainly synchronic and some diachronic data. This comprehensive study shows the need to adjust previous analyses of Spanish inequality comparatives. The architecture of these comparatives is shown to be determined by: (i) the size of the comparison standard (phrasal or clausal) and (ii) the type of link between the standard and the comparative cluster (coordination or dependency). Finally, the syntactic and semantic selection restrictions of the Spanish standard markers are defined.

*Keywords:* comparatives, coordination, subordination, ellipsis, Spanish.

RECIBIDO: 13/05/2022

APROBADO: 01/06/2022

### 1. INTRODUCCIÓN

La habilidad humana de establecer comparaciones entre objetos o situaciones nos permite comprender el mundo y describir nuestras experiencias. Realizar comparaciones, es decir, reconocer y expresar las diferencias y semejanzas entre dos entidades, se considera una necesidad o capacidad cognitiva básica (Langacker 1987, Stassen 1985, Kennedy 2007).

Tratándose de una capacidad esencial para todo ser humano, entre las lenguas del mundo existe tanto una notable variación lingüística en cuanto a la expresión de relaciones de comparación (Stassen 1985, Beck, Oda y Sugisaki 2004, Kennedy 2007, Bhatt y Takahashi 2011, Bochnak 2013, entre otros) como una gran variedad intralingüística (Pancheva 2006, Merchant 2012, Sáez 1999, Vela-Plo 2018, p. ej.). En consecuencia, desde la década de los 70 numerosas/os filólogas/os y lingüistas han intentado determinar cuál es la representación sintáctica y categorización más adecuada de estas estructuras y se han propuesto clasificaciones tipológicas que definan los posibles patrones comunes en la expresión de comparación entre lenguas. Sin embargo, como se ha señalado en estudios recientes como Jäger (2019) o Lechner y Corver (2017), a pesar de décadas de estudio con notables avances (Bresnan 1973, Chomsky 1977, Hankamer 1973, Heim 1985, Kennedy 1999, Lechner 2004, entre otros), las estructuras comparativas siguen considerándose expresiones

complejas que dan lugar a múltiples problemas sintácticos y de categorización que continúan sin tener una solución satisfactoria.

En el caso específico del español, el estudio de las estructuras comparativas y, particularmente, de la comparación de desigualdad, es una cuestión especialmente interesante debido a que este idioma presenta dos marcadores diferentes para la coda comparativa (también denominada *segundo término de la comparación*). Como se ilustra en los ejemplos (1)-(2) con estructuras comparativas prototípicas, hay dos partículas diferentes que pueden introducir el estándar de comparación en esta lengua: la preposición *de* y el complementante o conjunción *que*.

(1)

Más	mujeres	que	hombres	asistieron al evento.
marcador comparativo + base nominal		marcador del estándar	estándar de comparación	
grupo comparativo		coda comparativa		

(2)

La película	era	más larga	de	lo que esperaba.
		marcador comparativo + base adjetival	marcador del estándar	estándar de comparación
		grupo comparativo	coda comparativa	

Sin embargo, estas partículas no pueden alternar libremente:

- (3) El libro tiene *más* { \* *que* / *de* } cien páginas<sup>3</sup>.  
 (4) Ha comprado *más* botellas de vino { *que* / \* *de* } { Ana / ayer }.

Además, la elección del marcador del estándar en español puede contribuir a una diferencia en el significado de la oración. El par mínimo en (5)-(6) evidencia esta propiedad:

- (5) Antonio comió *más de* un jabalí. (Quizá tres o cuatro jabalís).  
 (6) Antonio comió *más que* un jabalí. (Más comida de la que un jabalí podría comer.)

En el caso de las comparativas en español, encontramos numerosos análisis de carácter descriptivo principalmente (como Bolinger 1950, 1953, Solé 1982, Plann 1984, Gutiérrez Ordóñez 1994, Alarcos Llorach 1994, Romero Cambrón

3. El ejemplo (3) permite una interpretación de alteridad pseudocomparativa (Romero Cambrón 1997) en la siguiente línea: «El libro tiene algo más / más que cien páginas». Dado que este artículo se centra en comparativas con interpretaciones de desigualdad de cantidad o grado, dejaré de lado el análisis de las pseudocomparativas de alteridad por el momento (véase Vela-Plo 2020 donde se discuten las propiedades básicas de las pseudocomparativas con ejemplos en español, euskera e inglés).

1997, entre otros) además de algunos análisis teóricos interesantes con diferentes propuestas de categorización de estas estructuras (Sáez 1992, 1999, Brucart 2003, Sáez y Sánchez López 2013, Reglero 2007, p. ej.). Es importante señalar que en la mayoría de estas obras y también en la *Nueva gramática de la lengua española* de la RAE (2010), como veremos más adelante, se subraya de forma continua la falta de consenso entre autores a la hora de analizar y clasificar las estructuras comparativas en esta lengua.

Asimismo, a pesar de que las construcciones comparativas en español han recibido especial atención en los últimos años (véanse Sáez y Sánchez López 2013 o Mendia 2020), aún no se ha logrado un análisis adecuado de la diferencia entre las comparativas con *que* y *de* en la coda (así lo señalan Bolinger 1950, 1953, Solé 1982, Plann 1984, Gutiérrez Ordóñez 1994, Sáez 1999, Brucart 2003, Gallego 2013 y Sáez y Sánchez López 2013, entre otros). La existencia de dos marcadores del estándar diferentes y el tema sin resolver de su análisis gramatical y categorización se vuelve aun más interesante con la reciente publicación de investigaciones sobre la contribución tanto sintáctica como semántica de los morfemas comparativos y del estándar en lenguas como hindi, chino o inglés (Bhatt y Pancheva 2004, Pancheva 2006, Schwarzschild 2010, Alrenga, Kennedy y Merchant 2012, Bylinina y Lyer 2013).

Ante esta situación, en primer lugar, presentaré las preguntas teóricas esenciales que todavía no han sido resueltas sobre la clasificación y la arquitectura de las comparativas con *que* y *de* en español para, en segundo lugar, poder ofrecer una respuesta clara a cada una de estas polémicas cuestiones. Con ese fin, analizaré en profundidad las propiedades morfosintácticas de las comparativas en español actual y contrastaré sus características con algunas particularidades de las estructuras comparativas en latín y español medieval de las que derivan las construcciones actuales. Este análisis nos revelará que es necesaria una nueva forma de categorizar las comparativas en castellano y también podremos, por fin, aclarar las discrepancias y preguntas planteadas en relación a la naturaleza y estructura sintáctica de estas expresiones y sobre la selección del marcador del estándar *que* o *de*.

En cuanto a la organización del artículo, en la próxima Sección 2 sintetizo la problemática a la hora de clasificar y definir las oraciones comparativas de desigualdad en español y presento las insuficiencias de estudios previos en el abarque teórico de estas comparativas. En la Sección 3 examino en profundidad las características morfosintácticas de las comparativas en español actual teniendo en cuenta datos sincrónicos y algunos aspectos diacrónicos sobre las comparativas con *de* y *que* en

la coda comparativa. En la Sección 4 defiendo una propuesta de recategorización y reforma del análisis de las comparativas en español. Finalmente, en la Sección 5 resumiré las conclusiones principales del trabajo.

## 2. ESTADO DE LA CUESTIÓN SOBRE LAS COMPARATIVAS DE DESIGUALDAD EN ESPAÑOL

Respecto a la cuestión de la variación intralingüística y al caso del español, en concreto, todavía no está clara la naturaleza y complejidad de la coda comparativa de ejemplos como (7)-(8): algunas propuestas defienden la naturaleza sintagmática (es decir, no oracional) del estándar comparativo en estos ejemplos, otras defienden la presencia siempre de una coda clausal que podría verse reducida tras una operación de elipsis comparativa (Sáez y Sánchez 2013) y otros autores presentan argumentos sobre la coexistencia de ambos tipos de coda (Bácskai-Atkári 2014).

- (7) Más mujeres *que hombres* (\**vinieron ayer*) vinieron ayer.
- (8) Ha comprado más botellas de vino *que Ana* (\**compró*).

Respecto a la estructura interna del estándar de comparación, mientras que se ha propuesto que las oraciones comparativas están basadas en oraciones de relativo en lenguas como el japonés, este análisis no puede extenderse a todas las estructuras comparativas en castellano (véanse los ejemplos (7)-(8), donde no parece haber una estructura de relativo en la coda, como sí que es obvio en el caso de (2), en contraste). En concreto, como se argumenta en Vela-Plo (2018, p. 235; también Gutiérrez Ordóñez 1994, Romero Cambrón 1997, 1998, Sáez 1999, Brucart 2003, 2009, Mendia 2020), los estándares de comparación con el marcador *de* no pueden tomar complementos oracionales a menos que estén integrados en una cláusula relativa de cantidad o grado; tal y como se ilustra con el contraste entre las oraciones a-b de (9). Esta restricción no se aplica a todas las expresiones comparativas en español, como se muestra en el ejemplo (10), donde podemos observar que el marcador del estándar *que* sí puede, en este caso, tomar un complemento oracional.

- (9) a. \*Amelia ha conseguido más libros de [*Santi (necesitaba) (la semana pasada)*].  
b. Amelia ha conseguido más libros de [{*cuantos / los que*} *Santi necesitaba la semana pasada*].
- (10) Ayer, más mujeres fueron a la manifestación que [*hombres (fueron) (la semana pasada)*].

En virtud de este paradigma, Mencia (2020) concluye que los comparativos con *de* siempre se combinan con un nominal que denota un grado: ya sea un Sintagma de Medida como en los ejemplos (3)-(5), o un Sintagma Determinante que puede ser modificado por una cláusula de relativo como en (9b). Aunque tal restricción sintáctica o semántica parece ser válida para los comparativos con *de*, las restricciones de selección sintáctica y semántica del marcador del estándar *que* aún no están bien definidas en estudios previos. Según Brucart (2003, 2009), por ejemplo, los comparativos con *que* se combinan principalmente con elementos no cuantificacionales, que pueden ser individuos (como en los ejemplos (7)-(8)) o propiedades, aunque esta última descripción no se especifica más. Además, no hay consenso sobre si *que* puede seleccionar elementos sintagmáticos no oracionales o si solo subcategoriza y selecciona cláusulas que pueden haber sido reducidas por una operación de elipsis, como se ejemplifica en las oraciones (7)-(8) (véase la discusión en Sáez y Sánchez López 2013 sobre la denominada *Elipsis comparativa* en español). Ante esta situación, el presente estudio pretende superar las insuficiencias mencionadas y avanzar en el análisis formal de las estructuras comparativas de desigualdad en español y de los elementos relacionados con ellas.

Como hemos podido observar, son numerosas las dificultades que deben afrontarse a la hora de analizar las estructuras comparativas en español. Sáez y Sánchez López (2013), en su obra *Las construcciones comparativas*, resumen la controversia generada a la hora de clasificar y definir los marcadores del estándar de comparación y la dificultad que presenta clasificar estas estructuras de la siguiente manera:

Los distintos componentes de la cabeza comparativa o grupo cuantificativo (...) y de la coda comparativa (...) mantienen entre sí relaciones de selección y modificación que distan de ser transparentes desde el punto de vista semántico como sintáctico (Sáez y Sánchez López 2013, p. 74).

En su obra, tras resumir y criticar las principales propuestas de análisis sintáctico de las comparativas en español, Sáez y Sánchez López (2013, p. 110) cierran su trabajo con la idea de que aún «se requiere mucha mayor investigación si se desea desenmarañar finalmente este incierto panorama descriptivo» sobre las estructuras comparativas.

De forma similar, en la *Nueva gramática de la lengua española* de la RAE (2010) también se menciona otro punto polémico: el de la categorización de uno de los elementos que introducen la coda comparativa, concretamente, del *que* encargado de introducir el estándar en algunas comparativas en español.

Es polémica la clasificación de las conjunciones subordinantes. En primer lugar, son a veces difusos los límites entre las conjunciones coordinantes y las subordinantes, especialmente en lo que respecta a las conjunciones COMPARATIVAS (*que, como*) (...), incluidas por algunos gramáticos entre las coordinantes y por otros entre las subordinantes (RAE 2010, pp. 604-605).

Teniendo en cuenta las aproximaciones al problema de analizar las comparativas en español expuestas en Sáez y Sánchez López (2013), RAE (2010) y también las discrepancias esbozadas en trabajos previos (Bolinger 1950, 1953, Solé 1982, Plann 1984, Gutiérrez Ordóñez 1994, Romero Cambrón 1997, entre otros), a continuación, pasaré a aunar estos problemas y a agrupar en tres cuestiones esenciales las preguntas planteadas hasta ahora sobre la arquitectura y la composición de las estructuras comparativas:

- *Cuestión 1.* El tipo de nexo entre los dos objetos comparados: ¿se trata de una relación de coordinación o de dependencia (es decir, subordinación)?
- *Cuestión 2.* El asunto del tamaño del estándar de comparación: ¿puede haber codas sintagmáticas en origen en las comparativas en español? ¿Se trata de codas oracionales con elipsis? ¿Son ambas opciones (las codas sintagmáticas y las oracionales) posibles?
- *Cuestión 3.* ¿Cuál es la naturaleza del proceso o procesos (enmarcados bajo la etiqueta de *Elipsis Comparativa*) responsables de silenciar uno o varios constituyentes de la coda de las construcciones que presentan una oración en su estándar?

Estas son las preguntas esenciales a las que ofreceremos una respuesta en el presente trabajo. Para ello, en la siguiente sección presentaré, en primer lugar, mi réplica a las cuestiones 1-3 que acabamos de plantear y, en las siguientes subsecciones, motivaré esta nueva propuesta de categorización de las estructuras comparativas en español mediante un estudio sistemático y en profundidad de las características morfosintácticas de estas expresiones en la actualidad. Además, en algunos casos contrastaré estas particularidades con las características de las comparativas empleadas en español medieval y latín, presentando así una aproximación sincrónica y algunos aspectos diacrónicos de estas complejas construcciones.

### 3. ANÁLISIS MORFOSINTÁCTICO Y SEMÁNTICO: APROXIMACIÓN SINCRÓNICA Y ALGUNOS ASPECTOS DIACRÓNICOS

En respuesta a la *Cuestión 1* descrita arriba, en las próximas subsecciones aplicaremos diversas pruebas sintácticas para demostrar de forma metódica la validez de la propuesta que paso a desarrollar a continuación. En español, mostraré que un subgrupo de construcciones comparativas tiene la estructura subyacente de una coordinación, mientras otro subgrupo de comparativas presenta un estándar de comparación análogo a una estructura subordinada o dependiente. En las siguientes secciones comprobaremos la idoneidad de esta propuesta, ya que, en primer lugar, ofreceré pruebas sintácticas claras que abogan por una primera división de las comparativas españolas en dos tipos de construcciones según la relación entre la coda y el correlato comparativo: *comparativas con coda coordinada* o *comparativas con coda dependiente*.

En segundo lugar, es importante destacar que la división entre dos tipos de comparativas según el nexo de unión entre coda y correlato que defiendo en este trabajo aportará una serie de ventajas relevantes en relación a otra de las cuestiones teóricas que examinamos en este estudio. Con la propuesta de división entre comparativas con coda coordinada y con coda dependiente podremos olvidarnos del proceso de borrado específico para comparativas que se ha denominado *Elipsis Comparativa* (nuestra *Cuestión 3* anteriormente descrita). Esta operación se propuso únicamente para explicar algunas características distintivas de las estructuras comparativas. En lugar de esta descripción *ad hoc*, siguiendo la propuesta de diferenciación entre comparativas coordinadas y dependientes, se entiende que los procesos de borrado que se producen en las comparaciones con coda coordinada son los mismos que se aplican a las estructuras coordinadas de forma más general, mientras que las operaciones de borrado que se observan en las comparativas con coda dependiente son las mismas que operan sobre otros tipos de constituyentes dependientes o subordinados. Por lo tanto, con la presente propuesta de división entre dos tipos de comparativas según la relación entre la coda y el correlato comparativo podremos reducir el conjunto de operaciones de elipsis o borrado que se pueden aplicar sobre cierta cadena lingüística. De esta manera, nuestro análisis seguirá el principio de economía del Programa Minimalista (Chomsky 1995).

En tercer lugar, y en respuesta a la *Cuestión 2* arriba descrita, mostraré que existe otra variable independiente que determina la estructura sintáctica y la elección del marcador del estándar en las comparativas en español: la presencia de una *coda*



*oracional* (con elementos silenciados o no) o la presencia de una *coda sintagmática*, que no se deriva de una oración.

Así, atendiendo a la propuesta formal que acabo de definir, nuestra hipótesis de partida en este trabajo será la que aparece en (11):

- (11) Hipótesis de partida: Deben tenerse en cuenta dos variables morfosintácticas independientes a la hora de definir, analizar y categorizar las estructuras comparativas en español: (i) si la coda es sintagmática u oracional y (ii) si la coda está coordinada o subordinada al grupo comparativo. La selección del marcador del estándar *de* o *que* dependerá del tipo de estándar que encontremos atendiendo a estas dos variables.

Siguiendo esta hipótesis, las oraciones comparativas en español con *que* y *de* podrán categorizarse teniendo en cuenta la primera variable del tipo de relación entre la coda y el correlato (de coordinación o dependencia) y la segunda variable que alude a la estructura subyacente oracional o sintagmática del estándar de comparación. De esta manera, en este trabajo veremos que podemos dividir las comparativas de desigualdad en español en cuatro subtipos principales.

- (12) Principales subtipos de comparativas de desigualdad en español según su estructura subyacente:
- (i) comparativas con coda coordinada oracional;
  - (ii) comparativas con coda coordinada sintagmática;
  - (iii) comparativas con coda dependiente sintagmática;
  - (iv) y comparativas con coda dependiente oracional.

En las siguientes subsecciones examinaré la idoneidad de esta nueva clasificación y ofreceré pruebas sintácticas claras que demuestran la necesidad de esta nueva categorización de las estructuras comparativas con *que* y *de* en español.

### 3.1. *Comparativas con que: propiedades de tipo coordinativo y estándar oracional*

Aunque numerosos estudios han demostrado que las estructuras comparativas tienen una estrecha relación en diversos aspectos con las estructuras coordinadas y/o con los elementos dependientes o subordinados (véase Pinkham 1982, Napoli 1983, Nespór y Napoli 1983, Emonds 1985, Hendriks 1991, Sáez 1992, 1999, Lechner 1999, 2001, 2004, Osborne 2009, entre otros), rara vez encontramos análisis sintácticos o semánticos completos de estas estructuras que, a simple

vista, parecen presentar propiedades contradictorias. Además, en la mayoría de los análisis que encontramos en la bibliografía, se han obviado las características de tipo coordinación que muestran las comparaciones (a excepción de trabajos como Corver 1993 o Lechner 2004 sobre el inglés o Sáez 1999 para el español). Por ejemplo, en el apartado sobre las construcciones comparativas de la *Nueva gramática de la lengua española*, el único comentario sobre las propiedades de tipo coordinativo de las comparativas españolas es el siguiente: «sí reconocer en la conjunción comparativa *que* algunas propiedades de las conjunciones coordinantes» (RAE 2010, p. 863), sin ofrecer más detalles.

Para avanzar en su análisis, a continuación, presentaré de manera exhaustiva las propiedades que hacen que las comparativas de desigualdad en español se parezcan a estructuras coordinadas.

### 3.1.1. *Vaciado y Elevación al nodo derecho*

Las operaciones de elipsis que son características de las estructuras coordinadas, a saber, las operaciones de reducción de segmentos coordinados denominadas *Vaciado* y *Elevación al nodo derecho* (conocidas como *Gapping* y *Right-Node Raising*, en inglés), también están disponibles en las comparativas con *que* en español, como muestro en los siguientes ejemplos:

- (13) a. *Vaciado* en oraciones coordinadas:  
A Jed le gustó Banja Luka y a Svenja ~~le gustó~~ Sarajevo.
- b. *Elevación al nodo derecho* en oraciones coordinadas:  
A muchas personas les gustó el ~~lugar~~ pero a otras les disgustó, el lugar.
- (14) a. *Vaciado* en comparativas con *que*:  
A Jed le gustó más Banja Luka que a Svenja ~~le gustó~~ Sarajevo.
- b. *Elevación al nodo derecho* en comparativas con *que*:  
A más personas les gustó el ~~lugar~~ que les disgustó, el lugar.

La observación de que algunas estructuras comparativas presentan características propias de la coordinación ha sido ampliamente descrita en la bibliografía sobre estructuras comparativas en inglés (Chomsky y Lasnik 1977, p. 495; Napoli 1983, p. 676; Emonds 1985, p. 329; McCawley 1988, p. 282; Hendriks 1991, p. 42; Corver 1993, p. 777; Lechner 2004, p. 91) o alemán (Lechner 1999, 2001; Osborne 2009), pero en menor medida en español (Sáez 1992).

Es importante subrayar que las operaciones de *Vaciado* y *Elevación al nodo derecho* no son gramaticales con comparativas con *de* en español, como muestro en el ejemplo (15), y solo pueden operar en comparativas con *que*. Fundamentalmente, estas dos operaciones de elipsis están prohibidas en los constituyentes dependientes y solo se legitiman en estructuras coordinadas (ver Huang 1977; Corver 1990; Hendriks 1991):

(15) a. *Vaciado* en comparativas con *de*:

\*A Jed le gustó más Banja Luka de lo que a Svenja \_ Sarajevo.

b. *Elevación al nodo derecho* en comparativas con *de*:

\*/?A más personas de las que les disgustó, les gustó, el lugar.

Las oraciones de (15) muestran que las operaciones de *Vaciado* y *Elevación al nodo derecho* ofrecen resultados agramaticales en estructuras dependientes. En contraste, cabe recalcar que las comparativas que permiten estas operaciones de elipsis propias de la coordinación deben tener necesariamente una coda introducida por el marcador del estándar *que*. Como también defiende Sáez (1999) en el caso de las comparativas en español, la disponibilidad de estas operaciones de elipsis de reducción de segmentos en estructuras comparativas es uno de los principales argumentos que apoyan el análisis de un subgrupo de comparativas con *que* como si implicaran una estructura subyacente de coordinación.

### 3.1.2. La Restricción de estructuras coordinadas 1 y el Movimiento transversal

A continuación, observaremos que un subgrupo de estructuras comparativas con *que* muestra las mismas restricciones de movimiento que afectan a las estructuras coordinadas. En concreto, Ross (1967) afirma en su *Restricción de estructuras coordinadas* que estas estructuras deben obedecer dos condiciones:

(16) En una estructura coordinada:

(i) *Restricción de estructuras coordinadas 1*: no se puede mover ningún segmento coordinado,

(ii) *Restricción de estructuras coordinadas 2*: ni ningún elemento contenido en un segmento coordinado se puede mover fuera de ese segmento.

En lo que sigue, primero comprobaré la aplicabilidad de la segunda restricción en comparativas con *que* y *de* en español y luego pasaré a examinar la *Restricción de estructuras coordinadas 1* por primera vez en las comparativas en español, y

contrastaré sus propiedades con el comportamiento de las comparativas en latín y español medieval.

Probar la *Restricción de estructuras coordinadas 2* sobre comparativas con *de* y *que* ofrece algunos hallazgos reveladores que apoyan (i) un análisis dependiente de las codas comparativas con *de* y (ii) un análisis coordinado de las comparativas con *que*. Para ser más precisos, esta restricción bloquea la extracción asimétrica de un constituyente de uno de los segmentos de una coordinación, como se puede observar en el ejemplo (17a) de Sáez (1999, p. 1145). De forma paralela, la operación comparativa en el ejemplo (17b) muestra que esta misma restricción se aplica también a comparativas cuya coda está introducida por *que*:

- (17) Extracción asimétrica: produce agramaticalidad en oraciones coordinadas y comparativas con *que*.
- a. \*¿Dónde<sub>i</sub> compró Juan dos libros t<sub>i</sub> y Luis dos discos en Madrid?
- b. \*¿Dónde<sub>i</sub> compró Juan más libros t<sub>i</sub> que Luis discos en Madrid?

No obstante, es conocido que la *Restricción de estructuras coordinadas 2* puede evitarse si un constituyente se mueve de ambos segmentos simultáneamente, es decir, realizando una operación denominada *Movimiento transversal* en la terminología de Williams (1978; *Across-the-board movement* en inglés). Compare los ejemplos en (17) con los de (18)<sup>4</sup>:

- (18) Extracción simétrica o *Movimiento transversal*: posible en oraciones coordinadas y comparativas con *que*:
- a. ¿Dónde<sub>i</sub> compró Juan dos libros t<sub>i</sub> y Luis dos discos t<sub>i</sub>?
- b. ¿Dónde<sub>i</sub> compró Juan más libros t<sub>i</sub> que Luis discos t<sub>i</sub>?

El contraste entre la agramaticalidad de las oraciones coordinadas y comparativas con *que* en (17) y la gramaticalidad de la extracción simétrica en los ejemplos de (18) pone en evidencia que las mismas restricciones que se aplican a las estructuras coordinadas comunes operan también en al menos un subgrupo de comparativas con *que*. Como ilustro en los siguientes ejemplos, este contraste se mantiene si modificamos el tipo de base comparativa y analizamos comparativas adverbiales

---

4. Al analizar las comparativas en otras lenguas, como en inglés, Napoli (1983, p. 682), Hendriks (1991, p. 45), Corver (1993, p. 777) y Lechner (2004, p. 19), *inter alia*, también observaron un contraste entre extraer un constituyente de forma asimétrica o de manera simétrica en un subgrupo de comparativas en esta lengua. Estos autores concluyeron que al menos un tipo de comparativas en inglés tienen la estructura subyacente de las construcciones coordinadas.

con coda oracional como (19) o comparativas nominales con coda aparentemente no oracional como (20):

- (19) a. \*¿A quién<sub>i</sub> ha visitado Krzys a Isabel más que Sebastian t<sub>i</sub>?  
 b. ¿A quién<sub>i</sub> ha visitado Krzys t<sub>i</sub> más que Sebastian t<sub>i</sub>?  
 (20) a. \*¿[De qué músico]<sub>i</sub> ha escuchado Alex más ensayos de Arzalluz que conciertos t<sub>i</sub>?  
 b. ¿[De qué músico]<sub>i</sub> ha escuchado Alex más ensayos t<sub>i</sub> que conciertos t<sub>i</sub>?

Ante estos resultados, parece que existe un claro paralelismo entre el comportamiento de las estructuras coordinadas y las comparativas con *que* en español respecto a las limitaciones de extracción de constituyentes. Ambos tipos de construcciones prohíben las extracciones asimétricas mientras permiten el *Movimiento transversal*. Por el contrario, las comparativas con *de* en español no imponen una restricción a la extracción asimétrica. Esto se ilustra en el ejemplo (21), donde el pronombre interrogativo *dónde* se extrae asimétricamente de la oración matriz sin causar agramaticalidad (véase también Sáez 1992):

- (21) ¿Dónde<sub>i</sub> compró Juan aún más libros t<sub>i</sub> *de* cuantos Luis compró en Madrid?

### 3.1.2. La Restricción de estructuras coordinadas 2: español actual, español medieval y latín

Con respecto a la *Restricción de estructuras coordinadas 2*, esta condición establece que ningún segmento de una serie de segmentos coordinados puede ser dislocado sin causar agramaticalidad. Las oraciones de (22a-b) ilustran este contraste:

- (22) a. Lucía compró dos libros *y* Aitor compró tres CDs.  
 b. \**Y* Aitor compró tres CDs, Lucía compró dos libros.

Por el contrario, algunas oraciones subordinadas o dependientes se pueden mover, por ejemplo, a la posición inicial de la oración, sin que ello conlleve un cambio en la aceptabilidad de la oración (véase el ejemplo (23)). Debemos tener en cuenta que no es posible evaluar el estatus dependiente o coordinado de las cláusulas que no se pueden mover, ya que tanto las cláusulas coordinadas como algunas dependientes son inamovibles. Sin embargo, aquellas cláusulas que pueden ser dislocadas tienen un estatus inequívoco de dependencia con respecto a la cláusula matriz:

- (23) a. Javi compró pasteles *porque era su cumpleaños*.  
 b. *Porque era su cumpleaños*, Javi compró pasteles.

Como señala Hendriks (1991, p. 45) en relación a la lengua inglesa, las codas comparativas oracionales en inglés no pueden dislocarse al inicio de la oración. El ejemplo (24) pone en evidencia esta observación. Asimismo, en los ejemplos de (25a-b) muestro como la misma restricción parece estar operativa también en español, ya que, en la actualidad, no es posible dislocar las codas comparativas con *que* ni las que comienzan por *de* al inicio de la oración:

- (24) \**Than Mary bought records*, John bought more books. (Hendriks 1991, p. 45)  
 (25) a. \*Que Aitor CDs Lucía compró más libros.  
 b. \**De los que esperaba* Lucía compró más libros.

Es particularmente interesante para nuestra discusión puntualizar que, a pesar de que ni las codas comparativas con *de* ni con *que* pueden moverse al inicio de la oración en español moderno, Romero Cambrón (1998) señala que en el período medieval las codas comparativas sí que podían anteceder al grupo comparativo, como se muestra en los ejemplos (26)-(27) de los versos de Gonzalo de Berceo (en Romero Cambrón 1998, pp. 111 y 49, respectivamente):

- (26) *De quanto nos dezimos* el mundo mucho mejor era. (SDomingo)  
 (27) Don renegado malo, *de Judas* muy peor, no sé por ti qui quiera rogar al Criador.  
 (*Milagros*)

Es importante destacar que Romero Cambrón (1998) no encuentra ninguna coda comparativa con *que* antepuesta en los corpus de español antiguo. La autora sugiere que esta ausencia no es accidental, sino que la dislocación de las codas con *que* ha sido ilícita en todas las etapas de esta lengua y que, ni siquiera en verso (donde los poetas con frecuencia toman algunas licencias estilísticas), fue posible esta dislocación.

Además, continuando con una aproximación diacrónica a las comparativas objeto de estudio, Romero Cambrón (1998) también señala la relevancia del paralelismo entre el español y el latín con respecto a la dislocación de codas comparativas. Como puede observarse en los ejemplos en (28) y (29), los estándares de comparación en latín podían estar marcados con caso ablativo o introducidos por la partícula *quam* ‘que, cuánto (cuánto); tanto como’:

- (28) Latín: Comparativa con coda en ablativo:  
*Phoebo pulchrior*. (Petronio, *Satiricón*, II, 91)  
 (Lit.) «Que Febo más hermoso.»

- (29) Latín: Comparativas con coda introducida por *quam*:
- a. ... rem ausus plus famae habituram ad posteros *quam fidei*. (Liu. 2, 10,11)  
«un acto de atrevimiento más famoso *que creíble* en la posteridad»
  - b. Vtilius sit vobis audisse *quam nobis dixisse*. (Tac. *Hist.* 4, 73, 1)  
«puede ser más útil para ti escuchar *que para mí haber hablado*»

Al igual que en el caso del español medieval, estas dos clases de comparativas en latín también tenían patrones de linealización contrastivos. Como expone Romero Cambrón (1998, p. 71), en latín la coda en ablativo (es decir, la forma de la que habrían derivado las actuales comparativas con *de* como marcador del estándar en español) iba generalmente antepuesta al grupo comparativo. Sin embargo, las codas introducidas por *quam* (la forma de la que habría derivado el actual *que* comparativo en español<sup>5</sup>) debían ir obligatoriamente pospuestas.

El contraste en la linealización de las comparativas con coda en caso ablativo y las comparativas con *quam* en latín es paralelo a las diferencias en la linealización de las comparativas con *de* y *que* en español medieval. Si, como comentábamos al comienzo de la subsección, el estatus dependiente o coordinado de las cláusulas que no se pueden mover no es posible de evaluar, pero las codas que pueden ser dislocadas tienen un estatus inequívoco de dependencia, podemos concluir que las comparativas con coda en caso ablativo en latín y las comparativas con *de* en español medieval, al menos, tenían un estatus de dependencia o subordinación al grupo comparativo.

Aunque probar la *Restricción de estructuras coordinadas 2* sobre las comparativas en español en su uso actual no ofrece resultados concluyentes, la presencia de codas antepuestas introducidas por *de* en español medieval ((26)-(27)) y con caso ablativo en latín ((28); de donde derivarían las comparativas con *de* actuales, según Romero Cambrón 1998) apunta hacia un estatus dependiente de las codas en estas comparativas. Los resultados del análisis en la Sección 3.1.1. también apoyan esta

5. Para ser más exactos, en el diccionario etimológico de Corominas (1985, pp. 704-705) aparece recogida la siguiente explicación:

QUE (...); en su función de conjunción comparativa procede de una confusión del que relativo y encabezador con la conjunción latina QUAM. 1.<sup>a</sup> doc.: orígenes del idioma (Glosas Emilianenses y Silenses). (...) En lo referente a la conjunción comparativa que (él es mejor que tú, saldré antes que llegue), el punto de partida es el lat. QUAM, convertido primero en ca y luego confundido con la conjunción anterior.

observación: las comparativas con *de* en español no presentan las restricciones a la extracción propias de las estructuras coordinadas y parecen tener, por lo tanto, un estatus dependiente.

Por el contrario, existe un claro paralelismo entre el comportamiento de las estructuras coordinadas y el de algunas comparativas con *que* en español actual respecto a: (i) las limitaciones de extracción de constituyentes analizadas en la Sección 3.1.1. y (ii) la restricción a la hora de poder anteponer la coda con *que*, limitación propia de las estructuras coordinadas comunes. A la vista de estos resultados, podemos concluir que al menos un subgrupo de comparativas con *que* en español tiene las propiedades características de las estructuras coordinadas<sup>6</sup>.

### 3.2. Comparativas con *que*: propiedades de tipo coordinativo y estándar sintagmático

Las conjunciones coordinadas pueden introducir diferentes categorías sintácticas tanto a nivel oracional como a nivel sintagmático, no oracional. Por este motivo, si la propuesta defendida en la anterior sección es correcta y el marcador del estándar *que* se comporta como una conjunción coordinante en algunas comparativas en español, la predicción que se deriva de nuestra hipótesis es que el marcador coordinante *que* debería poder seleccionar tanto complementos oracionales como sintagmáticos en las comparativas<sup>7</sup>. Pasemos a examinar esta predicción.

Las conjunciones coordinantes generalmente unen dos sintagmas del mismo tipo o con cierto grado de paralelismo o identidad entre ellos. Basándose en el paralelismo entre la base de la comparación y el estándar de comparación, Sáez (1992, 1999) argumenta que el siguiente tipo de comparativas en español, denominadas

---

6. En la Sección 3.4. veremos que la coda de otro subgrupo de comparativas con *que* en español actual presenta un comportamiento claramente dependiente. Caracterizaremos este subgrupo para poder diferenciarlo fácilmente de las comparativas con *que* y coda coordinada.

7. Napoli (1983) también defendió un análisis de tipo coordinación para las comparativas nominales, adjetivales y adverbiales del tipo ilustrado en (i) en lengua inglesa en el que propuso que, en estos casos, el marcador del estándar con función coordinante podía seleccionar tanto oraciones como complementos sintagmáticos:

- (i) a. The team made *more noise than headway*.
- b. Mary is *more clever than smart*.
- c. Mary sings *more sweetly than beautifully*.

Napoli (1983) sostiene que, en casos como los de (i), el marcador del estándar en inglés *than* se comporta como una conjunción coordinante que une el segmento que lo sigue (el estándar) con un ítem del mismo tipo que lo precede (el grupo comparativo).



subcomparativas<sup>8</sup>, tienen un estándar no oracional que debe ser obligatoriamente introducido por *que* e involucran una estructura coordinada subyacente:

- (30) *Más mujeres que hombres* vinieron a la reunión.  
 (31) Conozco *más partidarios que detractores* de Mao. (Sáez 1999, p. 1148 (108))

Sáez (1992, 1999) observa que las subcomparativas con base nominal como (31) se comportan de la misma manera que las estructuras coordinadas como (32), ya que admiten la presencia de complementos preposicionales compartidos (es decir, *de Mao*, en estos ejemplos). Tanto en la coordinación sintagmática en (32), como en la comparativa con *que* en (31), el Sintagma Preposicional *de Mao* funciona como complemento de dos sustantivos al mismo tiempo: *partidarios* y *detractores*.

- (32) Conozco partidarios y detractores *de Mao*. (Sáez 1999, p. 1148 (106))

En contraste, Sáez (1999) observa que, si los dos sustantivos que seleccionan un complemento no están en una relación de coordinación, como se ejemplifica en (33), el complemento de *detractores* no puede comportarse a su vez como complemento de *partidarios* y, como consecuencia, decae la aceptabilidad de la oración:

- (33) \*Conozco partidarios considerados como detractores *de Mao*. (Sáez 1999, p. 1148 (107))

Para mostrar el paradigma completo, en el siguiente conjunto de datos muestro que las comparativas con *de* en la coda, a diferencia de las comparativas con *que* anteriormente descritas, no mantienen el mismo patrón que las estructuras coordinadas en el sentido de que no permiten complementos compartidos:

- (34) \*Más partidarios de los que Antonio esperaba *de Mao* fueron a la reunión.

La presencia de complementos preposicionales compartidos es una propiedad característica de las estructuras coordinadas sintagmáticas y también de las

---

8. El término *subcomparativa* fue introducido por Bresnan (1973, 1975) para referirse a oraciones que comparan cantidades o grados de diferentes tipos de cosas y en las que se elimina o silencia un Sintagma de Medida dentro del estándar de comparación. En las subcomparativas en (i), el guion bajo (\_) indica la posición en la que aparentemente falta un elemento de medida (véase Bresnan 1973, 1975; Napoli 1983; Corver 1993, 2006; Kennedy 2002, *inter alia*):

- (i) a. They have more enemies than we have \_ friends.  
 b. More girls than \_ boys suffer from anxiety disorders.  
 c. This box is wider than it is \_ tall.  
 (ii) Ana se sintió más avergonzada que \_ orgullosa de su hermano. (Brucart 2003; Sáez 1992, 1999)

subcomparativas con *que* y estándar sintagmático en español. En ambos contextos, el nexos de unión relaciona dos sintagmas paralelos (los dos segmentos comparados) y la interpretación que obtenemos es una en la que el mismo complemento está modificando ambos segmentos, es decir, los dos elementos comparados. De esta forma, el comportamiento idéntico entre las comparativas con *que* analizadas y la coordinación sintagmática con respecto a los complementos compartidos pone de manifiesto que este tipo de comparativas en español tiene una estructura coordinada subyacente en la que se unen dos elementos no oracionales.

Por lo tanto, como adelantábamos en nuestra predicción al inicio de la sección, el marcador del estándar *que* se comporta como una conjunción coordinante en algunas comparativas en español y puede seleccionar tanto complementos oracionales (véanse los ejemplos (14a-b), (18b), (13), (19b) o (20b) analizados en la Sección 3.1.) como sintagmáticos (véanse los ejemplos (30)-(31)).

### 3.2.1. *Comparativas con coda coordinada con quam en latín*

Pasamos a realizar ahora una aproximación diacrónica novedosa, ya que observaremos las propiedades de las comparativas con *quam* en latín para apoyar nuestra propuesta de que un subgrupo de comparativas con *que* en español tiene coda coordinada, y que esta conjunción coordinante puede seleccionar tanto complementos oracionales como sintagmáticos.

En el caso de la arquitectura de las comparativas con *quam* en latín, la investigación llevada a cabo por Suárez Martínez (2002) defiende que, en estas expresiones de desigualdad, *quam* tenía una función coordinadora:

Siendo esto así, no puedo, sin embargo, estar de acuerdo con Fontana en que «en estos casos el comportamiento sintáctico de *quam* evidencia que este elemento no es adscribirle a ninguna de las categorías tradicionales» (1997, p. 73); es cierto que no es conjunción subordinante porque no introduce una oración, ni relativo, ni preposición, como sospecha Fontana, pero no puede no ser algo, es decir, debe poder ser «adscribible» a alguna categoría, que, en mi opinión, es la de las conjunciones coordinantes, y eso, tanto cuando introduce oración como, con mayor razón y claridad, cuando no la introduce (Suárez Martínez 2002, p. 236).

Además, en cuanto al tipo de complemento que selecciona *quam* (la forma de la que habría derivado el actual *que* comparativo en español), según hacen notar Baños (2002) en su obra sobre este tipo de comparativas en latín y Suárez Martínez (2002), las construcciones con *quam* podían introducir una oración o seleccionar

un complemento sintagmático no oracional. Por lo tanto, podemos confirmar que tanto las comparativas con *quam* en latín como algunas comparativas con *que* en español presentan una coda coordinada y, en esos casos, el marcador del estándar con función coordinante (*quam* en latín y *que* en español actual) puede seleccionar complementos oracionales o sintagmáticos.

### 3.2. Comparativas con *de*: propiedades de tipo dependiente y estándar sintagmático

Como hemos visto en la Sección 2.2 basándonos en gramáticas y análisis teóricos sobre la comparación en español y en las anteriores subsecciones 3.1 y 3.2, en cuanto a su estructura subyacente, las comparativas con *de* (ejemplificadas en (35)-(36)) presentan un caso claro de estructuras con coda dependiente y estándar sintagmático, no oracional.

- (35) a. El libro tiene más de [*cien páginas*].  
 b. Antonio comió más de [*un jabalí*]. (Quizá tres o cuatro jabalís.)
- (36) Amelia ha conseguido más libros de [[*cuantos Santi necesitaba la semana pasada*] / [*los que Santi necesitaba la semana pasada*]].

El hecho de que en estas comparativas se tome la preposición *de* como marcador del estándar, en vez de otro tipo de núcleo sintagmático, pone en evidencia que, en estos casos, la coda no se encuentra coordinada, sino que tiene un estatus dependiente propio de los elementos introducidos por una preposición en español. Además, las comparativas con *de* tampoco muestran un paralelismo entre los elementos que aparecen flanqueando al marcador del estándar, como sucede en las comparativas con *que* con propiedades de tipo coordinación, como las de los ejemplos (14b), (18), (30) o (31).

Por otro lado, en cuanto a sus restricciones semánticas, el marcador del estándar *de* siempre se combina con un nominal que denota un grado: ya sea un Sintagma de Medida como en los ejemplos en (35), una relativa libre de cantidad o grado (como *cuantos Santi necesitaba la semana pasada* en el ejemplo (36)) o un Sintagma Determinante modificado por una cláusula de relativo que le confiere un significado cuantitativo (como *los que Santi necesitaba la semana pasada* en el ejemplo (36), que hace referencia a la *cantidad* de libros que Santi necesitaba (véanse Gutiérrez Ordóñez 1994; Romero Cambrón 1997, 1998; Sáez 1999; Brucart 2003, 2009; RAE 2010; Vela-Plo 2018; o Mencia 2020).

Debido a que ya hemos examinado las propiedades de dependencia de las comparativas cuya coda está introducida por *de* en español en las secciones 3.1. y 3.2, y que existe un mayor consenso entre autores a la hora de caracterizar estas construcciones, pasaremos a analizar otro subgrupo de comparativas cuya naturaleza y categorización es más polémica en la siguiente subsección.

### 3.4. *Un caso especial: comparativas con que y un estándar dependiente oracional*

Con el objetivo de finalizar el análisis de los principales tipos de comparativas que encontramos en español según su estructura subyacente, voy a recordar la hipótesis que nos servía de guía al comenzar este análisis. Según nuestra hipótesis de partida de (11), debemos tener en cuenta dos variables morfosintácticas independientes a la hora de definir y categorizar las estructuras comparativas en español: (i) si la coda es sintagmática u oracional y (ii) si la coda está coordinada o subordinada al grupo comparativo. Siguiendo esta hipótesis, en las anteriores subsecciones hemos demostrado que en español tenemos: A) comparativas con coda coordinada oracional, analizadas en la Sección 3.1; B) comparativas con coda coordinada sintagmática, examinadas en la Sección 3.2; y C) comparativas con coda dependiente sintagmática, examinadas en la Sección 3.3. En esta subsección discutiré las propiedades fundamentales del último tipo de comparativas que encontramos en español: D) las comparativas con una coda dependiente oracional.

En las secciones anteriores hemos observado que siempre que una comparativa muestra propiedades típicas de las estructuras coordinadas, en español se emplea el marcador *que* para introducir la coda comparativa (ver secciones 3.1. y 3.2.). El ejemplo de (37) ilustra una comparación que evidencia el resultado de una operación de *Vaciado en la coda* (es decir, una operación de elipsis propia de las oraciones coordinadas) en la que el estándar debe ir obligatoriamente introducido por *que*:

(37) En un mes, Maitane compró más aguacates *que Oihana* \_ libros.

Por el contrario, como hemos visto anteriormente, las comparativas con *de* en la coda presentan los rasgos característicos de los constituyentes dependientes.

Este relevante contraste entre comparativas con *que* y *de* podría llevarnos a la conclusión precipitada de que todas las comparativas con *que* implican una estructura subyacente coordinada y que solo las comparativas con *de* presentan un estándar de comparación dependiente. Sin embargo, a continuación examino las propiedades de un tipo particular de construcción comparativa con *que* en español y los resultados de este

análisis: (i) nos llevarán a rechazar la conclusión apresurada que acabo de describir y (ii) respaldarán la propuesta de que el tamaño del estándar (sintagmático u oracional) y el tipo de nexo de unión entre coda y correlato (coordinado o dependiente) constituyen dos variables sintácticas independientes. Veremos así que son estas dos variables las que determinan la estructura subyacente de la comparativa, así como la elección entre el marcador del estándar *que* o *de* en español.

Como señalan Sáez (1992) y Vela-Plo (2018), existe un tipo específico de comparativa con *que* en español que no manifiesta las propiedades características de las estructuras coordinadas. Más concretamente, se trata de comparativas con *que* con un estándar oracional no reducido por operaciones de elipsis, como en los ejemplos (38)-(39):

(38) Esta mesa es más larga *que* [*ancha es esa puerta*].

(39) En un mes, Maitane compró más aguacates *que* [*libros*, *me dijo Alejandro* [*que leyó Oihana t<sub>i</sub>*]].

Podemos observar que la coda de estas comparativas se comporta claramente como un constituyente dependiente, en lugar de manifestar propiedades de coordinación, si examinamos sus características sintácticas. Una de estas características se manifiesta por el orden de las palabras de los constituyentes del estándar oracional. En este particular tipo de comparativas, en lugar del orden de palabras básico del español, es decir, S(ujeto)-V(erbo)-O(bjeto o complementos), el predicado nominal o graduable comparado dentro de la coda oracional aparece necesariamente desplazado hacia la periferia izquierda de esa cláusula (véanse (38) y el caso de (39) con movimiento del elemento comparado a larga distancia)<sup>9</sup>.

Además, como ilustran los ejemplos (38) y (39), el estándar oracional de estas comparativas manifiesta una inversión obligatoria entre el sujeto y el verbo conjugado. El alejamiento del orden SVO básico del español y la inversión SV en el estándar clausal de este tipo de comparativas ofrecen como resultado el orden lineal característico de las construcciones que implican movimiento de algún operador hacia la periferia izquierda de una cláusula en español (*cfr.* Torrego 1984; Uribe-Etxebarria 1992; Suñer 1994; Barbosa 2001; Mendia 2017, entre otros). Entre los ejemplos de construcciones con inversión obligatoria de sujeto-verbo se incluyen las construcciones con un sintagma focalizado o la formación de preguntas con

9. En los ejemplos (38)-(39) el predicado nominal o graduable comparado en el estándar aparece subrayado.

movimiento-*qu* en español. Ilustro estos dos tipos de ejemplos que exigen una inversión SV obligatoria de la misma manera que algunas comparativas con *que* en los pares mínimos en (40)-(41):

- (40) a. ¿Cuántos libros leyó Oihana?  
 b. \*¿Cuántos libros Oihana leyó?
- (41) a. No recordaba [cuántos libros me dijo Alejandro que leyó Oihana].  
 b. \*No recordaba [cuántos libros Alejandro me dijo que Oihana leyó].

Otro rasgo distintivo que encontramos en este tipo de comparativas con coda oracional con inversión sujeto-verbo en español es que estas construcciones no están sujetas a las restricciones características que operan sobre las estructuras coordinadas. Por ejemplo, la *Restricción de estructuras coordinadas* discutida en la Sección 3.1.2 no se les aplica, como muestro en el ejemplo (42):

- (42) ¿[A quién]<sub>i</sub> compró Janire más manzanas t<sub>i</sub> que peras vendió Ivan a Sarai?

Aunque las comparativas con coda oracional con inversión sujeto-verbo en español hacen uso del marcador del estándar *que* (el cual, en otras circunstancias, aparece en comparativas con estructura coordinada donde la *Restricción de estructuras coordinadas* está operativa, *cfr.* Sección 3.1 y 3.2), el estándar en estas comparativas se comporta como una cláusula dependiente con movimiento-*qu* en español (véanse Sáez 1992 y Vela-Plo 2018 para una discusión completa sobre este punto y las propiedades de tipo dependiente de estas comparativas con inversión SV en el estándar).

Estas y algunas otras observaciones llevaron a Sáez (1992) y Vela-Plo (2018) a concluir que las comparativas con coda oracional con inversión SV en español implican movimiento de un operador cuantitativo dentro del estándar paralelo al que exhiben las oraciones de relativo libre. En la Sección 4 tendremos en cuenta esta propuesta para la caracterización similar de comparativas con *de*, y comparativas con *que* y coda oracional con inversión SV en español.

Teniendo en cuenta la hipótesis de partida (11) que estamos comprobando en este trabajo, es importante tener en cuenta que las comparativas con *que* y coda oracional con inversión SV en español no muestran propiedades dependientes y coordinadas mixtas o híbridas, sino únicamente características propias de las estructuras dependientes. Por ejemplo, en cuanto a las operaciones de elipsis que se les puede aplicar. Como muestra el ejemplo en (43), las comparativas con *que*

y coda oracional con inversión no permiten aplicar operaciones de elipsis de segmentos propias de estructuras coordinadas como Elevación al nodo derecho, mientras que las comparativas con *que* sin inversión sí muestran este tipo de elipsis (es decir, las analizadas en la Sección 3.1 como el ejemplo de (14)b).

(43) \*Mari envió más juguetes que ropa compró Zuriñe, a Zeian.

Por tanto, no es cierto que la existencia de este segundo tipo de comparativas con *que* favorezca un enfoque híbrido de las comparativas. Más bien, las comparativas con *que* y coda con inversión SV pueden clasificarse como comparativas con coda oracional dependiente siguiendo nuestra propuesta de análisis de las comparativas en español (véase (12)). Por el contrario, las comparativas con *que* sin un estándar con inversión SV deben clasificarse como comparativas con estructura coordinada y estándar oracional o sintagmático (como hemos mostrado en las secciones 3.1. y 3.2). A su vez, esto significa que el marcador del estándar *que* en español es homófono entre un *que* cuya función es la de una conjunción subordinante en las comparativas con estándar oracional e inversión SV y un *que* cuya función es la de una conjunción coordinante en las comparativas con *que* restantes.

Aunque autores como Sáez (1992), Vela-Plo (2018) y también Reglero (2004) habían examinado las propiedades de las comparativas con *que* y estándar oracional con inversión, en este trabajo, por primera vez, hemos tomado una perspectiva más amplia y hemos tenido en cuenta otros tipos de comparativas que hay en español según su estructura subyacente y su significado para motivar una nueva clasificación de las comparativas con *de* y *que* en español, es decir, la propuesta presentada al comienzo de la Sección 3 en (11) y (12) y que detallo en el siguiente apartado.

#### 4. NUEVA CATEGORIZACIÓN Y DEFINICIÓN DE LAS RESTRICCIONES SINTÁCTICAS Y SEMÁNTICAS DE LAS COMPARATIVAS DE DESIGUALDAD CON *QUE* Y *DE* EN ESPAÑOL

Aunque las construcciones comparativas prototípicas en español han recibido una atención notable en los últimos años, la diferencia entre las comparativas con *que* y *de* aún no había recibido un análisis y caracterización adecuados (ver Sáez y Sánchez López 2013 para un resumen reciente).

Basándonos en las conclusiones sobre esta alternancia alcanzadas en la Sección 3, ahora podemos caracterizar la alternancia *que/de* en la elección del marcador del estándar en español como una interacción entre dos variables independientes. Concretamente, hemos observado que la arquitectura de las comparativas de desigualdad en español está determinada por: (i) el tamaño del estándar de comparación (sintagmático u oracional) y (ii) el tipo de vínculo entre la coda y el correlato comparativo (una estructura de coordinación o de dependencia). En la Tabla 1 resumimos la propuesta de caracterización de las comparativas con *que* y *de* defendida en este trabajo:

		TIPO DE NEXO	
		COORDINACIÓN	DEPENDENCIA
TAMAÑO DEL ESTÁNDAR	SINTAGMÁTICO	[1] <i>que</i> <sub>&amp;</sub> <ul style="list-style-type: none"> <li>• <b>Categoría gramatical:</b> Conjunción coordinante.</li> <li>• <b>Restricción semántica:</b> selecciona dos sintagmas del mismo tipo semántico.</li> </ul>	[3] <i>de</i> <sub>dep</sub> <ul style="list-style-type: none"> <li>• <b>Categoría gramatical:</b> Preposición.</li> <li>• <b>Restricción sintáctica:</b> selecciona un estándar sintagmático (o un Sintagma de Medida o un Sintagma Determinante modificado por una relativa de cantidad o grado).</li> <li>• <b>Restricción semántica:</b> selecciona un predicado graduable/cuantificable.</li> </ul>
		(44) Había más partidarios <i>que</i> <sub>&amp;</sub> <i>detractores de Mao</i> en aquella sala.	(45) El libro tiene más <i>de</i> <sub>dep</sub> <i>cien páginas</i> . (46) La película era más larga <i>de</i> <sub>dep</sub> <i>lo que esperaba</i> .
	ORACIONAL	[2] <i>que</i> <sub>&amp;</sub> <ul style="list-style-type: none"> <li>• <b>Categoría gramatical:</b> Conjunción coordinante.</li> <li>• <b>Restricción semántica:</b> selecciona dos sintagmas del mismo tipo semántico (dos proposiciones).</li> </ul>	[4] <i>que</i> <sub>dep</sub> <ul style="list-style-type: none"> <li>• <b>Categoría gramatical:</b> Conjunción subordinante.</li> <li>• <b>Restricción sintáctica:</b> selecciona un estándar oracional.</li> <li>• <b>Restricción semántica:</b> selecciona un predicado graduable/cuantificable.</li> </ul>
		(47) Maitane compró más aguacates <i>que Oihana</i> _ libros.	(48) Maitane compró más aguacates <i>que</i> <sub>dep</sub> libros me dijo Alejandro <i>que</i> leyó Oihana.

Tabla 1: Categorización y propiedades sintácticas y semánticas de las comparativas de desigualdad con *que* y *de* en español



Teniendo en cuenta la variación intralingüística en las comparativas en español, obtenemos cuatro tipos de estructuras básicas, cada una ejemplificada con una oración comparativa prototípica en la Tabla 1: [1] comparativas con coda coordinada sintagmática; [2] comparativas con coda coordinada oracional; [3] comparativas con coda dependiente sintagmática; y [4] comparativas con coda dependiente oracional. A continuación, ofrezco un ejemplo de cada tipo comparativo con una representación esquemática de la estructura subyacente del estándar comparativo:

- (49) Comparativas con coda dependiente sintagmática:  
 a. La película era más larga **de** *lo que esperaba*.  
 b. La película era *más* larga [*de*<sub>dep</sub> lo [<sub>SComp</sub> Op<sub>i</sub> que [<sub>REL</sub>]<sup>10</sup> esperaba ~~que la película fuera t<sub>i</sub>~~]]
- (50) Comparativa con coda dependiente oracional:  
 a. Esta mesa es más larga **que** *ancha es esa puerta*.  
 b. Esta mesa es *más* larga [*que*<sub>dep</sub> [<sub>SComp</sub> Op ancha<sub>i</sub> [<sub>QU</sub>] es esa puerta t<sub>i</sub>]].
- (51) Comparativas con coda coordinada sintagmática<sup>11</sup>:  
 a. más partidarios **que** *detractores* ...  
 b. [[<sub>SMedida</sub> más partidarios] que<sub>&</sub> [<sub>SMedida</sub> detractores]] ...
- (52) Comparativa con coda coordinada oracional:  
 a. A Marina le gustan más los bizcochos **que** *a Maider \_ las pizzas*.  
 b. [[<sub>ST</sub> A Marina le gustan más los bizcochos] que<sub>&</sub> [<sub>ST</sub> a Maider \_ las pizzas]]

¿En qué se diferencia la propuesta aquí presentada de análisis previos sobre las comparativas en español? La idea de que la alternancia *que/de* española es el resultado de una interacción entre dos variables independientes nos recuerda a la propuesta de Mendi (2020) sobre esta alternancia. Para este autor, la selección de *que* o *de* como marcador del estándar depende de dos factores: (i) un criterio

10. Para Rizzi (1997), las palabras interrogativas (palabras-*qu*) y el foco son elementos cuantificativos, a diferencia de los operadores relativos, por ejemplo, *que* se mueven y desencadenan una estructura con abstracción en términos de Heim y Kratzer (1998). Según este autor, las relaciones A' deben dividirse en aquellas que involucran una cuantificación genuina (con un cuantificador ascendente que vincula una variable, como se representa en la oración (49), p. ej.) y aquellas que involucran una vinculación A' no cuantificacional, como en el caso de los operadores relativos, que se mueven y desencadenan una abstracción al estilo de Heim y Kratzer (1998) sin cuantificación (véase la oración (50)). Por esta razón, en las comparativas oracionales con inversión SV marco el predicado graduable trasladado al inicio del estándar oracional con el subíndice [QU], mientras que marco los operadores relativos con el subíndice [REL], como en el ejemplo en (49) que involucra una cláusula relativa en el estándar de comparación.

11. Podemos encontrar una propuesta de análisis sintáctico y semántico completamente composicional de las comparativas con estructura coordinada en Vela-Plo (2020, en prensa).

sintáctico sobre el tamaño del estándar y (ii) un criterio semántico que determina si el estándar denota directamente un grado o no:

- (53) Propuesta de análisis de la alternancia *que/de* en español en Mendia (2020):
- |                                      |   |
|--------------------------------------|---|
| a. Marcador del estándar <i>de</i> : | b. Marcador del estándar <i>que</i> :   |
| Categoría gramatical: Preposición    | Categoría gramatical: Complementante    |
| Selecciona: Sintagma de Medida       | Selecciona: Oración (Sintagma Temporal) |

La mayoría de los análisis de las comparativas en español concluyen que el marcador del estándar *de* siempre se combina con un nominal que denota un grado (ya sea un Sintagma de Medida o una cláusula relativa que hace referencia a una cantidad o grado). Por lo tanto, la descripción en Mendia (2020) y mi propuesta (esbozada en la Tabla 1) de que el marcador del estándar *de* toma un argumento sintagmático que denota un grado o un predicado cuantificable no suponen una gran sorpresa.

No obstante, las restricciones de selección sintácticas y/o semánticas de *que* son menos claras en los trabajos previos sobre comparativas en español (véase la Sección 2). La propuesta en Mendia (2020) y la defendida en este trabajo sobre la alternancia *que/de* difieren principalmente en la caracterización del marcador del estándar *que*. Mendia (2020) asume que este marcador se comporta siempre como un complementante que selecciona una oración subordinada. Por lo tanto, su propuesta no puede explicar el comportamiento de tipo coordinativo, ni la naturaleza sintagmática (no oracional) del estándar en un subconjunto de comparativas con *que*, como he mostrado en las secciones 3.1 y 3.2, respectivamente. En contraste, mi propuesta presentada en la Tabla 1 aclara, por fin, las restricciones de selección sintáctica y semántica de los marcadores del estándar *que* y *de*, y da cuenta de las propiedades diferenciales de los tipos de estándares que selecciona cada marcador.

## 5. CONCLUSIÓN

Este trabajo ha examinado en profundidad las características morfosintácticas de las comparativas en español actual teniendo en cuenta datos sincrónicos y algunos aspectos diacrónicos sobre las comparativas con *de* y *que* en la coda comparativa, y ha defendido una propuesta de recategorización y reforma del análisis de las comparativas en español.

Tras sintetizar algunas de las principales preguntas teóricas que todavía no habían sido resueltas sobre la clasificación y la arquitectura de las comparativas de

desigualdad con *que* y *de* en español, hemos ofrecido una respuesta clara a cada una de estas polémicas cuestiones. Para ello, además de un análisis sincrónico de las comparativas con *que* y *de* en español, hemos contrastado sus características con las particularidades de las estructuras comparativas en latín y español medieval de las que derivan las construcciones actuales. Este análisis nos ha revelado que es necesaria una nueva forma de categorizar las comparativas en castellano y también hemos podido aclarar las discrepancias y preguntas planteadas en relación a la naturaleza y estructura sintáctica de estas expresiones y sobre la selección del marcador del estándar *que* o *de*.

En primer lugar, nos planteábamos la cuestión del tipo denexo entre los dos objetos comparados: ¿se trata de una relación de coordinación o de dependencia (o subordinación)? Hemos visto que las pruebas sintácticas ofrecidas abogan claramente por una primera división de las comparativas españolas en dos tipos de construcciones según la relación entre la coda y el correlato comparativo: *comparativas con coda coordinada* o *comparativas con coda dependiente*. Además, en relación a las comparativas con *que* en particular, hemos mostrado que el marcador del estándar *que* en español es homófono entre un *que* cuya función es la de una conjunción subordinante en las comparativas con estándar oracional e inversión SV y un *que* cuya función es la de una conjunción coordinante en las comparativas con *que* restantes.

En segundo lugar, la división entre dos tipos de comparativas según elnexo de unión entre coda y correlato defendida en este trabajo aporta una ventaja relevante en relación a otra de las cuestiones teóricas examinadas en este estudio. Más concretamente, hemos respondido a la pregunta sobre cuál es la naturaleza del proceso denominado *Elipsis Comparativa* responsable de silenciar uno o varios constituyentes de la coda de las construcciones que presentan una oración en su estándar. Con la propuesta de división entre comparativas con coda coordinada y con coda dependiente podemos olvidarnos del proceso *ad hoc* de borrado específico para comparativas denominado *Elipsis Comparativa*, ya que hemos visto que los procesos de borrado que se producen en las comparaciones con coda coordinada son los mismos que se aplican a las estructuras coordinadas de forma más general, mientras que las operaciones de borrado que se observan en las comparativas con coda dependiente son las mismas que operan sobre otros tipos de constituyentes dependientes o subordinados. De esta manera, podemos reducir el conjunto de operaciones de elipsis o borrado que se pueden aplicar sobre cierta cadena lingüística siguiendo un enfoque minimalista.

En tercer lugar, hemos respondido a la cuestión del tamaño del estándar de comparación, ya que hemos analizado datos sincrónicos y diacrónicos que muestran que existe otra variable independiente que determina la estructura sintáctica y la elección del marcador del estándar en las comparativas en español: la presencia de una *coda oracional* (con elementos silenciados o no) o la presencia de una *coda sintagmática*, que no se deriva de una oración. Así, la principal contribución del presente trabajo es la nueva propuesta de caracterización de la arquitectura de las comparativas de desigualdad en español como determinada por: (i) el tamaño del estándar de comparación (sintagmático u oracional) y (ii) el tipo de vínculo entre la coda y el correlato comparativo (una estructura de coordinación o de dependencia). Teniendo en cuenta estos importantes resultados, finalmente, se ha logrado definir las restricciones de selección sintáctica y semántica de los marcadores del estándar *que* y *de* atendiendo a las propiedades diferenciales de los tipos de estándares que selecciona cada marcador.

Departamento de Filología e Historia.

Oficina 3.7, Hizkuntzalaritza Teorikorako Taldea, Micaela Portilla  
(Vitoria-Gasteiz)

## REFERENCIAS BIBLIOGRÁFICAS

- Aarcos Llorach, E. (1994). *Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- Alrenga, P., Kennedy, C. y Merchant, J. (2012). A new standard of comparison, en N. Arnett y R. Bennett (Eds.), *Proceedings of the 30<sup>th</sup> West Coast Conference on Formal Linguistics* 30, pp. 32-42, Somerville, MA: Cascadilla Proceedings Project.
- Bácskai-Atkári, J. (2014). *The Syntax of Comparative Constructions*. Potsdam: University of Potsdam.
- Baños, J. M. (2002). Comparativas con *quam* y verbo personal e latín, en Espinilla, E., Quetglas, P. J. y M. E. Torrego (Eds.), *La comparación en latín*, Barcelona, UAM/UB, pp. 39-62.
- Barbosa, P. (2001). On inversion in wh-questions in Romance, en A. Belletti, A. Hulk y J.-Y. Pollock (Eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*, pp. 20-59, Oxford: Oxford University Press.
- Beck, S., Oda, T. y Sugisaki, K. (2004). Parametric variation in the semantics of comparison: Japanese vs. English, *Journal of East Asian Linguistics* 13, pp. 289-344.
- Bhatt, R. y Takahashi, S. (2011). Reduced and unreduced phrasal comparatives, *Natural Language and Linguistic Theory* 29(3), pp. 581-620. <https://doi.org/10.1007/s11049-011-9137-1>
- Bochnak, M. R. (2013). *Cross-linguistic variation in the semantics of comparatives*. Chicago: University of Chicago, tesis doctoral.

- Bolinger, D. L. (1950). The comparison of inequality in Spanish, *Language* 26(1), pp. 28–62. <https://doi.org/10.2307/410407>
- Bolinger, D. L. (1953). Addenda to the comparison of inequality in Spanish, *Language* 29(1), pp. 62–66. <https://doi.org/10.2307/410455>
- Bresnan, J. W. (1973). Syntax of the comparative clause construction in English, *Linguistic Inquiry* 4(3), pp. 275–343.
- Brucart, J. M. (2003). Adición, sustracción y comparación: un análisis composicional de las construcciones aditivo-sustractivas del español, en F. Sánchez-Miret (Ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*, 1(1), pp. 11–60, Verlag, Tübinga: Max Niemeyer. <https://doi.org/10.1515/9783110946277-004>
- Brucart, J. M. (2009). Patrones formales e interpretación: el funcionamiento de más en español, *ELUA: Estudios de Lingüística. Universidad de Alicante* 3, pp. 13–44.
- Bylinina, L. y Lyer, Y. (2013). ‘Than’=‘More’+ Exhaustivity: Evidence from Circassian, en M. Aloni, M. Franke y F. Roelofsen (Eds.), *Proceedings of the 19<sup>th</sup> Amsterdam Colloquium*, pp. 75–82. <https://publications.hse.ru/pubs/share/folder/x62hgh3w8l/106224297.pdf>
- Chomsky, N. y Lasnik, H. (1977). Filters and control, *Linguistic Inquiry* 8, pp. 425–504.
- Chomsky, N. (1977). On Wh-Movement, en P. W. Culicover, T. Wasow y A. Akmajian (Eds.), *Formal Syntax*, pp. 71–132, Nueva York: Academic Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Corominas, J. y Pascual, J. A. (1991–1997). *Diccionario crítico etimológico castellano e hispánico*. Obra completa. Madrid: Editorial Gredos.
- Corver, N. (1990). *The Syntax of Left Branch Extractions*. Tilburg: KUB Tilburg University, tesis doctoral.
- Corver, N. (1993). A note on subcomparatives, *Linguistic Inquiry* 24(4), pp. 773–781.
- Emonds, J. (1985). *A unified theory of syntactic categories*. Dordrecht: Foris.
- Fontana Elboj, G. (1997). *Las construcciones comparativas latinas: aspectos sincrónicos y diacrónicos*. Zaragoza: Universidad de Zaragoza.
- Gallego, Á. (2013). Preposiciones de trayectoria y estructuras comparativas, en L. Sáez y C. Sánchez López (Eds.), *Las construcciones comparativas*, pp. 225–256, Madrid: Visor.
- Gutiérrez Ordóñez, S. (1994). *Estructuras comparativas*. Madrid: Arco Libros.
- Hankamer, J. (1973). Why there are two *than*'s in English, en C. Corum, y T. C. Smith-Stark (Eds.), *Papers from the 9<sup>th</sup> Regional Meeting of the Chicago Linguistic Society*, pp. 179–191, Chicago, Illinois: Chicago Linguistic Society.
- Heim, I. y Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Heim, I. (1985). *Notes on comparatives y related matters*. Austin: University of Texas, Austin, manuscrito.
- Hendriks, P. (1991). The coordination-like structure of comparatives, *Linguistics in the Netherlys* 8(1), pp. 41–50. <https://doi.org/10.1075/avt.8.07hen>
- Huang, P. K. (1977). *Wh-fronting and related processes*. Connecticut: University of Connecticut, tesis doctoral.

- Jäger, A. (2019). The syntax of comparison constructions in diachronic y dialectal perspective, *Glossa: a journal of general linguistics* 4(1): 70, pp. 1-51. <https://doi.org/10.5334/gjgl.651>
- Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. Nueva York: Garly.
- Kennedy, C. (2002). Comparative deletion and optimality in syntax, *Natural Language and Linguistic Theory* 20(3), pp. 553-621.
- Kennedy, C. (2005). *Variation in the expression of comparison*. Artículo presentado en Cornell University, 11 de noviembre.
- Kennedy, C. (2007). Modes of comparison, en M. Elliott (Ed.), *Proceedings of the Chicago Linguistic Society*, 43, 1(1), pp. 141-165. <https://www.ingentaconnect.com/content/cls/pcls/2007/00000043/00000001>
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford: Stanford University Press.
- Lechner, W. y Corver, N. (2017). Comparative deletion and comparative subdeletion, en Martin Everaert, y Henk van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax*, John Wiley & Sons. <https://doi.org/10.1002/9781118358733.wbysyncom079>
- Lechner, W. (1999). *Comparatives and DP-structure*. Amherst, MA, University of Massachusetts, Amherst, tesis doctoral.
- Lechner, W. (2001). Reduced and phrasal comparatives, *Natural Language and Linguistic Theory* 19(4), pp. 683-735. <https://doi.org/10.1007/s11049-011-9137-1>
- Lechner, W. (2004). *Ellipsis in comparatives*. Berlín: Walter de Gruyter. <https://doi.org/10.1515/9783110197402>
- McCawley, J. D. (1998). *The syntactic phenomena of English*. Chicago y Londres: University of Chicago Press.
- Mendia, J. A. (2017). *Amount Relatives Redux*. Amherst, Massachusetts: University of Massachusetts Amherst, tesis doctoral.
- Mendia, J. A. (2020). One More Comparative, *Natural Language and Linguist Theory* 38, pp. 581-626. <https://doi.org/10.1007/s11049-019-09454-x>
- Merchant, J. (2012). *The two phrasal comparatives of Greek*. Chicago, manuscrito, University of Chicago. <http://home.uchicago.edu/merchant/pubs/gk.comps.rev.pdf>
- Napoli, D. J. y Nespors, M. (1986). Comparative structures in Italian, *Language*, pp. 622-653. <https://doi.org/10.2307/415481>
- Napoli, D. J. (1983). Comparative ellipsis: A phrase structure analysis, *Linguistic Inquiry* 14(4), pp. 675-694.
- Osborne, T. (2009). Comparative coordination vs. comparative subordination, *Natural Language and Linguistic Theory* 27(2), pp. 427-454. <https://doi.org/10.1007/s11049-009-9068-2>
- Pancheva, R. (2006). Phrasal and clausal comparatives in Slavic, en Y. Oseki, M. Esipova, y S. Harves (Eds.), *Proceedings of Formal approaches to Slavic linguistics*, 14, pp. 236-257, University of Michigan, Ann Arbor: Michigan Slavic Publications.
- Parra-Guinaldo, V. (2011). Typological overview of comparatives of inequality with an emphasis on old English *more than* expressions, *Anuario de Estudios Filológicos* 34, pp. 139-156.

- Pinkham, J. (1982). *The Formation of Comparative Clauses in French & English*, Indiana: Indiana University, tesis doctoral.
- Plann, S. (1984). The syntax and semantics of *más/menos ... que* versus *más/menos ... de...* in comparatives of inequality, *Hispanic Linguistics* 1(2), pp. 191-213.
- Real Academia Española (2010). *Nueva gramática de la lengua española*. Madrid: Espasa Libros.
- Reglero, L. (2007). On Spanish comparative subdeletion constructions, *Studia linguistica* 61(2), pp. 130-169. <https://doi.org/10.1111/j.1467-9582.2007.00133.x>
- Rizzi, L. (1997). The fine structure of the left periphery, en L. Haegeman (Ed.), *Elements of grammar*, pp. 281-337, Dordrecht: Kluwer Academic Publishers.
- Romero Cambrón, Á. (1997). De nuevo sobre la alternancia 'que'/de' en el segundo término comparativo, *Verba: Anuario Galego de Filoloxía* 24, pp. 335-348.
- Romero Cambrón, Á. (1998). *Historia sintáctica de las construcciones comparativas de desigualdad*. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.
- Ross, J. R. (1967). *Constraints on variables in syntax*. Cambridge, Massachusetts: Massachusetts Institute of Technology, tesis doctoral.
- Sáez, L. y Sánchez López, C. (2013). *Las construcciones comparativas*. Madrid: Visor.
- Sáez, L. (1992). Comparison and Coordination, en J. A. Lakarra y J. Ortiz de Urbina (Eds.), *Supplements of the Anuario Del Seminario de Filología Vasca «Julio de Urquijo», Syntactic Theory and Basque Syntax*, pp. 387-416, Donostia: Gipuzkoako Foro Aldundia.
- Sáez, L. (1999). Los cuantificadores: las construcciones comparativas y superlativas, en Ignacio Bosque y Violeta Demonte (Ed.), *Gramática descriptiva de la lengua española*, pp. 1129-1188, Madrid: Colección Nebrija y Bello, Espasa Calpe. <https://doi.org/10.4312/linguistica.41.1.176-178>
- Schwarzschild, R. (2010). Comparative Markers and Styrard Markers, en M. Y. Erlewine y Y. Sudo (Eds.), *Proceedings of the MIT Workshop on Comparatives*, 69, pp. 87-105, Cambridge, Massachusetts: MIT Working Papers in Linguistics.
- Seuren, P. A. M. (1984). The comparative revisited, *Journal of Semantics*, 3.1, pp. 109-141.
- Solé, Y. R. (1982). On *más/menos ... que* versus *más/menos ... de* comparatives, *Hispania* 65(4), pp. 614-619. <https://doi.org/10.2307/342385>
- Stassen, L. (1985). *Comparison and Universal Grammar*. Oxford: Blackwell.
- Suárez Martínez, P. M. (2002). ¿Subordinación o coordinación con *quam*?, en E. Espinilla Buisán, P. J. Quetglas y E. Torrego (Eds.), *La comparación en latín*, pp. 229-250, Madrid: Espasa. <http://hdl.handle.net/10651/23525>.
- Suñer, M. (1994). V-movement y the licensing of argumental wh-phrases in Spanish, *Natural Language and Linguistic Theory* 12, pp. 335-372.
- Torrego, E. (1984). On inversion in Spanish and some of its effects, *Linguistic Inquiry* 15(1), pp. 103-129.
- Uribe-Etxebarria, M. (1992). On the structural positions of the subject in Spanish, their nature and their consequences for quantification, *Anuario del Seminario de Filología Vasca «Julio de Urquijo»*, pp. 447-491.

- Vela-Plo, L. (2018). Standards of comparison and the case of Spanish «*que - de* alternation», en K. Bellamy, A. Ionova, y G. Saad (Eds.), *Proceedings of the 25th Conference of the Student Organization of Linguistics in Europe*, pp. 229–249, Leiden: Leiden University Centre for Linguistics. <http://doi.org/10.5281/zenodo.11167803>
- Vela-Plo, L. (2020). *Drawing comparisons: A syntactic and semantic approach to Basque, Spanish and English comparative structures*. Vitoria-Gasteiz: Universidad del País Vasco UPV/EHU, tesis doctoral. <http://hdl.handle.net/10810/46292>
- Vela-Plo, L. (en prensa). Phrasal subcomparatives: A comparative coordination analysis based on evidence from Basque, Spanish, and English, *Linguistics*. De Gruyter Mouton.
- Williams, E. (1978). Across-the-board Rule Application, *Linguistic Inquiry* 9, pp. 31-43.



## NOTAS



## LA PALABRA PARA 'HIJA' EN INDOEUROPEO: UN ANÁLISIS COMPARATIVO

Esteban NGOMO FERNÁNDEZ<sup>1</sup>  
*Universidad Complutense de Madrid*

### *Resumen*

Nos disponemos a realizar un análisis histórico-comparado del término para 'hija' en indoeuropeo. El estudio de las correspondencias fonéticas regulares no ofrece dudas para la reconstrucción de una raíz común que sustenta la formación de este término de parentesco en las distintas lenguas indoeuropeas. Las diversas propuestas etimológicas serán objeto de reflexión.

*Palabras clave:* hija; indoeuropeo; método comparativo; léxico de parentesco; fonética; diacronía.

## THE WORD FOR 'DAUGHTER' IN INDO-EUROPEAN: A COMPARATIVE ANALYSIS

### *Abstract*


We aim to make a comparative analysis of the words which mean 'daughter' in Indo-European. The study of the phonetic correspondences assures the reconstruction of a common root which is the basis of this kinship term in the different Indo-European languages. The main etymological proposals for the subject will be discussed.

*Keywords:* daughter; Indo-European linguistics; comparative method; kinship terminology; phonetics; diachrony.

RECIBIDO: 11/05/2021

APROBADO: 30/05/2022

---

1. enfernandez@ucm.es.  <https://orcid.org/0000-0002-9313-4344>

2. Este trabajo ha sido realizado con la ayuda del Proyecto PID2019-106606GB-C31, financiado por el Ministerio de Ciencia e Innovación del Gobierno de España. Forma parte de los trabajos del Grupo de Investigación Consolidado «Textos epigráficos antiguos de la Península Ibérica y del Mediterráneo griego» de la Universidad Complutense de Madrid

## 1. INTRODUCCIÓN

Es posible reconstruir el término antiguo para ‘hija’ en indoeuropeo, dado que cuenta con numerosos cognados en muchas lenguas (Mallory y Adams, 2006, pp. 210, 213). Sin embargo, el detalle de dicha reconstrucción plantea problemas, en tanto que no encaja con las leyes fonéticas comúnmente aceptadas. Nos disponemos a revisar y reflexionar sobre los conflictos que genera la reconstrucción fonética de este étimo indoeuropeo. La identificación de las anomalías y su discusión podría servir de ayuda para su esclarecimiento o como base para futuros estudios.

Algunos de los paralelos lingüísticos que sustentan la reconstrucción indoeuropea de este antiguo étimo son: θυγάτηρ θυγατρός, ἡ ‘hija’, mic. tu-ka-te, tu-ka-te-re, tu-ka-ta-si /*thugatēr*, -eres, -arsi ¿?/ ‘hija’; arm. *dustr* ‘hija’; osc. *futír* [Nom. sg.] ‘hija’, *futrei* [Dat. sg.], *futre* [ís [Gen. sg. ?]]; gót. *dauhtar* ‘hija’; a. isl. *dóttir* ‘hija’; lit. *duktė* ‘hija’; a. esl. *dъštī* ‘hija’, celtib. *tuateres* [Nom.pl.], *tuateros* [Gen. sg.] ‘hija’; galo *duxtir* ‘hija’; ai. *duhitár-* ‘hija’; a. av. *dugədar-* ‘hija’ av. *dugdar-* ‘hija’; Toc. B *tkācer* ‘hija’; Toc. A *ckacar* ‘hija’ (*IEW*, p. 277).

En el grupo anatolio se documenta una serie de formas que quizá podrían estar en relación con la raíz que nos ocupa: luv. *tuwatrali* ‘hija’; lic. *kbatra-* ‘hija’. Beekes, quien en cierta medida sigue a Kloekhorst, considera que estas formas apuntan a un antiguo grado pleno *-e* de la raíz y ofrece la siguiente reconstrucción: *\*dweṭr-* < *\*dwegtr-* < *\*dhuegH2tr-* (Beekes, 2010, p. 561; Kloekhorst, 2008, pp. 902-904). Chantraine (1968, p. 445) estimaba que la explicación más verosímil para la acentuación griega es proponer una extensión analógica a partir del vocativo θύγατερ = a. i. *dúhitar-*. En cambio, Beekes (2010, p. 561) afirma que no hay necesidad de explicar θυγάτηρ por analogía a través el vocativo θύγατερ. Asimismo, este autor y otros como Chantraine (1968, p. 445) y Frisk (Frisk, 1960, p. 690) sostienen que la forma osca *futír* ‘hija’ proviene de *\*dhug(H2)-ter*. Los mismos especialistas señalan que en la forma griega se observa el típico sufijo para nombres de parentesco en indoeuropeo: *\*-ter* (cf. gr. πατήρ, μήτηρ, φράτηρ, etc.), una idea generalmente aceptada (Chantraine, 1968, p. 445; Frisk, 1960, p. 690; Untermann, 2000, pp. 306-307).

La especulación en torno a la raíz del étimo que significa ‘hija’ en indoeuropeo se produjo desde época muy temprana. Se dudaba de si esta provenía del PIE *\*dhewgh-* ‘ordeñar’ o *\*dhewgh-* ‘ser útil’ (*IEW*, p. 271). Lehmann explica el segundo significado en la idea de que ‘llevar comida para completar la caza’ era la función de las mujeres más jóvenes en las sociedades primitivas (Lehmann, 1986, p. 88; Wodtko, Irslinger y Schneider, 2008, p. 127).

En este sentido, la interpretación etimológica de la forma *\*dhugH2-tér* se encuadra dentro de una discusión de mayor envergadura: hasta qué punto se puede establecer un corte morfológico que nos permita analizar elementos tales como: lexema + sufijo en los nombres de parentesco (Clackson, 2007, pp. 70, 188). Ciertamente, la comparación con otras formas indoeuropeas (cf. *\*pH2ter* ‘padre’, *\*meH2ter* ‘madre’) invita a aislar un sufijo *\*-ter* de parentesco que se habría integrado sobre una base radical *\*dhugH2-*. Asimismo, parece que en origen pudiera haber existido una vinculación estrecha entre *\*-ter* como sufijo de parentesco y el sufijo nominal *\*-ter*, *\*-tor*, *\*-tr-*, sujeto a la alternancia vocálica, que posee valor de agente y un carácter primario. Nótese que la hipótesis de reconstrucción de un sufijo *\*-H2ter* para los términos de parentesco es plausible, de manera que en la forma que nos ocupa se produciría una reinterpretación: *\*dhug-H2tér* frente a *\*dhugH2-tér*. Realmente esta segmentación daría lugar a una estructura radical que podemos expresar de la siguiente manera, en el caso de que alguna vez hubiera participado de la alternancia vocálica: *C(V)RC*. Clackson (2007, pp. 70, 188) expresa las múltiples inconsistencias y el escaso éxito que implica abordar la etimología de términos de parentesco en indoeuropeo en general y de *\*meH2ter-* ‘madre’ y *\*pH2ter-* ‘padre’ en particular. Afirma que la formación de estos se produjo en un momento demasiado lejano en la prehistoria de la lengua como para ser restituida. Y, por otro lado, que la escasez de material convierte en algo fútil la reconstrucción indoeuropea de dichos términos. El autor concluye, por tanto, que estos deberían ser analizados como bases léxicas y reconocer que no todas las palabras son susceptibles de una segmentación morfológica del tipo: raíz + sufijo (Clackson, 2007, pp. 70, 188).

En definitiva, la forma *\*dhugH2-tér*, sea cual sea su análisis estructural, evidencia que existe un término muy antiguo y de gran raigambre en casi todas las lenguas indoeuropeas. Los paralelos aducidos aportan correspondencias fonéticas con un alto grado de coincidencia y sugieren que, desde el punto de vista semántico, no se ha producido una extensión de significado del étimo heredado (Clackson, 2007, p. 202). No obstante, el análisis morfológico de esta forma no está exento de conflicto, puesto que la oclusiva velar sonora no sigue el cambio fonético esperable en algunas lenguas (cf. arm. *dustr* ‘hija’). Esto inclina a ciertos expertos a proponer una reconstrucción alternativa donde aparece una aspiración alternante: *\*dhug(h)H2-ter* (Chantraine, 1968, p. 445; Frisk, 1960, p. 690). También la vocalización o pérdida de la laringal *\*H2*, posiblemente determinada por la oposición acentual, confiere ciertas anomalías a los resultados en los que se basa nuestro análisis comparativo. Dicha alternancia

da lugar a formas sin vocal y formas con vocal de timbre abierto /a/ (cf. avéstico *dugdar-* frente a celtibérico *tuateres*) (Jordán Cólera, 2004, p. 62; Jordán Cólera, 2019, pp. 105-106; Rubio Orecilla, 1999, pp. 359-360; Wodtko, Irslinger y Schneider, 2008, p. 127). Las únicas lenguas que conservan la laríngeal \*H2 vocalizada son a. indio, griego, tocario, celtibérico y, más adelante, veremos específicamente si el en el grupo anatolio también se produce. Por el contrario, la laríngeal no tiene presencia en iranio, armenio, eslavo, báltico, germánico, itálico, ni –dentro de la rama celta– en galo. Esta disparidad de resultados seguramente sea producto de la existencia de un grupo consonántico CC tras la laríngeal, es decir: CHCC, dependiente de la alternancia vocálica del sufijo de parentesco \*-ter y de la oposición acentual, conforme a la hipótesis de Schmidt (Rubio Orecilla, 1999, pp. 359-360; Wodtko, Irslinger y Schneider, 2008, p. 127). La propuesta de reconstrucción realizada tiene en cuenta la diferencia en los datos que nos aportan las lenguas indoeuropeas sometidas a análisis.

## 2. EXPLICACIÓN MORFO-FONÉTICA

La palabra griega θυγάτηρ proviene de la raíz indoeuropea en grado Ø más el sufijo de parentesco \*-ter: \*dhugH2-ter. La oclusiva sonora aspirada inicial experimenta la evolución esperada en griego y se ensordece, la laríngeal \*H2 ejerce como núcleo silábico y vocaliza en una vocal de timbre abierto /a/. A propósito del acento, es imprescindible que este cayera sobre el sufijo, ya que de lo contrario la laríngeal \*H2 habría desaparecido sin vocalizar. En el caso de que el acento no recaiga en la sílaba inmediatamente posterior a la laríngeal, esta en principio desaparece (Jordán Cólera, 2004, p. 62). Ello nos lleva a afirmar que la fijación del acento de la forma griega θυγάτηρ es por cronología relativa posterior a la vocalización de \*H2. Una cuestión que, a juicio del autor, supone un argumento más para sostener la teoría tradicional de que dicho cambio de acentuación se debe a una extensión analógica a partir del vocativo, fenómeno bien documentado en la lengua griega (cf. Φαῖδρος [nombre propio] frente a φαίδρός [adj.]).

El término osco *futír* 'hija' proviene de la raíz en grado Ø, a la que se añadiría el sufijo de parentesco indoeuropeo \*-ter que también reconstruimos para el griego: \*dhug(h)H2-ter. Esta forma apunta a /futr-/ < \*fuxtr-. Aparentemente, la laríngeal indoeuropea permanece sin vocalizar en los dialectos sabélicos, tal vez como consecuencia del grupo consonántico que le sigue: -tr- (Vaan, 2011, p. 253). De ser así, habría que postular para el resto del paradigma una extensión analógica del

sufijo de parentesco en grado Ø, presente en los casos oblicuos, tal y como ocurre en el galo *duxtir* 'hija' (Untermann, 2000, p. 306). Pese a que el acento recae en la sílaba inmediatamente posterior a la laríngea \*H2, la presencia entre ambas sílabas del grupo consonántico *-tr-* impide su conservación. Debemos suponer que este tratamiento se extiende por analogía a los casos rectos, donde el sufijo está en grado pleno *-e* (Jordán Cólera, 2004, p. 62; Rubio Orecilla, 1999, pp. 359-360). Observamos la fricativización incondicionada de la oclusiva sonora aspirada en posición inicial, fenómeno general en toda la familia itálica: \*dh- > /f/. Llama la atención que Clackson (2007, p. 6) incluya de forma restringida esta isoglosa en su estudio de las lenguas sabélicas, dado que solo menciona el cambio en interior de palabra \*dh y \*bh > -f-. El término *futír* [Nom. sg.], <fuutrei> [Dat. sg.] revela claramente que en osco opera la fricativización de la dental sonora aspirada originaria en posición inicial, al igual que en latín y falisco (cf. lat. *facio*; umbro *feitu*, *fetu*; osc. *factud*; falisc. *faced*; gr. τίθημι \*dhi-dheH1-mi < PIE \*dheH1- 'establecer, crear')<sup>3</sup> (Luján Martínez y Berenguer Sánchez, 2005, pp. 197, 200; Untermann, 2000, p. 256). La oclusiva velar sonora final de la raíz probablemente fuese en este caso aspirada, de ahí que en interior de palabra: \*gh > /h/. M. de Vaan (2011, p. 253) manifiesta que la ausencia de /h/ en todas las formas oscas documentadas es llamativa, aunque no demasiado conflictiva como para albergar dudas en cuanto a su etimología. Clackson (2007, p. 11) expresa que «el vocalismo largo de *fūtrei* se encuentra garantizado (la lectura es <fuutrei>) y que es resultado de un alargamiento compensatorio tras la pérdida de *-h-* en el grupo *-kt-*». Sin embargo, dada la complejidad en la reconstrucción de este término de parentesco en indoeuropeo y la inexistencia de mayor especificación, solo podemos suponer que este mismo autor asume la evolución: \*gt- > -kt-. Es decir, tal vez se produjo una asimilación en el modo de articulación de la velar sonora originaria y la aspiración de esta *a posteriori*. El problema no reside para nosotros tanto en la ausencia de aspiración en las formas oscas –como señala M. de Vaan (*vide supra*)– sino en determinar cuál fue el origen de dicha aspiración. Desde el punto de vista fonético, la oclusiva velar sorda ante oclusiva dental sorda presenta una posición bastante estable que no justificaría su desaparición. En cambio, la única posibilidad para no reconstruir una oclusiva velar sorda originaria /k/, hipótesis que no resiste al análisis

3. Añadimos a continuación una relación de cognados griegos, latinos y osco-umbros que podrían evidenciar esta evolución fonética: osc. *faamat*, *famatted*; lat. *for*; gr. φημί < PIE \*bheH2- 'hablar' (Untermann, 2000, pp. 253-254). También, aunque con más problemas: umbro. *feliuf* 'hijos'; lat. *filius* 'hijo'; gr. θήλυς, -εια, -υ 'femenino'; irl. *del* 'pezón' < PIE \*dhH1-/ \*dheH1- 'mamar, chupar'.

comparativo, sería plantear que: \*gH2 > /h/. En resumen, la aspiración de la velar sonora originaria se podría deber a la laringal \*H2 y esta teoría nos llevaría a un escenario parecido al que se sugiere para el gótico (*vide infra*). En cualquier caso, la laringal \*H2 se habría perdido sin vocalizar pese a que el acento recayese en la sílaba inmediatamente posterior a ella, por las razones aducidas anteriormente: grupo -tr- posterior a \*H2. Y, finalmente, la aspiración se disimularía frente a la dental sorda del sufijo \*-ter (Clackson, 2015, p. 11).

El gótico *dauhtar* ‘hija’<sup>4</sup> podría provenir de la misma raíz en grado Ø que el osco *futír* ‘hija’ < PIE \*dhug(h)H2-ter. Se postula la existencia de una forma germánica intermedia \*đuhter-. La sonora aspirada indoeuropea sufre la rotación fonética del germánico y se convierte en una dental sonora /d/. La vocal posterior de timbre cerrado /u/ sufre una apertura de su timbre y es notada mediante el dígrafo <au>, que marca el alófono abierto de la vocal de timbre medio [ɔ]. Gracias a esta alteración del vocalismo se puede establecer una cronología relativa por la que hubiera tenido que producirse necesariamente antes la aspiración de la velar \*g > /h/ o bien \*gh > /g/ > /h/, para que la vocal /u/ experimentara una apertura de su timbre /u+/h/ > [ɔ]+/h/. Es probable que el acento recayese sobre la sílaba anterior a la laringal \*H2 y que esa hubiese sido la razón de su desaparición (Jordán Cólera, 2004, p. 62; Rubio Orecilla, 1999, pp. 359-360). Sin embargo, Wodtko, Irslinger y Schneider (2008, p. 126) postulan una reconstrucción en la que el acento se encontraría sobre el sufijo \*-ter. Esto permitiría explicar la conservación de la laringal \*H2 como la responsable de la aspiración irregular que se produce en la oclusiva velar sonora originaria: \*gH2 > /h/ (cf. gr. μέγας; ai. mahá- ‘grande’). La aparición en la forma gótica *dauhtar* de /h/ en lugar de /g/ es anómala. La evolución que parece experimentar la oclusiva velar sonora originaria tras la rotación consonántica, sea esta una sonora propia o una sonora aspirada, es irregular. También se reconoce la dificultad a la hora de esclarecer por qué el vocalismo del sufijo \*-ter en *dauhtar* es de timbre abierto /a/. Una hipótesis plausible es que nos encontremos ante un doble grado Ø tanto en la raíz como en el sufijo de parentesco: \*dhug-tr-, habiéndose producido una extensión analógica del grado Ø de los casos oblicuos al resto de la flexión. La vocal de apoyo de las sonantes en germánico es de timbre medio /o/, pero eventualmente se produce una irregularidad: vocalizaciones en /a/ delante o

---

4. La lista de los paralelos documentados dentro de la rama lingüística germánica puede ser consultada en Kroonen (2013, p. 107).



detrás de la líquida o nasal. En este caso, la vibrante definida como núcleo silábico podría haber desarrollado una vocal de timbre abierto /a/: \*-tr > -tar.

Entre las palabras que designan las relaciones familiares en celtibérico tenemos testimonio de aquella que significa 'hija': *tuateres* [Nom. pl.] y *tuateros* [Gen. sg.] (Wodtko, 2000, pp. 414-417; Delamarre, 2003, p. 159; Jordán Cólera, 2004, p. 62; Jordán Cólera, 2019, pp. 105-106; Luján Martínez, 2017, pp. 194-195; Matasović, 2009, pp. 109-110; Prósper, 2002, p. 243). El celtibérico *tuateres* 'hijas' es un término que causa cierta controversia entre los expertos de la rama celta, debido a la diferencia de resultados que encontramos en las lenguas pertenecientes a esta subfamilia (cf. galo *duxtir* 'hija', a. irl. *der* 'chica' y *der-* en términos compuestos y nombres propios). Tal vez haya que partir de una proto-forma con la raíz en grado Ø: *\*dhug(h)H2-ter*. Según Matasović (2009, pp. 109-110), las formas celtas provienen de una forma proto-celta *\*duxtir* que sería equivalente a la que constatamos en galo. En cuanto a la forma celtibérica el autor reconoce que la pérdida de *\*x* y el desarrollo de una vocal de timbre abierto /a/ resulta inexplicable. Algunos expertos han propuesto que –como adelantábamos para el griego– dichos cambios fonéticos se explican en relación con la oposición acentual y la presencia de laringal *\*H2* en las formas derivadas de esta raíz indoeuropea. Así pues, dado que en las lenguas celtas se produce la confluencia de ambas series de sonoras indoeuropeas –las sonoras propias y las aspiradas– la dental sonora aspirada inicial experimenta la evolución regular y se desfonologiza: *\*dh > d*. El acento de esta palabra recae en la sílaba inmediatamente posterior a la laringal *\*H2*, de manera que esta vocaliza con un timbre abierto /a/: *\*dhug(h)H2-ter > \*duga-tér*. En celtibérico se extiende analógicamente el vocalismo de los casos fuertes a los débiles, donde el sufijo de parentesco en principio debería ir en grado Ø *\*-tr-* y la desinencia en grado pleno *\*-os*: *tuateros* [Gen. sg.] /*duateros*/ frente a *\*\*tuatoros* en el semisilabario paleohispánico, que reflejaría una secuencia fónica: /*duatros*/. En galo, por el contrario, para explicar la forma *\*duxtir* habría que partir de la forma de los casos débiles: *\*dhug(h)H2-tr-ós* (la aparición de dos consonantes, es decir, *\*-tr-* entre la laringal y la sílaba posterior acentuada impide su conservación), con pérdida de la laringal y la extensión a todo el paradigma de la forma sin vocal de timbre abierto /a/ (cf. av. *dugdar-* 'hija'; arm. *dustr* 'hija') (Delamarre, 2003, p. 159; Jordán Cólera, 2019, pp. 105-106; Martirosyan, 2009, pp. 244-245; Ngomo Fernández, 2019, p. 9; Rubio Orcilla, 1999, pp. 359-360; Wodtko, Irslinger y Schneider, 2008, p. 127). Por último, la lenición celta en el término *tuateres* resulta clara si la reconstrucción es correcta. Sin embargo, ciertos contraejemplos encontrados demuestran el mantenimiento

de la velar sonora indoeuropea en celtibérico (cf. *sekontios*, *mezukenos*) y hacen pensar a los expertos que tal vez el contexto vocálico de *tuateres* sea propicio para la lenición en celtibérico: tras /u/ y quizá solo ante /a/ (Wodtko, 2000, p. 415; Jordán Cólera, 2004, p. 75; Jordán Cólera, 2019, pp. 105-106; Martirosyan, 2009, pp. 244-245; Rubio Orecilla, 1999, pp. 359-360; Stifter, 2006, p. 238).

El a. indio *duhitár-* ‘hija’ proviene de la misma raíz indoeuropea en grado  $\emptyset$  *\*dhug(h)H2-ter*, que postulábamos para las formas anteriormente tratadas (Mayrhofer, 1986, p. 56). Probablemente, a la luz del resultado con aspiración en interior de palabra *-h-* que muestra el término, la velar sonora originaria en este lexema debió de aspirarse en algún momento. La disimilación de la dental sonora aspirada en posición inicial es fruto de un condicionamiento interno del a. indio: la Ley de Grassmann. Sin embargo, se hace necesario establecer cronologías relativas para arrojar luz sobre esta cuestión en la que intervendrían diversos procesos fonéticos: la conservación de la sonora aspirada en antiguo indio, la satemización y la ya mencionada disimilación de aspiradas. La evolución de la velar sonora a una aspiración en posición interior de palabra nos obliga a reconstruir una velar sonora aspirada, que sufriría la primera palatalización y posteriormente se simplificaría con el resultado de aspiración: *\*gh > jh > h*. Si tomamos como punto de partida la reconstrucción fonética que proponemos para esta forma, no es posible fechar la Ley de Grassmann con relación a la satemización. La disimilación de aspiradas encontraría el mismo contexto para su aplicación tanto antes como después de la primera palatalización. Es decir, aquella pudo aplicarse bien en la fase más antigua *\*dhug(h)H2-ter > \*dug(h)H2-ter*, bien *a posteriori* *\*dhug(h)H2-ter > \*duhi-tar*. Sin embargo, lo que sí resulta claro es que la Ley de Grassmann es posterior a la creación de nuevas sordas aspiradas en antiguo indio por la intrusión –según la teoría tradicional– de fonemas glotales procedentes de dialectos dravídicos, que se asimilaron al resultado de sorda + laringal y se fonologizaron. La integración de estos nuevos fonemas que dio solución al desequilibrio interno del subsistema de oclusivas indoeuropeo heredado en antiguo indio es, en realidad, responsable de la conservación de las sonoras aspiradas en esta lengua. Si esta conservación no fuera anterior, si la reestructuración del sistema de oclusivas en antiguo indio no se hubiera producido antes, la disimilación de aspiradas no habría afectado a esta forma. Por otro parte, se puede constatar la vocalización de la laringal *\*H2* en una vocal cerrada de timbre anterior /i/ y la apertura de los timbres medios en antiguo indio, que daría cuenta del vocalismo /a/ que hallamos en el sufijo de parentesco *\*-ter*. Una vez más nos hacemos eco de la conjetura que sostienen algunos expertos para explicar cómo la vocalización de la laringal *\*H2*

en este caso es coherente con la posición del acento, pues este también recae en la sílaba inmediatamente posterior a la laríngea (Jordán Cólera, 2004, p. 62; Wodtke, Irslinger y Schneider, 2008, p. 127). Sin embargo, reconocemos que la aspiración de la oclusiva velar sonora sigue siendo anómala, es difícil considerar que sea etimológica, dado que esta no ha sido restituida en nuestra explicación del grupo celta ni en germánico.

Tal y como adelantábamos con respecto al término osco *futír* 'hija', una de las hipótesis para justificar la forma que acabamos de analizar es interpretar que la laríngea \*H2 provoca la aspiración de la oclusiva velar que la precede (Cheung, 2007, p. 121). En la introducción hemos mencionado cómo solemos realizar reconstrucciones fonéticas con la presencia de una aspiración alternante para soslayar este problema que los datos extraídos analítica y comparativamente nos plantean. Por nuestra parte, hemos tratado de formular una reconstrucción alternativa que no tuviera en cuenta la laríngea, puesto que resulta un tanto excesivo atribuirle la aspiración de la oclusiva velar precedente sin que haya desaparecido (cf. gót. *dauhtar*). Es posible aceptar un fenómeno doble de aspiración + vocalización y conferir dos efectos a la presencia de la laríngea \*H2. No obstante, es importante tener en cuenta que se trata de una hipótesis directamente planteada para clarificar la gran irregularidad, que existe en algunos paralelos lingüísticos, si se parte de una única reconstrucción.

### 3. LOS PARALELOS LINGÜÍSTICOS DEL GRUPO ANATOLIO

Por razones que ya adelantábamos en líneas anteriores parece aconsejable tratar de manera independiente los posibles paralelos lingüísticos del grupo anatolio: hit. *duttariyata/i*; luv. *tuwatra/i*; lic. *kbatra-*. La forma reconstruida que daría cuenta de todas las diferentes denominaciones para 'hija' en la mayoría de las lenguas indoeuropeas, a saber: \**dhugH2ter* ofrece ciertas dificultades en esta subfamilia lingüística. Tradicionalmente se interpretaba que el vocalismo \**a* presente tanto en *tuwatra* como *kbatra* era resultado de la vocalización de la laríngea \*H2, por tanto, se postulaba esta evolución: \**dhugH2tr-* > \**tugatr* > *tuwatr-*. La desaspiración de las oclusivas sonoras aspiradas afectaría a la oclusiva dental sonora aspirada originaria: \**dh-* > \**d-*. Asimismo, parece producirse la lenición de la oclusiva velar sonora intervocálica en licio y en luvita. En este caso, dada la menor tensión articulatoria de los fonemas sonoros, constituye un fenómeno que tal vez se deba encuadrar en la tipología lingüística más que en posibles evoluciones en paralelo. Sin embargo, hoy en día expertos como Kloekhorst (2008, p. 903) nos avisan de

que en ningún caso es asumible que en anatolio se produzca la vocalización en /a/ de una laringal en grado Ø: \*H2. Así pues, la única alternativa sería postular que se haya producido la aparición de una vocal anaptíctica en \*dhugH2tr- > \*dugtr- > \*dugətr- > \*dugatr- para favorecer la pronunciación del grupo consonántico. Y, tras la desaparición de \*g en proto-anatolio, la forma \*dugatr- habría originado las formas del luvita *tuwatra-* y licio *kbatra-*. La objeción de Kloekhorst (2008, p. 903) al respecto es que no se trata de una reconstrucción demasiado atractiva. Su crítica es que, si hubiera sido necesaria una vocal anaptíctica en algún momento, lo esperable habría sido que se desarrollara una vocal de apoyo para la vibrante: \*Vgtr- > \*Vgtər- mejor que \*Vgətr-. Kloekhorst añade que, si *duttariyata/i* verdaderamente constituye un paralelo, evidenciaría una forma sin dicha vocal anaptíctica. Esta forma más bien parece notar lo que fonéticamente sería: \*dugt(a) *ryada* < \*dhugH2t(e)r-iH2-ata.

Dados estos argumentos, Kloekhorst opina que la única opción para poner en relación las formas anatolias y el resto de los testimonios en otras lenguas indoeuropeas es reconstruir un grado pleno \*-e: \*dhuégH2tr [Nom. sg.], \*dhuegH2tér-m [Ac. sg.], \*dhuegH2tr-ós [Gen. sg.]. De esta manera, se postula la existencia de una antigua alternancia vocálica en la raíz, que justificaría la presencia de la vocal abierta /a/ en anatolio. La forma *duttariyata* provendría de \*dhugH2t(e)r- o \*dhugH2tr- más una adición de sufijos poco claros: \*-io- o quizá \*-iH2-, y \*ata-/i-, palabra que sería un préstamo al hitita procedente del luvita. Por otro lado, luv. *tuwatra/i-* y lic. *kbatra* vendrían de \*dwetr- < \*dwegtr- > \*dhuegH2tr- con raíz en grado pleno -e. A fin de esclarecer el cambio de vocalismo en licio se alude a una asimilación vocálica a partir de una forma más antigua \*kβetra- > *kbatra*. Y en ambos paralelos aducidos se reconoce la pérdida de \*g ante consonante (Kloekhorst, 2008, p. 904). La mayoría de los grupos consonánticos formados por Cb se explican porque en proto-luvita se puede reconstruir la presencia de una semiconsonante \*w, donde observamos la presencia de b en licio (Kloekhorst, 2009, p. 127). Es paradigmática esta evolución fonética para la forma *kbatra*, dado que kb- parece provenir de un más antiguo \*dw-. En la secuencia \*dw- la semivocal /w/ se encuentra definida como margen silábico y forma un grupo consonántico inicial con la oclusiva dental sonora d < \*dh- (cf. *kbi* ‘dos’ < \*dwi) (Kloekhorst, 2008, p. 902).

Posteriormente Kloekhorst (2008, p. 904) deja claro que la pérdida de la \*g habría producido de manera regular el alargamiento y también el cambio de timbre de la vocal alternante del grado pleno -e en contacto con la laringal: \*-eH2- > ā. Esto habría dado lugar al luv. *tuwatra/i-* y lic. *kbatra*. En tal caso, esta evolución no

se corresponde exactamente con la que se expone unas líneas antes, donde hemos visto que el vocalismo /e/ se mantiene en todas las fases de dicha reconstrucción. Se tuvo que producir antes la pérdida de la velar sonora \*g que la desaparición de la laringal, para poder admitir que la vocal alternante /e/ se vio afectada por la sonante laringal: cambio de timbre y alargamiento. En otro orden de cosas, si esto fue realmente así y es la hipótesis de reconstrucción que se plantea, no se entiende por qué se propone que la forma *kbatra* es producto de una asimilación de timbre: \**kbetra*- > *kbatra*. En teoría, la responsable del vocalismo /a/ en *kbatra* debería ser la secuencia \*-eH2- en \**dwe(g)H2tr*- < \**dhuegH2tr*-, según se trataba de demostrar. Asimismo, si se argumenta que se produjo la pérdida de la velar sonora \*g ante consonante, es decir, ante la dental sorda \*t en \**dwegtr*- > \**dwet*r, nos encontramos ante el problema de que la laringal \*H2 habría tenido que desaparecer previamente. Esto imposibilitaría que la laringal alargase y cambiase el timbre de la vocal alternante del grado pleno -e, dado que nunca habría estado en contacto con ella. No obstante, en el caso de que, por alguna razón que desconocemos la pérdida de \*g se hubiera producido ante la laringal \*H2 sin vocalizar: \**dwegH2tr*- > \**dweH2tr*-, sería innecesario apelar a la asimilación vocálica a partir de \**kbetra* para explicar *kbatra*. Creemos que, al seguir esta línea de razonamiento, es correcto decir que formas reconstruidas como \*\**dwegtr*- y \*\**dwet*r- jamás habrían podido existir.

Con respecto a la hipótesis de Kloekhorst (2008, p. 904; 2009, p. 127), esta subrayaría la presencia de rasgos más arcaicos en el grupo anatolio que en el resto de las lenguas indoeuropeas, dado que la forma más antigua sería precisamente aquella sujeta a alternancia vocálica de la raíz: \**dhuégH2tr*. Con posterioridad, una vez separado el grupo anatolio, se habría sustituido por el nominativo singular \**dhugH2tér* de manera secundaria en el resto de las subfamilias lingüísticas indoeuropeas, sobre el esquema de la alternancia vocálica en el sufijo y con oposición entre caso recto/oblicuo.

Esta teoría es muy interesante en tanto que da cuenta de las dificultades ante las que nos hallamos para explicar las formas anatolias frente a los paralelos que nos ofrecen el resto de las lenguas indoeuropeas. Sin embargo, creemos que apelar a la retrotracción del acento del vocativo en IE y a la anomalía que este presenta en la forma griega θυγάτηρ no apoya la reconstrucción del grado pleno -e en anatolio. Ya hemos expuesto que la acentuación de esta forma griega tal vez se deba a la analogía con el vocativo. Esto ocurre en otras formas del griego donde el acento de nombres propios, relacionables a veces con adjetivos de acentuación oxítone, se desplaza definitivamente a una posición anterior por el desmedido uso del vocativo

en este tipo de palabras. Evidentemente, la palabra para ‘hija’ no tiene estatuto de nombre propio, sino que es un término de parentesco. Hemos de considerar que este constituye un apelativo de uso común y reiterado en el ámbito familiar, y, por tanto, muy utilizado como vocativo. Además —y esto es lo determinante— nuestro análisis morfo-fonético de los paralelos lingüísticos fuera del grupo anatolio apoya que el acento debe recaer en la sílaba inmediatamente posterior a la laringal *\*H2*, es decir, sobre el sufijo *\*-ter*, para que esta vocalice. Ese es el caso, como veíamos anteriormente, tanto del griego como del a. indio y, por tanto, es difícil reconstruir un nominativo singular *\*θύγατηρ*. En el caso de que hubiera recaído el acento en sílaba inicial, la laringal *\*H2* no habría vocalizado en /a/ y deberíamos poder encontrar en griego una forma parecida a la del galo *duxtir* y a la del av. *dugdar*. Parece ser más consistente con los resultados que nos proporciona la aplicación del método histórico-comparado que, por influencia del vocativo: *θύγατερ* (cf. a. i. *dúhitar* ‘hija’), hubiese un desplazamiento del acento: *\*θυγατήρ > θυγάτηρ*. Entendemos que el fenómeno de la retrotracción del acento en IE es una marca de oposición muy antigua. Chantraine (1984, p. 445) ya explicaba la diferencia de acentuación en la forma *θυγάτηρ* [Nom. sg.] con respecto del a. i. *duhitár* ‘hija’ como una consecuencia de la analogía con la forma del vocativo, que sufre retrotracción del acento. La vocalización o pérdida de la laringal, derivada de la posición del acento en la forma reconstruida y la presencia o no de un grupo consonántico posterior a ella: *\*dhugH2ter*, pone en duda que la prosodia griega pueda ser esgrimida para sustentar la reconstrucción de un grado pleno *-e* en las lenguas anatólias.

#### 4. CONCLUSIONES

Las conclusiones a las que nos puede conducir el examen realizado sobre los paralelos lingüísticos presentes en este estudio son las siguientes:

1. El análisis comparativo abunda en la idea de que el término para ‘hija’ heredado del indoeuropeo se encuentra ampliamente documentado en la mayoría de las ramas pertenecientes a esta macrofamilia lingüística.
2. La interacción entre la posición del acento y los grupos consonánticos que se encuentran en contacto con una laringal resulta determinante para el resultado fonético que pueda dar su evolución en las diferentes lenguas indoeuropeas, tal y como nos ilustra el término para ‘hija’.
3. La acentuación griega del término *θυγάτηρ* ‘hija’ no parece un argumento plausible para defender la presencia de un grado pleno *-e*: *\*dhu(e)gH2-t(e)r*

en la subfamilia lingüística anatolia. Es probable, como ya hemos indicado, que el acento en griego sea secundario y analógico, debido a la forma del vocativo singular.

4. Hemos podido observar en la forma gótica *dauhtar* la manera en que la aspiración que produce la laríngea en la oclusiva sonora precedente —fenómeno inusual y que generalmente se postula para el a. indio— podría tener lugar también en germánico (Wodtke, Irslinger y Schneider, 2008, p. 126). Esta teoría podría asimismo arrojar luz sobre el itálico, a juzgar por la anomalía que ostentan las formas oscas: *futír*, *futrejis* y *fütrei*.

## REFERENCIAS BIBLIOGRÁFICAS

- Beekes, R. (2010). *Etymological Dictionary of Greek*. Leiden: Brill.
- Chantraine, P. (1968). *Dictionnaire étymologique de la langue grecque: Histoire des mots. vol. 1, A-K*. París: Klincksieck.
- Chantraine, P. (1984). *Dictionnaire étymologique de la langue grecque: Histoire des mots. Vol. 2, L-Omega*. París: Klincksieck.
- Cheung, J. (2007). *Etymological Dictionary of the Iranian Verb*. Leiden: Brill.
- Clackson, J. (2007). *Indo-European linguistics: An introduction*. Cambridge: Cambridge University Press.
- Clackson, J. (2015). Subgrouping in the Sabellian branch of Indo-European. *Transactions of the Philological Society*, 113(1), 4-37. DOI: <https://10.1111/1467-968X.12034>
- Delamarre, X. (2003). *Dictionnaire de la langue gauloise*. París: Ed. Errance.
- Frisk, H. (1960). *Griechisches etymologisches Wörterbuch. vol. 1, A-ko*. Heidelberg: Winter.
- Jordán Cólera, C. (2004). *Celtibérico*. Zaragoza: Ediciones del Departamento de Ciencias de la Antigüedad, Área de Filología Griega, Universidad de Zaragoza.
- Jordán Cólera, C. (2019). *Lengua y Epigrafía Celtibéricas*. Zaragoza: Universidad de Zaragoza.
- Kloekhorst, A. (2008). *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden: Brill.
- Kloekhorst, A. (2009). Studies in Lycian and Carian phonology and morphology. *Kadmos*, 47, 117-146. DOI: <https://doi.org/10.1515/KADMOS.2008.011>
- Kroonen, G. (2013). *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- Lehmann, W. P. (1986). *A Gothic Etymological Dictionary*. Leiden: Brill.
- Luján Martínez, E. R. (2017). Sobre los nombres de las unidades familiares indígenas en la hispania antigua (2ª parte). *Veleia*, (34), 189-20. DOI: <https://doi.org/10.1387/veleia.18082>
- Luján Martínez, E. R. y Berenguer Sánchez, J. A. (2005). Falisco «faced» y el perfecto de \*dheH -k- «hacer» en las lenguas itálicas. *Emerita*, 73(2), 197-215. DOI: <https://doi.org/10.3989/emerita.2005.v73.i2.42>

- Mallory, J. P. y Adams D. Q. (2006). *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Nueva York: Oxford University Press.
- Martirosyan, H. (2009). *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.
- Matasović, R. (2009). *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
- Mayrhofer, M. (1986). *Etymologisches Wörterbuch des altindoeuropäischen*. Heidelberg: C. Winter.
- Ngomo Fernández, E. (2019). A propósito de *matrubos* y los términos de parentesco en celtibérico. *Boletín del Archivo Epigráfico*, (4), 5-15.
- Pokorny, J. (1951-1959). *Indogermanisches etymologisches Wörterbuch I-III [= IEW]*, Berna / Múnich: Francke Verlag.
- Prósper, B. M. (2002). *Lenguas y religiones prerromanas del occidente de la Península Ibérica*. Salamanca: Ediciones Universidad de Salamanca.
- Rubio Orcilla, F. J. (1999). Celtibérico *tuateres*, galo *duxtir*, irlandés *der°*, la palabra indoeuropea para «hija». *Kalathos: Revista del Seminario de Arqueología y Etnología Turolense*, (18), 359-371.
- Stifter, D. (2006). Contributions to Celtiberian Etymology II. *Palaeohispanica*, (6), 237-245. DOI: <https://ifc.dpz.es/ojs/index.php/palaeohispanica/article/view/297>
- Untermann, J. (2000). *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- Vaan, M. de (2011). *Etymological Dictionary of Latin and the other Italic Languages*. Leiden: Brill.
- Wodtko, D. S. (Eds.) (2000). *Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Wodtko, D. S., Irslinger, B. y Schneider, C. (2008). *Nomina im indogermanischen Lexikon*. Heidelberg: Winter.



# RESEÑAS



## RESEÑAS

CONDE VALVERDE, Mercedes, *El lenguaje. En busca de las primeras palabras. Origen. Cuadernos de Atapuerca 24*, Burgos, Editorial Atapuerca, 2022, 32 pp., ISBN 978-84-123715-6-7.

*El lenguaje. En busca de las primeras palabras* reúne virtudes más que suficientes para poder concluir que la autora ha alcanzado con éxito la misión divulgativa que se proponía. Es, pues, un libro muy recomendable, particularmente para estudiantes que afrontan por primera vez el estudio del lenguaje desde una perspectiva evolutiva. Una economía extrema de citas y referencias especializadas, un lenguaje claro y una buena articulación argumentativa y narrativa son virtudes que creo que garantizan su éxito como herramienta pedagógica, a lo que ayuda también su reducida extensión. Figurará, desde luego, entre las recomendaciones bibliográficas sobre lingüística evolutiva en mis clases de lingüística general. Por todo esto, es decir, por la eficacia que le concedo como guía, tanto para el estudiante como para el profesor no especializado en los aspectos biológicos y evolutivos del lenguaje, creo que resulta necesario no dejar sin comentar algunas de las cosas que se dicen o se dejan de decir en el cuaderno.

El trabajo se centra, principalmente, en aspectos de la anatomía dedicada a la producción oral y a la recepción auditiva del habla. No es nada de extrañar, ni un defecto en sí mismo. Los datos paleoantropológicos más sólidos con relación a las capacidades relacionadas con el lenguaje proceden del registro fósil de dichos sistemas anatómicos. En este sentido, nada que oponer y sí mucho que elogiar a la autora por sus actualizadas y claras explicaciones. Ahora bien, el predominio de esta perspectiva, al que se suma un desliz argumentativo no trivial, creo que acaba por dar una imagen distorsionada del lenguaje como resultado de la evolución natural de la especie.

El desliz es el que contiene este fragmento:

La Paleoantropología se ha ocupado en rastrear en el registro fósil de la evolución humana la presencia de las capacidades anatómicas necesarias para producir y percibir con precisión los sonidos del habla humana, lo que podríamos denominar como el «hardware» del lenguaje (Conde Valverde, 2022, p. 8).

La perspectiva computacional basada en la distinción «software/hardware» de la ingeniería informática es, en líneas generales, un buen punto de apoyo para introducirse en el estudio de las complejidades de la cognición natural. Sin embargo, el uso de la idea de «hardware» que se hace en este fragmento resulta bastante insólita. En la aplicación común del símil, la noción de «software» se aplica a las diferentes funciones cerebrales, en tanto que caracterizables algorítmicamente y en un sentido abstracto o neutral respecto a la base material en que se ejecutan, mientras que la noción complementaria de «hardware» hace referencia a la base material misma, el sistema nervioso central, en tanto que facilitadora de los procesos propios de dichas funcionalidades. En las aplicaciones habituales de estas categorías en ciencia cognitiva, lo que Conde Valverde llama «hardware» no recibe tal denominación, sino, en todo caso, la de «periférico», que se aplica a todo tipo de dispositivos conectados al sistema a efectos de la comunicación con el usuario o con otros sistemas.

En trabajos realizados conjuntamente, Robert Berwick y Noam Chomsky utilizan, en concreto, la metáfora de la impresora, uno de los periféricos estándar de cualquier sistema informático, para conceptualizar la relación entre la función lingüística propiamente dicha de la mente y las capacidades anatómicas encargadas de producir y percibir con precisión los sonidos del habla humana, que estos autores denominan conjuntamente, desde el punto de vista de su funcionalidad, como el sistema de exteriorización (o, simplemente, Exteriorización) (Berwick y Chomsky, 2011, 2016). Una de las motivaciones más fuertes para atribuir este estatus periférico a la Exteriorización se basa en el hecho de que la función lingüística no está determinada a hacer uso de la vía oral/auditiva, sino que puede servirse con naturalidad de la gestual/visual, como demuestran las numerosas lenguas gestuales habladas por comunidades de personas sordas y adquiribles por las no sordas. Es decir, existen conexiones periféricas alternativas al alcance de la función lingüística, la cual, en aspectos esenciales, responde en cambio a un formato organizativo equiparable (Lillo-Martin y Gajewski, 2014).

El texto de Conde Valverde no consigue evitar, en este sentido, un efecto muy común a casi todas las aproximaciones paleoantropológicas a la evolución del lenguaje: un sesgo desmesurado hacia lo periférico, que, a menos a ojos del lingüista, provoca la sensación de que el lenguaje acaba siendo pasado por alto (véase, ej., el

por lo demás excelente Martínez y Arsuaga, 2009). En el caso del texto reseñado, el efecto se ve multiplicado por la simplificación extrema de la composición interna de una señal lingüística en que se basa la autora. Esta afirma que una señal se compone de sonidos básicos no significativos –fonemas: x, que se agrupan en otros aún no significativos –sílabas: xx, que componen ya símbolos significativos –palabras: (xx-xx). Nada más se nos dice, pero se supone que lo siguiente será una mera concatenación lineal de palabras: (xx-xx) > (xx-xx) > (xx-xx). Lo cierto es que se trata de una visión coherente para quien se centra exclusivamente en los mecanismos periféricos de exteriorización. Sin embargo, las señales lingüísticas consisten, además, en agrupaciones de palabras en frases –[(xx-xx)(xx-xx)(xx-xx)], que pueden contener otras agrupaciones semejantes dentro de sí mismas –[(xx-xx)(xx-xx) [(xx-xx) (xx-xx) (xx-xx)] (xx-xx)], y manifestar dependencias entre sus componentes a larga distancia y entre los niveles de estructura así construidos – [(xx-xx)<sub>j</sub> (xx-xx)<sub>i</sub> [(xx-xx) (xx-xx)<sub>j</sub> (xx-xx)<sub>i</sub>] (xx-xx)<sub>i</sub>].

El componente central del «software» del lenguaje es el programa que permite construir ese tipo de estructuras anidadas o jerarquizadas y establecer relaciones a distancia y cruzadas entre los elementos así estructurados. Y el mayor enigma, desde el punto de vista evolutivo, el de desvelar la evolución de un «hardware» capaz de sustentar las severas exigencias computacionales para su sostenimiento (Balari y Lorenzo, 2013). También, naturalmente, su conexión con un sistema de exteriorización que obliga a transformar esas representaciones en un estímulo lineal, recortando u ocultando una de sus dimensiones de partida.

Propone Conde Valverde un proceso gradual y continuo de evolución del lenguaje que habría tenido un punto de inflexión importante en el *H. erectus* y un relanzamiento en las dos especies herederas de *H. sapiens arcaico*. Critica, por ello, «la idea, apoyada por algunos investigadores muy influyentes, de que la mente y el lenguaje humanos aparecieron de forma súbita y solo en nuestra especie» (Conde Valverde, 2022, p. 29). Imagino que piensa, entre otros, en los referidos más arriba: Robert Berwick y Noam Chomsky, principalmente en el segundo. Al respecto, es importante hacer estas puntualizaciones. La primera es que la cronología de Conde Valverde se aplica, en realidad, solo al dispositivo periférico de exteriorización. La segunda es que Chomsky no defiende realmente una aparición abrupta o radicalmente saltacionista del lenguaje. Su idea, que ha ido concretando hasta cierto punto en los últimos años, es que al sistema computacional o central y al sistema periférico de exteriorización cabe atribuirles largas historias evolutivas independientes (al primero de ellos vinculado a la representación interna del pensamiento). Lo que

habría sido un acontecimiento evolutivo reciente habría sido la conexión entre ambos sistemas (ej. Chomsky, 2010). En concreto, Berwick y Chomsky sugieren en los trabajos ya referidos que seguramente haya sido en este proceso en el que hayan podido tener un papel causal relevante las mutaciones de FOXP2.

Al final, uno se queda con la amarga sensación de que la deseable comunicación y comprensión mutua entre biolingüistas y paleoantropólogos sigue siendo un ideal que no acaba de cumplirse.

#### REFERENCIAS BIBLIOGRÁFICAS

- Balari, S. y Lorenzo, G. (2013). *Computational phenotypes. Towards an evolutionary developmental biolinguistics*. Oxford: Oxford University Press.
- Berwick, R.C. y Chomsky, N. (2011). The biolinguistics program: the current state of its development, en A.M. di Sciullo y C. Boeckx (Eds.), *The biolinguistics enterprise. New perspectives on the evolution and nature of the human language faculty* (pp. 19-41). Oxford: Oxford University Press.
- Berwick, R.C. y Chomsky, N. (2016). *Why only us. Language and evolution*. Cambridge, MA: MIT Press.
- Chomsky, N. (2010). Some simple evo-devo theses: how might they be true for language?, en R.K Larson, V. Dèprez y H. Yamakido (Eds.), *The evolution of human language* (pp. 45-62), Cambridge: Cambridge University Press.
- Lillo-Martin, D.C. y Gajewski, J. (2014). One grammar or two? Sign languages and the nature of human language. *WIREs Cognitive Science* 5, pp. 387-401. DOI: <https://doi.org/10.1002/wcs.1297>
- Martínez, I. y Arsuaga, J.L. (2009). El origen del lenguaje: la evidencia paleontológica. *Munibe. Antropología – Arkeología* 60, pp. 5-16. <https://www.aranzadi.eus/fileadmin/docs/Munibe/2009005016AA.pdf>

GUILLERMO LORENZO  
Universidad de Oviedo

CERRUDO, Alba; GALLEGO, Ángel. J.; ROCA URGELL, Francesc (Eds.), *Syntactic geolectal variation. Traditional approaches, current challenges and new tools*, Ámsterdam/Filadelfia, John Benjamins Publishing Company, 2021, 385 pp., ISBN 978 90 272 1051 7.

El presente volumen ofrece un compendio renovado de estudios acerca de la variedad diatópica sintáctica del español, así como un acercamiento a lo ya conocido por la bibliografía tradicional al respecto. Este recorrido se hace desde un enfoque innovador, centrado en la extracción del máximo potencial del estudio de los atlas lingüísticos, a través de corpus —de correspondencias y de otro calado

(McEnery y Wilson, 2001)—, y de diversas herramientas incluso tecnológicas. Esta metodología se revela cada vez más precisa y necesaria para completar la información fonético-fonológica y morfológica (Demonte, 2003) que, principalmente, se ha trabajado hasta la fecha.

Se trata esta de una publicación reciente (del año 2021) que recoge trabajos teóricos y de corte analítico que observan datos de los últimos 24 meses. El volumen se organiza en 3 secciones temáticas. Una primera sección se dedica a la dialectología y a los materiales que, de manera más tradicional, sirven hasta hoy para su manejo (atlas, bases de datos y anotaciones). La segunda sección recoge estudios específicos de sintaxis aplicados a la dialectología (concretamente, en español europeo y ecuatoriano, en dialectos de origen asturiano y en euskera). Y la tercera sección se centra en las herramientas más novedosas que se están comenzando a implementar para rentabilizar el estudio de la sintaxis: es el caso de algunos corpus (desde académicos hasta *ASinEs*) o compilaciones de neogéneros (como *Twitter*) y de utensilios que pueden ayudar en su tratamiento (*FireAnt*, *Elastic Stack*, etc.). En lo que sigue, realizamos una revisión de estas secciones, poniendo el foco en los contenidos específicos de cada capítulo, con el fin de facilitar al lector interesado la información que precise para su investigación.

En primer lugar, la *Introducción* de los editores (Alba Cerrudo, Ángel J. Gallego y Francesc Roca Urgell) repasa el trabajo más tradicional confeccionado desde hace años, principalmente en áreas rurales, y a partir del cual se desarrolló la sociolingüística. Estos trabajos de campo supusieron el nacimiento de atlas y diccionarios como aglutinación de datos obtenidos manualmente. Son algunos ejemplos europeos el *Sprachatlas des Deutschen Reichs*, el *Atlas Linguistique de la France* o el *Atlas Lingüístico de la Península Ibérica*. Generalmente, estos materiales han sido la base de estudios sincrónicos y diacrónicos de procesos fónicos y morfológicos (Kabatek, 2012), pero han olvidado durante décadas la sintaxis. Los editores explican que el origen de dicho proceder puede haberse heredado de las corrientes estructuralistas, lo que hace que se sigan buscando parámetros de uniformidad frente a principios de variación, sin dejar margen a que la casuística de esta última pueda ser mucho más diversa y “variada” de lo que inicialmente proyectan las redes de principios y parámetros (Kayne, 1996). Así pues, deviene necesaria una aproximación novedosa, que exprima en lo posible el material sintáctico a partir de los datos con los que ya se cuenta, pero que también permita ampliar, desde las nuevas herramientas y tecnologías existentes, otras fuentes y observaciones de datos.

La sección I recoge 4 capítulos, los cuales sirven de base teórica y reflexión sobre lo ya trabajado y lo que falta por acometer en el terreno de la variación sintáctica.

En primer lugar, Pilar García Mouton nos recuerda cómo se configuró el *Atlas Lingüístico de la Península Ibérica (ALPI)*, dirigido por Tomás Navarro Tomás, a partir del acervo de respuestas a cuestionarios, los cuales se recogían de forma manual. De este atlas apenas puede extraerse alguna información sintáctica a partir de frases anotadas sobre los informantes de origen rural, si bien este no era su fin fundacional. Del mismo modo ocurre con la *Cartografía lingüística en la actualidad*, de Sanchis Guarner; el *Atlante linguistico-etnografico italiano della Corsica (ALEIC)*, de Bottigliani; el *Atlas Lingüístico y Etnográfico de Andalucía (ALEA)*, de Manuel Alvar; etc. Por esta razón, aunque pueden extraerse de ello algunas informaciones útiles, García Mouton (1994) abre la sección haciendo hincapié en la necesidad de revisar materiales y aproximaciones complementarias.

En el segundo capítulo, Mar Massanell i Messalles revisa cómo se realiza la selección de verbos auxiliares en lengua catalana, a partir del *Atles Lingüístic del Domini Català* (Joan Veny y Lúdia Pons). Explica sus diferentes regiones dialectales y las motivaciones de la alternancia *ésser/haver*. Tras ello, Maria Pilar Perea trabaja sobre el mismo atlas y el *ALPI*, si bien con un foco en las notas que se registraron para su confección. Detecta que en estas notas es más fácil acercarse a la configuración sintáctica desde una perspectiva contrastiva, como también anota previamente García Mouton.

En el cuarto capítulo, Ares Llop Naya hace un estudio microsintáctico de la polaridad negativa en los diferentes dialectos pirenaicos, de base española, catalana, francesa..., lo que le confiere a su análisis una transversalidad no solo dialectal, sino translingüística.

Veamos ahora, en la sección II, cuatro casos específicos de variación sintáctica. El primero lo presentan Bruno Camus y Edita Gutiérrez con un repaso de los principales atlas españoles (*ALEA*, *ALEANR*, *ALEICan*, *ALECant*, *ALCyL* y *ALECMAN*). Observan sus cuestionarios y cómo, a través de ellos, pueden reconocerse límites entre fenómenos lingüísticos concretos (ej. laísmo) más allá de las tradicionales isoglosas establecidas. En el siguiente capítulo, también firmado por Edita Gutiérrez, esta vez con Irene Gil, se describen y diferencian los usos del género neutro en el español y las lenguas asturianas a partir de los rasgos [-animado], [- individual] y [+proposicional].

Si avanzamos, en el séptimo capítulo del volumen M. Pilar Colomina se fija en la combinación de pronombres clíticos de tercera persona dativo y acusativo



en las variedades del español, del catalán y del aragonés (Rigau, 1998). Observa cómo la variación sintáctica también impone cambios morfológicos, y lo hace a partir de la combinación de materiales como gramáticas teóricas, atlas y corpus.

Para finalizar esta sección, por un lado, Montserrat Batllori, M. Lluïsa Hernanz y Carlos Rubio-Alcalá estudian el uso de los gerundios en el español andino de Ecuador a partir de un análisis de corpus diacrónico (*CORDIAM*) y sincrónico (*CREA*). Por otro lado, Maia Duguine y Aritz Irurtzun se centran en la prosodia de algunas construcciones interrogativas en vasco desde un estudio experimental que aplica cuestionarios al dialecto navarro-labortano, el cual ofreció 162 frases analizables.

La sección III es la última y más breve del volumen. Presenta las herramientas más novedosas y menos explotadas desde el análisis de la variación sintáctica.

En primera instancia, Alba Cerrudo y Anna Pineda introducen y describen la creación, organización y utilidades del nuevo *Atlas Sintáctico del Español (ASinEs)*, el cual trata de cubrir el hueco de información que autores previos observan en los materiales existentes al respecto. Este contiene una base de datos de fenómenos y construcciones sintácticas del español, así como un atlas interactivo de la lengua. A continuación, Cristina Buenafuentes de la Mata y Carlos Sánchez Lancis realizan una lectura de las posibilidades de estudio sintáctico contrastivo (Conque y Kayne, 2005) que puede ofrecer el corpus académico, aún en fase de ampliación, *Corpus del español del siglo XXI (CORPES XXI)*, desde los filtros o parámetros de búsqueda que este permite. Finalmente, Antonio Ruiz Tinoco explica cómo usar diferentes herramientas para la recolección de cantidades grandes de datos como *FireAnt*, *140dev Streaming API Framework* o *Elastic Stak*. Para ello, ejemplifica el uso de corpus masivos como las publicaciones de *Twitter*, a partir de las cuales pueden reconocerse distintas expresiones sintácticas contrastivas, si bien teniendo en cuenta que el origen de los productores de los mensajes puede ser un dato alterado en este tipo de neogéneros.

Con todo ello, el trabajo editado por Alba Cerrudo, Ángel J. Gallego y Francesc Roca Urgell resulta de consulta obligada en la aproximación sincrónica y diacrónica a la variación sintáctica del español. Lo es tanto para el especialista en el tema como para el investigador que quiere formarse en la actualidad en la dialectología de la lengua en Europa y América. Por ello, es estimable un trabajo de detalle y tan pormenorizado que no solo repasa la tradición lingüística, sino que también se adentra en las nuevas herramientas desde las que puede trabajarse la sintaxis, no a mayor escala, sino también de una manera más sencilla, desde el punto de vista contrastivo. Los doce capítulos reseñados, pero también su introducción

y la bibliografía que selecciona cada trabajo son útiles para la actualización de los estudios vigentes sobre la materia, y son una base para la ampliación de las investigaciones que, de este modo, quedan anotadas y potencialmente abiertas.

#### REFERENCIAS BIBLIOGRÁFICAS

- Conque, G. y Kayne, R. S. (Eds.) (2005). *The Oxford handbook of comparative syntax*. Oxford: Oxford University Press.
- Demonte, V. (2003). Microvariación sintáctica en español. Rasgos, categorías y virus, en P. Álvarez Santullano, M. Rosas, M. Contreras y P. Jiménez (Eds.). *Texto, lingüística y cultura* (pp. 9-40), Osorno: Universidad de Los Lagos/Sociedad Chilena de Lingüística.
- García Mouton, P. (1994). *Lenguas y dialectos de España*. Madrid: Arco/Libros.
- Kabatek, J. (2012). Nuevos rumbos de la sintaxis histórica del español, en *Actas del VIII Congreso de la AHLE* (pp. 77-100), Santiago de Compostela: AHLE/Meubook.
- Kayne, R. S. (1996). Microparametric syntax. Some introductory remarks, en J. R. Black y V. Motapanyane (Eds.), *Microparametric syntax and dialect variation* (pp. IX-XVIII), Ámsterdam: John Benjamins.
- McEnery, A. M. y Wilson, A. (2001). *Corpus linguistics: An introduction*. Edimburgo: Edinburgh University Press.
- Rigau, G. (1998). La variació sintáctica: Uniformitat en la diversitat, *Caplletra*, 25, pp. 63-81.

M.<sup>a</sup> AMPARO SOLER BONAFONT  
*Universidad Complutense de Madrid*

