



Research Article

Lars Fant*, Inge Bartning and Rakel Österberg

The left and right periphery in native and non-native speech – A comparative study between French L1/L2, Spanish L1/L2 and Swedish L1

<https://doi.org/10.1515/iral-2019-0095>

Abstract: High-proficient Swedish users of L2 French and Spanish were compared with native speakers of French, Spanish and Swedish with regard to how the syntactic peripheries in natural colloquial speech are structured. Two different though interrelated aspects were included: thus a cross-linguistic analysis comparing the three native speaker groups is combined with an analysis addressing the question of the upper limits of L2 acquisition. All left peripheral (LP) and right peripheral (RP) constituents of a corpus of 110,759 words were classified in a taxonomy relying both on syntactic and functional-pragmatic criteria. The cross-linguistic analysis showed that French and Spanish L1 speakers produced significantly longer LP sequences than the Swedish L1 speakers, who, in turn, put conspicuously more weight on the RP. Significant differences between the three languages were also found with regard to several LP and RP constituent categories. The L2 acquisition-oriented analysis showed that with few exceptions the high-proficient L2 users behaved like L1 users. Although a few interlanguage-related phenomena could be observed, no instances of any clear transfer from the speaker's L1 appeared in either L2 speaker group.

Keywords: left periphery, right periphery, pragmatic functions, cross-linguistic perspective, high-level proficiency in second language use

*Corresponding author: Lars Fant, Romance Languages and Classics, Humanistiska fakulteten, Stockholms Universitet, Stockholm, Sweden, E-mail: lars.fant@su.se

Inge Bartning and Rakel Österberg, Romance Languages and Classics, Stockholm University, Stockholm, Sweden

1 Introduction

1.1 Aim

The aim of the present study is to investigate the organization and use of the left and right periphery of utterances in French, Spanish and Swedish spontaneous dialog, and their acquisition by highly proficient Swedish speakers of L2 French and Spanish.

We assume that understanding the nature of syntactic peripheries is crucial for describing and understanding how languages structure information. The differences between Scandinavian and Romance languages in this respect have hardly been addressed by earlier scholarly work, one exception being Larsson-Ringqvist (2010). Notwithstanding, anyone with experience from teaching Romance languages in Scandinavia would agree that this is a domain in which speakers of Scandinavian and Romance languages behave differently.

This is why we have also looked at L2 production of peripheral structures. Again, we know from our own teaching experience that idiomatic command of peripheral structures constitutes an important challenge for Scandinavian learners of French and Spanish L2 at lower or even intermediate levels of proficiency. What has so far remained less well known is to what extent this difficulty subsists even at advanced or very advanced levels.¹

We considered it important to make explicit what constituent categories can be part of the left periphery (LP), the right periphery (RP), or both, and to what extent. Hence, a comprehensive classification of peripheral constituents is proposed for L1 French and Spanish (Section 4). This analysis was also deemed necessary in order to provide a basis for the Second Language Acquisition-related part of the study.

The present study originates in the research program *High Level Proficiency in Second Language Use* conducted at Stockholm University during 2006–2013 and involving the most proficient levels of second language use (Hyltenstam et al. 2018). A basic objective for this program in general as well as for the present study is to see to what extent the high-proficient non-native use aligns with native use.

The past few decades have seen an increasing interest in the so-called sentence peripheries, in a syntactic, semantic, pragmatic and discourse-organizational

¹ From earlier research (Bartning et al. 2012) we know that complex peripheral structures are difficult, if not impossible, to produce even for very advanced users of French. Our objective has been to replicate this research on a larger scale and also to involve another Romance language, namely Spanish.

perspective, and also from the stance of different and competing linguistic schools. On the one hand, generativists interested in the subject have used it to further develop generative theory (e.g., Cinque and Rizzi 2008) and, on the other hand, representatives of functional grammar have contributed to the field (Lambrecht 1994). In recent years, a seminal volume, Beeching and Detges 2014, has thrown light on various related aspects, including both general and language-specific phenomena. In their introductory chapter, after quoting De Saussure's (1916: 105) remark that discourse unfolds in time and that this fact is represented in Western linguistics as progression from left to right, the authors/editors continue:

A logical implication of this is that the left and right margins of units of language have different functions. For example, in dialogical conversation, the left margin of the most basic unit, the turn, is the place where the speaker takes the right to speak, whereas at the right margin the floor is handed over to the hearer (Beeching and Detges 2014: 1).

The question then arises to which unit or units the peripheries belong: whether the turn, the utterance or the sentence. This issue will be dealt with in Section 2.2.

An important assumption regarding the left and right peripheries is the existence of a functional labor division between them. The question, then, arises what this division consists of, whether it is absolute or partial, and to what extent it depends on universal or language-specific rules. These issues will be addressed in Section 4.3.

1.2 Comparative dimensions and research questions

On the basis of our proposal for a cross-linguistic classification grid for left- and right-peripheral constituent categories (Sections 4.1–4.2), a logically ensuing research question reads:

- In which ways do left periphery (LP) categories differ systematically from right periphery (RP) categories with regard to both their form and their functions? (RQ1)

Most scholars seem to agree on the view that there is indeed a fundamental functional asymmetry to be found between the two peripheries. Would it be correct, then, to assume, as several scholars do (cf. articles in Beeching and Detges 2014), that the difference consists of the fact that the LP mainly serves to establish subjective meanings whereas the RP basically specializes in creating intersubjective meanings? Or is the assumed functional asymmetry to be

interpreted in other terms? Is it for instance correct to claim – as do Detges and Waltereit (2014: 42) – that the LP constituents serve to anchor the upcoming utterance, in contrast to the RP which is the locus where the completed utterance can be negotiated? Our tentative response to this question will be found in Section 4.3.

Our study includes two comparative dimensions, the first of which takes a cross-linguistic perspective. Here, the two Romance languages French and Spanish will be compared with Swedish with regard to the distribution of the various peripheral constituent types; also, comparisons will be made between French and Spanish. The corresponding research question reads:

- To what extent do the LP and the RP constituent categories appear in similar or different proportions in the three languages, and what interpretations could be given to the results? (RQ2)

The second and ultimate comparative dimension takes a Second Language Acquisition perspective, comparing the performance of native speakers of French and Spanish with that of high-proficient L2 speakers having Swedish as their L1. In relation to this, our research question reads:

- To what extent do advanced non-native speakers behave like natives and, in case systematic differences should appear, with regard to what parameters? (RQ3)

The issues raised by the second and third research questions will be addressed in Section 5.

2 Theoretical background

2.1 Earlier research

The left periphery of the sentence was extensively explored in functionalist linguistics in the 1970's and 1980's from the functional sentence perspective and, later on, in the 1990's and 2000's, also from a generative universal grammar perspective. Focus was laid on information-structuring devices, on left dislocation, and on phenomena such as topicalization, clefting and focalization (e.g., Cinque and Rizzi 2008; Demonte and Fernández Soriano 2007; Kotschi 1996; Lambrecht 1994; Rizzi 1997; Sedano 2012). The right periphery, by contrast, has been comparatively less investigated, and with a later starting point (Beeching and Detges 2014: 21).

Research on discourse functions of LP and RP is also an intensive and growing field. Representative studies are to be found in Morel and Danon-Boileau (1998), Beeching et al. (2009), Traugott (2012) and the various contributions to Beeching

and Detges (2014). In general linguistics, in text linguistics and in usage-based studies, these theories have been highly influential.

It has often been assumed that the LP mainly involves subjective functions whereas intersubjective functions are taken care of by the RP. This position has lately been questioned, and in the afore-mentioned volume by Beeching and Detges 2014, contributions such as Beeching and Detges (2014), Detges and Waltereit (2014) and Traugott (2014) provide a more nuanced picture of the topic. Thus, in the introductory chapter of their volume, Beeching and Detges (2014) maintain that the left periphery is mainly concerned with discourse structuring and the right with modalizing (stance, subjectivity, intersubjectivity), although this division of functions cannot be upheld in an exclusive way. According to the authors, the ‘asymmetry claim’, rather than being categorical, is based on frequency (Beeching and Detges 2014: 18–19). As will be seen in Sections 4.3 and 5.3, our data seem to corroborate this view.

A more radical approach has been suggested by representatives of the Val.Es.Co group at the University of Valencia. Thus, work by Estellés Arguedas and Pons Bordería (2014), Salameh Jiménez et al. (2018) and by Pons Bordería (2018) call into question the theoretical basis of the periphery research, which consists in positing a fundamental division – artificial in their view – between core unit and peripheries. Instead, they advocate the application of the segmentation model for discourse units elaborated by their research group (Briz Gómez and Grupo Val.Es.Co 2003).² Their position will be further commented in Section 2.2 with regard to the current study (see also Fant 2019).

As regards the Second Language Acquisition field – the second comparative dimension of our study – there has been an increasing interest in the upper limits of acquisition, in particular as regards the mastery of the syntax-discourse interface in successful end-state learners. Below, some work relevant for our own study of peripheries in discourse organization is shortly presented.

In order to capture features in near-native grammar and the characteristics of final state grammars, Sorace (2003) analyzed optionality as an aspect of grammatical competence (Sorace 2003: 136) exhibited by English near-native speakers of Italian with regard to null vs. overt subject pronouns. The results indicated that the distribution of overt pronoun subjects in near-native Italian grammar is broader than in native Italian while the distribution of null subjects is correspondingly more restricted. Consequently, the non-native speakers used more overt pronouns than the natives. The possible overuse of pronouns in L2 speakers will be addressed in Section 5.1.2.

² In similar terms, Borreguero Zuloaga (2014) proposes an encompassing ‘Frame unit’ which extends beyond the utterance level and is relevant for the placement of discourse particles.

Another important field of investigation is discourse complexity which has been studied by, among others, Conway (2005) and Hancock (2007), within the context of a model proposed by Morel and Danon Boileau (1998), which includes the analytical tools of “preambles” (the thematic part of an utterance, grossly corresponding to our notion of left periphery) and rhemes. In a longitudinal qualitative study of two Swedish learners of French L2 before and after a longer study abroad period in France, Conway (2005) examined the modal elements of the preamble. Results showed, among other things, that the preambles in L2 speech were shorter and less developed than in L1 speech. The study thus indicated influence from L1 and possibly from interlanguage development.

Hancock (2007) used the model proposed by Morel and Danon Boileau (1998) in analyzing highly frequent modalizing adverbs such as *peut-être* (‘maybe’) and *vraiment* (‘really’) in L1 and very advanced or near-native L2 speech. It was shown that these adverbs appear in the preamble mostly in accordance with preferred native patterns.

In Bartning et al. (2012) the same model was used in demonstrating that the mean number of words per preamble is higher for the native than for the non-native groups. On the other hand, the non-native speakers followed native patterns in producing a comparable number of preambles with five or more constituents. In spite of this token of high proficiency, the non-native speakers produced a few morpho-syntactic deviances, which had been wrongly assumed to be absent at these advanced levels. The present study is intended to connect with the findings of Bartning (2012) and Bartning et al. (2012).

Left-dislocation – a mechanism which is an essential component of what we label as left periphery – is an important tool for structuring speech by marking a topic. Managing left-dislocation requires speakers to coordinate syntactic and pragmatic/discursive knowledge, and L2 learners have earlier been shown to have difficulties in handling this interface. However, Donaldson (2011) studying a database consisting of recordings of 10 dyadic conversations between near-natives and natives, came to interesting conclusions: the use of the most advanced non-native groups came very close to, and sometimes even converged with, native use in this regard. Left-dislocation structures were also analyzed by Engel (2010) in a longitudinal developmental study of L2 French, running from beginner to near-native levels. The results from interviews and retellings from 68 speakers at different levels, including a control group of natives, showed that the near-native speaker performance came close to that of the natives.

In dealing with left-periphery structures, recent findings from the active field of syntactic complexity in SLA deserve being mentioned. Thus, Kuiken et al. (2019) present results from studies on variation in syntactic complexity, addressing

inter-learner variation as well as the interface between syntax and morphology and how syntactic and lexical complexity interact.

The results of the afore-mentioned studies will here be addressed from a wider perspective, namely that of left/right peripheries. We thereby hope to fill a gap, by letting a cross-linguistic perspective become the departure point for investigating the Second Language Acquisition perspective as regards the use of syntactic peripheries.

2.2 The notion of syntactic peripheries

At this point, a closer look should be taken at the notion of periphery itself and what it actually means to say that a given constituent is part of the right or left periphery rather than the sentence or utterance core. The problematic point is that the periphery may be defined in relation to discourse units of varying scope: the clause, the sentence, the utterance or the turn.

A brief review of the corresponding notions would go as follows:

- the turn is the basic unit of dialog; a turn may consist of one or several utterances;
- the utterance is the basic unit of discourse and includes a core sentence, complete or incomplete, plus optional constituents preceding or following the core;
- sentences are made up by a verb (overt or elliptic) and its associated arguments along with additional adjuncts;
- a sentence may consist of one or several clauses, each of which will possess a sentential core structure that can be associated with peripheral elements. Constituents of the clause may also appear as “dislocated” to the left or right hand side of the core.

In summary, it can be seen that all these kinds of units may contain peripheral components in addition to their core.

In contrast to this multi-layered, semantically- and pragmatically-based analysis, turns, utterances and sentences/clauses can be viewed as simple surface structures. Here, various kinds of left-peripheral constituents cluster together regardless of whether they relate to the turn, the utterance, the sentence or even the subordinated clause, and the same goes for right-peripheral constituents. Moreover, the internal order in which the peripheral elements are placed seems to obey tendencies rather than rules and to be comparatively free³.

³ *Prosody* seems to play an important part in teasing cores and peripheries apart. Thus, in all three languages considered, the limit between the left periphery and the core is usually marked by a

Extended utterance		
Left periphery	Core unit	Right periphery
i. diferencias culturales ↑ ii. bueno → iii. nosotros Chile ↑	tiene mucho mejor nivel educacional que Brasil ↓	i. a nivel macro → ii. evidente- mente
i. après ↑ ii. bon → iii. voilà →	je me reconnais pas spécialement dans ce modèle de vie-là ↓	i. quoi → ii. principalement

Figure 1: Left periphery + core unit + right periphery: Examples from Spanish and French.

To illustrate our reasoning, let us look at two rather complex utterances taken from our corpus, one in Spanish and one in French (Figure 1).

Here, the response markers Sp. *bueno* and Fr. *bon* ‘well’ relate to their respective ‘outermost’ unit, viz. the *turn*, since they are both to be understood as addressing a preceding turn. The same goes for Fr. *voilà* ‘there you are’. On the other hand, the Fr. connective marker *après* ‘then again’ relates to the *utterance* that follows, which it connects to a preceding utterance by establishing a logical link between both. In the Spanish example, the constituent *diferencias culturales* ‘/as for/talking about/ cultural differences’ also relates to the whole utterance that follows and could thus be labeled a ‘discourse topic’, whereas the (grammatically deviant) expression *nosotros Chile* (‘we /in/ Chile’ i.e., ‘we Chilean people’), being a subject constituent located outside the core, should be regarded as related to the *sentence* unit and could thus be labeled a ‘sentence topic’.

As regards the constituents placed in the right periphery in our example, Fr. *quoi* ‘right?’ is a tag question and a confirmation claimer, and should therefore be classified as relating to the *turn* unit. On the other hand, Fr. *principalement* ‘in principle’/‘mainly’ and Sp. *evidentemente* ‘obviously’ are discourse markers intended to modify their respective *utterance* unit. Finally, Sp. *a nivel macro* ‘on a macro level’ is to be seen as an adjunct dislocated from, and added post hoc, to the core, and thus relates to the *sentence* unit represented in the core.

boundary tone, generally a rising or level tone signaling continuity. Likewise, a prosodic boundary tone normally marks the limit between the core unit and the right periphery. If the left or right periphery includes several constituents, these also tend to be separated by boundary tones (cf. regarding French, Morel and Danon-Boileau 1998, and also Engel 2010; for Spanish, Elordieta et al. 2005).

In contradistinction to the semantic/pragmatic perspective taken in the examples above, we have chosen to apply a “surface structure perspective” on our data. Thus, both examples above are regarded as consisting in three left-periphery constituents, plus the core sentence constituent, plus two right-periphery constituents. In our data analysis, each periphery constituent has then been attributed to a LP or RP category (for the classification, see Sections 4.1–4.2) and statistics have been applied to each of the five data sets, allowing for comparisons to be made.

3 Data and method

3.1 Database

The present study draws mainly on data from the corpus “Multi-task London-Paris-Santiago” (involving native and non-native users of English, French and Spanish), for the French data partly on the InterFra corpus (<https://spraakbanken.gu.se/eng/resurs/interfra>) and, for the Swedish data, on the SweDia database (<http://swedia.ling.gu.se>).

Our data consists of semi-structured and open-ended interviews in which the respondents are asked questions about their life and aspects perceived by them as relevant. The speech genre could best be labeled “Self-presentation”. The discourse represented is neither typically dialogic nor typically monologic. The interviewees respond mostly with long turns including more or less extensive narratives. The interviewers were all native speakers of their respective tongues. Their task was to elicit discourse while supporting and encouraging the respondent in producing talk. The interviews lasted on average 30 min. ranging between 2,000 and 3,000 words.

In sum, our database consists of five sets:

- Interviews in Metropolitan French with (a) 10 native speakers, and with (b) 10 non-native speakers of French.
- Interviews in Chilean Spanish with (a) 10 native speakers, and with (b) 10 non-native speakers of Spanish.
- Interviews in Standard Swedish with five native speakers of Swedish. These were included after considering the need for a control group, both for the cross-linguistic study and for the SLA-oriented one.

Table 1 shows the distribution of groups, participants and amount of speech recorded for each group.

Table 1: Number of participants and total number of words produced by each participant group.

Group	FR-NS	FR-NNS	SP-NS	SP-NNS	SWE-NS	TOT
Participants	10	10	10	10	5	45
Words	25,940	25,956	20,253	24,443	14,167	110,759

Table 2: Background variables in five participant groups.

Groups	FR L1	FR L2	SP L1	SP L2	SW L1	<i>All in all</i>
Average age	27.3	29.0	38.8	39.8	25.8	<i>Means: 34.1</i>
Range of age	23–34	25–33	23–65	27–59	21–32	21–65
Mean length of stay	–	10.3 years	–	9.9 years	–	<i>Means for the L2 users: 10.1 years</i>
Gender	6F/4M	10F/0M	4F/6M	6F/4M	3F/2M	29F/16M

3.2 Participants

The non-native speakers of French and Spanish were all Swedes who had lived in France and Chile, respectively, for between 5 and 15 years. Their command of L2 French and Spanish, respectively, was very advanced attaining in many cases near-native proficiency levels (cf. Hyltenstam et al. 2018), while their oral and interactional proficiency was estimated at C2 level of the CEFR scale (no measurement was done of their proficiency in writing). The age and gender of the participants is shown in Table 2⁴. It should be added that the non-native participants were matched with the native participants also with regard to socio-economic level (middle class) and educational level (as a minimum, at least initiated higher education).

⁴ The NS and NNS groups have been closely matched for each of the languages French and Spanish. However, the participants of the Spanish-speaking groups are about 10 years older on average than the French-speaking. As for gender, the French NNS group stands out as consisting of women only. This, however, has had little or no impact on the results, since no significant difference has been found between the group of French NNS females and the French NS females or between the French NNS females and the French NS speakers as a whole.

3.3 Procedure of analysis

The first step of our analysis was to sort out, and define the range of, each LP and RP sequence of the five datasets. This was done, first introspectively and separately by each of the authors involved, then intersubjectively, by the three authors, reaching consensus.

The next step consisted in establishing a comprehensive taxonomy for the categorization of the LP and RP constituents that appeared in our database. This endeavor was also carried out intersubjectively by the three authors, reaching consensus. It meant creating a tool for carrying out the analysis corresponding to RQ 2 and 3 (see Sections 4.1–4.2).

The third step implied the coding of the LP and RP constituents encountered. As in the first step, this was done, first introspectively by each author, then intersubjectively by all three together, reaching consensus. The LP sequences found included between one and four identifiable constituents as a rule, albeit not infrequently more, and in extreme cases up to eight constituents. The RP sequences, on the other hand, mostly included only one constituent, sometimes two, and occasionally as many as three constituents.

4 Categorization of the LP and RP constituents

Our primary approach was to establish a comprehensive surface-structure-based classification of all LP and RP constituents that appeared in our trilingual data. It consisted in listing different categories of constituents that typically appeared to the left and to the right of a core unit. In order to highlight the deeper dependencies of the peripheral constituents, we also attributed the categories to three discourse units of different scope, viz. the sentence/clause, the utterance and the turn. Only the first approach will be directly reflected in the account of the results in Section 5.

4.1 LP constituents

The flow chart in Figure 2 shows how the LP constituent categories are related to each other, in our taxonomy.

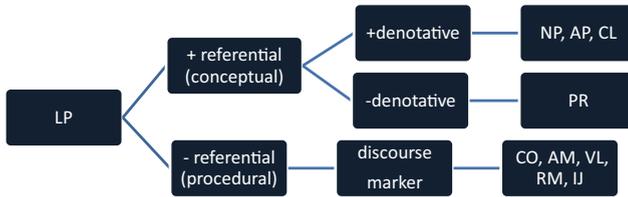


Figure 2: LP categories: LP = left periphery; NP = noun phrase; AP = adverbial phrase; CL = clause; PR = pronoun; CO = connective; AM = adverbial marker; VL = verb-based utterance launcher; RM = response marker; IJ = interjection.

4.1.1 Referential LP constituents

One group of LP constituent candidates were initially placed noun phrases (NPs). Among these we could single out dislocated objects, topical subjects and independent topics. The following examples are illustrative:⁵

- (1) **Le tango** [je trouve ça très très sympa].
Tangoes I find /that/ very very nice. (=dislocated object)

- (2) **El aspecto positivo del trabajo** primero que todo [es porque no es un trabajo común y corriente].
The positive side of the job, more than anything, it's because it's not a common everyday job. (=topical subject)

- (3) **Mon travail** [je travaille au service de l'Action Pédagogique].
My work, I work for Educational Action. (=independent topic)

It can be seen that dislocated object and topical subject NPs are constituents which relate to the clause or sentence level whereas an independent topic NP rather relates to the whole utterance. However, the limits between these types are fuzzy, and a dislocated object or topical subject could perhaps be thought of as belonging to the sentence and the utterance level at the same time.

Initially placed “full” (=non-clitic) pronouns (PR) can be found in the same functions as NPs:

⁵ In the examples of this Section, the highlighted constituent is marked in boldface and the core unit is placed within brackets. The English approximate translation follows in italics. All examples are directly taken from our data, except for (8), (12) and (14), which are partial re-phrasings of authentic examples.

- (4) **A mí** [los ingleses me caen el descueve].
I like (literally: ‘me please’) *the English a hell of a lot*. (=dislocated object)

Initially placed adjuncts make up another group of LP constituent candidates. In syntactic terms they may correspond to an adverbial phrase (AP) or a clause (CL). Their role seems to correspond to a temporal, spatial, causal or otherwise circumstantial framing of the core unit:

- (5) El aspecto positivo del trabajo **primero que todo** [es porque no es un trabajo común y corriente].
The positive side of the job, more than anything, it’s because it’s not a common everyday job. (=circumstantial AP)

- (6) **Comme lui il est cuisinier** [il va monter une affaire].
Since he is a chef, he will set up a business. (=causal CL)

4.1.2 Non-referential (procedural) LP constituents

The categories NP, AP and CL are basically “content expressions”, which is to say they are referential and denotative units (in contradistinction to pronouns, which are referential and non-denotative). A majority of LP constituents, however, are non-referential (procedural) expressions and could be grouped together under the heading “discourse markers”, this term being understood in a broad sense.

The first non-referential category is connectives (CO) which are perhaps the most thoroughly described class of discourse markers and are also the most widely represented kind in our data. Their function is to connect the unit they precede to previous discourse, regardless of whether they introduce a clause, a sentence or a whole utterance.

Structurally speaking, there are two kinds of connectives: clause-initial conjunctions and sentence- or utterance-initial adverbial connectives. Although it is debatable whether conjunctions should be regarded as true LP constituents, considering their embeddedness in the clause, we have chosen to include them due to their pragmatic function which is the same as that of connective adverbials: thus, for instance, a connective adverbial such as Sp. *entonces* does the same job as the conjunctive phrase *así que* (both = En. ‘so’)⁶:

⁶ A noticeable difference between them is that adverbial connectives are generally followed by a boundary tone whereas conjunctions are generally not.

- (7) **Entonces** el nivel sociocultural de la gente [es muy importante].
So people's socio-cultural status is very important.
- (8) **Así que** el nivel sociocultural de la gente [es muy importante].
So people's socio-cultural status is very important.

The second category of non-referential LP constituents consists of other types of adverbial discourse markers (AM)⁷. A frequent type is markers with an epistemic or an evidential function. These relate to either the sentence or the utterance:

- (9) **Peut-être** [je suis trop jeune].
Maybe I'm too young.
- (10) Entonces **obviamente** [tengo una imagen bastante poco objetiva].
So obviously I have a picture that is very little objective.

The third kind of non-referential LP constituents is verb-based utterance launchers (VL), which from a formal point of view are incomplete clauses that precede and introduce an utterance. From a pragmatic perspective, verb-based utterance launchers fill functions analogous with certain kinds of adverbial markers, in particular evidential expressions. Here utterance pairs such as (11–12) and (13–14) could be compared⁸:

- (11) **Je crois qu'**[il se trompe].
I believe he is mistaken. (=verb-based utterance launcher)
- (12) **Probablement** [il se trompe].
He is probably mistaken. (=adverbial)

7 There is an important distinction between what has been tagged as CO and as AM. Thus, only connective elements placed in the initial position have been tagged as COs, whereas connective expressions that follow are regarded as AMs. Thus, in the example *Entonces por lo tanto el nivel sociocultural de la gente es muy importante* ('so because of this, people's sociocultural status is very important'), *entonces* is tagged as a CO and *por lo tanto* as an AM.

8 As was the case with conjunctions vs. adverbial connectives, there is a difference in the prosodic marking between VLs and AMs: the latter are generally followed by a boundary tone whereas the former are not.

- (13) **Es que** yo [he hablado con mucha gente aquí].
Thing is, *I have spoken to many people around.* (=verb-based utterance launcher)
- (14) **Claro** [he hablado con mucha gente aquí].
Of course *I have spoken to many people around.* (=adverbial)

Finally, we found two types of LP constituents which relate to the turn rather than to the utterance unit, namely response markers and interjections.

Response markers (RM), the fourth kind of non-referential LP constituents, are reactions to the previous turn (produced by the interlocutor or by the speaker) and serve as an introduction to and setting for the turn that follows:

- (15) **Oui** parce que moi en fait [c'est soit la vie parisienne soit la province].
Yes, *because in my case it's actually either living in Paris or in the province.*⁹

A response marker need not stand out as directly positive or negative but can, as in the case of *bueno/bon* 'well', express an implicit objection:

- (16) **Bueno** ahí [tienes otra cosa que me llamó la atención].
Well *there you've got something else that drew my attention.*

Finally, interjections (IJ), the fifth kind of non-referential LP constituents, are affect-laden expressions that normally constitute responses to the preceding turn. In the same way as response markers, they serve as an introduction to and setting for the turn that follows:

- (17) **Uf mira** como te digo [no he alcanzado a conocer a tantos ingleses].
Oof look, *as I said, I haven't got to know that many English people.*

4.1.3 LP constituents in summary

All in all, in our total corpus of 110,759 words, 5,169 LP constituents were encountered. Table 3 provides the overall figures for the average number of occurrences per 100 words for all LP categories considered across the five participant

⁹ This utterance is a good example of the rather infrequently occurring LP sequences consisting of four separate constituents: *oui* ('yes', a response marker) + *parce que* ('because', a conjunction) + *moi* ('I myself', a topical subject pronoun) + *en fait* ('actually'), an evidential adverbial marker.

Table 3: LP categories across the five datasets: Mean number of constituents/100W.

LP constituents	FR L1	SP L1	SW L1	FR L2	SP L2
Adverbial markers	1.25 (<i>SD</i> = 0.51) (Range = 1.80)	1.74 (<i>SD</i> = 0.70) (Range = 1.90)	1.38 (<i>SD</i> = 0.33) (Range = 0.90)	1.20 (<i>SD</i> = 0.44) (Range = 1.40)	2.24 (<i>SD</i> = 0.33) (Range = 0.90)
Adverbial phrases	0.78 (<i>SD</i> = 0.34) (Range = 1.10)	0.95 (<i>SD</i> = 0.35) (Range = 1)	0.42 (<i>SD</i> = 0.13) (Range = 0.30)	0.85 (<i>SD</i> = 0.29) (Range = 1)	0.92 (<i>SD</i> = 0.34) (Range = 1)
Clauses	0.51 (<i>SD</i> = 0.26) (Range = 0.70)	0.62 (<i>SD</i> = 0.36) (Range = 1.10)	0.30 (<i>SD</i> = 0.14) (Range = 0.30)	0.63 (<i>SD</i> = 0.17) (Range = 0.50)	0.63 (<i>SD</i> = 0.21) (Range = 0.70)
Connectives	3.50 (<i>SD</i> = 0.79) (Range = 2.60)	2.58 (<i>SD</i> = 0.96) (Range = 3.10)	3.86 (<i>SD</i> = 1.12) (Range = 2.80)	3.46 (<i>SD</i> = 0.79) (Range = 2.80)	2.84 (<i>SD</i> = 0.39) (Range = 1.10)
Interjections	0.16 (<i>SD</i> = 0.18) (Range = 0.60)	0.09 (<i>SD</i> = 0.08) (Range = 0.20)	0.04 (<i>SD</i> = 0.09) (Range = 0.20)	0.13 (<i>SD</i> = 0.15) (Range = 0.40)	0.07 (<i>SD</i> = 0.09) (Range = 0.30)
Noun phrases	0.65 (<i>SD</i> = 0.22) (Range = 0.60)	1.62 (<i>SD</i> = 0.49) (Range = 1.50)	0.68 (<i>SD</i> = 0.18) (Range = 0.40)	0.52 (<i>SD</i> = 0.25) (Range = 0.70)	1.25 (<i>SD</i> = 0.39) (Range = 1.10)
Pronouns	0.66 (<i>SD</i> = 0.34) (Range = 1.10)	2.05 (<i>SD</i> = 0.45) (Range = 1.50)	0.62 (<i>SD</i> = 0.29) (Range = 0.70)	0.95 (<i>SD</i> = 0.31) (Range = 0.90)	2.48 (<i>SD</i> = 0.82) (Range = 2.50)
Response markers	2.04 (<i>SD</i> = 0.89) (Range = 2.70)	0.53 (<i>SD</i> = 0.27) (Range = 1)	1.02 (<i>SD</i> = 0.29) (Range = 0.80)	1.88 (<i>SD</i> = 1) (Range = 3)	0.67 (<i>SD</i> = 0.34) (Range = 0.90)
V-based utterance launchers	0.77 (<i>SD</i> = 0.41) (Range = 1.20)	0.49 (<i>SD</i> = 0.26) (Range = 0.90)	0.12 (<i>SD</i> = 0.16) (Range = 0.40)	1.01 (<i>SD</i> = 0.36) (Range = 1)	0.60 (<i>SD</i> = 0.99) (Range = 3.80)
Total:	10.32 (<i>SD</i> = 1.6) (Range = 4.9)	10.60 (<i>SD</i> = 1.4) (Range = 5)	8.44 (<i>SD</i> = 1) (Range = 2.60)	10.63 (<i>SD</i> = 1.7) (Range = 5.46)	11.70 (<i>SD</i> = 1.8) (Range = 5.5)

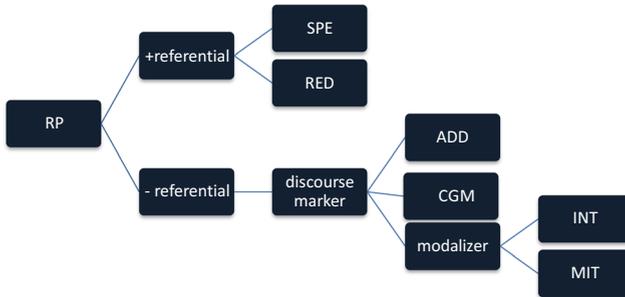


Figure 3: RP categories: RP = right periphery; SPE = specifier; RED = reduplicator; ADD = additive marker; CGM = common ground marker; INT = intensifier; MIT = mitigator.

groups (the rightmost columns present the averages across the board, and, within parentheses, the size of the respective categories by falling order).

Table 3 shows that the most widely represented LP category by far, across all five sets, is connectives (CO), followed by adverbial markers (AM); taken together, these two constituent types amount to 46% of all occurrences. Adding the remaining discourse marker types (i.e., response markers, RM, verb-based utterance launchers, VL, and interjections, IJ), the non-referential markers taken together represent 65% of the LP occurrences. The referential categories, thus, cover 35 %, of which 15% are pronouns (non-denotative) and only 20 % belong to denotative (i.e., semantically “heavy”) categories, viz. noun phrases (NP), adverbial phrases (AP) and clauses (CL).

4.2 RP constituents

Although our classification of LP constituents seemingly started by singling out traditional syntactic categories such as NP or AP, in reality semantic-pragmatic properties such as “topical” or “connective” soon proved more relevant. Admittedly, the criteria used for the classification of LP constituents are mixed, which is due to our aim to set up a practical data-driven, rather than an exclusively theory-driven taxonomy.

The flow chart in Figure 3 shows how the RP constituent categories are related to each other, in the proposed taxonomy.

As regards our classification of RP constituents, applying traditional syntactic criteria proved clearly irrelevant. In fact, although certain constituents such as right-dislocated subject NPs or pronouns could be seen as grammatically defined, they are included in semantically and pragmatically defined categories, viz. ‘reduplication’ and ‘specification’.

The first division line, as in the case of the LP constituents, runs between referential and non-referential (procedural) expressions.

4.2.1 Referential RP constituents

Two types of referential RP constituents are distinguished, both attributable to the sentence unit. One type consists in mere reduplications (RED) of constituents that already appeared in the core unit or the LP. Their general pragmatic function seems to be to add emphasis to the reduplicated constituent. An example of this is the construction in which the subject first person clitic *je* is repeated in the RP in the shape of the full pronoun *moi*, as in the following utterance¹⁰:

- (18) Et donc [je propose trois visites↓] **moi**.
And so I am proposing three visits myself.

The second type of referential expressions encountered in the RP is specifications (SPE) of information omitted in the core unit, i.e., a mechanism either for repair or for “fleshing out” content. The latter seems to be the case in the following utterance:

- (19) Y ahora [trabajo en la Comisión Europea↓] **en la embajada en Santiago**.
And now I'm working for the European Commission, in their embassy in Santiago.

4.2.2 Non-referential (procedural) RP constituents

Four out of six RP constituent categories are non-referential and function as discourse markers in a broad sense. In grammatical terms they represent a variety of structures: complete or elliptic clauses, APs, adverbials, particles, etc.

Many RP constituents can be classified as modalizers (Fant 2016: 20–21) or, more precisely, as expressions that either attenuate or aggravate the pragmatic force of the string they operate on. Modalizers relate to the utterance rather than to the sentence unit.

¹⁰ This construction, described by Detges and Waltereit (2014) is typical of spoken French and has no equivalent in spoken Spanish.

For illustrating attenuating expressions, or so-called mitigators (MIT), the following utterance may serve as an example:

- (20) [Y son los encargados del comercio exterior de Chile↓] **digamos**.
*And they're in charge of Chile's foreign trade, **let's say**.*

The counterpart of mitigators is aggravating constituents, referred to as intensifiers (INT), as illustrated in the following utterance:

- (21) [Je fais également aussi un master de sociologie↓] **voilà**.
*I'm also doing a master in sociology, **there you are**.*

A frequently occurring type of RP constituents is common ground markers (CGM), which relate to the turn unit and express intersubjectivity (cf. Fant 2007, 2011) by either establishing what has been previously said as common ground, or by requesting confirmation for something to be accepted as such. Typically, the common-ground-establishing and affirming type is marked by a falling tone and the requesting and common-ground-eliciting type, by a rising tone.

In the following example two common-ground-affirming markers are included:

- (22) [La Normandie] **ouais**↓. [On s'est rencontré à seize ans à Cherbourg] **donc**↓.
*Normandy, **right**. We met when we were sixteen in Cherbourg, **you see**.*

The following utterance may serve as an example of a common-ground-eliciting marker:

- (23) [Necesitas saber la dirección exacta] **o no**↑.
*You need to know the full address, **don't you?***

The fourth non-referential category of our taxonomy is additive markers (ADD). This is, in fact, a hybrid category including mainly markers akin with certain LP connectives, in particular expressions corresponding to 'also', 'for instance' or 'by the way'. They typically relate to the utterance unit:

- (24) [On a des amis] **d'ailleurs** là-bas en Suède.
*We have friends, **by the way**, over there in Sweden.*

A sub-type of additive markers is sentence-final pseudo-connectives. These are seemingly connective expressions that do not, however, introduce another utterance but are left hanging, thereby signaling open-endedness:

- (25) [Y teníamos muchos amigos comunes↓] **así que** →
*And we had many common friends, **so...***

Table 4: RP categories across the five data sets: mean number of constituents/100W.

RP constituents	French L1	Spanish L1	Swedish L1	French L2	Spanish L2
Additive markers	0.11 (<i>SD</i> = 0.09) (Range = 0.27)	0.06 (<i>SD</i> = 0.04) (Range = 0.11)	0.27 (<i>SD</i> = 0.21) (Range = 0.43)	0.13 (<i>SD</i> = 0.10) (Range = 0.32)	0.08 (<i>SD</i> = 0.08) (Range = 0.27)
C G markers	0.77 (<i>SD</i> = 0.46) (Range = 1.70)	0.49 (<i>SD</i> = 0.35) (Range = 1)	1.9 (<i>SD</i> = 1.76) (Range = 4.35)	0.38 (<i>SD</i> = 0.31) (Range = 0.81)	0.40 (<i>SD</i> = 0.42) (Range = 1.44)
Intensifiers	0.64 (<i>SD</i> = 0.41) (Range = 1.36)	0.10 (<i>SD</i> = 0.07) (Range = 0.24)	0.52 (<i>SD</i> = 0.70) (Range = 1.70)	0.47 (<i>SD</i> = 0.44) (Range = 1.24)	0.11 (<i>SD</i> = 0.12) (Range = 0.39)
Mitigators	0.27 (<i>SD</i> = 0.16) (Range = 0.48)	0.37 (<i>SD</i> = 0.16) (Range = 0.48)	1.23 (<i>SD</i> = 0.67) (Range = 1.71)	0.47 (<i>SD</i> = 0.19) (Range = 0.54)	0.46 (<i>SD</i> = 0.17) (Range = 0.52)
Reduplicators	0.03 (<i>SD</i> = 0.03) (Range = 0.07)	0.03 (<i>SD</i> = 0.03) (Range = 0.07)	0.05 (<i>SD</i> = 0.08) (Range = 0.20)	0.04 (<i>SD</i> = 0.05) (Range = 0.17)	0.03 (<i>SD</i> = 0.03) (Range = 0.07)
Specifiers	0.18 (<i>SD</i> = 0.41) (Range = 1.20)	0.27 (<i>SD</i> = 0.41) (Range = 1.20)	0.30 (<i>SD</i> = 0.16) (Range = 0.40)	0.22 (<i>SD</i> = 0.36) (Range = 1)	0.33 (<i>SD</i> = 0.46) (Range = 1.50)
Total:	1.99 (<i>SD</i> = 0.41) (Range = 1.45)	1.32 (<i>SD</i> = 0.32) (Range = 1.45)	4.38 (<i>SD</i> = 0.96) (Range = 4.79)	1.71 (<i>SD</i> = 0.39) (Range = 1.50)	1.42 (<i>SD</i> = 0.33) (Range = 1.70)

4.2.3 RP constituents in summary

All in all, in our total corpus of 110,759 words, 1,305 RP constituent tokens were encountered. Table 4 provides the average numbers of instances per 100 words for all RP categories considered across the five participant groups (the rightmost columns present the averages across the board, and, within parentheses, the size of the respective categories by falling order).

First of all it should be noticed that the total number of RP constituents, across the board, represent only one fourth of that of LP constituents (see Section 4.1.3). There are, however, important divergences across the five datasets. These will be further commented on in Section 5.2.

The most frequent RP category is that of common ground markers (GGM), followed by the two modalizer categories, the mitigators (MIT) and the intensifiers (INT). These three categories taken together make up 81% of the total number of occurrences. Together with the additive markers, the proportion of non-referential RP constituents amounts to more than 86%. Thus, the referential categories, viz. the reduplicators (RED) and the specifiers (SPE), represent less than 14% of all RP constituents. This is a clearly smaller proportion than the corresponding proportion of referential LP categories (=35%).

4.3 Commonalities and differences between the left and the right periphery (RQ1)

With regard to RQ1, and returning to the issue of functional asymmetry between the peripheries, it can be seen that their relation is clearly asymmetrical but also that given pragmatic functions are represented in both peripheries. Thus, tokens of intersubjectivity can be expressed both in the left periphery by means of response markers (RM), and in the right periphery by means of common ground markers (CGM). Also, tokens of subjectivity appear to be expressed both in the left periphery, by means of various types of discourse markers (connectives, evidential/epistemic adverbial markers, interjections), and in the right periphery, especially by means of modalizing markers (INT and MIT).

Still, there is a noticeable labor division between the peripheries, which supports the principle of asymmetry as suggested by Beeching and Detges (2014). Thus, the left periphery connects with previous discourse and thereby creates coherence. This is mainly done by means of connectives (CO), but also by means of the turn-initial categories response markers (RM) and interjections (IJ). At the same time, LP constituents also draw a roadmap for the interpretation of upcoming

discourse. This is done by means of topicality markers (NP) and circumstantial markers (AP and CL) but also by means of all kinds of discourse markers.

The description of the right periphery as the locus for negotiation of meaning (Detges and Waltereit 2014: 42) appears to be totally accurate. Thus, the common ground markers open up for an interactional negotiation, either by asserting the truth of a string (to be confirmed or denied by the interlocutor) or by requesting confirmation (which can, or cannot, be conceded). However, they are interactional in a different way than the response markers of the left periphery: while the latter are response acts, the former are acts of initiative (Briz Gómez and GrupoVal.Es.Co 2003; Linell et al. 1988).

As regards subjectivity signals, it seems to be the case that when they appear in the LP they are meant as an orientation to be observed by the receiver, whereas when they occur in the RP, they are understood as strengthening, or weakening, or otherwise qualifying, the content expressed in the core unit. It should be noted that sentence-initial first person singular pronouns often convey the subjective meaning ‘I hereby express my own opinion’.¹¹

With regard to the so-called dislocated NPs, which occur in both the LP and the RP, it is clear that in the LP they are topical, thus telling the receiver what the upcoming string is about, whereas when occurring in the RP, they serve the purpose of specifying or reminding what the previous string was about (in the case of specifiers) or for highlighting a referent (in the case of reduplicators).¹²

5 Comparisons across the datasets

The present section will address Research questions 2 and 3. RQ 2 concerns cross-linguistic differences between Swedish, French and Spanish and explores how frequently the various types of LP and RP constituents appear in all three languages, whereas RQ 3 addresses similarities and divergences between L1 and L2 performance in French and Spanish as well as the possibility of transfer effects of L1 Swedish on L2 French/Spanish performance.

¹¹ In particular regarding Spanish sentence-initial *yo*, it can be discussed whether it is just a default marker accompanying acts of opinion or perception and thus an expression of subjectivity (e.g., *yo pienso que S* ‘I think that S’; *yo he visto que S* ‘I have seen that S’). Cf. below, Section 5.1.1.

¹² As regards the pragmatic functions of left- vs. right-dislocated *moi* in French, see Detges and Waltereit (2014).

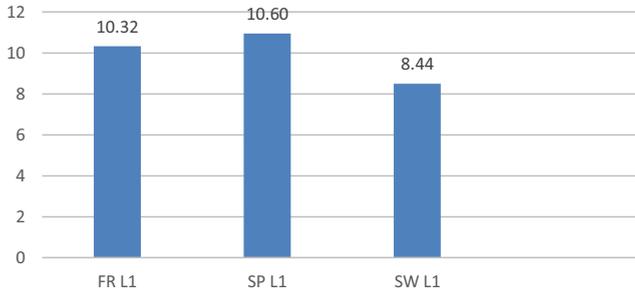


Figure 4: Number of LP constituents per 100 words in the L1 data.

5.1 The left periphery

5.1.1 Comparisons between L1 French, Spanish and Swedish (RQ2)

The total average of LP constituents was compared in the three different languages showing that Swedish L1 speakers produce significantly fewer LP constituents than French and Spanish L1 speakers (Figure 4). Thus, significant differences were shown, by means of an independent samples *t*-test, between French L1 and Swedish L1 ($t[13] = 2.76, p = 0.016^*$) and to an even higher degree between Spanish L1 and Swedish L1 ($t[13] = 3.67, p = 0.003^{***}$)¹³.

The comparatively high LP figures in French and Spanish L1 speakers are not surprising, considering certain earlier research (cf. Morel and Danon Boileau 1998, for French, and Briz Gómez and Grupo Val.Es.Co 2003, for Spanish) and also considering the typological similarity between Spanish and French. The corresponding low figures for Swedish L1 speakers are also in line with our expectations, although little corresponding quantitative or comparative research has been carried out so far.

Apart from this general tendency, certain differences between Swedish L1 compared to French L1 and Spanish L1 with regard to the LP categories can be found, most of which are statistically significant.

- The SW-L1 speakers produce fewer *adverbial phrases* (0.42/100W) in comparison to both the FR-L1 (0.78/100W; $t[13] = -2.49, p = 0.03^*$) and SP-L1 speakers (0.95/100W; $t[12.54] = -4.26, p = 0.001^{***}$).
- The SW-L1 speakers also produce fewer *clauses* (0.30/100W) than the FR-L1 (0.51/100W, $t[12.76] = -2.02, p = 0.06$) and SP-L1 speakers (0.62/100W; $t[12.68] = -2.45, p = 0.03^{**}$).

¹³ See Table 3 for statistical measures on standard deviation.

The fact that fewer constituents of these two denotative categories are used in Swedish accounts for most of the gap between Swedish and the two Romance languages as regards LP constituents. The most plausible explanation to account for this discrepancy would be that the AP and CL categories, which both convey *circumstantial* information, are more likely to be fronted in French and Spanish colloquial talk than in Swedish, which tends to prefer a position inside the core sentence, to the right of the verb (cf. Auer and Lindström 2016).

No salient differences can be observed between the languages regarding the use of discourse (i.e., non-referential, non-denotative) markers, although with one clear exception:

- The SW-L1 speakers produce strikingly fewer *verb-based utterance launchers* (0.12/100W) than the FR-L1 (0.77/100W; $t[9.84] = -5.44, p = 0.001^{**}$) and SP-L1 speakers (0.49/100W; $t[12] = -3.33, p = 0.006^{**}$).

One possible explanation could be that Swedish prefers placing the corresponding information, which to a great extent corresponds to mitigators (such as ‘I guess’) or intensifiers (such as ‘I’m sure’), in the RP (cf. Auer and Lindström 2016).

So far, only differences between FR-L1 and SP-L1 on the one hand, and SW-L1, on the other, have been highlighted, since the typological proximity between the two Romance languages has been taken for granted. However, with regard to three LP constituent categories, a few conspicuous divergences between the two Romance languages can be found in our data:

- The SP-L1 speakers use more than twice as many *pronouns* in the LP as do the FR-L1 speakers (1.62 vs. 0.66 instances per 100W), and three times as many *NP* constituents (2.05 vs. 0.65/100W).

Some tentative explanations for these differences are called for, even though they fall outside the scope of the present study. As regards the frequent Spanish use of initial pronouns, the first person singular *yo* is strikingly overrepresented in the SP data; it appears far more frequently than its pro-drop alternative, and it widely overshadows the seemingly analogous use of *moi je* in the FR data (cf. Detges and Waltreit 2014). It could on these grounds be questioned whether Sp. *yo* should actually count as a true LP constituent rather than as an element on the verge of developing into a pronominal clitic, like French *je*.

As for the divergence between Spanish and French in the number of fronted NPs, one explanation that deserves being further investigated is whether Spanish makes use of an initial NP in cases where French would prefer a cleft *c’est NP que/qui VP* ‘it’s NP that VP’ construction, which places the highlighted NP (focus or topic constituent) in the core clause. This structural difference could at least partly account for the observed gap.

Table 5: Types of sequences per 100 words in the LP.

	FR L1	SP L1	SW L1	FR L2	SP L2
1-Constituent sequences/100W	4.01	4.58	5.00	4.17	4.70
2-Constituent sequences/100W	1.80	1.80	1.30	1.72	2.00
3-Constituent sequences/100W	0.49	0.63	0.25	0.71	0.70
4+ Constituent sequences/100W	0.33	0.19	0.03	0.22	0.20
Total sequences/100W:	6.63	7.20	6.58	6.82	7.60

There are other cross-linguistic divergences where the difference between French and Spanish is conspicuous. Thus, an intriguing difference was found in the production of *response markers* (RM):

- The FR-NSs produce twice as many instances of response markers as do the SW-NSs (2.04 vs. 1.02 instances per 100 words), who in turn produce twice as many as the SP-NSs (0.53 instances per 100 words). We can find no satisfactory explanation that would account for these differences, which would seemingly reflect a varying degree of “interactive involvement”, although that interpretation could hardly find support through other results in our study.

Finally, there is an interesting difference observed between all three languages concerning the *complexity* of the LP (for a recent overview of studies of variation in complexity in L1/L2 see Kuiken et al. 2019).

- The FR-L1 speakers produce more sequences with *four or more constituents* than any of the other observed groups. Interestingly, although the Spanish L1 speakers produce a somewhat larger number of LP units overall (SP-L1, 11.7 units/100W, compared to 10.6 units/100W for FR-L1), when it comes to highly elaborated LPs with four constituents or more, the French L1 figure is significantly higher (Table 5). Thus, the French L1 speakers produce on average 0.33 instances/100W, i.e., almost twice the number of the Spanish L1 group (0.19 instances/100W, $t[18] = 2.07, p = 0.05^*$), and 11 times more than the Swedish L1 speakers (0.03/100W, $t[13] = 3.54, p = 0.004^*$). No other than speculative answers could be given to the question why French native speakers tend to produce such long preambles. Although one factor could be cultural expectations of a high competitiveness in turn-taking, this remains an open question.

5.1.2 Comparisons between L1 and L2 speakers of French and of Spanish (RQ3)

The general answer to RQ3 would be to say that a very high degree of alignment to native use can be observed in the L2 speakers of our data. Seemingly, the advanced non-native speakers behave very much in a target-like way, and display virtually no influence from their Swedish L1.

However, two interesting differences can be observed between the L1 and L2 use in our data:

- Firstly, both groups of L2 speakers tend to overuse *pronouns* in comparison with native speakers. Thus, the SP-L2 speakers use 2.48 pronouns/100W in comparison to 2.05 pronouns /100W in SP-L1 speakers ($t[18] = -1.45, p = 0.16$). The same comparison in the French data yields 0.95/100W for FR-L2 vs. 0.66/100W for FR-L1 ($t[27] = -1.99, p = 0.06$). This overuse cannot be explained in terms of transfer since the corresponding level in the SW-L1 group¹⁴ is even lower (0.62 pronouns/100W). The overuse is more likely to be explained as a phenomenon due to independent interlanguage mechanisms. (For an interesting synthesis of overuse/underuse of pronouns and NPs in different groups of L2 users, see Chini 2005).
- The second difference between L1 and L2 use concerns the French data. As was observed above in Section 5.1.1, a strikingly high proportion of LP sequences with more than four constituents was found among the French natives (0.33/100W). This high level was not attained by the FR-L2 speakers (0.22/100W; see Table 5), let alone by the any of the SP and SW groups.

RQ 3 also aims at investigating whether there is any trace of L1 transfer in the L2 speakers' performance. Here, our data shows that French and Spanish L2 speakers exhibit similar patterns to those of French and Spanish L1 speakers and quite different patterns to those of Swedish L1 speakers. In fact, no clear instances of transfer from L1 Swedish could be detected in our data.

¹⁴ The same tendency to overuse personal pronouns by French L2 speakers has previously been observed by Engel (2010), who shows how French non-native speakers use pronominal dislocations in a non-targetlike way, in order, for example, to introduce new topics (Engel 2010: 20, 43).

5.2 The right periphery

5.2.1 Comparisons between L1 French, Spanish and Swedish (RQ2)

The total average of the RP constituents in the three datasets was compared, showing a significantly higher number of RP constituents being used by the Swedish L1 speakers than by the French and Spanish L1 speakers:

- As seen in Figure 5, the SW-L1 speakers produce 4.38 RP constituents /100W, as compared to 1.99 for the FR-L1 speakers ($t[88] = 2.58, p = 0.01^*$), and 1.34 for the SP-L1 speakers ($t[88] = 3.49, p = 0.001^{**}$)¹⁵.
- The comparison between the Swedish L1, French L1 and Spanish L1 data as regards the specific RP categories yields a mixed picture. As regards the referential RP categories (i.e., specifiers and reduplicators), differences between the groups are negligible. As for the non-referential categories, however, there are clear divergences between the Swedish L1 data, on the one hand, and the French and Spanish L1 data, on the other. For the following differences we can find, as yet, no clear explanation (although they could certainly generate new research questions):
- The SW-L1 speakers use more *common ground markers* (1.9 instances/100W) than both the FR-L1 (0.8 instances/100W) and the SP-L1 speakers (0.5 instances/100W). The difference is statistically significant for the SW-L1/SP-L1 pairing ($t[13] = 2.55, p = 0.02^*$) though not for the SW-L1/FR-L1 pairing ($t[13] = 1.99, p = 0.07$).
- The SW-L1 speakers use more *mitigators* (1.23 instances/100W) than both the FR-L1 (0.27 instances/100W; $t[13] = 4.41, p = 0.001^{***}$) and the SP-L1 speakers (0.37 instances/100W; $t[1] = 2.96, p = 0.01^*$).
- The SW-L1 speakers use more *intensifiers* (0.5 instances/100W) than the SP-L1 speakers (0.1 instances/100W) (this difference is not, however, statistically significant: $t[13] = 1.93, p = 0.08$). Here, the FR-L1 speakers produce roughly the same number of instances (0.6/100W) as do the SW-L1 speakers ($t[13] = -0.41, p = 0.68$).
- The SW-L1 speakers use more *additive markers* (0.27 instances/100W) than the FR-L1 speakers (0.11 instances/100W) and clearly more than the SP-L1 speakers (only 0.06 instances/100W). Again, the difference proves statistically significant only for the SW-L1/SP-L1 pairing ($t[13] = 2.97, p = 0.01^{**}$).

¹⁵ See Table 4 for statistical measures of standard deviation.

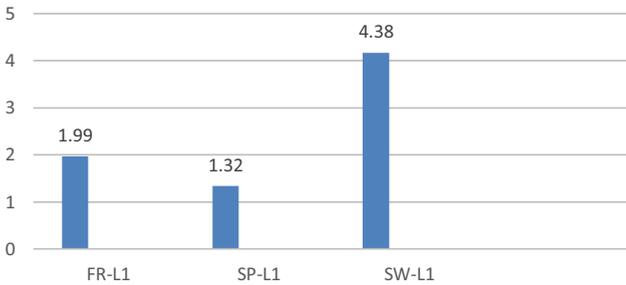


Figure 5: Number of RP constituents per 100 words in the L1 data.

5.2.2 Comparisons between L1 and L2 speakers of French and Spanish (RQ3)

As concerns the third research question, our analysis showed few clear discrepancies between the L1 and L2 groups when it comes to the use of RP constituents. Thus, the general conclusion is the same as for LP, namely that a high degree of alignment to native use can be observed. This goes both for the referential categories (specifiers and reduplicators) and for the non-referential ones (discourse marker categories).

However, two significant discrepancies between L1 and L2 speakers can be found in the French data. One regards the category of *common ground markers*, for which a higher level is encountered in the L1 than in the L2 group (0.77/100W vs. 0.38/100W; $t[38] = -2.17, p = 0.04^*$). This difference cannot be attributable to any transfer from L1 Swedish, considering the fact that the SW-L1 speakers produce a much larger number of CGMs (1.9 instances/ 100W).

The second discrepancy runs in the opposite direction: the L2 French speakers produce a significantly higher number of *mitigators* than the L1 French speakers (0.47/100W for the L2 group, as compared to 0.27/100W for the natives ($t[18] = 2.56, p = 0.02^*$). Here, a transfer effect from the speakers' L1 Swedish could theoretically account for the divergence. It should be noted, however, that no corresponding difference can be found between the Spanish L1 and L2 groups.

In conclusion, the French and Spanish L2 speakers produce significantly fewer RP constituents than the SW-L1 speakers, which is to say that the Swedish L1 tendency towards right-peripherization has virtually no impact on our groups of highly proficient L2 users.

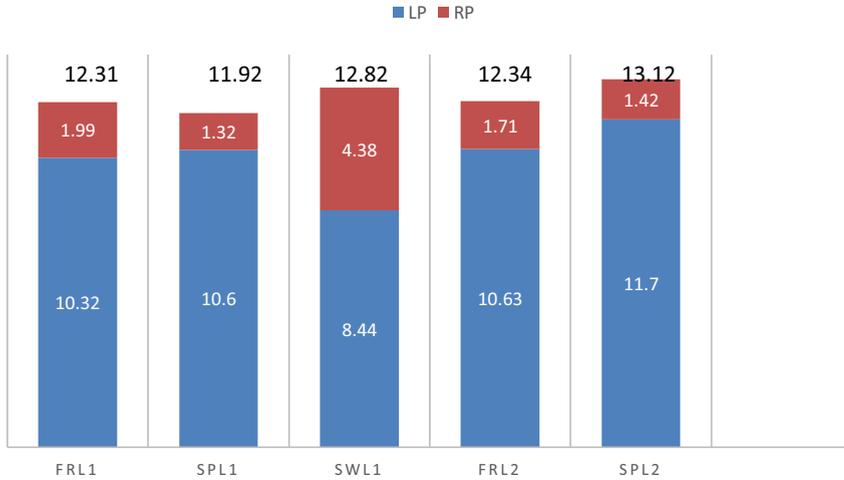


Figure 6: Number of RP + LP constituents per 100 words in the five participant groups.

5.3 The left versus the right periphery across the five participant groups

Our results seemingly reflect a typological difference between languages that tend towards left-peripherization, such as Spanish and French, and those which tend to peripherize to the right, such as Swedish. However, this conclusion needs to be qualified. Thus, Swedish L1 speakers right-peripherize more in comparison to French and Spanish L1 speakers than French and Spanish L1 speakers left-peripherize in comparison to Swedish L1 speakers. More precisely, Swedish L1 speakers produce about 80% of the LP constituents produced by French or by Spanish L1 speakers, whereas the two latter groups produce only about one third of the RP constituents produced by the Swedes.

A more intricate question concerns what kind of information is placed in the left vs. the right periphery in the two Romance languages, as compared to Swedish. Could it, for instance, be the case that much of the information typically occurring in the LP in French and Spanish would tend to appear in the RP in Swedish? This seems to be the case as regards the category of verb-based utterance launcher category (VL) in the Romance languages, which contains a kind of information that speakers of colloquial Swedish supposedly would prefer placing on the right hand side of the core sentence (Fr. *je pense que [ce serait inutile]*, Sp. *creo que [sería inútil]*, Sw. [*det vore lönlöst*] *tror jag* ‘I think it would be pointless’ vs. ‘it would be pointless I think’).

Part of the explanation of why French and Spanish left-peripherize more than Swedish could consist in different preferences in discourse organization, such as mentioning the circumstances before the fact, in the case of Romance languages (cf. Morel and Danon-Boileau 1998: 37), or the opposite tendency in Swedish (cf. Auer and Lindström 2016: 70).

Also, we have so far neglected the question of placement in the core unit: how much and what kind of information will be placed inside the core unit vs. in the sentence peripheries? This is another research question which would need to be investigated in order to reach a deeper understanding of the LP/RP and the Romance/Swedish asymmetries.

A final point: interestingly, if we add the number of RP constituents to that of LP constituents in each participant group, thus obtaining a measure of the total size of peripherally placed constituents, we can see that the figures for each group resemble each other fairly closely (Figure 6).

This observation in fact lends support to the idea of (a certain degree of) interchangeability between LP and RP functions. This is a question that deserves further investigation.

6 Summary

As regards what types of LP and RP constituents would qualify as relevant categories, a taxonomy was established on the basis of the 110,759 words that make up our corpus. The taxonomy afforded an exhaustive tagging of the altogether 6474 peripheral constituents (5169 LP and 1305 RP constituents) encountered in the data.

Concerning our RQ1 (Sections 1.2, 4.3 and 5.3) – in which ways LP categories differ systematically from RP categories –, most claims made by Beeching and Detges (2014) were confirmed. Both subjectivity and intersubjectivity marking occurred in both peripheries, albeit with different functions, the LP connecting with previous discourse and constituting a road map for the interpretations of upcoming discourse, while the RP basically constitutes the locus for meaning qualification and negotiation. However, the fact that the sum of LP and RP constituents proved very similar across the five participant groups, in spite of clear differences as regards the proportions of each type, lends support to the idea of partial interchangeability between LP and RP functions.

Our RQ2 had a cross-linguistic scope, addressing commonalities and divergences between French, Swedish and Spanish as native tongues (Sections 1.2, 5.1.1 and 5.2.1)

The main findings were:

- There are important differences in periphery structure between both Romance languages, on the one hand, and Swedish, on the other.
- Swedish L1 speakers right-peripherize to a remarkably higher extent than French and Spanish L1 speakers, whereas Swedish L1 speakers left-peripherize clearly less than French and Spanish L1 speakers. We attribute these divergences to different discourse-organizational preferences (Sections 5.1.1 and 5.2.1).
- The French L1 speakers produced a higher proportion of very long LP sequences (including four or more constituents) than any other speaker group including the L2 groups (Section 5.1.1). We attribute this divergence to discourse-organizational preferences specific to French (Bartning et al. 2012; Morel and Danon-Boileau 1998), possibly in connection with culture-specific turn-taking patterns.
- The Swedish L1 speakers used significantly fewer verb-based utterance launchers (VL), adverbial phrases (AP) and clauses (CL) than the other L1 groups. We attribute this divergence mainly to different syntactic-structural preferences (Section 5.1.1).
- Not only L1 but also L2 users of French produced surprisingly more response markers (RM) than the other speaker groups (Sections 5.2.2 and 5.3). This divergence is most likely attributable to different conversation-organizational preferences, which in turn may be culturally grounded.
- The Swedish L1 speakers use considerably more common ground markers (CGM) and mitigators (MIT) than any other speaker group (Sections 5.2.1 and 5.3). This may be due to discourse- or conversation-organizational preferences, which in turn may be culturally grounded.
- Between French and Spanish periphery structure there are also a few far from negligible differences. Thus the French L1 speakers produced almost four times as many verb-based utterance launchers as the Spanish L1 speakers, and these, in turn, produced about three times as many pronouns and noun phrases as the French L1 speakers (Section 5.1.1). The latter divergences may partly be due to structural factors, whereas the former seem rather to depend on stylistic preferences.

By RQ3, which could be said to constitute the ultimate aim of this study, we wanted to find out about possible systematic differences between native speakers and highly proficient L2 speakers of French and Spanish in the domain of syntactic peripheries and, in particular, whether any detectable divergence from L1 use could be attributable to transfer from the L2 speakers' Swedish L1 (Section 5.3). Our main findings were the following:

- Highly proficient French and Spanish L2 users having Swedish as their L1 show a high or almost total alignment with native use as far as peripheral structures are concerned. This finding is in line with those of Donaldson (2011) and Engel (2010).
- Highly proficient French and Spanish L2 users having Swedish as their L1 behaved very differently from Swedish L1 users. In contradistinction to findings by Conway (2005), hardly any transfer effects from the L1 were to be found in our data.
- However, there were two clear-cut divergences encountered between L1 and L2 use in French. Firstly, L2 speakers produced significantly more mitigators (MIT) than native speakers, which could in fact be regarded as an L1 transfer effect. Secondly, L2 speakers produced significantly fewer common ground markers (CGM) than native speakers, for which we see at present no clear explanation. Neither of these divergences was to be found between L1 and L2 speakers of Spanish.
- Finally, an interesting divergence in L2 use could be detected: in the LP, both the French and Spanish L2 speakers overused pronouns in comparison to native use. We have taken this to be an interlanguage feature, not transferable from the speakers' L1.

References

- Auer, Peter & Jan Lindström. 2016. Left/right asymmetries and the grammar of pre- vs. postpositioning in German and Swedish talk-in-interaction. *Language Sciences* 56. 68–92.
- Bartning, Inge. 2012. High level proficiency in second language use – morphosyntax and discourse. In Marzena Watorek, Sandra Benazzo & Maya Hickman (eds.), *Comparative perspectives in language acquisition. A tribute to Clive Perdue*, 170–187. Bristol: Multilingual Matters.
- Bartning, Inge, Fanny Forsberg Lundell & Victorine Hancock. 2012. On the role of contextual factors for morpho-syntactic stabilization in high level L2 French. *Studies in Second Language Acquisition* 34(2). 243–267.
- Beeching, Kate, Liesbeth Degand, Ulrich Detges, Elisabeth Closs Traugott & Richard WALTEREIT. 2009. *Summary of the workshop on meaning in diachrony at the conference on meaning in interaction*. Bristol: University of the West of England.
- Beeching, Kate & Ulrich Detges. 2014a. Introduction. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery. Crosslinguistic investigations of language use and language change*, 1–23. Leiden: Brill.
- Beeching, Kate & Ulrich Detges (eds.). 2014. *Discourse functions at the left and right periphery. Crosslinguistic investigations of language use and language change*. Leiden: Brill.
- Borreguero Zuloaga, Margarita. 2014. Chapter 12. Left periphery in discourse. Frame units and discourse markers. In Andreas Dufter & Álvaro S. Octavio de Toledo y Huerta (eds.), *Left*

- sentence peripheries in Spanish: Diachronic, variationist and comparative perspectives*, 383–418. Amsterdam/Philadelphia: John Benjamins.
- Briz Gómez, Antonio & Grupo Val.Es.Co. 2003. Un sistema de unidades para el estudio del lenguaje coloquial. *Oralia* 6. 7–61.
- Cinque, Guglielmo & Luigi Rizzi. 2008. The cartography of syntactic structures. *Studies in Linguistics* 2. 42–58.
- Chini, Marina. 2005. Reference to person in learner discourse. In Henriette Hendriks (ed.), *The structure of learner varieties*, 65–110. Berlin: De Gruyter.
- Conway, Åsa. 2005. *Le paragraphe oral en français L1, en suédois L1 et en français L2. Étude syntaxique, prosodique et discursive (Études romanes de Lund 73)*. Lund: Lund University.
- Detges, Ulrich & Richard Waltereit. 2014. ‘Moi je ne sais pas’ vs ‘je ne sais pas moi’: French Disjoint Pronouns in the left vs right periphery. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery. Crosslinguistic investigations of language use and language change*, 24–46. Leiden: Brill.
- Demonte, Victoria & Olga Fernández Soriano. 2007. La periferia izquierda oracional y los complementantes del español. In Juan Cuartero Otal & Martina Emsel (eds.), *Vernetzungen: Bedeutung in Wort, Satz und Text*, 1–13. Frankfurt: Peter Lang.
- Donaldson, Bryan. 2011. Left dislocation in near-native French. *Studies in Second Language Acquisition* 33. 399–432.
- Elordieta, Gorka, Sónia Frota & Marina Vigário. 2005. Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica* 59(2/3). 110–143.
- Engel, Hugues. 2010. *Dislocation et référence aux entités en français L2 (Cahiers de la recherche 43)*. Stockholm: Stockholm University.
- Estellés Arguedas, María & Salvador Pons Bordería. 2014. Absolute initial position. In Salvador Pons Bordería (ed.), *Discourse segmentation in Romance languages*, 121–155. Amsterdam/ Philadelphia: John Benjamins.
- Fant, Lars. 2007. Gestión de la intersubjetividad en una conversación chilena de carácter confrontativo. In Martina Schrader-Kniffki & Laura Morgenthaler García (eds.), *La Romania en Interacción: Entre historia, contacto y política. Ensayos en homenaje a Klaus Zimmermann*, 215–234. Frankfurt am Main/Madrid: Vervuert/ Iberoamericana.
- Fant, Lars. 2011. Managing intersubjectivity and argumentation in problem-solving. In Nuria Lorenzo Dus (ed.), *Spanish at work*, 172–191. London: Palgrave MacMillan.
- Fant, Lars. 2016. Pragmatic markers in high-level second language use. In Kenneth Hyltenstam (ed.), *Advanced proficiency and exceptional ability in second languages*, 15–42. Berlin/New York: De Gruyter Mouton.
- Fant, Lars. 2019. Las periferias sintácticas en el habla coloquial de español L1 y L2 y de sueco L1: un estudio tipológico y adquisicional. *Normas Revista de estudios lingüísticos hispánicos* 9(1). 85–100.
- Hancock, Victorine. 2007. Quelques éléments modaux dissociés dans le paragraphe oral dans des interviews en français L2 et L1. *Journal of French Language Studies* 17. 21–47.
- Hyltenstam, Kenneth, Inge Bartning & Lars Fant (eds.). 2018. *High level proficiency in second languages and multilingual contexts*. Cambridge: CUP.
- Kotschi, Thomas. 1996. Procedimientos de producción y estructura informacional en el lenguaje hablado. In Thomas Kotschi, Wolf Oesterreicher & Klaus Zimmermann (eds.), *El español hablado y la cultura oral en España e Hispanoamérica*, 185–206. Frankfurt am Main: Vervuert/ Madrid: Iberoamericana.

- Kuiken, Folkert, Ineke Vedder, Alex Housen & Bastien De Clerq. 2019. Variation in syntactic complexity: Introduction. *International Journal of Applied Linguistics* 29(2) [Special issue on syntactic complexity]. 161–170.
- Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: CUP.
- Larsson Ringqvist, Eva. 2010. La dislocation en français et en suédois: Aspects contrastifs et acquisitionnels. *Association sans-public* 2010/1(13–14). 69–80.
- Linell, Per, Lennart Gustavsson & Päivi Juvonen. 1988. Interactional dominance in dyadic communication: A presentation of initiative-response analysis. *Linguistics* 26. 415–442.
- Morel, Mary-Annick & Laurent Danon Boileau. 1998. *Grammaire de l'intonation. L'exemple du français*. Paris: Ophrys.
- Pons Bordería, Salvador. 2018. Paths of grammaticalization: Beyond the LP/RP debate. In Salvador Pons Bordería & Óscar Loureda Lamas (eds.), *Beyond grammaticalization and discourse markers. New issues in the study of language change (Studies in pragmatics 18)*, 334–383. Leiden: Brill.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.
- Salameh Jiménez, Shima, María Estellés Arguedas & Salvador Pons Bordería. 2018. Beyond the notion of periphery: An account of polyfunctional discourse markers within the Val.Es.Co. model of discourse segmentation. In Kate Beeching, Chiara Ghezzi & Piera Molinelli (eds.), *Positioning the self and others: Linguistic perspectives*, 105–125. Amsterdam/Philadelphia: John Benjamins.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Lausanne: Payot.
- Sedano, Mercedes. 2012. Dislocación a la izquierda y a la derecha: Semejanzas y diferencias. *Homenaxes* 58, Departamento de lingua española, servizo de publicacións e intercambio científico da Universidade de Santiago de Compostela. 791–805.
- Sorace, Antonella. 2003. Near-nativeness. In Catherine Doughty & Michael Long (eds.), *Handbook of second language acquisition*, 130–151. Oxford: Blackwell.
- Traugott, Elisabeth Closs. 2012. Intersubjectification and clause periphery. *English Text Construction* 5(1). 7–28.
- Traugott, Elisabeth Closs. 2014. On the function of the epistemic adverbs Surely and No doubt at the left and right peripheries of the clause. In Kate Beeching & Ulrich Detges (eds.), *Discourse functions at the left and right periphery*, 72–91. Leiden/Boston: Brill.