

Stressed Clitic Pronouns in Two Spanish Varieties: A perception study*

Laura Colantoni

University of Toronto
laura.colantoni@utoronto.ca

José-Ignacio Hualde
Ane Icardo Isasa

University of Illinois at Urbana-Champaign
jihualde@illinois.edu
icardo2@illinois.edu



Received: December 10, 2018

Accepted: June 6, 2019

Abstract

Clitic elements are prosodically attached to another phonological word. As such, they do not carry their own stress. In general, Spanish non-prepositional object pronouns (e.g. *me/te/le/lo/se*) meet the definition of clitic, whether preposed (proclitic) or postposed (enclitic) to the verb. Nevertheless, in several Romance varieties, including some Spanish varieties, the stress may surface on the clitic in verb+clitic(s) groups. We refer to this phenomenon as Enclitic Stress Shift (ESS). In this paper, we report on the perception of ESS by Spanish speakers from Argentina and Spain and explore the acoustic dimensions that may affect this perception. The discussion includes comparisons with other Romance languages and implications for sound change, such as the changes in stress patterns in the development of future and conditional tenses.

Keywords: stress perception; clitics; secondary stress; Spanish phonology; prominence; Argentine Spanish; Peninsular Spanish

Resum. *Clítics pronominals accentuats en dues varietats del castellà: un estudi de percepció*

Els elements clítics s'adjunten prosòdicament a una altra paraula fonològica, per tant, no porten accent propi. En general, els pronoms no preposicionals amb funció d'objecte (e.g. *me/te/le/lo/se*) compleixen la definició de clíctic, tant si van davant (proclíctic) com darrera (enclíctic) del verb. Però en algunes varietats romàniques, incloses algunes de l'espanyol, l'accent pot aparèixer sobre el clíctic del grup verb+clíctic(s). Ens referim a aquest fenomen com a desplaçament de l'accent enclíctic (ESS, en anglès). En aquest article, presentem un estudi de percepció entre parlants d'espanyol de l'Argentina i d'Espanya i examinem les dimensions acústiques que poden afectar aquesta percepció. La discussió inclou comparacions amb altres llengües romàniques i implicacions en

* We are thankful to Maria-Rosa Lloret and to a second, anonymous, reviewer for their very useful comments and observations. We also want to thank Marissa Barlaz for statistical advice and Vanina Machado Araujo for her assistance in testing the Argentine Spanish speakers.

relació amb el canvi fonètic històric, com el que presenten els patrons accentuals en l'evolució de les formes verbals de futur i condicional.

Paraules clau: percepció de l'accent; clítics; accent secundari; fonologia de l'espanyol; prominença; espanyol de l'Argentina; espanyol peninsular

Table of Contents

1. Introduction	References
2. Perception experiment	Appendix 1
3. Discussion	Appendix 2. List of stimuli

1. Introduction

Although clitics are by definition unstressed elements that prosodically attach to a host word, many studies have claimed that in verb+clitic sequences, as those illustrated in (1), the stress may shift to the clitic in a process that we call *Enclitic stress shift* (ESS):

- (1) *Dáme* 'give' – *dàmeló*¹ 'give it to me'
Dár 'to give' – *dárle* 'to give her/him' – *dàrselo* 'to give it to her/him'
Dándo 'giving' – *dàndolé* 'giving to her/him'

As we will discuss below, this process has been reported in a variety of Romance languages, including some Spanish varieties. In Argentine Spanish, ESS appears to be more frequent than its alternative (stress on the verbal host) and acoustic evidence suggests that the clitic bears primary rather than secondary stress (Colantoni & Cuervo 2013; Klassen & Patience 2016). It is not clear, though, to what extent this phenomenon is observed in other Spanish-speaking varieties and whether speakers of the different varieties are sensitive to ESS. Thus, our goal here is to investigate the perception of verb+clitic sequences by two groups of participants who represent the Argentine and Peninsular varieties. The latter variety may be particularly relevant, since incipient evidence (Colantoni, Cuervo & Hualde 2010) suggests that verb+clitic sequences are sometimes doubly stressed in this dialect as well.

In the remainder of this paper, we first describe the realization of verb+clitic sequences in Romance (section 1.1) including a discussion of whether the main prominence is on the verb, the clitic or both (section 1.2). In section 2, we report on a perception study, whose results are discussed in section 3, where we draw some parallels between our Spanish data and those reported in other varieties, and

1. Throughout this paper, we will use an acute accent mark to indicate (perceived) primary stress and a grave accent mark to indicate secondary stress. For clarity, we add accent marks even when they are not required orthographically.

suggest that the stress shift observed in clitic-bearing verbal forms resembles the diachronic evolution of future and conditional tenses in Romance.

1.1. Clitics in Spanish and Romance

Clitics are syntactic elements that prosodically attach to another phonological word. In Spanish direct and indirect object pronouns are prosodically attached to the verb. Clitic pronouns may appear either before the verb (as proclitic), as in (2a), or after the verb (enclitic), as in (2b). Proclisis is the rule except in affirmative imperatives, infinitives and gerunds. In periphrastic verbal forms containing a conjugated verb + infinitive or gerund (2c), the clitic can appear either before the conjugated verb or after the infinitive.²

(2a) *Me lo dan* ‘they are giving it to me’

(2b) *Dame* ‘give me’, *dámelo* ‘give it to me’
(but negative imperative *no me lo des* ‘don’t give it to me’)

Darme ‘to give me’, *dármelo* ‘to give it to me’

Dándome ‘giving me’, *dándomelo* ‘giving it to me’

(2c) *Me lo quieren dar* ~ *quieren dármelo* ‘they want to give it to me’

Me lo van a dar ~ *van a dármelo* ‘they are going to give it to me’

Me lo están dando ~ *están dándomelo* ‘they are giving it to me’

Whereas, in general, the location of the primary stress in Spanish is restricted to one of the last three syllables of the word (e.g., *camaleón* ‘chameleon’, *cucaracha* ‘cockroach’, *libélula* ‘dragonfly’, but not **libelula*, **cúcaracha*, with stress beyond the three-syllable window), verb+enclitic forms with two enclitics may commonly show stress four syllables from the end (*explicándomelo* ‘explaining it to me’, *explicámelo* ‘explain it to me!’). In the rare cases where a verb carries three enclitic pronouns, the stressed syllable may even be located five syllables from the end, e.g. *las polillas se me están comiendo los libros* ‘moths are eating up my books’ → *están comiéndoselos* ‘they are eating them up on me’. Enclitics are thus invisible to stress assignment on the verb and do not affect the location of the primary stress: *comiéndoselos*, except for the ESS phenomenon that we study here³ (but see below).

2. In addition to direct and indirect object clitics, the impersonal subject pronoun *se*, as in *Se come muy bien en este restaurant* ‘one eats very well in this restaurant’ can also be enclitic, e.g. *Sigue comiéndose bien en ese restaurant* ‘one still eats well in this restaurant’.
3. All Spanish verbs are stressed following the same pattern. Unlike in Italian, Spanish verbs do not contrast in the position of stress. The verb forms that may take enclitic pronouns are either paroxytonic (the gerund, e.g. *comiéndoselos* ‘eating’ and *tú* and *usted(es)* imperatives, e.g. *cóme (tú)*, *cóma (usted)*, *cóman (ustedes)* ‘eat!’) or oxytonic (the infinitive, e.g. *comér* ‘to eat’ and *vos* and *vosotros/as* imperatives, e.g. *comé (vos)*, *coméd (vosotros/as)* ‘eat’). *Usted* ‘you-sg formal’ and *ustedes* ‘you-pl’ forms are used in all Spanish dialects. *Vosotros/as* ‘you-pl familiar’ is only used in Spain. As a second person familiar form, some dialects, including Spain, use *tú* and other dialects, including Argentina, use *vos*. See Appendix 1.

Clitics are words that cannot stand on their own prosodically, but must “lean” on a host. This implies lack of stress. It has thus been observed that clitic pronouns in Spanish cannot receive stress “even if contrastive focus is intended” (Hualde 2005: 258). Thus, an English sentence like *I already told Mary and now I am going to tell YOU what I think* could be translated in Spanish as *Ya se lo he dicho a María y ahora te voy a decir a TI lo que pienso*, but not as **Ya se lo he dicho a María y ahora TE voy a decir lo que pienso*. Nevertheless, at least in some varieties, enclitics may in fact bear stress under certain conditions (Fernández Soriano 1999: 1256).

Both stress displacement to the clitic and stress displacement to another syllable in the verbal root triggered by the presence of a clitic have in fact been reported in a wide variety of Romance dialects, including some Spanish dialects. Ordóñez & Repetti (2006: 170) summarize five different patterns in verbs bearing enclitics in Romance varieties, which range from stress stability to systematic stress shift to the last syllable. The three intermediate patterns reported by these authors include generalized penultimate stress shift, two-clitic penultimate stress shift (i.e., the stress shifts to the first clitic in a cluster) and mixed penultimate stress shift (i.e., the stress variably shifts to the penultimate syllable). Final stress shift, which is of interest here, given the patterns that have been reported in Argentine Spanish, has been observed in some varieties of Sardinian, as illustrated in (3), as well as in Gascon and Majorcan Catalan.

- (3) Final stress shift in some varieties of Sardinian (Ordóñez & Repetti 2006: 170)
- Narra*. ‘say’
Narramí ‘tell me’
Narralá ‘say it’
Narramilá ‘say it to me’

Variation in stress patterns in verb+enclitic forms has been reported not only across languages but even among dialects of the same language. This is the situation, for instance, in Catalan (Torres Tamarit & Pons Moll 2018), where stress stability on the verb has been reported for Central Catalan, a tendency to shift the stress to the penultimate syllable has been described for Formentera Catalan and a strong tendency to shift to the final syllable is found in Minorcan and Majorcan Catalan (see also Nadeu et al. 2017).

Whereas in Spanish, in general, enclitics are taken to be invisible for stress assignment purposes (e.g., using the examples above in (3): *narra*, *narrame*, *narralla*, *narramela*), stress shift in verbs with enclitics has been reported in Argentine Spanish (Moyna 1999; Huidobro 2005; Colantoni & Cuervo 2013; Klassen & Patience 2016). In this variety, shift to the final clitic is available whenever there is at least one unstressed syllable between the stressed syllable of the verb and the clitic as in (4a).⁴ Sequences of two adjacent stressed syllables as in (4b) are

4. As an anonymous reviewer points out, it is interesting to note that, when ESS takes place, the stress always shifts to the final syllable, even though the unmarked pattern in Spanish is penultimate stress.

ungrammatical. It is important to highlight, however, that even when ESS is available, it does not categorically happen. Colantoni & Cuervo (2013) report that ESS occurred in 71% of the tokens in their corpus where the phonological conditions for the process obtained. The authors identify some morphological and phonological factors as favoring ESS. For example, ESS is more frequent in imperatives than in infinitives and gerunds, and when the target word is followed by a pause than when followed by another word.

- (4a) *Encontrándolo*. ‘finding it’
Dàmeló ‘give it to me’
Dàndomeló ‘giving it to me’
- (4b) **Encontràrló* ‘to find it’
 **Dàmé* ‘give me’

A question that remains to be explored is whether the likelihood of ESS also depends on the number of syllables counting from the stressed syllable of the verb to the end of the word; e.g. is stress shift more likely in *explicándomelo* ‘explaining it to me’ than in *explicándolo*?

Further evidence of ESS in Argentine Spanish is its pervasive documentation in literary works and comics. In written texts, it is not uncommon to see imperatives with an orthographic accent on the clitic, as illustrated in these examples taken from *La novela de Perón* by Tomás Eloy Martínez (1991): [...] *le dije “Quedesé tranquila porque...”* ‘I told her: stay calm, because...’ (p.33); *Esfuercesé* ‘make an effort’ (p. 52); *imaginesé* ‘imagine’(p. 44). (The standard forms are *quédese, esfuércese, imagínese*).

Very little is known about how widespread ESS is in other Spanish varieties. For Peninsular Spanish, Menéndez Pidal (1918: 255) documents its existence in 16th century poetry, as well as in contemporary speech. He gives examples of emphatic imperatives with two stresses: “*levántaté, entiéndemé*” (Menéndez Pidal 1918: 255), although not of infinitives or gerunds.⁵ Colantoni, Cuervo & Hualde (2010), in unpublished work, conducted a small pilot study and reported that ESS occurs in Peninsular Spanish, albeit less frequently than in Argentine Spanish (18% of possible instances in their Peninsular corpus, as perceived by the authors). The acoustic analysis of the data tentatively confirmed Menéndez Pidal’s intuition that these forms may receive two stresses, as in (5a). The authors also noticed that ESS appears to be restricted to phrase final position, as shown in (5b).

- (5a) *siéntatè ~ sièntatè* ‘sit down’
- (5b) *vàmonós* ‘let’s go’; *escùchamé* ‘listen to me’, but **vàmonós yá* ‘let’s go now’;
 **escùchamé tú* ‘listen to me’

5. Gili Gaya (1943) makes a similar comment regarding the double accentuation of these imperative forms.

Doubly stressed verb+clitic sequences have not been reported in other studies on Spanish, except for Peperkamp (1997), see below. Potentially double stressing constitutes a logical step in the change from stress on the verb to stress on the clitic (e.g. *vámonos* > *vámonós* > *vamonós*). The phenomenon may also be related to an emphatic intonational contour with a low tone on the verb and a prominent high accent on the last stressed syllable as one may find in utterances like *vàmonos yá* ‘let’s go now’ or *dàme el libro* ‘give me the book’.

1.2. Prominence and cliticization

The fact that doubly stressed verb+clitic sequences have been reported leads us to wonder whether ESS is better understood as a case of stress shift or as an instance of secondary stress. A variety of analyses and interpretations have in fact been proposed. Based on auditory transcriptions and comparisons with stress patterns in compounds, Moyna (1999) suggests that the case of Argentine Spanish represents an instance of secondary stress on the clitic, where the primary stress remains on the host.

Colantoni & Cuervo (2013) and Klassen & Patience (2016) conducted acoustic studies of a corpus of Argentine TV soap operas and, based on the results obtained for three acoustic parameters (duration, intensity and f0), concluded that the primary stress is on the clitic in cases of ESS. Colantoni & Cuervo (2013) based their conclusions on patterns of perceived prominence; namely, they observed that in those cases in which the stress was perceived on the clitic, the vowel of the clitic was longer in duration and higher in intensity. To avoid this shortcoming and to determine whether there was evidence that clitics behaved in the same way as inflectional affixes, Klassen and Patience (2016) compared stressed clitics to unstressed clitics and to stressed words (e.g. *mandameló* ‘send it to me’; vs. *mandámelo* vs. *mandámelo vos*) as well as to stressed and unstressed affixes. They largely replicated the results of Colantoni & Cuervo (2013) as concerns the realization of stressed clitics. Additionally, they found evidence in support that stressed vs. unstressed clitics pattern in the same way as stressed and unstressed affixes, in the sense that stressed elements in both cases were longer in duration, had higher intensity and higher pitch.

A recent analysis of Majorcan Catalan (Nadeu et al. 2017), a variety in which final stressed clitics had been reported, also found evidence consistent with the claim that the clitic bears primary rather than secondary stress in this Catalan variety. In their study, 8 speakers of Central Catalan, a variety with no ESS, and Majorcan Catalan, which is reported to have ESS, were compared in their production of verb+clitic and verb+non-clitic pronoun sequences using a reading task. The stressed vowel of the verb and those of the enclitic and the following non-clitic pronoun were compared in their duration, mean f0 and vowel quality. The latter feature is important in Catalan, a language with systematic unstressed vowel reduction. An additional sort of evidence of the stress shifting to the clitic would be a change in the quality of the vowel of the verb, which should be reduced. Results revealed that post-verbal pronouns in Majorcan Catalan differed from Central Catalan in their

increased duration and in the shape of the intonational contour associated with the clitic, which was convex in the former and concave in the latter. Additionally, the authors found evidence of vowel reduction associated with the verb in Majorcan but not in Central Catalan. This acoustic evidence led the authors to conclude that stress shifts to the clitic in Majorcan Catalan and that an interpretation of the facts as involving complete stress shift is more compatible with the acoustic data than an analysis involving either two prominences or a main prominence on the verb and a secondary prominence on the clitic.

Alternative analyses have been proposed for ESS in Romance varieties spoken in Italy. Peperkamp (1997) suggested that the stress shift to the penultimate observed in Neapolitan when a clitic cluster is encliticized to the verb (e.g. *narramilo*) could be analyzed as an instance of double prominence (one on the verb and one on the clitic), with the prominence on the clitic being higher. Kim & Repetti (2013), instead, interpreted the case of stress shift in Campidanese Sardinian, not as a change in word level stress, but as an association of the verb+clitic structure with an HL* pitch accent. Using a word translation task, the authors elicited verb+clitic sequences from 4 speakers of the dialect, who also spoke standard Italian. They analyzed the vowel duration and the pitch accents. As concerns the former, they found lengthening of the penultimate syllable. The pitch patterns are described as being rather consistent among stimuli and participants, with an H tone associated with the lexically stressed vowel of the verb and an L tone associated with the penultimate syllable of the sequence. Lengthening and a low tone consistently associated with the rightmost prominent syllables are interpreted as an indication that the pronominal element is a weak pronoun rather than a true clitic, and thus functions as a prosodic word. This prosodic word and the verb form a phonological phrase, and a bitonal pitch accent is associated with this phrase.

1.3. Stress shift in verb+clitic sequences and the formation of the future and conditional tenses

The studies that we have just reviewed have revealed that Romance varieties differ in the relative prominence patterns associated with verbs and pronouns in verb+clitic sequences. Acoustic studies have shown that when the stress is perceived to be on the clitic, duration is affected (Colantoni & Cuervo 2013; Kim & Repetti 2013; Nadeu et al. 2017) and pitch patterns may also differ from the patterns found in cases where the clitic is unstressed (Kim & Repetti 2013; Nadeu et al. 2017). This body of experimental evidence is consistent with early reports of doubly accented clitics in 16th century Spanish poetry and in literary and popular Argentine texts, where the orthographic accent is written on the clitic.

Stress shifts in verbal phrases are not new in Romance varieties. In the formation of future and conditional tenses, the main stress shifted from the root to the auxiliary in a process that was variable and long and had as its end point a paradigmatic stress on the last syllable. This process closely resembled ESS, and we are not the first ones to draw such parallel. In fact, in an article on the formation of future and conditional tenses in Romance, Nocentini (2001) argues

that the auxiliary, which had been shortened, did not have the same level of prominence as an independent word but was not deprived of an accent either. Nocentini explicitly compares the behavior of the auxiliary to that of a clitic (2001: 382). The auxiliary would, then, carry a secondary stress depending on the length of the verb+auxiliary compound. For example, it probably carried secondary stress in the first-person plural (e.g. *cantáre habémus* > *cantárèmus*, our example). In order to explain why we have a paradigmatic accent in the future tense, Nocentini invokes, among other factors, the power to attract stress that clitics have. He suggests that in sequences of root+(pronoun)+auxiliary, the increased distance between the primary and secondary stress, triggered a rhythmic pattern frequent in Latin (iambic inversion) and resulted in switching the primary stress to the auxiliary, with rhythmic retraction of the secondary stress on the infinitival root (2001: 385-386): *cantáre hábeo* > **cantár-àyo* > **càntaràyo* (> It *canteró*, Sp *cantaré*), *cantáre illud hábeo* > **cantár-lu-àyo* > **cantár-lu-áyo* (OSp *cantàr-lo-é*). Nocentini points out that the same tendency that operated in shifting the stress from the root to the auxiliary in the formation of the future is still active in many Italian varieties, which, as we have seen, show patterns of stress shift in verb+clitic sequences (see Ordoñez & Repetti 2006). Crucially, this process was not instantaneous. Although the first evidence that infinitive+HABEO sequences were used to express future tense is found already in the 3rd century, it took over 700 years for the synthetic form to fully replace the analytic one (Nocentini 2001). Moreover, Menéndez Pidal (1918: section 123) claims that, in Spanish, the future could be treated as a compound until well into the 17th century, as evidenced by the fact that pronouns were allowed between the verb and the auxiliary (e.g. *dar le has* ‘you will give him/her’; *venir vos edes* ‘you will come’ – Menéndez Pidal 1918: 326).

If Nocentini is right regarding the origin of stress shift in the development of Romance future and conditional tenses, we should expect that verb+clitic forms that at some point bore secondary stress on the clitic, as claimed for example by Kim & Repetti (2013), could eventually develop into one prosodic word with the main stress on the clitic (e.g. *dámelò* > *dameló* ‘give it to me’). Different varieties of a given language are expected to possibly be at different stages.

Regarding the two varieties of Spanish that we are concerned with in this paper, Argentine and Peninsular Spanish, there could be independent factors that may favor the pattern with main stress on the clitic in Argentine Spanish. One such factor may be the use of different second person pronouns and corresponding verb forms, e.g. Arg. (*vos*) *cantás* vs Pen. (*tú*) *cántas* ‘you-sg sing’ (see appendix I). The fact that *vos* imperatives have final stress (*cantá* ‘sing!’, *decí* ‘say!’, *explicá* ‘explain’, *introducí* ‘introduce!’) may lead speakers of Argentine Spanish to interpret doubly-stressed imperative forms as having final stress, e.g.: *decímelò* > *decimeló* ‘say it to me!’. A further step may have been the spread of final stress to nonfinite verb forms with enclitics (*cantármelo* > *cantármelò* > *cantarmeló*, *cantándomelo* > *cantándomelò* > *cantandomeló*).

Based on previous reports and experimental studies, we expect Argentine Spanish to be more advanced in this evolution towards final stress than Peninsular

Spanish. Here we report on a perception study where native speakers from these two Spanish dialects were asked to judge stress prominence in stimuli from both dialects. One question that we want to address is whether previous reports regarding the relative frequency in the production of possible stress patterns in each of the two dialects is reflected in the perception of native speakers.

2. Perception experiment

2.1. Research questions

Reports on ESS typically rely on the researcher's perception of prominence. If our goal is to document the existence of the phenomenon and its distribution, this may indeed be enough, since in canonical instances the shift of prominence is perceptually salient and unambiguous. If, on the other hand, we aim at quantifying ESS based on corpus data in order to be able to determine, for instance, differences among dialects or speech styles, as in the unpublished work by Colantoni, Cuervo & Hualde (2010) mentioned in section 1, it is less obvious that one person's perception may offer reliable data.

For comparison, Spanish speakers agree on which syllable of any word carries lexical stress. In citation form of words, this syllable is acoustically enhanced. In natural speech, however, many words are produced with very little or no acoustic evidence showing which syllable is lexically stressed, so that, in the case of minimal pairs, listeners may be unable to identify the meaning that the speaker intended (Torreira, Simonet & Hualde 2014). That is, for instance, *cánton* 'I sing' and *cantó* 's-he sang' are a minimal pair in Spanish, which native speakers clearly distinguish in citation form. However, within a phrase such as *siempre que cánton/cantó esta canción [...]* 'Every time I sing/s-he sang this song [...]', the position of the stress, which serves to identify the word, may be conveyed very weakly or not at all, so that the discourse context may be needed in order to identify the word.

Similarly, phrasal accent in English may be used for pragmatic purposes and phonologists agree on its expected position in different contexts (e.g. Ladd 2008). Judging phrasal prominence in actual discourse, however, becomes very subjective and unreliable. If consulted, naïve speakers do not show very high agreement in their judgments regarding which words in the phrase are prominent. With a sufficiently large number of participants, however, it is possible to determine which factors contribute to the perception of prominence for different groups of listeners, as shown in recent work by J. Cole and colleagues (Cole, Moe & Hasegawa-Johnson 2010; Cole, Mo & Baek 2010; Cole, Mahrt & Roy 2017; Roy, Cole & Mahrt 2017; for German, see also Baumann & Winter 2018).

Here we use a similar methodology (asking a relatively large group of naïve, educated speakers) to study the perception of ESS in Spanish. Our research questions are, thus, to what extent naïve speakers agree on the position of the stress, whether there are differences among native speakers from Argentina and Spain in this perception, and what acoustic features listeners may use in their judgment of prominence on different syllables.

2.2. Methods

2.2.1. Participants

A total of 36 native Spanish-speakers participated in this experiment: 15 speakers of Argentine Spanish recruited in Toronto and 21 speakers of Peninsular Spanish recruited in Irun, Gipuzkoa, Basque Country, Spain. Argentinian participants recruited in Toronto had spent no more than three years in Canada and used Spanish every day. Indeed, half of our Argentine participants lived in Argentina and were visiting relatives in Canada at the time of the experiment. Irun is a bilingual Spanish/Basque town, but where Spanish is by far the most common language of social interaction. Our Peninsular Spanish speakers all used Spanish daily and varied in their knowledge of Basque.⁶ Participants did not report having any hearing or speech impairments. The participants in this experiment may in general be characterized as educated speakers, but with no training in Linguistics and naïve as to the exact purpose and hypotheses of the experiment.

2.2.2. Stimuli

We extracted all verb+clitic sequences that could be potential targets for ESS from 12 hours of Spanish-language TV soap operas (*telenovelas*): 6 hours from a series produced in Spain (“Cuéntame cómo pasó”) and 6 hours from two series made in Argentina (“Floricienta” and “Socias”).⁷ As mentioned in the introduction, potential targets of ESS are verb+enclitic sequences where there is at least one syllable between the lexically stressed syllable of the verb and the last syllable of the sequence (e.g., *cantármelo*, but not *cantárllo*). After eliminating tokens with background music and other types of background noise, we obtained 100 tokens to be used as stimuli for our perception experiment, 65 from the TV show produced in Spain and 35 from the two Argentine TV shows.⁸

Figures 1-4 show examples of four types of pitch contours observable in our stimuli. They are all from the series produced in Spain. Figure 1 shows a double-peak contour, with upstepping of the second peak, interpretable as double stress, with primary stress on the clitic (given the upstep): *míranós* ‘look at us’. Figure 2 illustrates an example of main pitch prominence on the clitic: *calmaté* ‘calm down’. Figure 3 shows a pitch rise on the stressed syllable of the verb continuing as high plateau: *sácala* ‘take it out’. This contour seems more difficult to interpret in terms of prominence. There is prominence on the verb, but the clitic may or may not be perceived as prominent as well. Finally, Figure 4 shows a clear case with a single prominent pitch rise and fall on the verb: *vámonos*.

6. Given the great morphological differences between Basque and Spanish, we do not expect any transfer effects due to knowledge of Basque. Basque does not have clitic pronouns. Some Basque dialects have lexical pitch-accent, but that is not the case in the variety of Irun.
7. The actors, as far as we know, are all native speakers of Spanish from each of the two countries. The actors from which the Argentine tokens were obtained are all native speakers of Buenos Aires Spanish.
8. Our goal was to obtain an equal number of stimuli per variety. To do that, we selected the same number of hours of TV shows from each variety. However, in the Argentine shows, we obtained a smaller number of tokens and we had to discard many tokens, given the presence of music and overlapping conversations.

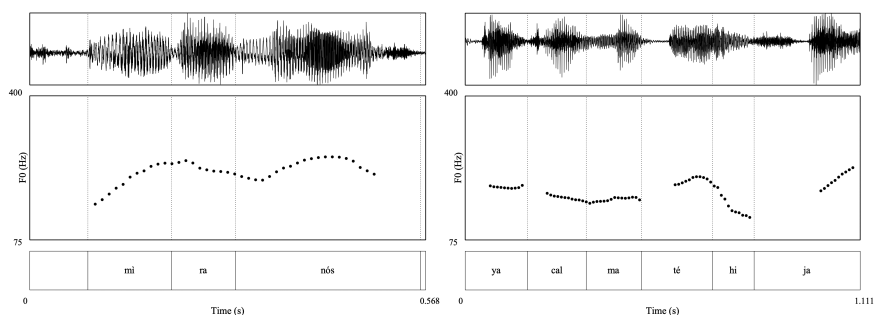


Figure 1 (left). Illustration of double tonal prominence in *mirános* ‘look at us’.

Figure 2 (right). Illustration of tonal prominence on the clitic in *yá calmaté* ‘calm down already’.

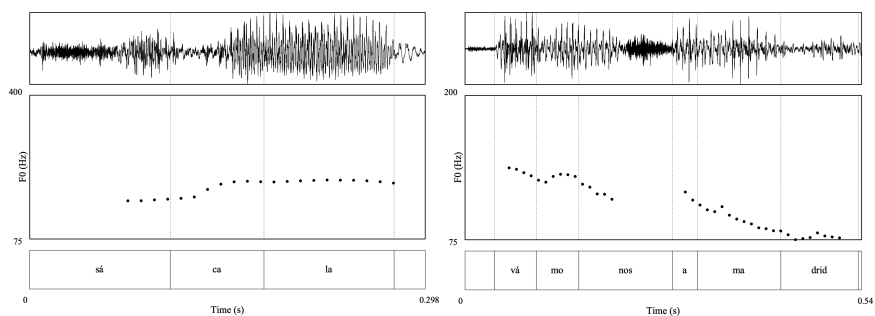


Figure 3 (left). Illustration of an instance of pitch rise on verb followed by a sustained high tone in *sácala* ‘take it out’.

Figure 4 (right). Illustration of a pitch rise and fall on the verb in *vámonos* ‘let’s go’.

Since besides pitch, other acoustic features of the speech signal, such as duration and intensity, may affect the perception of prominence on a given syllable, we acoustically analyzed our stimuli in Praat (Boersma & Weenick 2018) using ProsodyPro (Xu 2013). The three target vowels were segmented using TextGrids (see Figure 5) and labeled as follows: v1: lexically stressed vowel of the verb; v2: penultimate vowel of verb+enclitic unit; v3: final vowel of verb+enclitic unit.

We extracted the mean duration, mean intensity and mean f0 (in semitones) of each of the three vowels. For duration and intensity, we then calculated the ratios v1/v3 and v2/v3. For example, a v1 to v3 ratio of “1” in intensity indicates that the lexically stressed vowel of the verb and the vowel of the last clitic have the same intensity. For pitch, we subtracted the value of the third vowel to that of v1 or v2. Differences beyond 1.5 semitones are considered to be perceptible.⁹

9. There is evidence that differences of 3ST play a role in communication (t’Hart et al. 1990: 29). However, Rietveld and Gussenhoven (1985) report lower thresholds. They found that differences of 1.5 STs were perceptually salient.

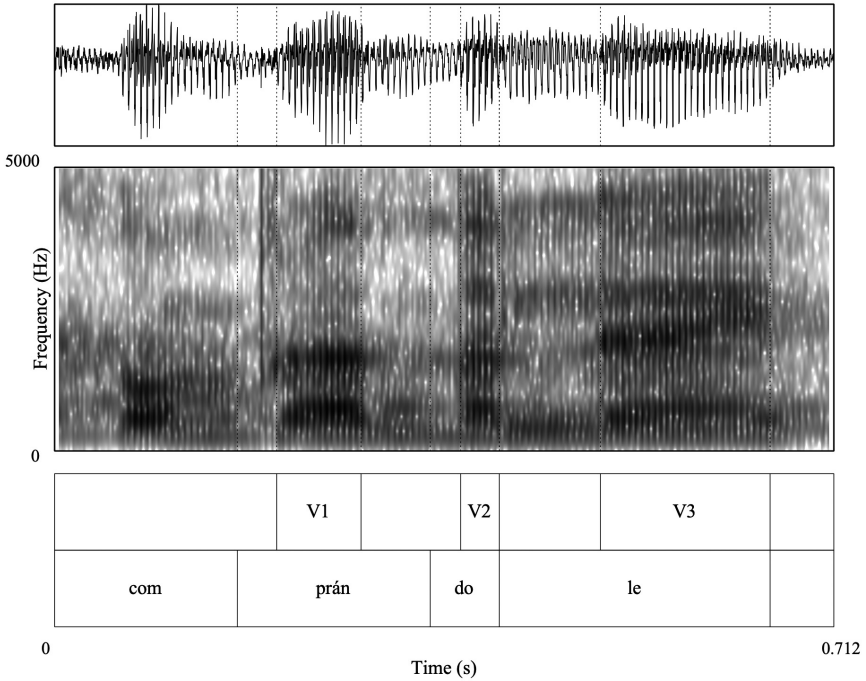


Figure 5. Example of segmentation of target vowels. Sound file: *Comprándole* ‘buying if for her/him’ (extracted from the Spanish soap opera *Cuéntame cómo pasó*).

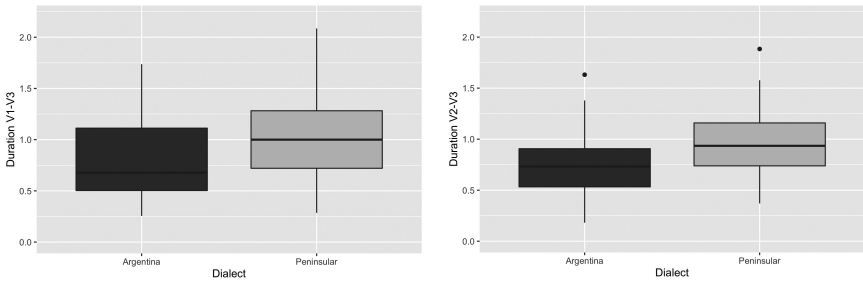


Figure 6. v1/v3 ratio (left) and v2/v3 ratio (right) in duration of all the stimuli in our sample.

Figures 6 and 7 (produced in R (R Core Team 2017) with the package *ggplot2* (Wickham 2016)) illustrate the between-dialect ratios in duration and differences in pitch, respectively.

In both cases, the ratio of the Argentine data is below 1, which suggests that the final vowel tends to be longer than both the lexically stressed vowel of the verb and

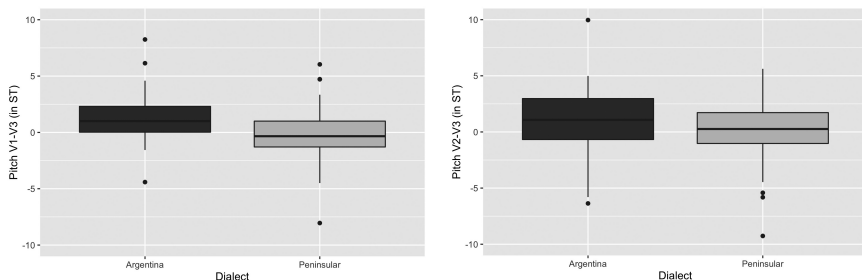


Figure 7. v1-v3 difference (left) and v2-v3 difference (right) in pitch (semitones) of all the stimuli in our sample.

the penultimate vowel. For the Peninsular stimuli, on the other hand, the mean for both ratios is close to 1, indicating relatively small differences in duration among the three vowels overall. To test whether the stimuli of the two dialects differed in their duration ratio, we ran two U-tests, which returned a significant difference for both ratios (v1/v3: $W = 1530.5$, $p\text{-value} = 0.004$; v2/v3: $W = 1553$, $p\text{-value} = 0.002$).

As concerns pitch, the median positive values obtained in the Argentine sample (see Figure 7) suggest that the final vowel is lower in pitch than the other two preceding vowels. The difference in pitch between v1-v3 and v2-v3 are likely to be perceptible, since these are close to the 1.5 semitone threshold (1.3 semitones). In the Peninsular data, the natural declination tendency seems not to apply and differences in semitones between the vowels compared hovered below 0 (v1-v3=-0.2st; v2-v3=-0.05). Once again, results of U-tests comparing the two samples revealed significant between-dialect differences in the v1-v3 difference ($W = 660$, $p\text{-value} = 0.001$) as well as in the v2-v3 difference ($W = 801$, $p\text{-value} = 0.02$).

Intensity differences between the two samples turned out to be non-significant for both ratios (v1/v3: $W = 1012$, $p\text{-value} = 0.3$; v2/v3: $W = 963$, $p\text{-value} = 0.2$).

Given the difference in the two sets of stimuli regarding pitch and durational cues, we expect our participants to perceive an accent on the clitic more frequently in the Argentine tokens.

2.2.3. Experimental procedure

The perception experiment was run in *Psychopy* (Pierce 2007) on a laptop computer. Stimuli were presented to the participants in a randomized order generated by the program (putting together all stimuli from both dialects). Participants wore headphones and did the experiment in the presence of the experimenter, in a quiet environment (not through the internet). They were asked to identify the position of the stress by pressing ‘1’ on the keyboard, if they perceived the stress as falling on the verb (e.g. *vámonos*); ‘2’, if they perceived two stresses (*vámonós*); or ‘3’, if they perceived stress at the end of the word (*vamonós*). Participants had a short training session and they could listen to the stimuli only once. A screen shot illustrating the procedure is shown in Figure 8.

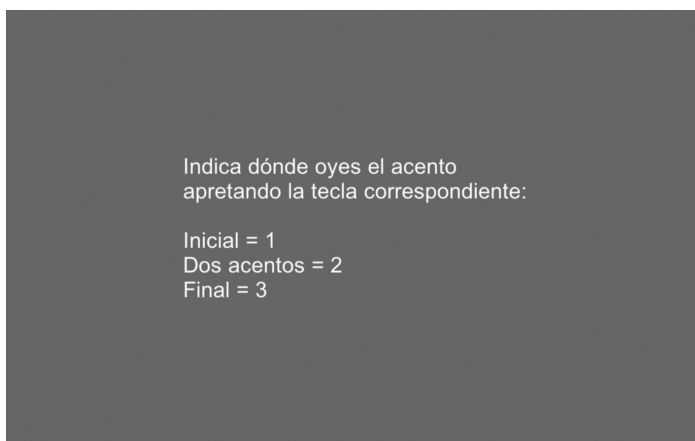


Figure 8. Screen shot of example presented in Psychopy.

2.2.4. Data analysis

Both the results of the acoustic analysis of the signal performed in *ProsodyPro* (Xu 2013) and the participants' responses from *Psychopy* (Pierce 2007) were exported to *Excel*, and statistical analysis was performed in *R* (R Core Team 2017) within the *RStudio* (RStudio Team 2016) environment. Binomial mixed-effects logistic regressions were fitted to the participants' responses with the package *lme4* (Bates et al. 2015), and p-values were obtained with *lmerTest* (Kuznetsova et al. 2017). We explore whether both groups of participants, from Argentina (ARG) and Spain (PEN), reacted in a similar way to the stimuli, and what acoustic factors determined their perception of stress on the clitic.

2.3. Results

As mentioned in section 2.2.3, participants in the experiment could provide one of three responses: 1 = stress on verb, 2 = two stresses or 3 = stress on the clitic. In the first section of our results, §2.3.1, we focus on the perception of ESS. Thus, for the statistical analysis, we will group together the two answers that indicate ESS, that is, perception of prominence on the clitic (2 and 3) vs. prominence only on the verb (response 1). Then, in section 2.3.2, we examine the distribution of double stress, and, thus, we compare response 2 to the two responses that indicate that only one syllable was perceived as prominent (i.e., 1 and 3).

2.3.1. Perception of ESS

In general, agreement for all 36 participants on the rating of 100 tokens was only 'slight' (Fleiss' Kappa = 0.189, $z = 47.3$, $p < 0.0001$). If we consider only the 21 participants in the Peninsular group, agreement rises to 'fair' (Fleiss' Kappa = 0.227, $z = 32.9$, $p < 0.0001$). Considering only the 15 Argentine participants there is no improvement in agreement.

Table 1. Perception of prominence pattern by participant group

Participant group	1 = stress on verb	2 = double stress	3 = stress on clitic
ARG	44.4%	29.3%	26.3%
PEN	33.1%	35.2%	31.8%

Overall, the Peninsular group was more likely to perceive both stress on the clitic and double stress than the group from Argentina. As shown in Table 1, whereas ARG participants perceived a single stress on the verb 44.4% of the time, for the PEN group only 33.1% of the responses fall into this category.

When we group together answers ‘2’ (= two stresses perceived) and ‘3’ (= stressed perceived on the clitic), given that in both cases there is perceived prominence on the clitic (that is, ESS), and separate the responses by the dialect of the stimulus, we obtain the density plots in Figure 9.¹⁰ Here, the left panel shows the distribution of responses of our 15 participants from Argentina and the right panel that of our 21 participants from Spain.

As shown in the left panel of Figure 9, for the Argentine participants, responses to ‘arg’ (= Argentina) stimuli were grouped towards the right of ‘pen’ (= Spain) stimuli, indicating that ‘arg’ stimuli were perceived as either stressed on the clitic or having double stress more often. In these plots, the x axis indicates the number of participants who agreed on perceiving ESS. For instance, for the ‘arg’ stimuli, there were 9 tokens for which 14 participants perceived ESS and 5 tokens for which all 15 Argentine participants agreed on perceiving ESS. This is to be compared with only 2 ‘pen’ tokens for which 14 of the participants agreed on ESS and 1 ‘pen’ token where they all agreed. The gray line representing ‘pen’ stimuli reaches its peak at 6 on the x-axis because, for 11 ‘pen’ tokens, 6 Argentine participants gave a response consistent with ESS and this was the highest agreement that was reached for ‘pen’ stimuli (6/15).

The participants tested in Spain (Figure 9, right) show a somewhat different distribution, with more perception of prominence on the clitic for both sets of stimuli, but much more so for ‘arg’ stimuli. That is, both groups of participants perceived more instances of ESS in the stimuli extracted from Argentine soap operas and the participants from Spain provided more frequent ‘2’ and ‘3’ responses.

To determine statistical significance, we ran a mixed-effect binomial regression model with perceived prominence (‘2’ and ‘3’ combined vs. ‘1’) as the dependent variable and dialect of the stimulus (Argentine, Peninsular) and participants’ dialect (Argentine, Peninsular) as fixed factors. Item and participant were included as random effects. Both fixed factors turned out to be significant (stimulus dialect, $p < 0.001$, participants’ dialect, $p < 0.01$). As shown in Figure 10,¹¹ the regression indicates that both groups of participants tended to perceive stress on the clitic more often in Argentine than in Peninsular stimuli. In addition, participants from Spain tended to perceive stress on the clitic more often in both kinds of stimuli.

10. Density plots were made with the package *ggplot2* (Wickham 2016).

11. This figure was created with the package *visireg* (Breheni & Burchett 2017).

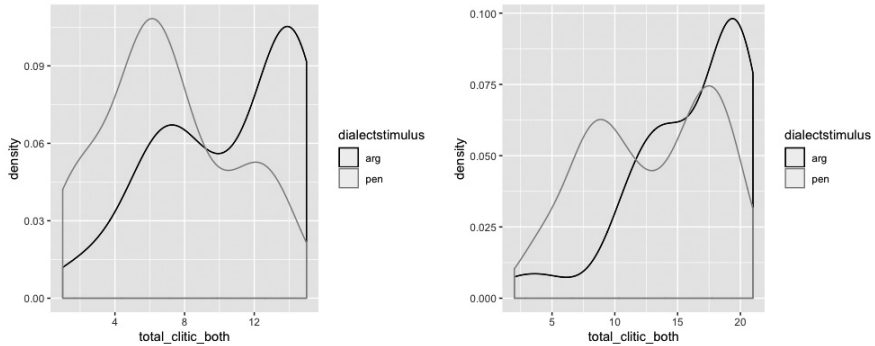


Figure 9. Density plots of responses ‘2’ (two stresses) and ‘3’ (stressed on the clitic) combined, separately for ARG (left panel) and PEN (right panel) participants. The black line shows the distribution of responses to Argentinian stimuli and the gray line shows responses to Peninsular stimuli.

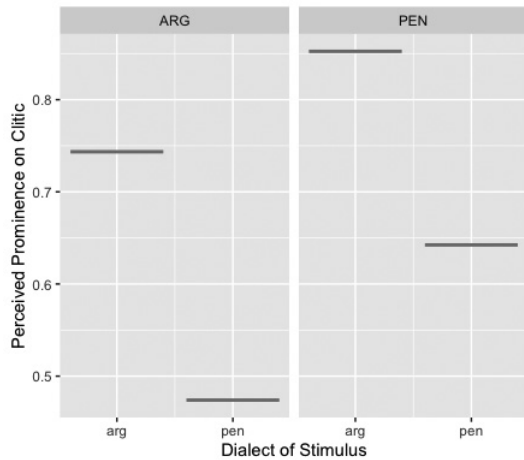


Figure 10. Effect of participants’ dialect and of stimuli dialect on the perception of prominence on the clitic (double prominence included). Left panel = responses of participants from Argentina (ARG); right panel = responses of participants from Spain (PEN).

Argentini- ans perceived stress on the clitic in stimuli from Spain with very low frequency.

Given this substantial difference in responses between the two groups of participants, in order to examine the effects of acoustic features on the perception of stress on the clitic, we decided to consider the responses of each group of participants separately. For each of the two groups, then, we fitted a binomial mixed-effects regression model with the structure indicated in Table 2.

Table 2. Structure of mixed-effects binomial logistic regression with acoustic predictors

Dependent variable	Binary Response (stress on the verb vs. stress on clitic/two stresses)
Fixed factors	Duration ratio between V1 and V3 Intensity ratio between V1 and V3 Pitch difference between V1 and V3
Random factors	Word Participant

Table 3. Output of mixed effect models with acoustic predictors only for ARG participants

	Estimate	z-value	p-value
Duration	-0.01	-3.23	= 0.001
Intensity	-0.15	-3.11	= 0.002
Pitch	0.003	0.81	= 0.4 n.s.

Table 4. Output of mixed effect models with acoustic predictors only for PEN participants

	Estimate	z-value	p-value
Duration	0.18	0.53	= 0.6 n.s.
Intensity	-4.93	-1.63	=0.1 n.s.
Pitch	0.008	2.35	=0.018

A summary of the output of the regression regarding the effect of the fixed factors for the Argentine participants (ARG) is shown in Table 3. For this group of participants, there is a significant effect of duration ($p < 0.001$) and intensity (< 0.01), but not of f_0 , on whether stress is perceived on the clitic (whether an additional stress is perceived on the verb or not).

For the participants tested in Spain (PEN), the regression returns the output summarized in Table 4. Interestingly, for this group of participants neither duration nor intensity (as measured by $v1/v3$ ratios) appears to have a significant effect on their response. On the other hand, pitch has a marginally significant effect.

2.3.2. Perception of double stress

In this final section of the results we explore the perception of double stress. The questions that we ask are the following: Is double stress perceived more often in the Peninsular than in the Argentine data? Do speakers of both geographical varieties show similar patterns of perception?

The data in Table 1 above showed that ARG speakers responded ‘double stress’ only slightly less often than PEN participants (29% vs 35%, respectively). In Table 5, we show the frequency with which the Argentine and Peninsular stimuli elicited each of the three possible responses, for all 36 participants together:

The stimuli taken from Argentine TV shows elicited many more ‘double-stress’ responses than those from Spanish shows (40% vs 28.8%). The difference in the responses to Argentine and Peninsular stimuli is consistent with

Table 5. Perception of prominence pattern by dialect of stimulus (all 36 participants)

Stimulus	1= stress on verb	2 = double stress	3 = stress on clitic
Argentine	328 = 26%	503= 40%	429 = 34%
Peninsular	1040= 44.4%	674= 28.8%	627= 26.8%

some of the differences in the acoustic properties of the stimuli reported above in section 2.2.2. In order to explore whether the dialect of the participant or the dialect of the stimuli had a significant effect on the type of response, we created a second binary variable putting together answers ‘1’ and ‘3’, which indicate that a single stress was perceived either on the verb or on the clitic, vs. answer ‘2’ (double stress perceived). This binary dependent variable was modeled using a binomial mixed effects regression with the same structure as the one used in the previous section to test the difference between a stressed perceived on the clitic vs. double stress.

The regression returned a significant effect of the dialect of the stimulus ($\beta = 0.54$, $z = 2.48$, $p = 0.012$), but no significant effect of the dialect of the participants. Both groups of participants perceived double stress significantly more often in the stimuli extracted from the Argentinian soap operas (see Figure 11).

Given the non-significant effect of participants’ dialect on the perception of double stress, we did not run a mixed effects model with the acoustic variables that characterize the stimuli, as we did in the previous section.

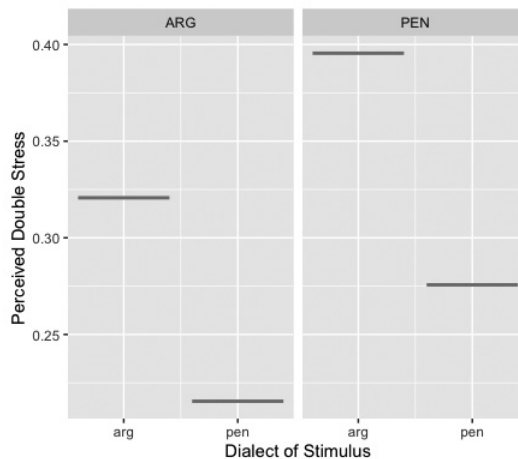


Figure 11. Effect of participants’ dialect and of stimuli dialect on the perception of double prominence vs. single prominence (either on the verb or on the clitic). Left panel = participants from Argentina; right panel = participants from Spain.

3. Discussion

3.1. *Summary of results and answers to our research questions*

The results of our perception experiment revealed a more complex pattern than the one suggested by previous experimental studies. As concerns our question regarding whether participants' dialect affected the perception of ESS in verb+clitic sequence, the answer is affirmative: we obtained different response patterns for Peninsular and for Argentine Spanish speakers. Whereas the former tended to have their responses almost equally distributed across the three categories (stress on the verb, on the clitic or both), the Argentine participants' most frequent response was that the main prominence was on the verb. This difference in perception between participants depending on their native dialect is possibly the most solid result of this perception study. Our results tentatively suggest that each group of participants used different cues in order to judge the prominence pattern. As shown in Table 3, duration and intensity turned out to be significant cues for our Argentine participants, whereas only pitch was significant for the Peninsular Spanish participants. This is an interesting result that was not predicted and would require further testing.

Notice that, in principle, the finding that our Peninsular participants provided more frequent ESS responses than our Argentine participants would appear to go against our predictions. However, in order to interpret these results, it is important to analyze the interaction between participant's dialect and the dialect of the stimuli. One interpretation could be that, because ESS is more advanced in Argentina, speakers from this country have a higher acoustic threshold for its perception.

Both possible ESS patterns, namely shift of the stress to the last syllable and double stress, were perceived more often in the examples from Argentinian TV shows than in those from Spain. This could be due to an accident of the specific stimuli that we happened to select for this study but is consistent with prior research indicating a higher frequency of ESS in Argentine than in Peninsular Spanish.

Finally, our participants frequently indicated perception of double stress (in 40% of the stimuli from Argentinian soap operas and in almost 30% of those from Spanish TV shows), but there were also many cases of stress perceived only on the syllable containing the (last) clitic pronoun, especially in responses to stimuli from Argentinian soap operas (34% of all responses to stimuli in this group). These results may be interpreted as showing a stage in a diachronic path towards shifting of the (production and perception of the) stress from the verb root to the enclitic, as happened in the evolution of the future tense.

3.2. *Potential task effects*

The methodology that we used, which has been previously applied to the perception of phrasal accent (see section 2.1), seems promising. This methodology is inherently inductive and seeks to determine whether participants can show some agreement in the way they categorize stimuli and whether the stimuli that belong to a given category share some acoustic properties. In this way, we tried to avoid some of the shortcomings of previous acoustic studies (e.g. Colantoni & Cuervo

2013; Klassen & Patience 2016), which relied on the categorization of the stimuli based on the researchers' perception.

Indeed, we observed in our data that individual participants had different patterns of response, but tendencies emerged at the group level. It should be noted that this was a difficult task, since participants could listen to each sound file only once, they had to rate 100 files and the sound files contained words in isolation. It is important to highlight that the stimuli differed to a great extent across many variables, such as the gender and age of the speaker, the speech rate and the loudness. Moreover, we used the examples that we obtained from 12 hours of recordings, and presented them out of context. Given the nature of the stimuli and the findings by Torreira et al. (2014) regarding the perception of lexical stress in running speech, we should expect some degree of uncertainty in participants' responses, a degree of uncertainty that we would not expect if the stimuli had been produced in isolation with the aforementioned variables being controlled for.

Our study did not include a control group, i.e. participants and stimuli from a dialect where ESS had not been reported (e.g. Mexican Spanish), as it was the case in Nadeu et al. (2017). We believe, however, that testing the perception of speakers from Argentina and Spain constitutes a valuable contribution.

In future research, the responses obtained with this methodology could be also fruitfully compared with the shared judgment of experts listening to verb forms in context.

3.3. *ESS and iambic inversion*

In our experiment the stimuli obtained from Argentine soap operas were judged more frequently to display ESS, that is, to either have the main prominence on the clitic or to have double prominence, than stimuli from Spain. This is consistent with previous reports that ESS is more frequent in Argentine than in Peninsular Spanish and with the results of the acoustic analysis of the stimuli (although, again, it also could be an accident in the selection of tokens).

On the other hand, based on our own observations and the scarce reports in the literature (e.g. Menéndez Pidal 1918), we expected to find a preference for a perception of double stress (as opposed to complete shift of the stress to the clitic) in the Peninsular stimuli and, instead, a preference for final stress in the stimuli from Argentina. Although the Argentine stimuli received more 'stress on clitic' responses than the Peninsular stimuli, they also received even more 'double stress' responses.

From an analytical point of view, it may be the case that, in stimuli where double stress tended to be perceived, the main stress had shifted to the clitic but a perceivable secondary stress remained on the verb stem; or it may also be the case that the primary stress remained on the stem, but the clitic had a secondary stress that was also perceived. A factor that needs to be taken into account is that we tested educated speakers who know where the stress is supposed to be, but still might have heard a prominence on the clitic, which might have led them to report two stresses rather than just one stress on the clitic.

Independently of the specific explanation, which should be explored in future studies, our results revealed a difference in the perception of the stimuli coming from each dialect and suggest that ESS in Argentine Spanish is in the mixed penultimate stress shift category in Ordoñez & Repetti's (2006) typology. Again, the differences in responses to the two groups of stimuli, although consistent with our expectations, may be due to the characteristics of the specific set of stimuli that we used. That is, it could be that by chance we happened to choose a greater proportion of ARG stimuli with ESS. For this reason, further experimentation is needed.

The variability observed in the Argentine participants' responses also suggests, first of all, that final stress shift is not systematic and, secondly, that speakers of Argentine Spanish have a higher threshold for perceiving ESS than speakers from Spain. They provided responses interpretable as ESS less frequently, especially to Peninsular tokens. Peninsular Spanish seems to be at a different point in the final stress shift typology, as evidenced mainly by the fact that the responses obtained from Peninsular speakers were equally distributed across the three possible categories included in the study. This suggests that at least there is some vacillation regarding the placement of the main prominence in these sequences.

In section 1.3, we suggested that the differences between varieties could be linked to differences in the second person verb paradigm observed in these two dialects. Although this is a potential factor that seems to be worth considering, both the acoustic analysis of our stimuli and the acoustic cues that participants from each dialect seem to be using as a cue to main prominence suggest an additional explanation: there may be differences in the cues that convey stress in the two dialects (with a greater weight of duration and intensity in Argentine Spanish and of pitch in Peninsular Spanish). This clearly requires further investigation.

What are then the parallels between ESS and the iambic inversion observed in the formation of Romance future and conditional tenses? In our stimuli, we can see that there is a tendency to perceiving some type of prominence on the clitic (either primary or secondary), as seems to have been the case in sequences of infinitive + HABEO in Late Latin. We have also observed that the tendency to perceive the prominence on the clitic is not categorical and shows cross-dialectal variation. What we do not know is whether this variability is a stable feature of the language, as it may be concluded from Menéndez Pidal's (1918) report on the presence of doubly stressed verb+clitic sequences in 16th century poetry, or if it is a case of a change in progress.

References

- Bates, Douglas, Maechler, Martin, Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1): 1-48.
- Baumann, Stephan & Winter, Bodo. 2018. What makes a word prominent? Predicting untrained German listeners perceptual judgments. *Journal of Phonetics* 70: 20-38.
- Boersma, Paul & Weenink, David. 2018. Praat: doing phonetics by computer, version 6.035, <www.praat.org> [computer program].

- Breheny, Patrick & Burchett, Woodrow. 2017. Visualization of Regression Models Using visreg. *The R Journal* 9: 56-71.
- Colantoni, Laura & Cuervo, María Cristina. 2013. Clíticos acentuados. In Laura Colantoni & Celeste Rodríguez Louro. *Perspectivas teóricas y experimentales sobre el español de la Argentina*, 143-157. Madrid: Iberoamericana.
- Colantoni, Laura, Cuervo, María Cristina & Hualde, José Ignacio. 2010. Stress as a symptom. Paper presented at the *Linguistic Symposium on Romance Languages 40*, University Washington, Seattle, March 26-28.
- Cole, Jennifer, Mo, Yoonsook & Hasegawa-Johnson, Mark. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1: 425-452.
- Cole, Jennifer, Mo, Yoonsook & Baek, Soondo. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes* 25(7): 1141-1177.
- Cole, Jennifer, Mahrt, Timothy & Roy, Joseph. 2017. Crowd-sourcing prosodic annotation. *Computer Speech & Language* 45: 300-325.
- Fernández Soriano, Olga. 1999. El pronombre personal. Formas y distribuciones. In Ignacio Bosque & Violeta Demonte (eds). *Gramática descriptiva de la lengua española*, 1209-1273. Madrid: Espasa.
- Fontanella de Weinberg, María Beatriz. 2000. El español bonaerense. En María Beatriz Fontanella de Weinberg & Nélica Donni de Mirande (eds.). *El español de la Argentina y sus variedades regionales*, 37-61. Buenos Aires: Edicial.
- Gili Gaya, Samuel. 1943. *Curso superior de sintaxis española*. Barcelona: Bibliograf.
- Kuznetsova, Alexandra, Brockhoff, Per & Christensen, Rune. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13): 1-26.
<<https://doi.org/10.18637/jss.v082.i13>>
- Hualde, José Ignacio. 2005. *The sounds of Spanish*. Cambridge: CUP.
- Huidobro, Susana. 2005. *Phonological constraints on verum focus in Argentinian Spanish*. Manuscrito, SUNY.
- Kim, Miran & Repetti, Lori. 2013. Bitonal pitch accents and phonological alignment in Sardinian. *Probus* 25, 267-300.
- Klassen, Gabrielle & Patience, Matthew. 2016. Stressed clitics in Argentine Spanish: Which way does the clitic lean? In Alejandro Cuza, Lori Czerwionka & Daniel Olson (eds.). *Inquiries in Hispanic Linguistics: From theory to empirical evidence*, 149-169. Amsterdam: John Benjamins.
- Martínez, Tomás E. 1991. *La novela de Perón*. Buenos Aires. La Nación.
- Moyna, María Irene. 1999. Pronominal clitic stress in Río de la Plata Spanish: an optimality account. *The SECOL Review* 23, 15-44.
- Nadeu, Mariana, Simonet, Miquel & Llopart, Miquel. 2017. Stress postverbal pronominals in Catalan. *Probus* 29(1): 119-162
- Nocentini, Alberto. 2001. La genesi del futuro e del condizionale sintetico romanzo. *Zeitschrift für romanische Philologie* 117(3): 367-401.
- Peperkamp, Shannon. 1997. *Prosodic words*. The Hague: HAG.
- Ordóñez, Francisco & Repetti, Lori. 2005. Stressed clitics? In Jean-Pierre Montreuil (ed.). *New perspectives on Romance linguistics. Volume II: Phonetics*, 167-181. Amsterdam: John Benjamins.

- Peirce, Jonathan. 2007. PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods* 162 (1-2): 8-13.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>
- Rietveld, Toni & Gussenhoven, Carlos. 1985. On the relation between pitch excursion size and prominence. *Journal of Phonetics* 13: 299-308.
- Roy, Joseph, Cole, Jennifer & Mahrt, Timothy. 2017. Individual differences and patterns of convergence in prosody perception. *Laboratory Phonology* 8(1): 22. <<http://doi.org/10.5334/labphon.108>>
- RStudio Team. 2016. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <<http://www.rstudio.com/>>.
- t'Hart, Johan, Collier, René & Cohen, Antonie. 1990. *A perceptual study of intonation: An experimental phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Torreira, Francisco, Simonet, Miquel & Hualde, José Ignacio. 2014. Quasi-neutralization of stress contrasts in Spanish. Campbell, Gibbon & Hirst (eds.). *Proceedings of Speech Prosody* 7, 197-201. Dublin.
- Torres Tamarit, Francisco & Pons-Moll, Claudia. 2018. Enclitic-induced stress shift in Catalan. *Journal of Linguistics* 55(2): 407-444. <<https://doi.org/10.1017/S0022226718000427>>
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Xu, Yi. 2013. ProsodyPro: A tool for large-scale systematic prosodic analysis. TRASP 2013, Aix-en-Provence. <http://www.homepages.ucl.ac.uk/~uclyyix/ProsodyPro/#old_version>.

Appendix 1

Second person Imperative forms in Argentine and Peninsular Spanish, illustrated with the verb *cantar* 'to sing'. Notice that the 2nd person singular form *vos*, used in Argentina, and the 2nd person plural form *vosotros/as*, used in Spain, are etymologically related. In the Imperative, *vos* and *vosotros/as* forms have final stress. They only differ in that the *vosotros/as* forms have a final /d/.

	Argentina	Spain
2nd person singular familiar	<i>cantá (vos)</i>	<i>cánta (tú)</i>
2nd person singular formal	<i>cánte (usted)</i>	<i>cánte (usted)</i>
2nd person plural	<i>cánten (ustedes)</i>	<i>cantád (vosotros/as)</i> <i>cánten (ustedes) formal</i>

Appendix 2. List of stimuli

Stimulus	Dialect
abonenme	Peninsular
acompañandole	Peninsular
acuerdate	Peninsular
adelantamelos	Peninsular
avisame	Peninsular
ayudandolo	Argentine
callate	Peninsular
calmate	Peninsular
comete	Peninsular
comportate	Peninsular
comprandole	Argentine
comprandole	Argentine
convirtiendonos	Argentine
creame	Peninsular
cuidandote	Argentine
cuidate	Peninsular
damelo	Argentine
decirmelo	Argentine
deciselo (2)	Argentine
dejale	Peninsular
dejame	Peninsular
dejame	Peninsular
dejamela	Argentine
dejate	Peninsular
dejeme	Argentine
digales	Peninsular
digame (6)	Peninsular
dimelo	Peninsular
disculpeme (2)	Argentine
diselo	Peninsular
disfrazandose	Argentine
embarazandome	Argentine
encierrame	Peninsular
ensename	Peninsular
escucheme (2)	Argentine
esperame	Peninsular
esperandole	Peninsular
explicamelo	Peninsular
fiate	Peninsular
imagine (2)	Peninsular
imaginese	Peninsular
incitandonos	Peninsular
llamemosla	Argentine

Stimulus	Dialect
llevate (2)	Peninsular
llevele	Argentine
miralo	Peninsular
miralos	Peninsular
miranos	Peninsular
mireme	Argentine
olvidate	Peninsular
pensemoslo	Argentine
perdoneme	Argentine
permitame	Peninsular
pidamelo	Argentine
piensalo	Peninsular
ponetelo (2)	Argentine
pongase	Argentine
portate	Peninsular
preparate	Peninsular
prometale	Argentine
quedate (2)	Peninsular
quedese	Peninsular
quedese	Argentine
recibiendole	Peninsular
sacala	Peninsular
secate	Peninsular
sientate (5)	Peninsular
sientese	Argentine
sintiendome	Peninsular
subame	Peninsular
tranoslo	Argentine
tranquilicese	Peninsular
tranquilicese (2)	Argentine
transmitales	Peninsular
vamonos (4)	Peninsular
vayase	Argentine
vendiendoselo	Argentine
viendolo (2)	Peninsular
vistase	Argentine
