THE FOCUS PROMINENCE RULE IN SPANISH FROM A PERCEPTION PERSPECTIVE[*]

Christoph Gabriel[a] & Steffen Heidinger[b]

[a]Johannes Gutenberg University Mainz, Germany; [b]University of Graz, Austria

ABSTRACT. The focus prominence rule (FPR) predicts that speakers articulate their utterances in such a way that the nuclear stress falls within the focus domain (*¿Qué compró Juan?* 'What did John buy?' → *Juan compró* [*una biciCLEta*]$_F$ 'John bought [a BIKE]$_F$' / #*JUAN compró* [*una bicicleta*]$_F$ 'JOHN bought [a bike]$_F$'). To examine the consequences of the FPR for focus interpretation, we carried out a perception experiment using oral production data produced by Argentinean speakers. Two groups of hearers representing either the Argentinean or the Peninsular variety of Spanish were tested. We examined whether the focus-background partition assigned by hearers to (contextless) SVO sentences coincides with the focus-background partition under which the sentences had originally been produced. The results show that the hearers' interpretations coincide with the original focus-background partition in 70% of the responses and that the accuracy rate strongly depends on three variables: *focus type* (contrastive (CF) > information focus (IF)), *focused constituent* (subject > direct object), and *variety spoken by participants* (Argentinean Spanish > Peninsular Spanish). The accuracy ranges from 94% ([subject]$_{CF}$, Argentinean participants) to 43% ([object]$_{IF}$, speakers of Peninsular Spanish). Besides the three above-mentioned factors, we discuss whether stress placement (and sentence form more generally) can be seen as focus marking devices in Spanish. We argue that sentence form is best viewed as a filter which rules out (or makes improbable) certain focus-background partitions. However, contextual cues are often necessary to identify the actual focus-background partition of a sentence.

Keywords. Argentinean Spanish; Peninsular Spanish; focus; word order; nuclear stress; syntax-prosody interface; perception

RESUMEN. La ley de correspondencia entre foco y acento (inglés *focus prominence rule*, FPR) predice que los hablantes articulan sus enunciados de tal manera que el acento nuclear cae dentro del dominio focal (*¿Qué compró Juan?* → *Juan compró* [*una biciCLEta*]$_F$ / #*JUAN compró* [*una bicicleta*]$_F$). Para examinar las consecuencias de la FPR para la interpretación del foco, realizamos un experimento de percepción usando datos orales producidos por hablantes argentinos. Participaron dos grupos de oyentes representando las variedades argentina y peninsular. Examinamos si la partición foco-trasfondo asignada por los oyentes a las oraciones SVO (sin contexto) coincide con la estructura informativa bajo la cual las oraciones habían sido producidas originalmente. Los resultados muestran que la interpretación por los oyentes coincide con la partición foco-trasfondo de los estímulos en el 70% de las respuestas y que la tasa de precisión depende fuertemente de tres variables: *tipo de foco* (foco contrastivo (FC) > foco informativo (FI)), *constituyente focalizado* (sujeto > objeto directo) y *variedad hablada por los participantes* (español argentino >

español peninsular). La precisión oscila entre el 94% ([sujeto]$_{FC}$, participantes argentinos) y el 43% ([objeto]$_{FI}$, participantes peninsulares). Además de los tres factores mencionados, discutimos si la colocación del acento nuclear (y la forma de la oración en general) pueden considerarse dispositivos para marcar el foco en español. Sostenemos que la forma de la oración se ve mejor como un filtro que descarta ciertas interpretaciones de la estructura informativa. Sin embargo, las pistas contextuales son a menudo necesarias para identificar la partición foco-trasfondo.

Palabras clave: español argentino; español peninsular; foco; orden de palabras; acento nuclear; interfaz sintaxis-prosodia; percepción

1. Introduction

This paper addresses a particular aspect of the focus-prosody interface in Spanish: the interaction between the nuclear accent and the information-structural category *focus* (as the locus of alternatives (Krifka 2008: 247)). The focus prominence rule (FPR; Jackendoff 1972) requires that speakers articulate their utterances in such a way that the nuclear stress falls within the focus domain (see (1b) vs. (1b')).[1]

(1)   a.   What did John buy?
        b.   John bought [a BIKE]$_F$.
        b'.   #JOHN bought [a bike]$_F$.

The FPR has been assumed irrespective of the focus type and should hold in the case of both information focus (see (1)) and contrastive focus (see (2)).

(2)   a.   John bought a house, right?
        b.   No, John bought [a BIKE]$_F$.
        b'.   #No, JOHN bought [a bike]$_F$.

Spanish is one of the many languages for which it is assumed that the FPR mediates the relation between focus and nuclear stress, as illustrated in (3b) (see, amongst others, Zubizarreta 1998, Gutiérrez-Bravo 2008, Bosque & Gutiérrez-Rexach 2009, Olarrea 2012).

(3)   a.   (Context: What did Pepín do?)
        b.   #[Llegó   tarde]$_F$ PePÍN.
               arrived   late   Pepín
               'Pepín arrived late.'
               (Bosque & Gutiérrez-Rexach 2009: 682)

Nevertheless, there are hints in the literature which suggest that violations of the FPR may occur in Spanish. First, Calhoun et al. (2018: 18) conducted a production experiment using a picture description task. They report for focused subjects in intransitives in Venezuelan Spanish the following three types of syntactic-prosodic structures: subject-verb order with nuclear stress on the verb ([S]$_F$-**V**; violating the FPR), subject-verb order with nuclear stress on the subject ([**S**]$_F$-V), and verb-subject order with nuclear stress on the subject (V-[**S**]$_F$). It is striking that the most frequently produced structure is the one that violates the FPR, namely [S]$_F$-**V** (see Calhoun et al. 2018: 18). Second,

---

[1] Throughout this paper nuclear stress is indicated by CAPS, unless in abbreviations where it is indicated by bold face (e.g., **dO**).
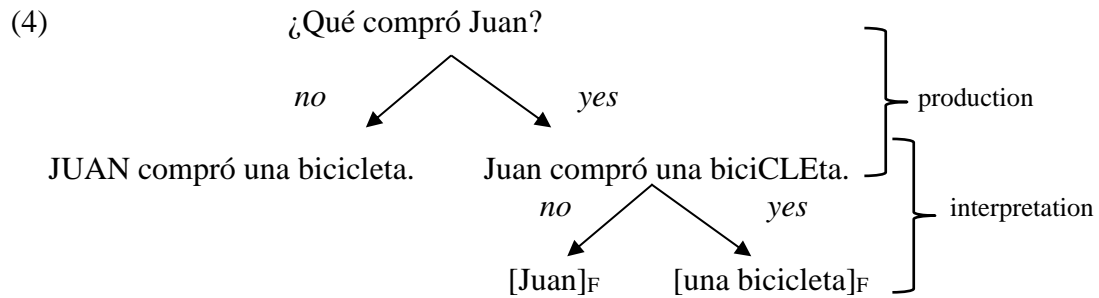
Feldhausen & Vanrell (2015: 48) present experimental data on the prosodic realization of Spanish clefts where the nuclear stress does not fall within the focus. It should be kept in mind, however, that the cases of a mismatch between focus and nuclear stress reported by Feldhausen & Vanrell (2015) concern cleft sentences (which, according to their analysis, consist of two intonational phrases (IP)), and not simple (i.e., non-cleft) sentences. Finally, Hoot (2016) collected acceptability judgements from Mexican raters for different combinations of word orders and stress patterns in the context of wh-questions targeting either the subject or the object as focus. Besides combinations where focus and nuclear stress coincide, he also tested the acceptability of mismatches, i.e., stimuli violating the FPR.[2] The results show that these stimuli receive relatively high scores (3.29 and 3.26 on a 5-point Likert scale with 5 as the best score) (see Hoot 2012, 2017 for similar results). Additionally, in the case of the focused subject, stimuli with a mismatch receive higher scores than stimuli where the focused subject ends up in sentence final position via p-movement. Although Hoot's (2012, 2016, 2017) perception experiments include interesting data on the acceptability scores of mismatches, they do not include information on where the nuclear stress is perceived and whether it guides the hearer's focus interpretation.[3] It is this latter aspect of the FPR that we investigate in our study.

Focus shapes sentence form in Spanish in many ways and we expect that sentence form signals, at least to some extent, the focus-background partition of a given sentence (see the notion *information packaging* (Chafe 1976; Vallduví 1990; Vallduví & Engdahl 1996)). After all one would expect that speakers choose among different possible sentence forms in such a way that they facilitate the hearer's focus interpretation. Therefore, the FPR does not only make a prediction about production but also about perception (here: interpretation) (see (4)).[4]

---

[2] Utterances violating the FPR are not consistently labeled in the literature. While some authors label them as ungrammatical (indicated by the asterisk: *) (e.g., Olarrea 2012: 606), others label them as contextually or pragmatically infelicitous (indicated by the pound sign: #) (e.g., Bosque & Gutiérrez-Rexach 2009: 682 in (3b)). What authors using the pound sign seem to have in mind is that the respective utterance – with its linear order and position of nuclear stress – is perfectly acceptable in other contexts. For example, the sentence in (1b') would be fine in the context of a wh-question targeting the subject. Note, however, that in an OT analysis the respective constraint (e.g., STRESSFOC) would be part of the grammar, and its violation would thus target a grammatical constraint, and violations may therefore lead to ungrammaticality. As concerns the relation between acceptability and contextual felicity, Hoot (2016: 354) interprets the rather high scores of mismatches as indicating a difference between acceptability and contextual felicity (where infelicitous stimuli still score in the mid-range of the acceptability scale, i.e., infelicity does not imply unacceptability).

[3] Another difference is that the stimuli used in Hoot's perception studies were specifically produced for these studies with rather strict criteria regarding their prosodic shape (Hoot 2012: 164–166). By contrast, we applied only minimal criteria in stimuli selection and are interested in the question of to what extent the nuclear stress is still a perceivable cue for focus in the utterances that speakers had produced in the underlying production experiment (see below).

[4] See (1) for translations.

(4)   ¿Qué compró Juan?

no   *yes*   — production

JUAN compró una bicicleta.   Juan compró una biciCLEta.

no   *yes*   — interpretation

[Juan]$_F$ [una bicicleta]$_F$

Given that the position of nuclear stress is relevant for the focus-background partition assigned to a given sentence, we do not only expect it to be realized within the focus domain, but also assume that it is perceivable as such by the hearer and that it guides the hearer's focus interpretation. The respective hypothesis is formulated in (5).

(5)   If a constituent X is the focus of a sentence produced by speaker A, then hearer B perceives the nuclear stress of the sentence within constituent X and discards focus interpretations in which X is not part of the focus.

To test this hypothesis, we conducted a perception experiment with Spanish SVO sentences, produced in contexts of focus on either the subject or the object. The stimuli were originally produced by Argentinean speakers (Gabriel 2012). Besides the overall accuracy (where *accurate* means that the focus-background partition of the participant's response in the perception experiment matches the focus-background partition of the stimulus), we are interested in the impact of three factors on the accuracy rate: (i) focus type (contrastive focus vs. information focus), (ii) focused constituent (subject vs. object), and (iii) the variety of Spanish spoken by the participants (Argentinean vs. Peninsular Spanish).

The paper is organized as follows. In Section 2, we provide the reader with the relevant background and motivate our predictions with respect to the three above-mentioned factors. Section 3 is devoted to the presentation of our empirical study. We first describe our material and methodology (Section 3.1.), before presenting (Section 3.2.) and discussing our results (Section 3.3.). Besides the three above-mentioned factors, we discuss whether stress placement (and sentence form more generally) can be considered a focus-marking device in Spanish. We argue that in terms of information packaging (Vallduví & Engdahl 1996) sentence form is best viewed as a filter which rules out (or makes improbable) certain focus-background partitions, but that contextual cues are often necessary to identify the actual focus-background partition of a given sentence. Section 4 offers some concluding remarks.

2. Background and predictions

The general expectation based on the FPR is that speakers articulate their utterances in such a way that the nuclear stress falls within the focus domain. We also make this assumption with respect to the production data used in our perception experiment. These data were collected in 2008 from a total of 30 speakers of Argentinean Spanish living in Buenos Aires and Neuquén (Northern Patagonia) using an elicited production task (Gabriel 2012). The varieties spoken in the two areas are part of what is commonly referred to as *argentino neutro* (Staudinger & Kailuweit 2018) and pattern alike regarding their overall prosodic shape, i.e., the inventory and use of pitch accents and

boundary tones.[5] While Peninsular Spanish considerably differs from Argentinean Spanish with respect to the alignment properties of pitch accents (see Section 2.3.), the FPR supposedly holds across Spanish varieties. Accordingly, the position of the nuclear stress in the utterance should be recognizable for the participants of the perception experiment and thereby guide their focus interpretation, independently of the variety they speak. We thus expect that the focus choice in the perception experiment correlates with the focus in the underlying stimuli. Although it is difficult to pin down the expected accuracy in quantitative terms, it should be significantly above the chance probability of .5 (in an experiment with two options, see Section 3.1.).

Besides the overall accuracy, we are interested in three factors which we assume to be relevant for the accuracy rate of focus detection: (i) focus type (contrastive focus vs. information focus), (ii) focused constituent (subject vs. object), and (iii) the variety of Spanish spoken by the participants (Argentinean vs. Peninsular Spanish). We discuss these factors successively in Sections 2.1.–2.3.

2.1. Focus type

As concerns the factor *focus type*, we predict that contrastive focus shows a higher accuracy rate than information focus. There is evidence from the literature, mainly on English and Spanish, that contrastive focus is realized with a greater degree of phonetic salience as compared to information focus. This, in turn, entails that listeners perceive contrastive foci as being phonetically more prominent than their neutrally focused counterparts. In an experimental study, Katz & Selkirk (2011) tested the phonetic prominence of contrastive foci and discourse-new constituents in English and showed that the former are phonetically more prominent in terms of duration, pitch, and intensity than the latter. According to the authors, this suggests that "this phonetic prominence relation must be what provides hearers with an essential cue to the presence and location of [contrastive] Focus in English sentences" (p. 806). This view is corroborated by Breen et al. (2010), who showed that speakers of Standard American English produced contrastively focused objects in simple SVO sentences "with a higher maximum intensity, a longer duration and silence, and higher maximum F0" (p. 1061) as compared to non-contrastively focused objects. As concerns the distinction between contrastively and non-contrastively focused subjects and verbs, however, no such acoustic differences were found.

Turning to Argentinean Spanish, which is the variety of the stimuli used in our empirical study, Feldhausen et al. (2011) provided evidence for a categorical scaling contrast in the nuclear position of utterances with narrowly focused clause-final direct objects by applying a categorical perception paradigm complemented with reaction time measurements: while the tritonal pitch accent L+H*+L, which "typically occurs in nuclear position and in utterances with a contrastive or emphatic reading" (Gabriel et al. 2010: 289), was perceived as signaling contrastiveness by the hearers (context question 1; see (6a)), the level tone L* was interpreted as signaling narrow information focus (context question 2; see (6b)).

---

[5] Minor prosodic differences concern the occurrence frequencies of certain surface realizations of high intermediate phrase (ip) boundaries (Feldhausen et al. 2010).

(6) a. Context question 1: Did she give a magazine to her brother?
   (No). Un DIArio.   '(No). A newspaper.'

L+H*L

L+H*+L
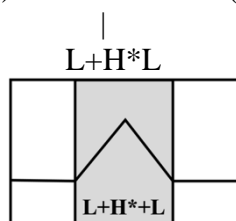
b. Context question 2: What did she give to her brother?
    Un DIArio.   'A newspaper.'

L*

L*
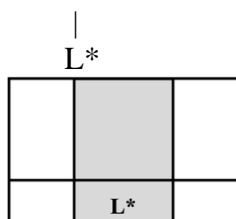
The authors used manipulated stimuli based on data gathered in an elicited production task as shown in Figure 1. Participants were confronted with the (contrastively or non-contrastively focused) object constituents only and asked to classify the stimuli as being felicitous answers in either context 1 (contrastive focus) or context 2 (information focus). Results showed that the stimuli with a higher scaling, i.e., the original "contrastive" production (context 1; (6a)) and its modestly downstepped variants (Figure 1, contours 1a, 1b) as well as the considerably upstepped versions of the original "neutral" production (Figure 1, 2c–e), were perceived as felicitous answers in the contrastive context 1 (6a). In turn, the stimuli exhibiting a rather flat contour, i.e., the original "neutral" production (context 2; (6b)) and its modestly upstepped variants (Figure 1; 2a, 2b) as well as the considerably downstepped variants of the original "contrastive" production (Figure 1; 1c–d), were perceived as appropriate answers in the non-contrastive context 2 (6b). Interestingly, listeners needed more time to make their decisions when presented with modified stimuli, i.e., although stimuli 2c–2e (Figure 1) were unanimously identified as signaling contrastiveness, reaction times were longer as compared to the original "contrastive" stimuli (context 1; (6a)) and its only modestly modified variants (Figure 1; 1a, 1b). This suggests that in addition to the tonal cue (high-scaled pitch accent L+H*+L), non-tonal cues such as voice quality (e.g., a greater degree of sonority) generally associated with emphasis also contribute to the perception of contrastiveness.

Figure 1: *Schematic representation of the manipulations performed on original data in Feldhausen et al.'s (2011) perception experiment. The bold dashed and solid lines symbolize the original contrastive and neutral stimuli, respectively; the solid lines in between represent the stimuli manipulated starting from the original contrastive (1a–e) and neutral (2a–e) recordings.*
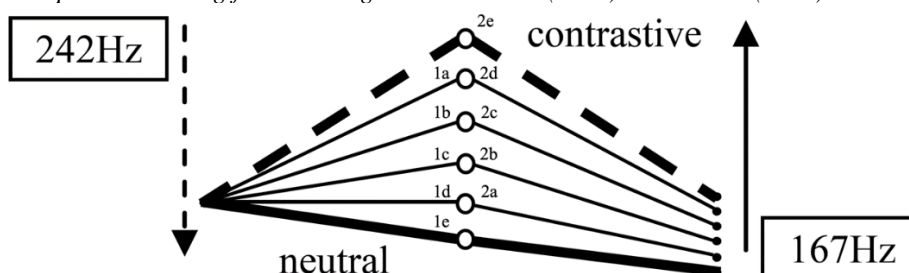
All in all, Feldhausen et al.'s (2011) outcomes suggest that the scaling contrast found in nuclear position (i.e., L+H*+L for a narrow contrastive focus on the object and L* for narrow information focus on the object) is rather a signal of emphatic prosody, which, in turn, is more likely to occur with contrastively focused items. This view is in line with Downing & Pompino-Marschall (2013), who claim, mainly based on data from the Bantu language Chichewa, "that the focus prosody reported is actually emphasis prosody" (p. 657). The findings of a quantitative study on pitch accent distribution in Buenos Aires Spanish by Gabriel et al. (2013: 107) point to the same direction: about two thirds (65%) of the 23 short exclamative sentences of the type exemplified in (7), which were gathered in an elicited production task, are produced with the nuclear configuration L+H*+L L% and thus pattern with declarative sentences with a contrastively focused XP in clause-final position.

(7) ¡Mmmh! ¡Qué ricas   mediaLUnas!
     yummy   what delicious croissants
     'Yummy! What delicious croissants!'

Contrastive focus seems to be nothing more than one of the factors besides obviousness, amazement, etc., which amount to using the expressive means of emphatic prosody. Given the assumption that contrastiveness is not obligatorily expressed through the means of emphatic prosody, but that its use is only more likely in contrastive than in non-contrastive contexts, it comes as no surprise that in some production studies no tonal differences were found between contrastive focus and (narrow) information focus (see, e.g., Hanssen et al. 2008 for focused objects in Dutch).

Note, however, that the information-structural distinction between clause-final narrow focus and broad focus (which also bears final nuclear stress, surfacing as an L* level tone in Argentinean Spanish) was not addressed by Feldhausen et al. (2011), nor was a possible prosodic distinction between narrow contrastive focus and information focus located in the canonical preverbal subject position considered. However, Vanrell et al. (2013) investigated the relevance of three prosodic parameters (alignment, duration, and scaling) in the conveyance of contrastive focus in Catalan, Italian, and Spanish, and showed that in Madrid Spanish the L+H pitch accent located on clause-initial subject XPs which convey a contrastive interpretation as in (8) is produced with a considerably earlier alignment by all speakers and with a higher scaling by at least two out of the five speakers recorded (Vanrell et al. 2013: 205, 207). Note that, although in these cases the F0 contour falls after the peak already within the temporal limits of the stressed syllable, the focal pitch accents are clearly L+H* accents and no tritonal one since the subsequent L target is reached only in the posttonic syllable (and not within the tonic one as is the case for the Argentinean L+H*+L accent, discussed above).

(8) (No.) MaRIna vendrá   mañana.
     No   Marina   will.come tomorrow
     '(No.) Marina will come tomorrow.'

A similar result had been achieved by Face (2001) in an earlier study on the intonational marking of contrastive focus in Madrid Spanish: the pitch accents produced on contrastively focused constituents in non-final position were consistently produced with early-peak alignment (L+H*), whereas in neutral cases prenuclear pitch accents surface

with the F0 peak located in the posttonic syllable (late-peak alignment, i.e., L+>H*).[6] The prosodic distinction between narrow information focus and contrastive focus in clause-initial position was corroborated from a perception perspective in earlier work by Vanrell et al. (2011), who, using a gating paradigm, showed that Catalan, Italian, and Spanish hearers need not listen to the post-focal region to distinguish between two focal conditions (information focus vs. contrastive focus): their participants were able to recognize the presumed contrastiveness of the pitch accent located on the non-final constituent just by listening to the portion going from the beginning of the sentence to three quarters of the way through the syllable bearing the focal pitch accent. Interestingly, in all of the pitch contours presented by Vanrell et al. (2013: 203), a clear intermediate phrasal boundary (L-) is visible after the clause-initial focused XP. The same has been shown for the prosodic shape of declaratives with contrastively focused subjects in clause-initial positions in Argentinean (Pešková et al. 2011: 92) and in Peninsular Spanish (Hualde 2014: 272). This strongly suggests that in the case of a non-final focused constituent post-focal compression and intermediate phrasing, including a clearly discernible prosodic break following the focused material, are more reliable cues of focus than pitch accent realization (on the relevance of post-focal compression see also Féry 2017: 152–156). Interestingly, this view is also confirmed by El Zarka et al. (2020) for Egyptian Arabic, where the tonal shape of post-focal stretch even allows to distinguish between preceding contrastive or information focus: "while contrastive foci themselves were scaled even lower than non-contrastive foci, their higher relative prominence seems to be achieved by post-focus compression and additional register lowering after focus" (El Zarka et al. 2020: 1886). However, as shown by Face (2002: 47–48), post-focal compression (or: deaccentuation) is not compulsory in Madrid Spanish, i.e., pitch accents may also surface as clearly perceivable pitch excursions in the post-focal domain, which in this case does not present an entirely flat F0 contour.

2.2. Focused constituent

As concerns the factor *focused constituent*, it is not obvious whether to expect subjects or direct objects to show a higher degree of accuracy. Below, we will discuss the potential impact of prosody and focus affinity on the accuracy of focused subjects and objects.

Two essential points follow from what was described in Section 2.1. regarding the prosodic marking of neutrally and contrastively focused subjects and objects: first, in both varieties considered in this paper, i.e., Argentinean and Peninsular Spanish, contrastively focused constituents are prosodically marked with higher scaling or earlier peak alignment as compared to the same constituents in narrow information focus and in broad focus (Argentinean Spanish: L+H*+L instead of L+H* for clause-initial subjects and L+H*+L instead of L* for clause-final objects; Peninsular Spanish: L+H* instead of L+<H* for subjects and L+H* instead of L* for objects). Considering only the shape of pitch accents, however, the contrast between the two focus types is expressed more saliently in clause-final (i.e., contour tone L+H*+L/L+H* vs. level tone L*) than in clause-initial position, where it is only mirrored in the alignment properties of a rising (or rising-falling) contour tone (see Section 2.3.). Second, the low intermediate phrasal boundary (L-) following the narrowly focused XP, regardless of whether we are dealing with contrastive focus or information focus, seems to be a quite reliable

---

[6] The interrelation between earlier early peak alignment, contrastively, and a greater degree of perceived salience is corroborated by Cole et al. (2019: 127) who showed for Peninsular Spanish that rising pitch accents with early peak alignment (L+H*) are more likely to be perceived as prominent than their late-aligned counterparts (L+<H*).

tonal surface cue of focus (facultatively followed by post-focal compression). This post-focal cue only appears with non-final constituents, such as the clause-initial subjects, but not with objects in clause-final position. Sentence-final objects can also be marked as narrow focus by optional intermediary phrasing (H-) preceding the focused object (see Hualde 2014: 269–271 on Peninsular Spanish and Pešková et al. 2011 on Argentinean Spanish). However, at least in Argentinean Spanish this procedure of prosodic focus marking (H- boundary) occurs much less frequently (31% of the cases analyzed) than the L- boundary following the clause-initial subject (100%; Pešková et al. 2011: 93–94). This suggests that in SVO constructions focused subjects are more easily recognized as such than is the case for focused objects.

In addition to the prosodic features of focused subjects and objects laid out above, we must also consider that subjects and objects differ with respect to focus affinity, i.e., their overall likelihood to be the focus of a sentence (see Heidinger 2018, 2021; Heidinger & Onea 2021). This difference in focus affinity, however, could have opposite effects in the perception experiment. Let us first have a look at the focus affinity of subjects and objects and then consider the possible consequences for the accuracy of focus choice.

In line with existing literature, we assume that direct objects are better candidates for focus than subjects, i.e., direct objects have a higher degree of focus affinity (Bossong 1984a, 1984b; Firbas 1992; Lambrecht 1994; Drubig 2003; Wunderlich 2006; Dufter 2007; Dufter & Gabriel 2016). Further, it has been argued that the VP or the predicate is the default focus of a sentence (Bossong 1984a; Lambrecht 1994; Drubig 2003; Dufter 2007; Dufter & Gabriel 2016); hence the unmarked focus-background partition of a clause with subject, verb, and object would be as in (9). This asymmetry again makes objects more likely foci than subjects.

(9)   John [bought a bike]$_F$.

Further evidence comes from research on *preferred argument structure* (Du Bois 1987), which refers to the information status and morphosyntactic form of arguments. In this line of research, it has been observed that objects express new information (and not given information) more often than subjects, and that objects are more often expressed as lexical NPs (and not as pronouns) than subjects (see Ashby & Bentivoglio 1993 for Spanish). Since new and lexical constituents are more often focus than given and pronominal constituents, we can take this as further evidence that objects are more likely to be focus than subjects.

As concerns the impact on the accuracy of focus choice, the difference in focus affinity between subjects and objects does not make a clear prediction as to which constituent type (subject or object) will have the higher degree of accuracy. Under the assumption that focus affinity negatively correlates with the effort to mark a grammatical role as focus (see also Leonetti & Escandell-Vidal 2021: 160), we expect that subject focus has a stronger prosodic effect than object focus. In terms of markedness, subject focus is conceptually more marked than object focus, and we therefore expect it to be also formally more marked, as outlined above and in Section 2.1 regarding intonation (see also the high frequency of cleft subjects as compared to cleft objects in the data analyzed by Gabriel 2010). This would predict that in SVO sentences subject focus is encoded in a more salient (or: recognizable) way than object focus. However, the asymmetry in focus affinity may also have the opposite effect if the hearer does not follow the prosodic cues, but instead interprets the focus-background partition based on the expectation that objects are generally more likely to be focus than subjects. In the latter

case, objects would be chosen more often as focus than subjects. As we would like to refrain from unmotivated *a priori* predictions, we acknowledge that our study is exploratory when it comes to the impact of the focused constituent on accuracy.

2.3. Variety spoken by participants

As outlined in Sections 2.1. and 2.2., the two varieties of the experiment's participants do not differ substantially as to the prosodic marking of narrowly focused constituents: in both Peninsular and Argentinean Spanish focused clause-initial subjects, which are obligatorily followed by a L- intermediate boundary, are more reliably marked as focus through prosodic phrasing (pattern: $([S]_F)_{L-}(VO)$) as compared to focused clause-final objects, which are facultatively separated by a H- boundary from the preceding background material (pattern: $(SV)_{H-}([O]_F)$). Peninsular and Argentinean Spanish also pattern alike as to the intonational marking of contrastive focus, which in both varieties exhibits a greater degree of salience due to higher scaling and earlier alignment of the focal pitch accent as compared to information focus. However, the overall inventory of pitch accents and boundary tones differs considerably between the two varieties – hence the question of whether speakers of different varieties of the same language show different rates in focus recognition when confronted with the same stimuli. While in Peninsular Spanish prenuclear accents, at least in neutral declaratives, regularly surface with late-peak alignment (L+<H*), their Argentinean counterparts are produced with an early peak (L+H*) in the unmarked case. For contrastively focused constituents, the pitch peak is shifted to the left in each case, resulting in L+H* for Peninsular Spanish and in L+H*+L for the Argentinean variety, regardless of whether the focused XP is located in clause-initial or in clause-final position (see the previous sections and Estebas-Vilaplana & Prieto 2010; Gabriel et al. 2010; Hualde & Prieto 2015). As to the tonal correlates of intermediate phrasing, the overwhelming majority of high ip phrasal boundaries (underlying /H-/) is realized as a continuation rise in Peninsular Spanish, i.e., as continuative F0 rise from the last pre-boundary pitch accent until the boundary (88% of the cases analyzed by Frota et al. 2007), whereas the variety of Argentinean Spanish (Buenos Aires) addressed by Gabriel et al. (2011) is characterized by an even distribution of the continuation rises (42%) and sustained pitches, where the F0 contour forms a high plateau that continues until the break (32%). A further major difference between Peninsular and Argentinean Spanish intonation refers to the tonal shape of neutral yes-no questions, which show a globally rising pattern in Peninsular and a globally falling one in Argentinean Spanish (Estebas-Vilaplana & Prieto 2010; Gabriel et al. 2010; Hualde & Prieto 2015). This aspect, however, is not relevant in the context of our perception experiment, which includes only declarative sentences as stimuli (see Section 3).

As concerns the variety spoken by the participants, we expect a higher accuracy rate for the Argentinean than for the Spanish subjects. This is motivated by the fact the underlying stimuli were produced by native speakers of Argentinean Spanish, and our expectation is that the underlying prosodic patterns signaling the focus-background partition of a given sentence are more easily recognized by listeners speaking the same variety (Argentinean Spanish) than by listeners speaking a variety in which different surface cues in terms of the alignment properties of pitch accents and surface correlates of underlying boundary tones are used (Peninsular Spanish). This assumption nevertheless needs empirical validation, especially in the face of studies like El Zarka & Hödl (2021), who report for a similar task using Arabic stimuli that German native speakers scored better than Arabic native speakers.

3. Empirical study

In the following, we present our empirical study. We first describe our method and material (Section 3.1.), before presenting the results in Section 3.2 and discussing them in Section 3.3.

3.1. Method and material

3.1.1. Task

We hypothesize that in narrow-focus sentences the perceived nuclear stress lies within the focus domain (see Section 1). To test this hypothesis, a forced-choice perception experiment with auditory stimuli was carried out. Participants listened to a target sentence as an audio file and were asked to indicate which of two wh-questions was more suitable as the preceding context of the sentence they had heard (see Figure 2). One wh-question targeted the subject as a narrow focus, the other one targeted the direct object. Participants were instructed to listen to the stimuli not more than three times, but response time was not limited (i.e., the experiment was self-paced). Participants were also instructed to use headphones.

Figure 2: Task and example stimulus.
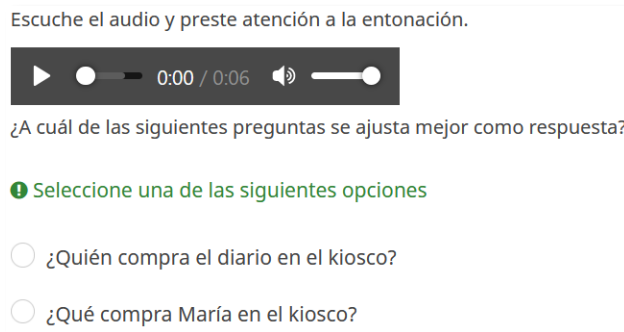
We assumed that the participants' choice between the two questions would indicate their focus interpretation of the target sentence, and hence the position of the perceived nuclear stress. Since we knew about the stimuli's focus-background partition we were able to examine whether the perceived nuclear stress in the participants' answers coincided with the focus of the target sentence. Given that the FPR predicts that the nuclear stress is perceived within the focus, we took the participants' choice of the respective question as indication of their focus interpretation, i.e., the *quién* question in the case of subject focus and the *qué* question in the case of object focus.

3.1.2. Design and conditions

Three factors (independent variables) were considered in the experimental design: *focus type* and *focused constituent* as within-subjects factors, and *variety spoken by participants* as a between-subjects factor. The factor *focus type* had two different levels, namely *contrastive focus* and *information focus*. The factor *focused constituent* also had two levels: *subject* and *direct object*. Therefore, the total number of experimental conditions amounts to four, as shown in Table 1.

Table 1:   Experimental conditions.

		focused constituent	
		subject	direct object
focus type	contrastive	1	2
	information	3	4

Since the order of constituents was kept constant in the unmarked SVO order, the four conditions had the structures as shown in (10).[7] Note, however, that in some stimuli for Conditions 1 and 3 the direct object is followed by some other postverbal constituent (locative adjunct or indirect object).

(10)   Condition 1: $[S]_{CF}$-V-dO
         Condition 2: S-V-$[dO]_{CF}$
         Condition 3: $[S]_{IF}$-V-dO
         Condition 4: S-V-$[dO]_{IF}$

Regarding the between-subjects factor *variety* (spoken by participants), we distinguished between two groups of participants based on the variety they spoke and their current country of residence, the two levels of this factor being *Peninsular Spanish* and *Argentinean Spanish* (in the following we will sometimes use *Spain* and *Argentina* as proxies for the two varieties spoken by the participants).

The dependent variable was the participants' choice between two wh-questions, based on which question he or she considered more suitable as the preceding context of the sentence (see Figure 2).[8] On the basis of this choice, we calculated the accuracy, i.e., the rate of correspondence between the focus in the respective stimulus and the focus as indicated by the participants' choice.

3.1.3. Stimuli

The auditory stimuli of the experiment stem from a picture-based production experiment carried out in December 2008 in Argentina as part of larger data collection (Hamburg Corpus of Argentinean Spanish (HaCASpa); PID: http://hdl.handle.net/11022/0000-0000-5F0B-B, see Gabriel 2012). The stimuli were produced either as answers to narrow wh-questions (see (11)) or as corrective replies in contrastive contexts (see (12))

---

[7] A reviewer raised the question of whether the non-final position of the subject might rule out the wh-question targeting the subject as a preceding context for the Peninsular participants (because they would prefer p-movement in the case of information focus on the subject). We considered this possibility but deem it very unlikely for the following two reasons. First, Vanrell & Fernández-Soriano (2018: 48) report *in situ* focus on the subject as the most frequent strategy in "Castilian Spanish". Admittedly, Feldhausen & Vanrell (2015: 13) present results showing that for subject information focus clefts and p-movement are more frequent than other strategies (e.g., preverbal *in situ*), but it is not clear whether they refer to sentences with lexical objects (i.e., the sentence type used in our experiment). Second, in pretesting (see below) participants from Spain had the chance to go with a contrastive context targeting the subject as focus, but still preferred the wh-question targeting the subject as the preceding context.

[8] Note that in pretesting we used a total of four contexts, i.e., one additional context for contrastive focus on the subject and one for contrastive focus on the object. Since the participants chose such contrastive preceding contexts in less the 5% of the answers in pretesting, we decided to reduce the options in the actual experiment to two wh-questions, as shown in Figure 2. Participants from pretesting did not participate in the final experiment reported in this paper.

by a total of 60 speakers (ages: 18–69), 32 from Buenos Aires and 28 from the Comahue area (Neuquén, Northern Patagonia).

(11)   ¿Quién   compró el   diario?   (wh, eliciting IF)
       who   bought   the   newspaper
       'Who bought the newspaper?'

(12)   Julia compró el   diario,   ¿verdad?   (correction, eliciting CF)
       Julia bought   the   newspaper right
       'Julia bought the newspaper, right?'

In the underlying production experiment, the participants could freely answer without restrictions on the syntactic or prosodic form of the answer (they were instructed, however, to answer in full sentences). Therefore, the answers show a great deal of syntactic and prosodic variation, and the number of stimuli that survived preselection was not the same for all four conditions. In particular, the data show great syntactic variation (focus-only answers, clefts, movement of focus to final position, focus fronting, etc.). For the present experiment, we only considered answers that had an SVO order with both subject and object being overtly expressed as lexical NPs. Besides this syntactic preselection, the potential stimuli (i.e., the answers from the production experiment) were preselected on prosodic criteria (items with strong hesitation were excluded). While sufficient stimuli were available for Conditions 3 and 4, we were left with 22 stimuli for Condition 1 and with only 11 for Condition 2. Based on these numbers we opted for a set up and distribution of stimuli among participants as described in Section 3.1.4.

A further clarification is necessary as concerns the stimuli produced in contrastive contexts (used in Conditions 1 and 2). We presented the stimuli which were produced in contrastive contexts in the underlying production experiments without the occasional *no*. Originally, we had planned to let participants choose from four types of preceding contexts: wh-question targeting subject, wh-question targeting object, contrastive context for subject, and contrastive context for object. Since the occasional *no* would already trigger a contrastive interpretation, we did not include it in the sound files used as stimuli. Pretesting showed, however, that contrastive contexts were generally avoided as preceding contexts and we therefore only presented the two wh-questions as preceding contexts. With only wh-questions as options for the preceding context, the *no* would not be appropriate in the answers.

3.1.4. Experimental set-up

A total of 90 persons participated in the experiment. 40 monolingual native speakers of Argentinean Spanish were recruited among the academic community (students and faculty members) of the Universidad Nacional del Comahue (Neuquén, Argentina), with the age ranging from 18 to 50 years (average = 28 years). In addition, 50 monolingual native speakers of Peninsular Spanish (21–55 years; average = 32 years) were recruited for the experiment via Prolific Academic (https://www.prolific.co/) and received a compensation of £ 2.5 for their participation. The participants were unfamiliar with the experiment's purpose and the underlying concepts.

As shown in Table 2, participants were assigned to one of two lists (A and B; 20 participants from Argentina saw list A and 20 saw list B; 25 participants from Spain saw list A and 25 saw list B). For Condition 1, all 22 stimuli were used and assigned randomly to list A or B. For Condition 2, all 11 stimuli were used and assigned to both

lists. For Condition 3, we randomly selected 11 stimuli with L- and 11 with H- and then randomly assigned them to A and B.[9] For Condition 4, finally, we selected all 5 stimuli with H- preceding the direct object and randomly selected 17 without ip-boundary. Hence, a total of 77 different stimuli were used in the experiment. Each participant saw each of the four conditions 11 times, which amounts to 44 stimuli for each participant.

Table 2:   Lists, conditions, and number of stimuli.

Condition		A	B
1	[S]$_{CF}$-V-dO	11	11
2	S-V-[dO]$_{CF}$	11 = 11	
3	[S]$_{IF}$-V-dO	11	11
4	S-V-[dO]$_{IF}$	11	11
		44	44

The experiment consisted of three parts and was presented in a web-based environment using the experimental software Limesurvey. In the first part, participants received explanations and legal information and answered demographic questions. The second part, a short training phase, allowed participants to get accustomed to the task of the actual experimental trials. Finally, in the third and crucial part, each participant saw 44 trials on separate pages. Experimental items were manipulated according to the factorial design discussed above, resulting in a total two different sets of trials, i.e., two different lists. The lists contained control trials but no distractors. Each participant saw one list, and each list was seen by 45 participants (20 from Argentina and 25 from Spain). The order of trials as well as the order of the two questions was randomized. After the last trial, participants from Spain returned to Prolific Academic to receive their reward and to get further information about the experiment.

3.2. Results
3.2.1. Descriptive statistics
We collected a total of 3960 judgments (990 for each of the four conditions) on which of the two questions was a more suitable preceding context. No participant had to be excluded based on the control items, and the data from all 90 participants is considered in the subsequent analysis.

First, we examined how often the focus of the stimulus (as controlled in the production experiment) and the focus according to the selected question coincide. Across all 90 participants, the average accuracy amounts to 70.3% (SD = 11.2). Thus, on average the perceived focus corresponds to the focus of the stimulus in 70.3% of the answers. However, the accuracy varies to a considerable degree once we take into consideration the factors *focused constituent*, *focus type*, and *variety* (Figure 3).

---

[9] Note that the original production data for information focus on subjects consists of 21 items with L- and 36 with H-. Hence, prosodic realizations with the unexpected (or even "non-canonical") H- are more frequent than realizations with L-. If we would have chosen only items with L- for the perception experiment, we would have promoted the less frequent type to the only type for this condition. By choosing 11 items for both L- and H- we wish to do justice to the variation in the prosodic realization of narrow information focus on subjects.

Figure 3: Accuracy of focus choice in % (three factors). (S = subject, dO = direct object; CF = contrastive focus, IF = information focus; Arg. = Argentina).
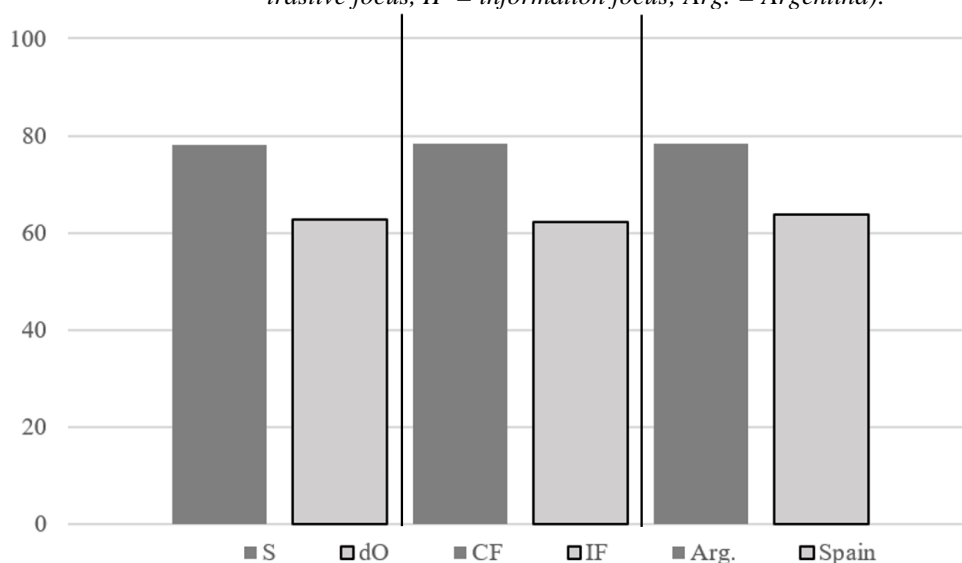
As concerns the factor *focused constituent*, the average accuracy is higher for stimuli with subject focus than for those with object focus (78.0% (SD = 17.8) vs. 62.6% (SD = 23.8)). For *focus type*, the average accuracy is higher for stimuli with contrastive focus than for those with information focus (78.3% (SD = 14.4) vs. 62.3% (SD = 10.7)). Finally, as concerns *variety*, the average accuracy is higher for Argentinean participants than for participants who are speakers of Peninsular Spanish (78.4% (SD = 7.8) vs. 63.8% (SD = 9.2)).

In the following, we present the results separately for the two groups of participants (Argentina vs. Spain). As shown in Figure 4 and Table 3, in three out of the four conditions the accuracy is considerably higher for Argentina than for Spain. Only in the condition with an information focus on the subject, the accuracy is slightly higher for Spain than for Argentina. The results also show a remarkable difference between the two groups. While Spanish participants generally show a higher accuracy in subject stimuli than in object stimuli, Argentinean participants generally show a higher accuracy in contrastive focus than in information focus stimuli (we will come back to this difference in Section 3.2.2.).[10]

---

[10] This difference between the two groups also becomes apparent if we rank the four conditions based on their accuracy (we will see in Section 3.2.2. that the difference between $[dO]_{IF}$ and $[S]_{IF}$ for Argentinean participants is not statistically significant):

(i)   Conditions ranked by accuracy

      Argentina: $[S]_{CF} >$ $[dO]_{CF} >$ $[dO]_{IF} >$ $[S]_{IF}$

      Spain: $[S]_{CF} >$ $[S]_{IF} >$ $[dO]_{CF} >$ $[dO]_{IF}$

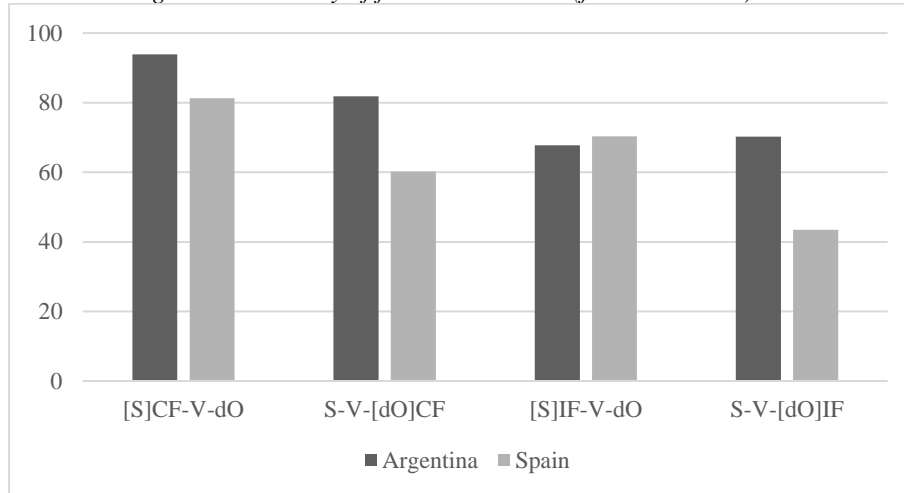*Figure 4: Accuracy of focus choice in % (four conditions).*


Table 3: Accuracy of focus choice in %.

Conditions	Argentina	Spain
1 $[S]_{CF}$-V-dO	94	81
2 S-V-$[dO]_{CF}$	82	60
3 $[S]_{IF}$-V-dO	68	70
4 S-V-$[dO]_{IF}$	70	43
average	78	64

As shown in Table 3, the conditions and two groups show considerable variation with respect to accuracy. Accuracy reaches from 94% in the case the of stimuli produced with narrow contrastive focus on the subject judged by the Argentinean participants to 43% in the case of the stimuli produced with narrow information focus on the direct object judged by the participants from Spain. The average percentage of accuracy across all four conditions and both groups of participants amounts to 70%. This shows that the stimuli are far from unequivocal with respect to focus interpretation, despite being produced in contexts of narrow focus. Given that the participants of the perception experiment only had two focus interpretations to choose from, the accuracy of only 70% still stands above the probability of .5. However, in terms of information packaging and the signaling of the focus to a hearer, it seems that the participants of the underlying production experiment had not done such a great job: in the SVO sentences used as stimuli, the prosodic form does not clearly signal the focus. We will come back to this point in Section 3.3.2. of the discussion.

3.2.2. Inferential statistics

To further analyze the impact of the three factors, i.e., *focus type* (contrastive vs. information), *focused constituent* (subject vs. direct object), and *variety* (Argentina vs. Spain), we conducted a mixed binominal regression analysis (model fitting was performed in R using lme4 package by Bates et al. (2015)). This regression analysis determines the impact of these factors on the dependent variable, i.e., the accuracy of the participants' choice (0 = mismatch, 1 = match). The model includes *focus type*, *focused constituent* and *variety* as fixed factors, and *subject* (= participant) and *item* as random factors. Further, we included *list* as a fixed factor to determine differences related to the two lists (A and B). However, since the lists did not have a significant impact, we ignore this factor in the following analysis.

A generalized linear mixed model (see Model 1 in appendix) with the factors *focus type*, *focused constituent*, and *variety* as fixed factors, and *subject* (= individual participant) and *item* as random factors shows a highly significant interaction (β = -1.55980, p < 0.0001), where *variety* interacts with the impact of *focus type* and *focused constituent*. As shown in Figure 3, Argentinean participants generally have higher accuracy rates than participants from Spain, contrastive focus shows higher accuracy than information focus, and subjects are accurately identified as focus more often than objects. However, the interaction between focus type and focused constituent is not the same in the two groups of participants (as illustrated by the distinct line patterns in Figure 5).

Figure 5: Accuracy of focus choice in % (Argentina vs. Spain).
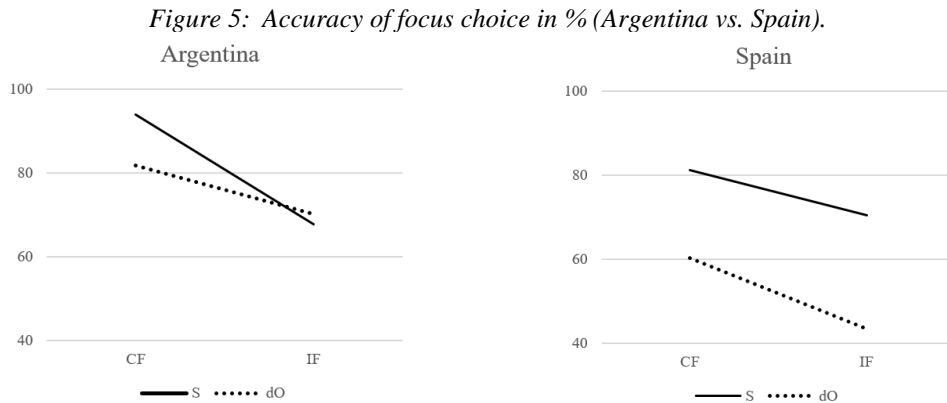
Given the difference between Argentinean and Spanish participants, we built separate generalized linear mixed models with the factors *focus type* and *focused constituent* for the two groups (see appendix for Model 2 (Argentina) and Model 3 (Spain)). As for the Argentinean participants, *focus type* and *focused constituent* interact in such a way that the probability of giving a matching response in the case of subject focus is significantly higher only in the case of contrastive focus (β = 1.4627, p < 0.0005) (see Model 2A in appendix) but not in the case of information focus (β = 0.1764, p = 0.69937) (see Model 2B in appendix); i.e., the quantitative difference between 68% for $[S]_{IF}$ and 70% for $[dO]_{IF}$ is not significant. As for participants from Spain, focus type and focused constituent do not interact (see the parallel lines in Figure 5), but there are highly significant main effects for focus type (CF > IF; β = 0.7781, p < 0.00567) and focused constituent (S > dO; β = 1.2661, p < 0.0001).

For the sake of completeness, we also report the results for a model that ignores the above-mentioned interactions (see Model 4 in appendix). We observe highly significant main effects of *focus type* (CF > IF; β = 0.9472, p < 0.0001), *focused constituent* (S > dO, β = 0.9385, p < 0.0001), and *variety* (Argentina > Spain; β = -0.9927, p < 0.0001). As concerns the strength of the effect, *focus type* has the highest coefficient (0.947), followed by *focused constituent* (0.9385) and *variety* (-0.883).

3.3. Discussion

The results show that all three factors (*focus type*, *focused constituent*, and *variety*) have an impact on the accuracy of focus choice. In this section, we will first discuss this result against the background our initial predictions (Section 3.3.1.), and then move on to two more specific issues: the relation between information packaging and the FPR (Section 3.3.2.), and the impact of the prosodic shape of individual stimuli on accuracy (Section 3.3.3.).

3.3.1. The three factors

As concerns the difference between contrastive and information focus, we predicted in Section 2.1. that the accuracy in focus choice will be higher with contrastive than with information focus. The results confirm this prediction as accuracy is indeed higher for stimuli with contrastive focus (78.3%) than for those with information focus (62.3%). Our prediction was motivated by existing empirical studies (see Section 2.1. for details) showing that contrastive focus is realized with greater phonetic salience, which in turn facilitates focus detection. The manifestation of the phonetic salience may vary from language to language and even among the varieties of individual languages. In English, for example, contrastive focus has an effect on duration, pitch, and intensity (Katz & Selkirk 2011). In Spanish, contrastive and non-contrastive accents differ with respect to scaling (Feldhausen et al. 2011; Vanrell et al. 2013) and alignment (Face 2001; Vanrell et al. 2013). These prosodic differences have an impact on perception in that hearers can distinguish between contrastive and non-contrastive foci based on prosodic cues (Vanrell et al. 2011).

Our results contribute to this line of research as they provide a new type of relevant data: in addition to studies showing that contrastive and non-contrastive foci have different prosodic features in production, which allow to discriminate between them in perception, our study shows that the two focus types also differ with respect to the accuracy of focus detection. The results thus solidify the status of Spanish as a language where the opposition between contrastive and non-contrastive foci has perceivable prosodic reflexes. This is a feature which distinguishes Spanish from languages such as Dutch where no such differences between contrastive and non-contrastive focus have been detected (Hanssen et al. 2008).

As concerns the focused constituent, recall from Section 2.2. that the difference in focus affinity between subjects and objects (objects being more likely to be focus than subjects) makes no clear predictions for accuracy rate. Our experimental results, however, show that the average accuracy is higher for stimuli with subject focus than for those with object focus (78.0% vs. 62.6%) (see Section 3.2.). This suggests an indirect effect of focus affinity on the recognition of focus: focus affinity negatively correlates with the effort to mark an XP bearing a specific grammatical role as focus, and therefore subject focus has a more salient encoding than object focus. The opposite effect, namely that the hearer/participant interprets the focus-background partition based on the expectation that objects are generally more likely to be focus than subjects (Bossong 1984a, 1984b; Firbas 1992; Lambrecht 1994; Drubig 2003; Wunderlich 2006; Dufter 2007; Dufter & Gabriel 2016) is not supported by our data. Crucially, focus affinity does not overwrite what is signaled by prosody (pitch accent types, frequent post-focal compression, and ip boundary tones in the case of prefinal foci vs. rare cases of pre-focal ip boundary tones in the case of clause-final foci).

There is one exception to the general tendency that subjects exhibit higher accuracy than objects: Argentinean participants show higher accuracy for $[dO]_{IF}$ than for $[S]_{IF}$. Although the difference between 68% for $[S]_{IF}$ and 70% for $[dO]_{IF}$ is not statistically significant (see Section 3.2.2.), already the lack of a significantly higher accuracy of $[S]_{IF}$ than $[dO]_{IF}$ is noteworthy. It comes as no surprise that this concerns information focus and not contrastive focus, given that the former is prosodically less salient than the latter (see Section 2.1. and the results of our experiment). In the absence of a clear prosodic cue for subject focus, the higher degree of focus affinity of direct objects might increase the choice of objects in the case of stimuli with underlying subject focus. It remains puzzling, however, why we find this effect only with Argentinean but not with Spanish participants.

The results also highlight methodological challenges. Focus candidates must always be presented in a certain syntactic position (in our case: clause-initial subjects and clause-final objects) and in a certain (morpho)syntactic shape (lexical NPs), but these choices are not innocent. For example, being expressed as a lexical NP presumably has different interpretational effects in the case of subjects than in the case of objects. The frequent use of pronouns as subjects (see Ashby & Bentivoglio 1993 for Spanish) makes a lexical subject stand out. Since we could not compare subjects and objects independently of their (morpho)syntactic form, we do not interpret our results as evidence for a higher degree of focus affinity of subjects as compared to objects. Instead, our data show that focused subjects are expressed in a way that makes them more easily recognizable as the focus than in the case of focused objects.

As concerns the last factor, namely *variety spoken by participants*, we predicted that the average accuracy would be higher for Argentinean than for Peninsular Spanish speakers. The results confirm this prediction (78.4% for Argentinean participants vs. 63.8% for participants from Spain). The only exception to the higher accuracy of Argentinean participants is found with Condition 3 ($[S]_{IF}$-V-dO) where participants from Spain show a slightly higher accuracy (the difference is, however, not statistically significant).

The fact that the overall difference in accuracy is statistically significant but not drastic is not surprising since the two varieties differ in the specific inventory of pitch accents and boundary tones but share basic features of prosodic focus marking (see Section 2.3.). The commonalities are a frequent L- intermediate boundary following clause-initial subjects, a facultative H- boundary preceding focused clause-final objects, and a stronger intonational marking of contrastive focus as compared to information focus. The differences concern peak alignment in neutral declaratives (L+<H* in Peninsular vs. L+H* in Argentinean Spanish) and in contrastively focused constituents (L+H* for Peninsular and L+H*+L for the Argentinean Spanish). The results from our study suggest that such differences in the prosodic inventory do have an impact on the recognition of prosodic prominence and therefore on the accuracy of focus choice.

3.3.2. Information packaging and the FPR

It is generally assumed that sentence form is strongly connected to information structure (see Drubig & Schaffar 2001; Drubig 2003; Büring 2009; Krifka & Musan 2012 for crosslinguistic evidence), both from a production and a processing perspective (see notions such as *information packaging*, Section 1). The speaker chooses a sentence form depending on the sentence's information structure. In turn, the hearer interprets the information structure of a sentence based on its form. In spoken Spanish, information packaging may take many forms: e.g., prosody, word order variation, clefts, and demotion or deletion of non-focal material (see the recent overview in Leonetti & Escandell-Vidal 2021, and especially Cassarà 2021 and Heidinger 2022 on the frequency of deletion). In our experiment, we only considered SVO constructions and thus kept the syntactic form of the test sentences constant. Hence, we cannot make general claims about information packaging in Spanish. Our data is nevertheless instructive with respect to prosody, and more specifically the position of nuclear stress as a means to indicate the focus of a sentence.

Recall from Section 1 that the FPR strongly restricts the mapping between focus-background partition and nuclear stress both in production and perception. In production, the nuclear stress must fall within the focused constituent (abstracting away from other restrictions). In a sentence with a focused subject, a verb, and a direct object (see

(13a)), the FPR requires that the nuclear stress be on the subject (see (13b)), while nuclear stress on the verb or the direct object are infelicitous (see (13b') and (13b'')).

(13) Production
a.   [John]$_F$ + bought + a bike

FPR

b.   [JOHN]$_F$ bought a bike.
b.'   #[John]$_F$ BOUGHT a bike.
b.''   #[John]$_F$ bought a BIKE.

In perception/interpretation, we assume the inverse process, but the FPR still restricts the available interpretations. In a sentence with a subject, a verb, and a direct object where the subject carries the nuclear stress (see (14a)), the FPR allows only focus interpretations where the nuclear stress falls within the focus, i.e., focus on the subject (see (14b)), but not focus on the verb or the direct object (see (14b') and (14b'')).

(14) Perception/Interpretation
a.   JOHN bought a bike

FPR

b.   [JOHN]$_F$ bought a bike.
b.'   #JOHN [bought]$_F$ a bike.
b.''   #JOHN bought a [bike]$_F$.

The FPR thus makes clear predictions for the results of our experiment. Speakers of the underlying production experiment put the nuclear stress on the focused constituent. Participants of the perception experiment assign the focus-background partition according to the position of the nuclear stress. This predicts an extremely high degree of accuracy in the perception experiment. If we compare this prediction to our outcomes, we must bear in mind that the participants only had two focus candidates (subject and direct object), while potential ambiguities due to focus projection (see (15)) were excluded.

(15) Nuria comprobó el   regulaDOR.
      Nuria controlled the regulator
      'Nuria controlled the regulator.'
a.   What happened?
        [Nuria comprobó el regulaDOR]$_F$.
b.   What did Nuria do?
        Nuria [comprobó el regulaDOR]$_F$.
c.   What did Nuria control?
        Nuria comprobó [el regulaDOR]$_F$.
        (Leonetti 2014: 9; modified)

Even in this restricted environment, the perceived focus corresponded to the focus of the stimulus in only 70.3% of the cases. Although accuracy varies considerably depending on the factors *focused constituent*, *focus type*, and *variety* (see Section 3.2.), the average accuracy of 70.3% shows that the position of nuclear stress does not

unambiguously indicate the focus-background partition. The explanation for this discrepancy between the predicted and the observed accuracy is that in the FPR, the notion of is *nuclear stress* is binary, in that one and only one syllable (and thus one constituent) carries the nuclear stress, while prosody is gradual. The binary, and thus clearly perceivable, distinction between the syllable carrying the nuclear stress and the rest of the sentence is an abstraction of actual prosodic realizations. One contribution of the present study is to give us an idea of the gap between the idealized representation of the FPR on the one hand and the gradualness of prosodic surface realizations on the other.

Based on the results of our study, we argue for a less mechanic relation between sentence form and focus interpretation: sentence form is a filter which rules out (or makes improbable) certain focus interpretations, but contextual cues are often necessary to identify for the actual focus-background partition (i.e., context functions as another filter making certain focus interpretations more or less likely). Context is mentioned as a decisive factor for focus interpretation by Féry (2017) who even gives priority to context over sentence form: the context typically tells the hearer (in spoken language) which element of a sentence is focused, and the grammatical reflexes of focus are due to attentional factors: "It is helpful to highlight important words and constituents, as this may help a listener to process what is being said and to extract the information in an efficient way" (Féry 2017: 142). Although we remain agnostic as to how the two filters (sentence form and context) are ranked, our results stress the importance of context for focus interpretation. Even in the restricted set-up of our experiment (with only two focus candidates) the overall accuracy in focus detection is only 70%.

3.3.3. Prosody as a factor

Our experiment relies on four conditions to test the impact of focus type and focused constituent on the accuracy of focus choice. The respective stimuli for these conditions are heterogeneous in several respects: they stem from different speakers, they have different prosodic shapes, and they also differ with respect to accuracy in focus choice. We first concentrate on the latter fact, before we come back to prosodic differences.

So far, we have mainly looked at the accuracy in terms of means per condition (but indicating the standard deviation, and sometimes differentiating between the two participant groups). At closer inspection, considerable differences in accuracy among the stimuli for one and the same condition become apparent. Figure 6 presents the accuracy for the stimuli split up between the two participant groups (each dot represents one stimulus). Besides visualizing once again the tendencies and preferences discussed in the previous sections, the graphs also show the range of the accuracy among the stimuli of one and the same condition. Especially striking is the range for stimuli of Condition 3 where a large number of stimuli have an accuracy rate far above and far below the mean (especially with Argentinean participants).

Figure 6: Accuracy (%) per stimulus for four conditions.

Condition 1: [S]$_{CF}$-V-dO

Condition 2: S-V-[dO]$_{CF}$

Condition 3: [S]$_{IF}$-V-dO

Condition 4: S-V-[dO]$_{IF}$
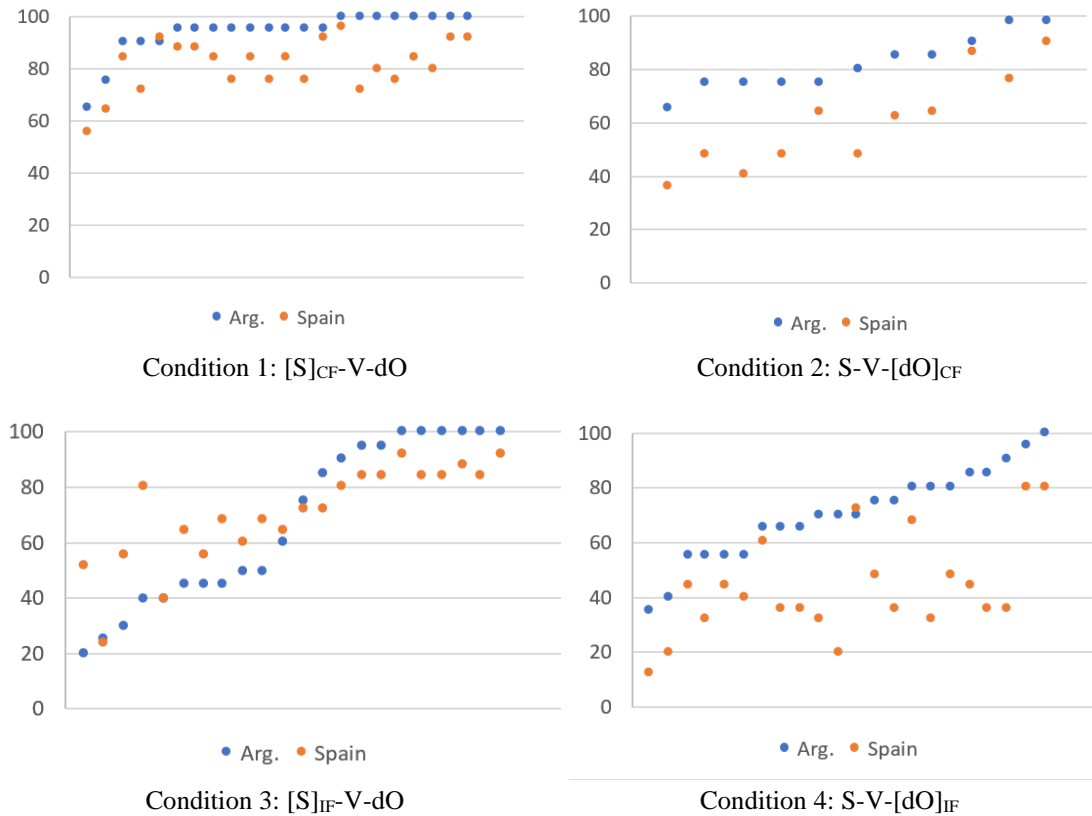
Given that the syntactic shape of the stimuli is fixed as SVO (with an additional postverbal constituent in some stimuli for subject focus), prosody is the obvious suspect for the variation in accuracy among the stimuli displayed in Figure 6. Although the prosodic properties of the stimuli were not part of the experimental design (as variables), they were considered in the selection of stimuli from the underlying production data, notably in the case of those conditions where more stimuli than needed were available (Section 3.1.3.). Based on the prosodic categorization of stimuli used for the selection, we can distinguish between canonical and non-canonical stimuli. For Condition 1 ([S]$_{CF}$), *canonical* means a low ip boundary tone (L-) following the focus domain as well as post-focal deaccentuation, as illustrated in Figure 7 (for this stimulus the accuracy rates are 95% (Argentina) and 88% (Spain)).

Figure 7: Canonical stimulus for Condition 1.

										L-			L-L%
	L+H*+L				*		*					*	
ma	ri	a	le	ða	el	ðja	rjwa		swer		ma	no	
	María		le	da	el		diario	a	su		hermano		
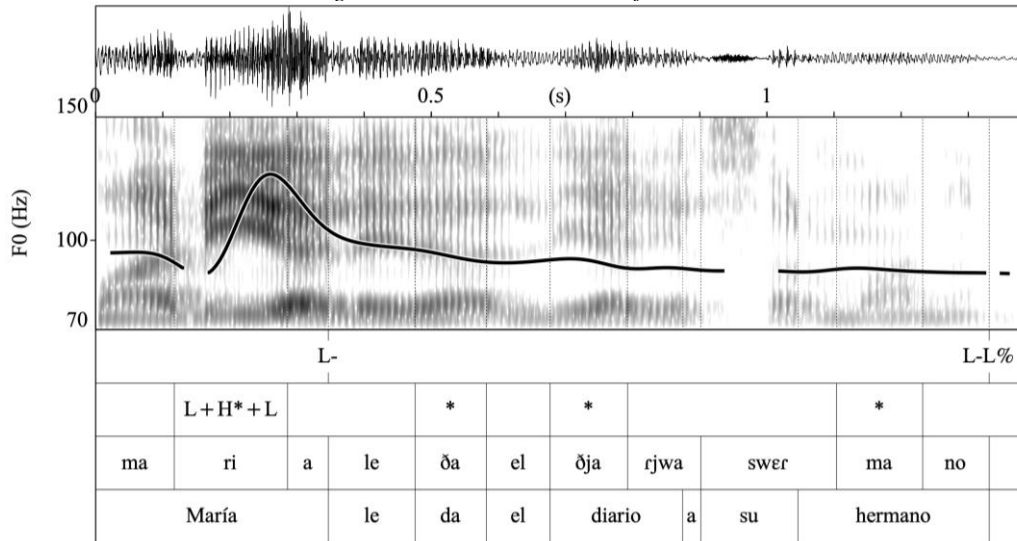
For Condition 2 ([dO]$_{CF}$), *canonical* means a tritonal pitch accent (L+H*+L) on the focused object. For Condition 3 ([S]$_{IF}$), *canonical* again means a low boundary tone L- after the focus. For Condition 4 ([dO]$_{IF}$), finally, *canonical* means a high boundary tone H- preceding the focus domain. All other prosodic shapes were considered non-canonical for the present purpose; see Figure 8 for a non-canonical stimulus for condition 1 (for this stimulus the accuracy rates are 75% (Argentina) and 64% (Spain)).

Figure 8: Non-canonical stimulus for Condition 1.

			H-										L-L%
	L+H*				L+H*		L*+H					L*	
ma	ɾi	a	le	ða	ɛl	ðja	rjoa̠		su	eɾ	ma	no̠	
	María		le	da	el		diario	a	su		hermano		
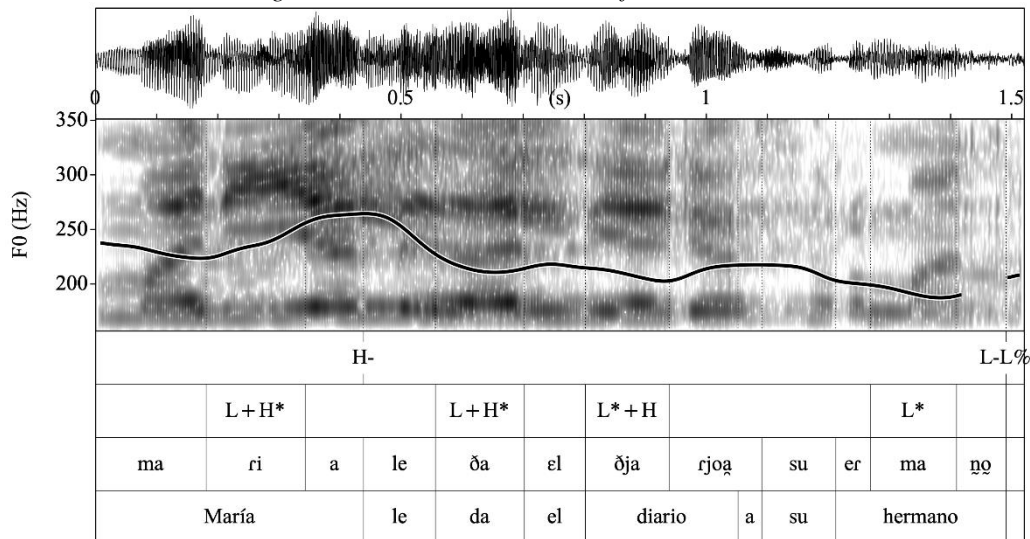
The opposition *canonical vs. non-canonical* was not part of the experimental design, because there were not sufficient items for a balanced set of stimuli. For example, among the 11 items for S-V-[dO]$_{CF}$, only one showed the canonical tritonal pitch accent L+H*+L. Despite these restrictions, we have investigated in a post-hoc analysis whether the distinction between canonical and non-canonical realizations has an impact on accuracy.

Across all four conditions the accuracy is considerably higher for canonical stimuli (82.88%) than for non-canonical ones (61.18%). This suggests that the prosodic criteria used in the categorization as canonical and non-canonical are linked to the signaling of

prosodic prominence, and therefore relevant for focus marking. Below, we provide further data showing the accuracy of canonical and non-canonical stimuli divided by participant groups (see Tables 4 and 5). The tables read as follows: in the responses for Condition 1 by Argentinean participants, the accuracy of focus choice amounts to 97% for canonical stimuli and 77% for non-canonical stimuli.

Table 4: Accuracy of focus choice in % for canonical and non-canonical stimuli (Argentina).

	Canonical	Non-canonical
Condition 1: $[S]_{CF}$-V-dO	97	77
Condition 2: S-V-$[dO]_{CF}$	85	82
Condition 3: $[S]_{IF}$-V-dO	85	50
Condition 4: S-V-$[dO]_{IF}$	67	71

Table 5: Accuracy of focus choice in % for canonical and non-canonical stimuli (Spain).

	Canonical	Non-canonical
Condition 1: $[S]_{CF}$-V-dO	84	64
Condition 2: S-V-$[dO]_{CF}$	62	60
Condition 3: $[S]_{IF}$-V-dO	81	59
Condition 4: S-V-$[dO]_{IF}$	56	40

When interpreting the accuracy rates for the canonical and non-canonical stimuli of individual conditions, we must bear in mind that canonical and non-canonical stimuli are not quantitatively balanced (see comments above) and as a result, the absolute number of stimuli might be very low. In Condition 2, there is only one canonical stimulus, but ten non-canonical ones. Hence, the lack of a difference between the two types in accuracy (see Table 4 and 5) is not instructive. The absolute number of stimuli is also low for the canonical realizations of condition 4 (5 out of 22 stimuli) and the non-canonical realizations of condition 1 (3 out of 22 stimuli).

Finally, the above distinction between canonical and non-canonical stimuli also allows to evaluate the representativity of the stimuli used in the perception experiment vis-à-vis the underlying data from the production experiment. Since for condition 1 and 2 all available stimuli from the production data were used, this is an issue only with the non-contrastive conditions 3 and 4. Due to our selection (as described in Section 3.1.3.), canonical stimuli were slightly overrepresented in the material of the perception experiment as compared to the underlying production data (see Table 6 in the appendix). This suggests that the difference in accuracy between information focus and contrastive focus would have been even greater if we would have used canonical and non-canonical stimuli according to the proportion in the data from the production experiment.

4. Concluding remarks

In this paper, we investigated a specific aspect of the interface between prosody and information structure in Spanish, namely the consequences of FPR for focus detection. In production, the nuclear stress must fall within the focused constituent. Since we assume the inverse process in perception/interpretation, the FPR is expected to heavily restrict the available focus interpretations. In a perception experiment, we tested whether the focus-background partition assigned by hearers to (contextless) SVO sentences corresponds to the focus-background partition of the underlying production data. We used a matching task with an auditorily presented SVO declarative and two different wh-questions as possible preceding contexts.

The results of the experiment have shown that the focus-background partition assigned by hearers corresponds in 70% of the cases to the original focus-background partition. However, the accuracy rate strongly depends on the factors *focus type*, *focused constituent*, and *variety spoken by participants*. The accuracy is higher for stimuli with contrastive focus than for those with information focus (78.3% vs. 62.3%); it is higher for stimuli with subject focus than for those with object focus (78.0% vs. 62.6%); it is higher for participants speaking Argentinean Spanish than for those speaking Peninsular Spanish (78.4% vs. 63.8%). Although the impact of the three factors corresponded to our predictions, the overall accuracy of only 70% needed further discussion. We argued that although nuclear stress placement helps in focus detection, it should not be seen as an unequivocal focus marking device in Spanish. Instead, prosody (and sentence form more generally) should be viewed as a filter which rules out certain focus-background partitions. However, contextual cues are often necessary to identify for the actual focus-background partition. In addition to the three above-mentioned factors, the prosodic features of the individual stimuli had an impact on the accuracy rate. Although the stimuli's prosodic properties were not part of the experimental design (as variables), we distinguished between canonical and non-canonical stimuli in a post-hoc analysis. Crucially, canonical realizations showed a higher rate of accuracy than non-canonical ones (82.88% vs. 61.18%). The underlying production data did not allow for a proper inclusion of the prosodic properties in the experimental design (since canonical realizations were surprisingly rare for some conditions) which makes this topic an obvious candidate for future research.

Christoph Gabriel
Johannes Gutenberg University Mainz
Department of Romance Studies
Jakob-Welder-Weg 18
55128 Mainz
Germany
christoph.gabriel@uni-mainz.de
+49 6131 39-24782

Steffen Heidinger
University of Graz
Department of Romance Studies
Merangasse 70/III
8010 Graz
Austria
steffen.heidinger@uni-graz.at
+43 316 380-8215

References

Ashby, W. J. & P. Bentivoglio (1993). Preferred argument structure in spoken French and Spanish. *Language Variation and Change* 5, pp. 61–76. https://doi.org/10.1017/S095439450000140X

Bates, D., M. Mächler, B. Bolker & S. Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67 (1), pp. 1–48. https://doi.org/10.18637/jss.v067.i01

Bosque, I. & J. Gutiérrez-Rexach (2009). *Fundamentos de sintaxis formal*. Madrid, Akal.

Bossong, G. (1984a). Diachronie und Pragmatik der spanischen Wortstellung. *Zeitschrift für romanische Philologie* 100, pp. 92–111. https://doi.org/10.1515/zrph.1984.100.1-2.92

Bossong, G. (1984b). Wortstellung, Satzperspektive und Textkonstitution im Ibero-Romanischen, dargestellt am Beispiel eines Textes von Juan Rulfo. *Iberoromania* 19, pp. 1–16. https://doi.org/10.1515/iber.1984.1984.19.1

Breen, M., E. Fedorenko, M. Wagner & E. Gibson (2010). Acoustic correlates of information structure. *Language and Cognitive Processes* 25 (7–9), pp. 1044–1098. https://doi.org/10.1080/01690965.2010.504378

Büring, D. (2009). Towards a typology of focus realization, in M. Zimmermann & C. Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*. Oxford, Oxford University Press, pp. 177–205. https://doi.org/10.1093/acprof:oso/9780199570959.003.0008

Calhoun, S., E. La Cruz & A. Olssen (2018). The interplay of information structure, semantics, prosody, and word ordering in Spanish intransitives. Laboratory Phonology. *Journal of the Association for Laboratory Phonology* 9 (1), p. 8. https://doi.org/10.5334/labphon.65

Cassarà, A. C. (2021). Subject focus in French and Spanish. Dissertation, Universität zu Köln.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view, in C. N. Li (ed.), Subject and topic. New York, Academic Press, pp. 25–55.

Cole, J., J. I. Hualde, C. L. Smith, C. Eager, T. Mahrt & R. Napoleão de Souza (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics* 75, pp. 113–147. https://doi.org/10.1016/j.wocn.2019.05.002

Downing, L. J. & B. Pompino-Marschall (2013). The focus prosody of Chichewa and the Stress-Focus constraint: A response to Samek-Lodovici (2005). *Natural Language & Linguistic Theory* 31 (3), pp. 647–681. https://doi.org/10.1007/s11049-013-9192-x

Drubig, H. B. (2003). Toward a typology of focus and focus constructions. Linguistics 41 (1), pp. 1–50. https://doi.org/10.1515/ling.2003.003

Drubig, H. B. & W. Schaffar (2001). Focus constructions, in M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds.), *Language typology and language universals. An international handbook*. Berlin, De Gruyter, pp. 1079–1104. https://doi.org/10.1515/9783110194036

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language* 63 (4), pp. 805–855. https://doi.org/10.2307/415719

Dufter, A. (2007). Zwischen Kompositionalität und Konventionalisierung: Satzspaltung mit *c'est* im Französischen der Gegenwart. *Romanistisches Jahrbuch* 57 (2006), pp. 31–59. https://doi.org/10.1515/9783110192889.1.31

Dufter, A. & C. Gabriel (2016). Information structure, prosody, and word order, in S. Fischer & C. Gabriel (eds.), *Manual of Grammatical Interfaces in Romance*. Berlin, De Gruyter, pp. 419–455. https://doi.org/10.1515/9783110311860-017

El Zarka, D., A. Kelterer & B. Schuppler (2020). An analysis of prosodic prominence cues to information structure in Egyptian Arabic, in *Proceedings of Interspeech 2020. Shanghai,* pp. 1883–1887. Available at: https://indico2.conference4me.

psnc.pl/event/35/contributions/3181/attachments/320/344/Tue-1-10-6.pdf https://doi.org/10.21437/Interspeech.2020-2322

El Zarka, D. & P. Hödl (2021). A study on the perception of prosodic cues to focus by Egyptian listeners: Some make use of them, but most of them don't. *Speech Communication* 132, pp. 55–69. https://doi.org/10.1016/j.specom.2021.05.002

Estebas-Vilaplana, E. & P. Prieto (2010). Castilian Spanish Intonation, in P. Prieto & P. Roseano (eds.), *Transcription of intonation of the Spanish language*. München, Lincom Europa, pp. 17–48.

Face, T. L. (2001). Focus and early peak alignment in Spanish intonation. *Probus* 13, pp. 223–246. https://doi.org/10.1515/prbs.2001.004

Face, T. L. (2002). El foco y la altura tonal. *Boletín de lingüística* 17, pp. 30–52. Available at: https://www.redalyc.org/pdf/347/34701703.pdf

Feldhausen, I., A. Pešková & C. Gabriel (2010). Prosodic phrasing in Argentinean Spanish: Buenos Aires and Neuquén, in M. Hasegawa-Johnson (ed.), *Proceedings of Speech Prosody 2010*. Chicago, IL, pp. 1–4. Available at: https://www.isca-speech.org/archive_v0/sp2010/papers/sp10_111.pdf

Feldhausen, I., A. Pešková, E. Kireva & C. Gabriel (2011). Categorical perception of Porteño nuclear accents, in W.-S. Lee & E. Zee (eds.), *Proceedings of the 17th In-ternational Congress of Phonetic Sciences 2011. Hong Kong, China*, pp. 116–119. Available at: http://icphs2011.hk.lt.cityu.edu.hk/resources/OnlinePro ceedings/SpecialSession/Session4/Feldhausen/Feldhausen.pdf

Feldhausen, I. & M. d. M. Vanrell (2015). Oraciones hendidas y marcación del foco estrecho en español: Una aproximación desde la Teoría de la Optimidad Esto-cástica. *Revista Internacional de Lingüística Iberoamericana* 26, pp. 39–60. DOI: https://www.jstor.org/stable/24769043

Féry, C. (2017). *Intonation and prosodic structure*. Cambridge, Cambridge University Press. https://doi.org/10.1017/9781139022064

Firbas, J. (1992). *Functional sentence perspective in written and spoken communication*. Cambridge, Cambridge University Press. https://doi.org/10.1017/CBO97 80511597817

Frota, S., M. D'Imperio, G. Elordieta & M. Vigário (2007). The phonetics and phonology of intonational phrasing in Romance, in P. Prieto, J. Mascaró & M.-J. Solé (eds.), *Prosodic and Segmental Issues in (Romance) Phonology*. Amsterdam, John Benjamins, pp. 131–153. https://doi.org/10.1075/cilt

Gabriel, C. (2010). On focus, prosody, and word order in Argentinean Spanish: A min-imalist OT account. *Revista Virtual de Estudos da Linguagem – ReVEL* Special edition n. 4, pp. 183–222. Available at: http://www.revel.inf.br/files/artigos/revel_special_4_on_focus_prosody_and_word_order.pdf

Gabriel, C. (2012). The Hamburg Corpus of Argentinean Spanish (HaCASpa), in T. Schmidt & K. Wörner (eds.), *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam, John Benjamins, pp. 183–197. https://doi.org/10.1075/hsm.14.13gab

Gabriel, C., I. Feldhausen & A. Pešková (2011). Prosodic phrasing in Porteño Spanish, in C. Gabriel & C. Lleó (eds.), *Intonational phrasing in Romance and Ger-manic: Cross-linguistic and bilingual studies*. Amsterdam, John Benjamins, pp. 153–182. https://doi.org/10.1075/hsm.10.09gab

Gabriel, C., I. Feldhausen, A. Pešková, L. Colantoni, S.-A. Lee, V. Arana & L. O. Labastía (2010). Argentinian Spanish Intonation, in P. Prieto & P. Roseano

(eds.), Transcription of intonation of the Spanish language. München, Lincom Europa, pp. 285–317.

Gabriel, C., A. Pešková, L. Labastía & B. Blázquez (2013). La entonación en el español de Buenos Aires, in L. Colantoni & C. Rodríguez Louro (eds.), Perspectivas teóricas y experimentales sobre el español de la Argentina. Frankfurt/Madrid, Vervuert, pp. 99–115. https://doi.org/10.31819/9783954871971-007

Gutiérrez-Bravo, R. (2008). La identificación de los tópicos y de los focos. *Nueva Revista de Filología Hispánica* 56 (2), pp. 363–401. Available at: https://www.jstor.org/stable/40300721 https://doi.org/10.24201/nrfh.v56i2.969

Hanssen, J., J. Peters & C. Gussenhoven (2008). Prosodic effects of focus in Dutch declaratives, in *Proceedings of Speech Prosody 2008*. Campinas, Brazil, pp. 609–612. Available at: http://www.sprosig.org/sp2008/papers/id153.pdf

Heidinger, S. (2018). Sekundäre Prädikation und Informationsstruktur: Fokus und Informationsstatus bei spanischen Depiktiven. Berlin, Peter Lang. https://doi.org/10.3726/b14723

Heidinger, S. (2021). El foco informativo y la afinidad con el foco, in M. Leonetti & M. V. Escandell-Vidal (eds.), *La estructura informativa*. Madrid, Visor Libros, pp. 319–357.

Heidinger, S. (2022). Corpus data and the position of information focus in Spanish. *Studies in Hispanic and Lusophone Linguistics*. 15(1), pp. 67–109. https://doi.org/10.1515/shll-2022-2056

Heidinger, S. & E. Onea (2021). Focus affinity in Spanish: An experimental study. *Journal of Pragmatics* 186, pp. 100–116. https://doi.org/10.1016/j.pragma.2021.09.015

Hoot, B. (2012). *Presentational focus in heritage and monolingual Spanish*. Dissertation, University of Illinois at Chicago, Chicago. Available at: https://indigo.uic.edu/articles/thesis/Presentational_Focus_in_Heritage_and_Monolingual_Spanish/10832456

Hoot, B. (2016). Narrow presentational focus in Mexican Spanish: Experimental evidence. *Probus* 28 (2), pp. 335–365. https://doi.org/10.1515/probus-2014-0004

Hoot, B. (2017). Narrow presentational focus in heritage Spanish and the syntax-discourse interface. *Linguistic Approaches to Bilingualism* 7, pp. 63–95. https://doi.org/10.1075/lab.14021.hoo

Hualde, J. I. (2014). *Los sonidos del español*. Cambridge, Cambridge University Press. https://doi.org/10.1017/CBO9780511719943

Hualde, J. I. & P. Prieto (2015). Intonational variation in Spanish: European and American varieties, in S. Frota & P. Prieto (eds.), *Intonation in Romance*. Oxford, Oxford University Press, pp. 350–391. https://doi.org/10.1093/acprof:oso/9780199685332.003.0010

Jackendoff, R. (1972). *Semantic interpretation in Generative Grammar*. Cambridge MA, MIT Press.

Katz, J. & E. Selkirk (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87 (4), pp. 771–816. https://doi.org/10.1353/lan.2011.0076

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica* 55, pp. 243–276. Available at: https://www.jstor.org/stable/26190905 https://doi.org/10.1556/ALing.55.2008.3-4.2

Krifka, M. & R. Musan (eds.) (2012). *The expression of information structure*. Berlin, De Gruyter. https://doi.org/10.1515/9783110261608

Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus, and the mental representations of discourse referents*. Cambridge, Cambridge University Press. https://doi.org/10.1017/CBO9780511620607

Leonetti, M. (2014). Gramática y pragmática en el orden de palabras. *Lingüística en la red* XII, pp. 1–25. Available at: https://www.linred.es/monograficos_pdf/LR_monografico12-articulo6.pdf

Leonetti, M. & M. V. Escandell-Vidal (2021). La estructura informativa: Preguntas frecuentes, in M. Leonetti & M. V. Escandell-Vidal (eds.), *La estructura informativa*. Madrid, Visor Libros, pp. 15–181.

Olarrea, A. (2012). Word order and information structure, in J. I. Hualde, A. Olarrea & E. O'Rourke (eds.), *The handbook of Hispanic linguistics*. Malden MA, Wiley-Blackwell, pp. 603–627. https://doi.org/10.1002/9781118228098.ch28

Pešková, A., I. Feldhausen & C. Gabriel (2011). Fraseo prosódico en el español porteño: Evidencia de datos leídos y semi-espontáneos, in R. Kailuweit & Á. L. Di Tullio (eds.), *El español rioplatense*. Madrid & Frankfurt/M., Vervuert, pp. 77–102. https://doi.org/10.31819/9783865278821-004

Staudinger, E. & R. Kailuweit (2018). Norma(s) pluricéntrica(s) y medios de comunicación: El caso del 'argentino neutro', in C. Strosetzki (ed.), Aspectos actuales del hispanismo mundial: Literatura – Cultura – Lengua. Berlin, De Gruyter, pp. 588–599. https://doi.org/10.1515/9783110450828-104

Vallduví, E. (1990). The informational component. Dissertation, University of Pennsylvania. Available at: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1189&context=ircs_reports

Vallduví, E. & E. Engdahl (1996). The linguistic realization of information packaging. Linguistics 34, pp. 459–519. https://doi.org/10.1515/ling.1996.34.3.459

Vanrell, M. d. M., B. Gili Fivela, A. Stella & P. Prieto (2011). Prosodic cues for the recognition of contrastive focus, in B. Gili Fivela, A. Stella, L. Garrapa & M. Grimaldi (eds.), *Contesto comunicativo e variabilità nella produzione e percezione della lingua. Atti del VII Convegno AISV 2011*. Roma, Bulzoni, pp. 224–235.

Vanrell, M. d. M., A. Stella, B. Gili Fivela & P. Prieto (2013). Prosodic manifestations of the Effort Code in Catalan, Italian and Spanish contrastive focus. *Journal of the International Phonetic Association* 43 (2), pp. 195–220. https://doi.org/10.1017/S0025100313000066

Vanrell, M. d. M. & O. Fernández Soriano (2018). Language variation at the prosody-syntax interface: Focus in European Spanish, in M. García García & M. Uth (eds.), Focus realization in Romance and beyond. Amsterdam, John Benjamins, pp. 33–70. https://doi.org/10.1075/slcs.201.02van

Wunderlich, D. (2006). Argument hierarchy and other factors determining argument realization, in I. Bornkessel, M. Schlesewsky, B. Comrie & A. D. Friederici (eds.), *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives*. Berlin, De Gruyter, pp. 15–52. https://doi.org/10.1515/9783110219272.15

Zubizarreta, M. L. (1998). *Prosody, focus, and word order*. Cambridge MA, MIT Press.

Appendix
Production data and stimuli of perception experiment

Table 6:   Canonical vs. non-canonical prosodic realizations in data from production experiment and in stimuli used in perception experiment (absolute frequencies).

		Production data	Perception experiment
Condition 1: [S]$_{CF}$-V-dO	canonical	19	19
	non-canonical	3	3
Condition 2: S-V-[dO]$_{CF}$	canonical	1	1
	non-canonical	10	10
Condition 3: [S]$_{IF}$-V-dO	canonical	21	11
	non-canonical	36	11
Condition 4: S-V-[dO]$_{IF}$	canonical	5	5
	non-canonical	24	17

Output R

Model 1

```
  Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
   Family: binomial  ( logit )
  Formula: Correct ~ Focus_Constituent * Focus_Type * Language + (1 | Subject) +
     (1 | Stimulus)

  Random effects:
   Groups    Name         Variance Std.Dev.
   Subject  (Intercept) 0.1291   0.3592
   Stimulus (Intercept) 0.7130   0.8444
 Number of obs: 3960, groups:  Subject, 90; Stimulus, 77

  Fixed effects:
                                                  Estimate Std. Error z value
 (Intercept)                                       0.98598    0.21955   4.491
 Focus_ConstituentS                               -0.04244    0.30134  -0.141
 Focus_TypeCF                                      0.73819    0.35675   2.069
 LanguageSpanish                                  -1.27929    0.16418  -7.792
 Focus_ConstituentS:Focus_TypeCF                   1.31404    0.49681   2.645
 Focus_ConstituentS:LanguageSpanish                1.43152    0.21105   6.783
 Focus_TypeCF:LanguageSpanish                      0.05638    0.21553   0.262
 Focus_ConstituentS:Focus_TypeCF:LanguageSpanish -1.55980    0.34969  -4.460
                                                  Pr(>|z|)
 (Intercept)                                      7.09e-06 ***
 Focus_ConstituentS                               0.88798
 Focus_TypeCF                                     0.03853 *
 LanguageSpanish                                  6.59e-15 ***
 Focus_ConstituentS:Focus_TypeCF                  0.00817 **
 Focus_ConstituentS:LanguageSpanish               1.18e-11 ***
 Focus_TypeCF:LanguageSpanish                     0.79363
 Focus_ConstituentS:Focus_TypeCF:LanguageSpanish 8.18e-06 ***
 ---
 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
 Family: binomial  ( logit )
Formula: Correct ~ Focus_Constituent * Focus_Type + (1 | Subject) + (1 |
    Stimulus)

Random effects:
 Groups    Name         Variance Std.Dev.
 Stimulus (Intercept) 1.4621   1.2092
 Subject  (Intercept) 0.1736   0.4166
Number of obs: 1760, groups:  Stimulus, 77; Subject, 40

Fixed effects:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                       1.0522     0.2929   3.592 0.000328 ***
Focus_ConstituentS                0.1251     0.4121   0.304 0.761429
Focus_TypeCF                      0.7224     0.4856   1.488 0.136819
Focus_ConstituentS:Focus_TypeCF   1.4502     0.6712   2.161 0.030723 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2A

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
 Family: binomial  ( logit )
Formula: Correct ~ Focus_Constituent + (1 | Subject) + (1 | Stimulus)

Random effects:
 Groups    Name         Variance Std.Dev.
 Subject  (Intercept) 0.5393   0.7344
 Stimulus (Intercept) 0.7257   0.8519
Number of obs: 880, groups:  Subject, 40; Stimulus, 33

Fixed effects:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)          1.8244     0.3248   5.617 1.94e-08 ***
Focus_ConstituentS   1.4627     0.4220   3.466 0.000529 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2B

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
 Family: binomial  ( logit )
Formula: Correct ~ Focus_Constituent + (1 | Subject) + (1 | Stimulus)

Random effects:
 Groups    Name         Variance Std.Dev.
 Stimulus (Intercept) 1.88546  1.3731
 Subject  (Intercept) 0.02269  0.1506
Number of obs: 880, groups:  Stimulus, 44; Subject, 40

Fixed effects:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)          1.0423     0.3176   3.282  0.00103 **
Focus_ConstituentS   0.1764     0.4568   0.386  0.69937
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 3

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
 Family: binomial  ( logit )
Formula: Correct ~ Focus_Constituent * Focus_Type + (1 | Subject) + (1 |     Stim-
ulus)

Random effects:
 Groups    Name          Variance Std.Dev.
 Stimulus (Intercept) 0.4524   0.6726
 Subject  (Intercept) 0.1116   0.3341
Number of obs: 2200, groups:  Stimulus, 77; Subject, 50

Fixed effects:
                               Estimate Std. Error z value Pr(>|z|)
(Intercept)                     -0.2918     0.1767  -1.652  0.09862 .
Focus_ConstituentS               1.2661     0.2446   5.176 2.27e-07 ***
Focus_TypeCF                     0.7781     0.2813   2.766  0.00567 **
Focus_ConstituentS:Focus_TypeCF -0.1503     0.3785  -0.397  0.69119
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 4

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
 Family: binomial  ( logit )
Formula: Correct ~ Focus_Constituent + Focus_Type + Language + (1 | Subject) +
    (1 | Stimulus)

Random effects:
 Groups    Name          Variance Std.Dev.
 Subject  (Intercept) 0.1220   0.3492
 Stimulus (Intercept) 0.7178   0.8473
Number of obs: 3960, groups:  Subject, 90; Stimulus, 77

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         0.6841     0.1882   3.634 0.000279 ***
Focus_ConstituentS  0.9385     0.2148   4.369 1.25e-05 ***
Focus_TypeCF        0.9472     0.2167   4.371 1.23e-05 ***
LanguageSpanish    -0.8827     0.1097  -8.048 8.41e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) Fcs_CS Fc_TKF
Fcs_CnsttnS -0.547
Focus_TypCF -0.389 -0.134
LangugSpnsh -0.332 -0.018 -0.019
```