

On the interplay of segmentation cues in adult OV-VO bilinguals: prosody, frequency and context language¹

Irene de la Cruz-Pavía*
University of Deusto & Ikerbasque

Gorka Elordieta, Itziar Laka
Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

ABSTRACT: The present study investigated the role of phrasal prosody in speech segmentation in adult bilingual speakers of two languages with opposite basic word orders: Basque and Spanish (Object-Verb and Verb-Object, respectively). We created a structurally ambiguous artificial language (AL) that allowed two possible parses, mimicking the order characteristic of OV and VO languages, and enriched it with the prosodic cue associated to languages with an OV order: a pitch contrast, in which the element receiving prominence within phrases has higher pitch than the non-prominent elements. We tested highly proficient L1Basque-L2Spanish and L1Spanish-L2Basque speakers on their segmentation preferences of the AL, addressing both groups in their respective L1 during the study. Analysis of their segmentation preferences revealed that the presence of OV prosody modulated—but did not determine—their parsing of the ambiguous AL, as both groups of bilinguals exhibited a significantly greater preference for the segmentation associated to OV languages than two similar groups presented with a prosodically flat variant of the same AL in de la Cruz-Pavía *et al.* (2022). These results confirm that

¹ We wish to thank Idoia Ros for her help running a subset of the participants. This research was supported by the Basque Government under Grants [BFI07.253, IT1537-22 and IT1439-22], Ikerbasque, Basque Foundation for Science, and the Spanish Ministry of Science and Innovation under Grants PID2022-142167NB-I00, and Grant RYC2021-033951-I funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

* **Corresponding author:** Irene de la Cruz-Pavía. University of Deusto & Basque Foundation for Science Ikerbasque, Facultad de Ciencias Sociales y Humanas Unibertsitate Etorbidea, 24 (48007 Bilbo). – irene.delacruz.pavia@deusto.es – <https://orcid.org/0000-0003-3425-0596>

How to cite: De la Cruz-Pavía, Irene; Elordieta, Gorka; Laka, Itziar (2023). «On the interplay of segmentation cues in adult OV-VO bilinguals: prosody, frequency and context language», *ASJU*, 57 (1-2), 247-260. (<https://doi.org/10.1387/asju.25921>).

Received: 2022-06-27; Accepted: 2022-10-23.

ISSN 0582-6152 - eISSN 2444-2992 / © UPV/EHU Press



This work is licensed under a

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

phrasal prosody is an available segmentation cue for adult bilinguals, but suggests that it is not a salient segmentation cue in this population.

KEYWORDS: bilingualism; word order; phrase segmentation; artificial languages; prosodic cues; statistical cues; segmental cues; hierarchy of phrase segmentation cues.

1. Introduction

Word order is a central feature of grammar, which comprises multiple phenomena that vary across languages (Hahn & Xu 2022). The world's languages vary in their order of verbs and objects, of relative clauses and nouns, of adjectives and nouns, etc. (Greenberg 1963), and (some of) these phenomena co-vary systematically (Dryer 1992; Greenberg 1963). One such example is basic word order, i.e. the relative order of verbs and objects, which correlates cross-linguistically with the order of functors (e.g. determiners, adpositions, verbal inflection: *the, in, -ing*), and content words (e.g. nouns, verbs, adjectives: *turtle, walk, slow*). Languages in which Verbs precede Objects (henceforth VO languages; e.g. English: *I feed_{Verb} the turtle_{Object}*), such as Spanish or English, are typically functor-initial, that is, functors tend to occur at the beginning of their syntactic phrases (e.g. English: *of the woman*). In turn, languages in which Verbs follow Objects (OV languages; e.g. Basque: *Dortoka_{Object} elikatzen dut_{Verb} — turtle feed aux-1pers-sg*), such as Basque or Hindi, are typically functor-final, that is, functors tend to occur at the end of syntactic phrases (e.g. Basque: *emakume-a-ren — woman-the-possessive*).

Functors are characterized by statistical and phonological properties that set them apart from content words (Morgan *et al.* 1996). Individual functors occur extremely frequently in the input and are perceptually minimal (e.g. they tend to be unstressed, short, and have simple syllabic structures), while individual content words have a much lower frequency of occurrence and are phonologically salient (e.g. they receive prosodic prominence, tend to be longer and have more complex syllabic structures). Because the statistical, prosodic and distributional properties setting functors and content words apart are directly observable in the speech signal, they have been proposed to reliably cue phrase boundaries and basic word order (Gervain *et al.* 2008; Mazuka 1996; Morgan *et al.* 1996).

Gervain and colleagues' seminal work provided evidence supporting this hypothesis, showing that both prelexical infants and adults are sensitive to the frequency of the elements in the input and use this cue to extract phrases from the input (Gervain *et al.* 2008, 2013). When presented with artificial languages that comprise frequent and infrequent elements in strict alternation, and which have an ambiguous structure that allows two possible parses (mimicking the two possible orders of functors and content words in natural languages), both infants and adults parse the language into phrase-like units that follow the order characteristic in their native language. Adult bilinguals can deploy the frequency-based strategies of their two languages (de la Cruz-Pavía *et al.* 2015, 2022; de la Cruz-Pavía, Werker, *et al.* 2020). Note that bilingual speakers of an OV and a VO language (e.g. Basque-Spanish bilinguals) hear both functor-initial *and* functor-final phrases in their input. And yet, these bilinguals modulate their parsing preferences of the ambiguous artificial language (AL) depending on the language of context, that is, the language in which they are ad-

dressed and receive the instructions of the study (de la Cruz-Pavía *et al.* 2015, 2022): bilinguals exhibit a greater frequent-final parsing preference of the AL when addressed in their OV language (Basque, in these studies) as compared with their VO language (Spanish).

In addition to this frequency-based cue, the speech signal contains another source of information signalling phrase boundaries and basic word order, namely phrasal prosody. Within phrases, phrasal prosodic prominence is carried by the content words, not by the functors, and its acoustic realization varies across languages in correlation with basic word order. In VO languages prominence is realized through increased duration, yielding a short-long pattern (English: *in Rome*), while in OV languages it is realized through an increase in pitch (or intensity), resulting a high-low (or loud-soft) pattern (Japanese: [^]*Tokyo ni — Tokyo in*; de la Cruz-Pavía, Gervain, *et al.* 2020; Gervain & Werker 2013; Nespor *et al.* 2008).

Sensitivity to this prosodic cue is observed from early infancy (Bernard & Gervain 2012; de la Cruz-Pavía *et al.* 2019; Gervain & Werker 2013). Seven-month-old bilingual learners of a VO language (English) and an OV language (Hindi, Japanese, Urdu, etc.) parse the structurally ambiguous AL designed by Gervain *et al.* (2008) into frequent-initial “phrases” when enriched with VO prosody (i.e. a contrast in duration: frequent elements are shorter than infrequent ones), but parse the same AL into frequent-final “phrases” when enriched with OV prosody (i.e. a contrast in pitch: frequent elements have lower pitch than infrequent ones; Gervain & Werker 2013). Further, infants integrate this prosodic cue with the available frequency information. When the two cues appear in conflict—for instance by exposing monolingual learners of a VO, functor-initial language with the AL containing OV prosody—their preference for the familiar frequent-initial word order disappears (Bernard & Gervain 2012; Gervain & Werker 2013), suggesting that infants weigh the available frequency and prosodic cues equally.

To date, a single study has examined adult listeners’ sensitivity to this prosodic cue. De la Cruz-Pavía, Werker, *et al.* (2020) presented monolingual speakers of English (VO, functor-initial) and highly proficient bilingual speakers of English and an OV language (Hindi, Japanese, Tamil, Korean, etc.) with three ALs that differed in the cues to phrase boundaries that they contained: (1) frequency-based cues, (2) frequency and VO prosodic cues, or (3) frequency and OV prosodic cues.²

As predicted for speakers of a VO, functor-initial language, English monolinguals parsed the AL containing only frequency information into frequent-initial “phrases”. English-OV bilinguals exhibited the same parsing preference. Although frequency is an ambiguous cue for these OV-VO bilinguals (which are exposed both to frequent-final and frequent-initial structures in their input), the language of context might have determined their parsing, as all participants (monolinguals and bilinguals) were addressed and received the instructions of the study in English (VO, functor-initial language). The addition of a familiar and redundant cue, namely VO prosody, did

² This study included other conditions manipulating visual cues to phrase boundaries. These conditions are not discussed here, as they are not relevant for the present research.

not heighten English monolinguals' parsing preference for frequent-initial chunks, but it significantly increased that of the bilinguals. The seemingly enhanced attention to this prosodic cue found in bilingual populations supports Gervain and Werker's (2013) proposal that phrasal prosody could allow OV-VO bilinguals to differentiate their two languages and, combined with frequency, discover the basic word order of their native languages during acquisition.

Finally, exposure to the unfamiliar OV prosody led to a reversal of English monolinguals' parsing preference into frequent-final "phrases". OV prosody is the prosodic pattern associated to languages with a functor-final (i.e. frequent-final) order, and was hence in conflict with frequency information, which signalled a frequent-initial order to the English monolinguals, as speakers of a VO, functor-initial language. Chunking elements containing contrasts in pitch into high-low groupings has been hypothesized to be a general auditory bias found in humans (Bion *et al.* 2011) and even rats (de la Mora *et al.* 2013). This result supports the auditory bias hypothesis and, in line with prior studies, suggests that adult listeners might weigh prosodic information more heavily than statistical cues (Shukla *et al.* 2007). By contrast, the presence of OV prosody did not impact English-OV bilinguals' segmentation: a preference for frequent-initial "phrases" obtained, similar to the one this population exhibited when exposed to the AL containing only frequency cues. Note that bilinguals were addressed exclusively in their VO language English during the study. De la Cruz-Pavía, Werker, *et al.* (2020) argue thus that the language of context overrode the simultaneously available prosodic cue, determining bilinguals' parsing preferences of the ambiguous AL.

Manipulating the language of context to put this hypothesis to test was not feasible for de la Cruz-Pavía, Werker, *et al.* (2020), as the bilinguals in their study were native speakers of 12 different OV languages. In the present study, we overcome this difficulty by examining a homogeneous bilingual population, namely Basque-Spanish (OV-VO) speakers. This population is ideal to establish the relative weight of prosodic cues within the hierarchy of bilingual segmentation cues, as previous studies with Basque-Spanish bilinguals have investigated the role and interplay of frequency, context language, and segmental cues.

De la Cruz-Pavía *et al.* (2015) examined highly proficient Basque-Spanish bilinguals' parsing preference of a structurally ambiguous AL containing frequency cues and synthesized in German, i.e. an unfamiliar language. They manipulated the language of context, addressing half of the participants in Basque, the remaining half in Spanish. Participants addressed in their OV language, Basque, showed a greater preference for the frequent-final parsing associated to OV languages, as compared with participants addressed in their VO language, Spanish. Thus, the language of context modulated bilinguals' parsing preferences of the ambiguous AL, although it did not fully reverse them. Indeed, all participants exhibited a general frequent-final parsing preference that the authors hypothesize resulted from acoustic-phonetic information provided by the German voice.

De la Cruz-Pavía *et al.* (2022) replicated this context language effect, in addition to further investigating its interplay with segmental cues. To this end, they designed an AL with the same ambiguous structure as the one in de la Cruz-Pavía *et al.*

(2015), using only phonemes shared by Spanish and Basque³ and synthesized using a Spanish voice. These changes to the AL's segmental information (i.e., the differences in the acoustic-phonetic information of the sounds used in the German vs. the Spanish ALs) led to a reversal of the bilinguals' segmentation preferences to a frequent-initial parsing preference (i.e. a parse associated to VO languages such as Spanish), which was once again modulated by the language of context. This pattern of results suggests a hierarchical organisation of segmentation cues in which segmental cues are more heavily weighed by adult bilinguals than the language of context.

The present study further investigates the relative position of phrasal prosody within the hierarchy of segmentation cues. We examined the parsing preferences of two populations of highly proficient Basque-Spanish bilinguals: Basque natives (henceforth L1Basque-L2Spanish) and Spanish natives (i.e. L1Spanish-L2Basque). We presented them with a structurally ambiguous AL, specifically, the "Spanish" AL originally designed by de la Cruz-Pavía *et al.* (2022), to which we added OV phrasal prosodic cues. Note that, while OV prosody signals a frequent-final parse of the AL, frequency information is ambiguous for these OV-VO bilinguals, who are exposed to both functor-initial and functor-final structures in their input. In contrast with de la Cruz-Pavía, Werker, *et al.* (2020), we addressed participants in their L1 during the study. Therefore, L1Basque bilinguals received the instructions of the study in Basque, while L1Spanish bilinguals were addressed in Spanish.

In de la Cruz-Pavía, Werker, *et al.* (2020) the presence of VO prosody modulated OV-VO bilinguals' parsing preferences of a structurally ambiguous AL, while the presence of OV prosody did not. Note however that these bilinguals were addressed in their VO language during the study. In light of these results, the following predictions can be drawn for the present study: the presence of OV prosody should not modulate the segmentation preference of L1Spanish bilinguals, as these OV-VO bilinguals were addressed in their VO language during the study. In turn, we examine whether OV prosody modulates the responses of L1Basque bilinguals when addressed in their OV language, that is, when language of context and prosody provide convergent information. Integrating these two cues might lead to a frequent-final parsing preference similar to the one observed in adult monolinguals in de la Cruz-Pavía, Werker, *et al.* (2020). Such a pattern would suggest that the integrated prosodic and context language cues override the segmental cues signalling a frequent-initial segmentation of the AL. If, by contrast, segmental cues still outweigh the combined prosodic and context language cues, a modulation towards a greater frequent-final segmentation might obtain, although a full reversal of the bilinguals' parsing preference is not predicted.

2. Methods

2.1. Participants

Participants consisted of 32 highly proficient bilingual speakers of Spanish and Basque (25 females, mean age 21.9, age range 18 to 37). Of these, 16 were native

³ Their phonetic inventories overlap largely, with the Spanish consonantal set being, with a few exceptions, a subset of the Basque one.

speakers of Basque, i.e. L1Basque-L2Spanish bilinguals (12 females, mean age 20.9, age range 18 to 29), while the remaining 16 participants were native speakers of Spanish, i.e. L1Spanish-L2Basque bilinguals (13 females, mean age 22.8, age range 18 to 37). Their linguistic background was assessed by means of a questionnaire developed by members of the research group The Bilingual Mind (UPV/EHU). This questionnaire (reproduced in de la Cruz-Pavía *et al.* 2015, 2022) measures the bilinguals' age of acquisition, self-reported proficiency, and use of their two languages in different contexts and at three different points in their lives (childhood, adolescence, adulthood). All L1Basque bilinguals had been raised in Basque-only speaking homes, and L1Spanish bilinguals in Spanish-only speaking homes during infancy. All bilinguals had had formal education in both their languages and reported being highly proficient in them. Participants received a small compensation for their participation.

2.2. Stimuli

We used the AL designed by de la Cruz-Pavía *et al.* (2022). This AL has an ambiguous structure similar to the AL originally created by Gervain *et al.* (2013) and the ones used in de la Cruz-Pavía *et al.* (2015) and de la Cruz-Pavía, Werker, *et al.* (2020). It comprises two types of categories: frequent and infrequent. Frequent categories consist of categories *a*, *b* and *c*, each containing a single Consonant-Vowel (CV) token ($a = fi$, $b = nu$, $c = pe$; see Figure 1). In turn, infrequent categories consist of categories *X*, *Y* and *Z*, each containing 9 CV tokens ($X = FE, KA...$, $Y = LI, SA...$, $Z = PO, TU...$). The AL has thus a total 3 frequent and 27 infrequent tokens, and all CV tokens contain only phonemes shared in Spanish and Basque, two languages that have largely overlapping inventories. The 6 categories are combined into a basic unit with the structure $aXbYcZ$, that is, in which frequent and infrequent categories strictly alternate. The basic unit is then concatenated 377 times ($aXbYcZaXbYcZaXbYc...$), creating a 9 minute and 3 second long familiarization stream. As a result of this design, tokens from frequent categories (1 token per category) occur 9-times more frequently than tokens of infrequent categories (9 tokens per category). The amplitude of the first and last minute of the stream is ramped, resulting in a structurally ambiguous stream that allows two possible parses: a parse that *starts* with a frequent element (i.e. frequent-initial): $[aXbYcZ]aXb...$, or a parse that *ends* in a frequent element (i.e. frequent-final): $a[XbYcZa]Xb...$

Test stimuli consist of 36 hexasyllabic sequences, half with a frequent-initial order (e.g. $aXbYcZ$: $fiLUnuSApeKI$), the remaining half with a frequent-final order (e.g. $XbYcZa$: $KAnuLIpePOfi$). The six categories occur with equal frequency in all positions within sequences.

The familiarization stream and 36 test items are synthesized with the es1 (Spanish male) voice of the MBROLA database (Dutoit 1997). All CV tokens have a constant duration of 120 ms per segment and a flat intensity level. Because the AL in de la Cruz-Pavía *et al.* (2022) contained only frequency-based information, CV tokens had a constant pitch of 100 Hz. In the present study we added a prosodic cue, specifically the pattern characteristic of OV, functor-final languages: a contrast in pitch (in which the content word has a higher pitch than the functor; de la Cruz-Pavía,

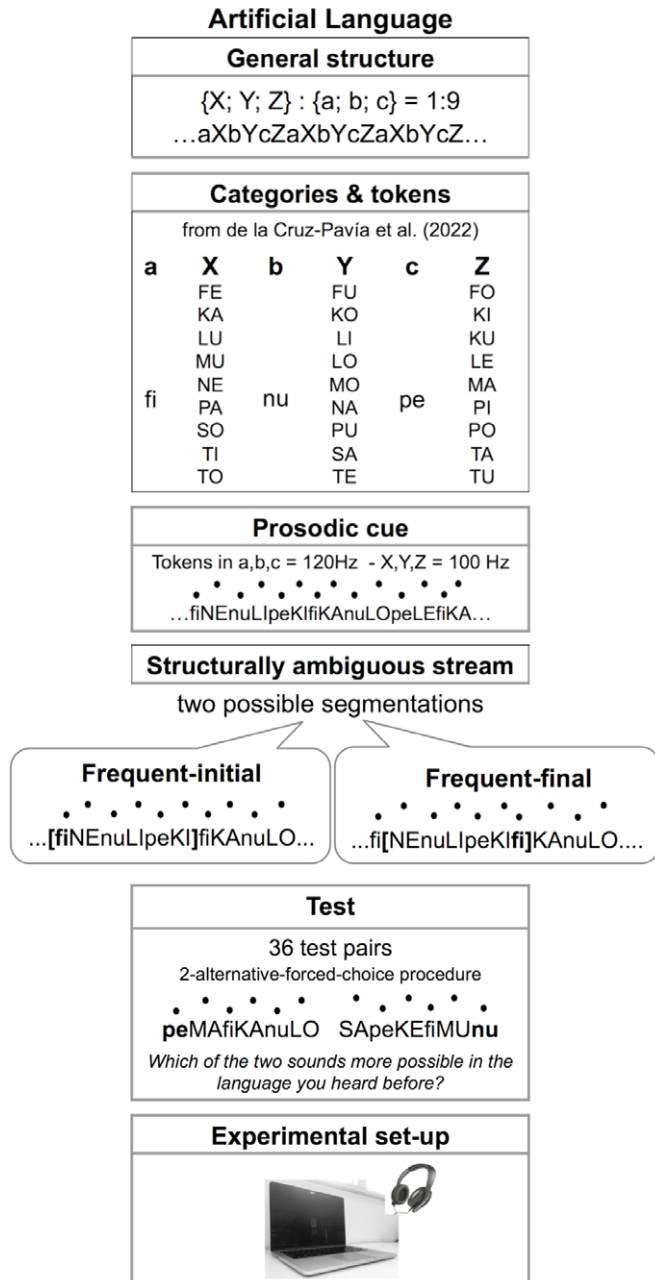


Figure 1

From top to bottom: general structure of the artificial language, categories and tokens, prosodic manipulation, the two possible word orders of the ambiguous stream, test items and task, and experimental set-up.

Gervain, *et al.* 2020; Gervain & Werker 2013; Nespor *et al.* 2008). We synthesized tokens from infrequent categories with a higher f0 than tokens in frequent categories (i.e. 120 vs. 100 Hz). The strict alternation of frequent and infrequent syllables tallied hence with the strict alternation of pitch falls and rises both in the familiarization stream (...aXbYcZaXbYcZaXbYc...: *fiNÉnuLÍpeKÍfiKÁnuLÓpeLÉfiKÁ...*) and test sequences (e.g. *fiLÚnuSÁpeKÍ*, *KÁnuLÍpePÓfi...*).

2.3. Procedure

Participants were tested individually in a quiet room at the University of the Basque Country UPV/EHU's psycholinguistics laboratory (Vitoria-Gasteiz, Spain). The experiment was displayed in a computer screen using DMDX software (Forster & Forster 2003), and participants were provided with high quality Logitech headphones. All participants were addressed in their L1. Thus, the group of L1Basque-L2Spanish bilinguals received the instructions of the study in Basque, whereas the group of L1Spanish-L2Basque bilinguals was instead addressed in Spanish. Participants were instructed that the study had two parts. First they would listen to an unknown language for 9 minutes, and were asked to pay close attention, as during the second part of the study they would answer questions about this language. Participants then completed a short training, in order to get familiarized with the procedure of the study. During training, participants heard six pairs of syllables (e.g. *me so*) and their task consisted on identifying a target syllable (i.e. *so*), and pressing one of two keys in the keyboard, depending on whether the target syllable had been heard first or second within the pair.

After the instruction and the training, participants listened to the 9-minute-long familiarization, immediately followed by the test phase, during which they listened to a total of 36 trials. In each test trial, participants heard two six-syllable long sequences, one with a frequent-initial order and the other one with a frequent-final order, separated by a 500 ms silence. The order of presentation of the two sequence types was counterbalanced across trials. Every test sequence appeared twice in the test phase, once as first member of a pair (e.g. *fiTInuTEpeKU* —500ms pause— *FUpeTAfiNÉnu*), and another time as second member of a different pair (e.g. *TAfiMUnuLÍpe* —500ms pause— *fiTInuTEpeKU*), although never in consecutive test trials. Participants' task consisted on choosing, for each pair, which sequence —the first or second in the pair— they thought sounded more like a possible sequence in the language heard during familiarization. To indicate this, they pressed one of two predefined keys in the keyboard. The session had a total duration of less than 30 minutes.

3. Results

DMDX recorded the number of frequent-final responses per participant out of the 36 test trials. Participants were presented with a two-alternative forced choice between frequent-initial and frequent-final sequences. Due to the binomial nature of their responses, and following previous studies (de la Cruz-Pavía *et al.* 2015, 2022; de la Cruz-Pavía, Werker, *et al.* 2020), we analysed them using binomial tests of proportions. All analyses were conducted in R (version 3.6.1., R Core Team 2019).

First, we examined whether the participants' responses differed significantly from chance. Within-group binomial tests of proportions revealed that both groups of bilinguals had a numerically frequent-initial parsing preference of the structurally ambiguous AL (see Table 1 and Figure 2). This preference only reached significance in the group of L1Spanish bilinguals (14.63/36, 40.63%; $p < .0001$), remaining near-significant in the group of L1Basque bilinguals (16.50/36, 45.83%; $p = .05$). Comparison of the two groups with a binomial test of proportions revealed that their parsing preferences did not differ significantly ($\chi^2(1, N = 32) = 3, p = .08$). In sum, the presence of a pitch contrast, prosodic pattern associated with OV, functor-final languages, did not yield a frequent-final parsing preference of the AL neither in native nor in non-native but highly proficient speakers of Basque. In other words, the presence of this prosodic cue did not determine adult bilinguals' parsing preference of the AL.

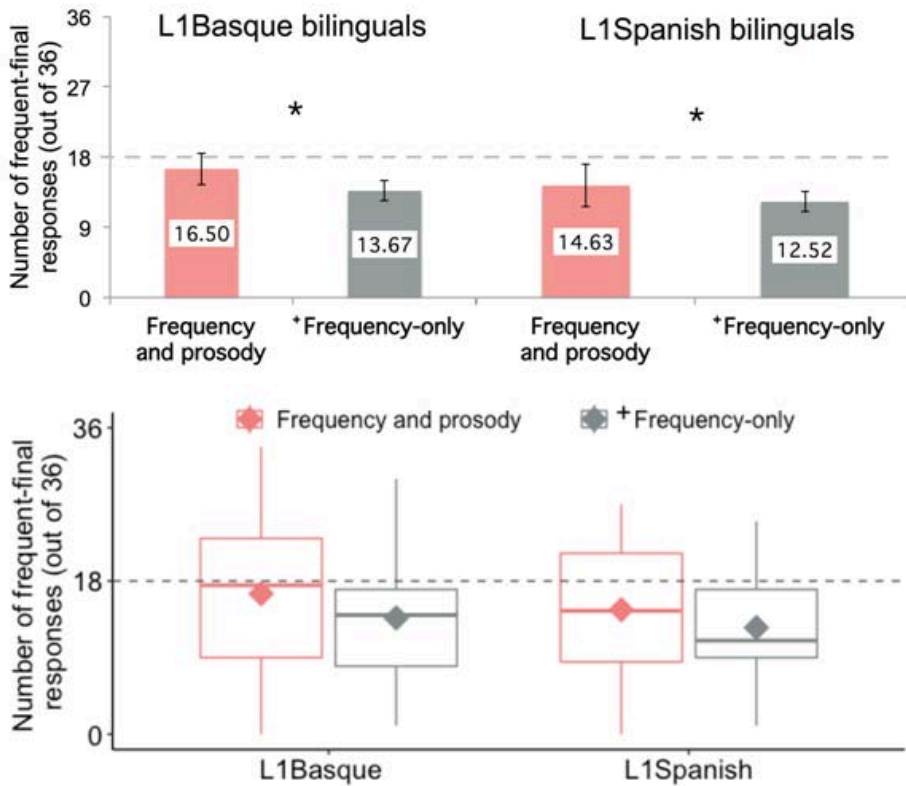
Table 1

Mean number of frequent-final responses out of the 36 test trials, standard error (SE), and 95% confidence intervals (CI) obtained in the two groups of Basque-Spanish bilinguals (L1Basque and L1Spanish) tested in the present study, and the two similar groups tested by de la Cruz-Pavía *et al.* (2022) (marked by a * preceding the values)

	Number of frequent-final responses out of the 36 trials	
	L1Basque-L2Spanish	L1Spanish-L2Basque
AL containing frequency-based and prosodic cues	16.50, ± 2.74 SE 10.67-22.33	14.63, ± 2.07 SE 10.21-19.04
AL containing only frequency-based cues	*13.67, ± 1.28 SE 11.03-16.30	*12.52, ± 1.35 SE 9.70-15.34

* From de la Cruz-Pavía *et al.* (2022).

In order to determine whether the addition of pitch cues nonetheless modulated the bilinguals' parsing preferences, we compared the results of the present study with the results of two similar groups tested in de la Cruz-Pavía *et al.* (2022) and presented only with frequency information (but with an otherwise identical AL). The two groups of L1Basque and L1Spanish bilinguals in de la Cruz-Pavía *et al.* (2022) had a significant frequent-initial parsing preference of the AL (L1Basque: 13.67/36, 37.96%; $p < .001$; L1Spanish: 12.52/36, 34.78%; $p < .001$; see Table 1 and Figure 2). Comparison of the two groups of L1Basque bilinguals (frequency-only: de la Cruz-Pavía *et al.* 2022 vs. frequency and prosody: present study) with a binomial test of proportions revealed that the addition of prosodic cues to the AL resulted in a significant increase in their number of frequent-final responses ($\chi^2(1, N = 43) = 9, p = .003$). A similar increase in frequent-final responses obtained in the comparison of the two groups of L1Spanish bilinguals ($\chi^2(1, N = 37) = 5, p = .003$). The addition of prosodic cues modulated hence the bilinguals' parsing preferences of the ambiguous AL.



* Frequency-only groups from De la Cruz-Pavía *et al.* (2022).

Figure 2

Parsing preferences of the structurally ambiguous ALs. The bar graphs with standard error (top) and the boxplots (bottom) depict the mean number and distribution of frequent-final responses out of the 36 test trials, in the two groups of Basque-Spanish bilinguals (L1 Basque and L1 Spanish) exposed to concurrent frequency and prosodic cues (coral) and the two similar groups presented only with frequency in de la Cruz-Pavía *et al.* (2022) (grey). A mean below 18 (chance) indicates a frequent-initial segmentation preference, whereas a mean above 18 indicates a frequent-final segmentation preference

4. Discussion

The present study examined the relative weight of phrasal prosody within the hierarchy of speech segmentation cues available to adult bilinguals. To that end, we presented Basque-Spanish bilinguals, that is, speakers of an OV and a VO language, with a structurally ambiguous AL enriched with the prosodic cue associated to OV languages, namely a contrast in pitch in which the element receiving prominence in a phrase has higher pitch).

The AL, originally designed by de la Cruz-Pavía *et al.* (2022), consisted of strictly alternating frequent and infrequent elements, mimicking the frequency distribution of functors and content words in natural languages. Because OV-VO bilinguals are exposed to both functor-initial and functor-final structures in their input, this frequency information was ambiguous for our participants. The AL was built only with phonemes shared by the Spanish and Basque inventories, and was synthesized using a Spanish voice. De la Cruz-Pavía *et al.* (2022) showed that, in the absence of prosodic cues, this segmental information led Basque-Spanish bilinguals to parse the AL into the frequent-initial order associated to their VO language, Spanish. De la Cruz-Pavía and colleagues showed that the language of context additionally modulated the parsing preferences of these bilinguals, as addressing them in their OV language, Basque, significantly increased their preference for the frequent-final order associated to OV, functor-final languages, as compared with addressing them in their VO language, Spanish. While the language of context impacted bilinguals' segmentation of the AL, it did not override the frequent-initial bias created by its segmental information. In the present study, we examined whether the addition of OV phrasal prosody modulated or determined the bilinguals' parsing preference of this AL, yielding a greater frequent-final segmentation.

We tested two groups of highly proficient Basque-Spanish bilinguals that differed in their native language —L1Basque-L2Spanish and L1Spanish-L2Basque speakers— and addressed both groups in their respective L1. Analysis of their segmentation preferences revealed that the presence of OV prosody modulated their parsing of the ambiguous AL. Thus, both groups of bilinguals exhibited a greater frequent-final segmentation preference (order associated to OV languages), as compared with two similar groups presented with a prosodically flat variant of the same AL (in De la Cruz-Pavía *et al.* 2022). However, phrasal prosodic cues did not *determine* the segmentation patterns, as both groups still exhibited a frequent-initial parsing preference (only marginally significant in the case of L1Basque bilinguals).

A single study to date had examined adult bilingual's use of phrasal prosody in speech segmentation. De la Cruz-Pavía, Werker *et al.* (2020) observed that, when addressed in their VO language (English) during the study, the presence of VO prosody modulated adult OV-VO bilinguals' parsing of a ambiguous AL similar to the one employed in the present study. By contrast, the presence of OV prosody, in conflict with the language of context, did not. L1Basque bilinguals in the present study were provided with congruent prosodic and context language cues (as they were addressed in their OV language, Basque during the study), both signalling a frequent-final parse of the AL. These converging cues led, as predicted, to a greater frequent-final segmentation preference. Integration of these two cues did however not suffice to overcome the frequent-initial bias created by segmental information. This pattern of results provides thus further evidence of the hierarchical organisation of the phrase segmentation cues available to adult bilinguals, and reveals that segmental cues are more heavily weighed than prosody and language of context (even when combined).

The results in de la Cruz-Pavía, Werker, *et al.* (2020), in which the presence of OV prosody did not modulate the bilinguals' parsing of the AL when addressed in their VO language, led us to predict that OV prosody would not modulate the parsing preference of the present study's group of L1Spanish-L2Basque bilin-

guals either, as they were also addressed in their VO language, Spanish, during the study. Contrary to prediction, the L1Spanish bilinguals chose a greater number of frequent-final items when exposed to OV prosody, as compared with the group of L1Spanish bilinguals exposed to prosodically flat variant in de la Cruz-Pavía *et al.* (2022). That is, while the presence of OV prosody did not modulate the parsing preference of English-OV bilinguals, it increased the L1Spanish-L2Basque bilinguals' preference for a frequent-final order, despite the fact that both groups of bilinguals were addressed in their VO language during the study. The origin of this discrepancy is unclear. However, we speculate that it might result from a methodological difference between the two studies. In de la Cruz-Pavía, Werker, *et al.* (2020), the familiarization stream was enriched with OV prosody, while test items were presented prosodically flat (that is, they had a constant pitch of 100 Hz). By contrast, both familiarization stream and test items carried a pitch contrast in the present study. The presence of pitch cues in test might thus have heightened our participants' sensitivity or use of this source of information. An experiment presenting Basque-Spanish bilinguals with prosodically flat test stimuli is pending, to put this hypothesis to test.

To conclude, the results of the present study confirm that phrasal prosody is an available cue to phrase boundaries for adult listeners, here specifically bilinguals, and show that prosody is a less weighed cue than the segmental information contained in the input. This research contributes new evidence of the hierarchical arrangement of phrase segmentation cues, and reveals an organisation that tallies with the one attested in the segmentation of smaller units, as segmental cues have also been found to outweigh prosodic information in word segmentation (Mattys *et al.* 2005).

References

- Bernard, Carline & Judit Gervain. 2012. Prosodic Cues to Word Order: What Level of Representation? *Frontiers in Psychology* 3. 451. <https://doi.org/10.3389/fpsyg.2012.00451>.
- Bion, Ricardo A. H., Silvia Benavides-Varela & Marina Nespor. 2011. Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech* 54(1). 123-140. <https://doi.org/10.1177/0023830910388018>.
- de la Cruz-Pavía, Irene, Gorka Elordieta, Nuria Sebastián-Gallés & Itziar Laka. 2015. On the role of frequency-based cues in the segmentation strategies of adult OV-VO bilinguals. *International Journal of Bilingual Education and Bilingualism* 18(2). 225-241. <https://doi.org/10.1080/13670050.2014.904839>.
- de la Cruz-Pavía, Irene, Gorka Elordieta, Julián Villegas, Judit Gervain & Itziar Laka. 2022. Segmental information drives adult bilingual phrase segmentation preference. *International Journal of Bilingual Education and Bilingualism* 25(2). 676-695. <https://doi.org/10.1080/13670050.2020.1713045>.
- de la Cruz-Pavía, Irene, Judit Gervain, Eric Vatikiotis-Bateson & Janet F. Werker. 2019. Finding phrases: On the role of co-verbal facial information in learning word order in infancy. *PLOS ONE* 14(11). e0224786. <https://doi.org/10.1371/journal.pone.0224786>.
- de la Cruz-Pavía, Irene, Judit Gervain, Eric Vatikiotis-Bateson & Janet F. Werker. 2020. Coverbal speech gestures signal phrase boundaries: A production study of Japanese and

- English infant- and adult-directed speech. *Language Acquisition* 27(2). 160-186. <https://doi.org/10.1080/10489223.2019.1659276>.
- de la Cruz-Pavía, Irene, Janet F. Werker, Eric Vatikiotis-Bateson & Judit Gervain. 2020. Finding Phrases: The Interplay of Word Frequency, Phrasal Prosody and Co-speech Visual Information in Chunking Speech by Monolingual and Bilingual Adults. *Language and Speech* 63(2). 264-291. <https://doi.org/10.1177/0023830919842353>.
- de la Mora, Daniela M., Marina Nespor & Juan M. Toro. 2013. Do humans and nonhuman animals share the grouping principles of the iambic–trochaic law? *Attention, Perception, & Psychophysics* 75(1). 92-100. <https://doi.org/10.3758/s13414-012-0371-3>.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81-138. <https://doi.org/10.1353/lan.1992.0028>.
- Dutoit, Thierry. 1997. *An introduction to text-to-speech synthesis*. Amsterdam: Kluwer Academic Publishers.
- Forster, Kenneth I. & Jonathan C. Forster. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35(1). 116-124. <https://doi.org/10.3758/BF03195503>.
- Gervain, Judit, Marina Nespor, Reiko Mazuka, Ryota Horie & Jacques Mehler. 2008. Bootstrapping word order in prelexical infants: A Japanese–Italian cross-linguistic study. *Cognitive Psychology* 57(1). 56-74. <https://doi.org/10.1016/j.cogpsych.2007.12.001>.
- Gervain, Judit, Nuria Sebastián-Gallés, Begoña Díaz, Itziar Laka, Reiko Mazuka, Naoto Yamane, Marina Nespor & Jacques Mehler. 2013. Word frequency cues word order in adults: Cross-linguistic evidence. *Frontiers in Psychology* 4. <https://doi.org/10.3389/fpsyg.2013.00689>.
- Gervain, Judit & Janet F. Werker. 2013. Prosody cues word order in 7-month-old bilingual infants. *Nature Communications* 4(1). 1490. <https://doi.org/10.1038/ncomms2430>.
- Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph H. Greenberg (ed.), *Universals of Human Language*, 73-113. Cambridge: MIT Press.
- Hahn, Michael & Yang Xu. 2022. Crosslinguistic word order variation reflects evolutionary pressures of dependency and information locality. *PNAS* 119(24). e2122604119. <https://doi.org/10.1073/pnas.2122604119>.
- Mattys, Sven L., Laurence White & James F. Melhorn. 2005. Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology: General* 134(4). 477-500. <https://doi.org/10.1037/0096-3445.134.4.477>.
- Mazuka, Reiko. 1996. Can a grammatical parameter be set before the first word? Prosodic contributions to early setting of a grammatical parameter. In James L. Morgan & Katherine Demuth (eds.), *Signal to Syntax*, 313-330. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Morgan, James L., Rushen Shi & Paul Allopenna. 1996. Perceptual Bases of Rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In James L. Morgan & Katherine Demuth (eds.), *Signal to Syntax*, 263-283. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Nespor, Marina, Mohinish Shukla, Ruben van de Vijver, Cinzia Avesani, Hanna Schraudolf & Caterina Donati. 2008. Different Phrasal Prominence Realizations in VO and OV Languages. *Lingue E Linguaggio* 7(2). 1-28.

- R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/> (29/12/2023).
- Shukla, Mohinish, Marina Nespor & Jacques Mehler. 2007. An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology* 54(1). 1-32. <https://doi.org/10.1016/j.cogpsych.2006.04.002>.