

QUANTITATIVE ANALYSIS OF FUNDAMENTAL FREQUENCY IN SPANISH (L2) AND BRAZILIAN PORTUGUESE (L1): EVIDENCE OF LEARNING AND LANGUAGE ATTRITION

SILVA, Cristiane Conceição^{1*}

ARANTES, Pablo²

¹Professor at the Departamento de Língua e Literatura Estrangeiras, Universidade Federal de Santa Catarina – ORCID: <https://orcid.org/0000-0002-2086-8523>

²Professor at the Departamento de Letras, Universidade Federal de São Carlos – ORCID: <https://orcid.org/0000-0001-9707-8493>

Abstract: *This paper analyzes the intonation of Spanish and Brazilian Portuguese (BP) produced by monolingual speakers of both languages and bilingual BP speakers that lived in Spain on average for 6 years. Bilinguals produced data in both Spanish L2 (BL2) and BP L1 (BL1). Speech materials are sentences in different modalities (declaratives, yes-no and wh-questions) and reading styles (isolated sentences and storytelling). Fundamental frequency (f₀) contours were analyzed to assess learning in Spanish L2 and language attrition in the L1 production of bilinguals. Variability in the f₀ contours of the four language conditions was gauged by means of three indices (peak rate, peak range and global standard deviation). Dynamic time warping (DTW) distances between pairs of f₀ contours were also measured as a way to measure within- and between-language differences in intonation patterns. The main results are: 1) BL2 and BL1 contours are significantly more variable than the monolingual ones both quantitatively and qualitatively; 2) BL2 contours partially converge towards the patterns of Spanish monolinguals, showing that there is learning; 3) there is evidence for language attrition in the form of transfer of Spanish patterns to BP contours produced by bilinguals; 4) Learning and attrition levels are different depending on sentence modality, such that learning is greater in modalities that differ less between BP and Spanish and attrition is greater in modalities that differ the most.*

Keywords: Bilingualism; Intonation; Language attrition; Spanish L2; Brazilian Portuguese L1.



1* Corresponding author: cristiane.conceicao@ufsc.br

1 Introduction

Intonation conveys many different communicative functions (1–5). In this paper we focus on the modality function and analyze intonational differences between fundamental frequency (henceforth f_0) contours in Spanish as a second language (L2) compared to Spanish as a mother tongue (L1) and Brazilian Portuguese (henceforth BP) as a mother tongue spoken by monolingual (from now on MP) and bilingual Brazilian speakers (from now on BL1) in declarative sentences, yes-no and wh-questions. We will look for evidence of learning in the production of L2 speakers by comparing it to the monolingual Spanish speakers and also for evidence of language attrition by comparing bilingual's L1 production to the control monolingual BP speakers. In the following subsections of this introduction, we discuss in detail the subjects that we bring together in this study. Firstly, we present the most relevant details of how both monolingual Spanish and BP speakers implement sentence modality in terms of intonation (sections 1.1 and 1.2). We also present results of studies that deal with Brazilian bilinguals learning Spanish L2 intonation (section 1.3). Then, we review references related to language attrition that the interpretation of the results we present here is based on (section 1.4). Lastly, we lay out the goals of the present study and the research questions we investigate (section 1.5).

1.1 Intonation of Spanish L1

There are many studies on the intonation of declarative and interrogative sentences in Castilian Spanish (6–10). Based on the Autosegmental metrical theory (AM), Estebas-Vilaplana and Prieto (10) described the basic intonation patterns of Castilian Spanish. The authors analyzed many different sentence modalities. However, we focus on their analysis of the broad-focus declaratives and yes-no and wh-questions intonation. The authors transcribe declaratives with a $L^+>H^*$ pitch accent in pre-nuclear position, indicating that the rising movement develops within the stressed syllable and culminates in the post-stressed syllable. The contour, then, progressively descends throughout the rest of the sentence up to its end at a low value. The final pitch accent does not show any relevant pitch movement, it is just part of the progressive descent in f_0 . Since the low f_0 is reached during the last stressed syllable, this final pattern is described by a L^* pitch accent followed by a boundary tone $L\%$. Yes-no questions are described as having a pre-nuclear L^*+H accent. The nuclear accent shows a f_0 drop which, according to the authors, can be interpreted as a L^* pitch accent. The final rising movement is described as $HH\%$ given the rapid rising movement at the end of the sentence. Finally, the authors observed in wh-question two different intonational patterns in nuclear position: a falling contour ($L^*L\%$) or a rising contour ($L^*HH\%$), the latter expressing “a nuance of interest and greater speaker involvement in the speech act” (10, p. 35). In pre-nuclear position, the pitch accent is the same (H^* , aligned to the interrogative pronoun) for both types of final contour.

Henriksen (11) carried out an experimental study to analyze only wh-questions. His study aimed at identifying how many different possible f_0 contours would be produced in peninsular Spanish. Six Spanish speakers from León took part in the production experiment. The experiment consisted of producing wh-question sentences under two different conditions: (1) participation in what the author describes as a short task-based dialogue, which the author called a person identification task and (2) a sentence reading task.

The author found four different f_0 contours: (1) final rise (f_0 peak at the offset of the first stressed syllable of the sentence, gradual fall throughout the last stressed syllable, a nuclear valley and a boundary rise); (2) nuclear circumflex (tonal plateau that starts at the onset of the sentence and lasts until the beginning of the nuclear stressed syllable or the initial tonal level is slightly higher and followed by a falling movement until the final position. In both cases, there is a rise-

fall movement aligned to the stressed syllable of the last content word); (3) global falling (the highest f_0 peak is aligned to the initial wh-word, the initial rise is followed by a gradual descent throughout the sentence until a valley at the nuclear stressed syllable, in some cases the f_0 continues its descent throughout the final syllable and did not fall to a valley); (4) nuclear falling (rise to a f_0 peak on the wh-words, a plateau that lasts until the nuclear stressed syllable with a falling movement and a tonal valley at the end of the sentence. The author showed that there was a significant effect of speech style on the use of different contour types. In the reading task, 70% of occurrences were of final rise, 13% of global falling, 10% of nuclear circumflex and 7% of nuclear falling contours. In the person identification task, there was much less variability, 57% produced a final rise and 43% a nuclear circumflex.

1.2 Intonation of Brazilian Portuguese L1

As is the case for Spanish L1, intonation of declarative and interrogative sentences in BP has also been the object of extensive research (12–18). According to Frota and Moraes (18), declarative sentences have one pitch accent per prosodic word in pre-nuclear position and non-final stressed syllables typically exhibit a rising melody (L+H). The nuclear contour, represented by a H+L* pitch accent, surfaces as a movement that starts in a peak and develops into a fall that extends through the stressed syllable towards a low target. The nuclear accent in yes-no questions presents a circumflex or rise-fall contour (L+H*L%) that surfaces a rising movement through the final stressed syllable with a late alignment of the f_0 peak followed by a falling movement in the post-stressed syllables (13). According to Frota and Moraes (18), this complex rise–fall final contour truncates in sentences that end in words with final stress. Lastly, in wh-questions the pre-nuclear contour has an extra high initial f_0 peak aligned to the wh-word (H+H*). Then, there is a gradual fall movement over the following syllables up the last stressed syllable in the sentence. The nuclear contour is the same as the declarative sentences, followed by a low boundary tone (H+L*L%) (13).

1.3 Intonation of Spanish L2

The first studies about the intonation of declarative and interrogative sentences in Spanish as an L2 spoken by Brazilians analyzed mainly the productions of Brazilians who learned Spanish in Brazil. Most of these studies used the AM theory to analyze the intonation in Spanish L2, typically describing and comparing the pre-nuclear and nuclear pitch accents in Spanish L2 compared to Spanish L1, especially the peninsular variety (19–22). In more recent research, the L2 intonation of Brazilians has been analyzed from different theoretical perspectives - see, for instance, (23–26).

Sá (19) analyzed the intonation in Spanish Foreign Language (SFL)² and Spanish L1 in reading sentences and in jokes. Two Brazilians and one Spanish participated in the study. She observed that the pitch accents in SFL and Spanish L1 were more similar only in declaratives. Regarding yes-no questions the native produced the initial L*+H pitch accent and, for the most part, the L*H% final pitch accent and learners used H*+H or H*+L tones in initial position and L+H*L% in final position. In wh-questions, native and learners produced similar patterns in pre-nuclear position, although differed in the final position, the native using a high tone and learners a low tone.

² “Spanish Foreign Language” is used by some authors to refer to learning contexts in which the L2 is not the dominant language and learning happens in more formal and institutionalized contexts, such as in school environments. This is in contrast to “Spanish L2”, an expression that usually refers to studies in which learning happens in more naturalistic situations (27).

Oliveira (23) studied L2 Spanish intonation using a model proposed by Cantero Serena (28).³ She analyzed data from twelve Brazilians who lived in Barcelona and Valladolid. The data were recorded during an interview in which participants talked about their experiences with the Spanish language and culture. She analyzed 152 declaratives, 28 yes-no and 15 wh-questions. Regarding the final contour of the declarative sentences, the author observed that 45% had a suspended melody (slight fall), 38% neutral melody (steeper fall) and 17% emphatic (circumflex-type contour). In yes-no questions, 57% of the cases had the final Spanish L1 contour (ascending), 33% BP contour (circumflex) and 8% suspended. For wh-questions, the contour type distribution observed was: 52% emphatic contour, 33% suspended and 20% neutral. The fact that the author investigated spontaneous speech might explain the presence of more than one contour form for each sentence modality, contrary to many of the previously cited studies, which mostly analyzed read speech.

Unlike studies that analyze transfer of contour features from BP to SFL as an all-or-nothing phenomenon (19,20), Oliveira's study points to a gradient variability in the production of the f_0 contours in Spanish L2. In each of the modalities analyzed, the author observed the occurrence of three types of final contours, which highlights the variability present in L2 production. This variability was also observed by Silva (26), who analyzed the production of 15 Brazilians who lived in Madrid. She analyzed the data using the PENTA framework (1). Her goal was to determine if the values for the PENTA model parameters inferred from the Spanish L2 f_0 contours in her corpus were more similar to those of Spanish L1, to those of BP L1, or if they have their own characteristics. In order to do that, the author analyzed three communicative functions: prominence, boundary and modality (declarative, yes-no and wh-questions).

Figure 1 shows f_0 contours over the prominent phonological word in final position in a declarative sentence in the three language conditions: Spanish L1, BP L1 and Spanish L2 (26). We can see that Spanish L1 speakers produce a smoother fall towards the low boundary. BP L1 contours, in comparison, tend to show sharper final falls. Spanish L2 productions show variability: it is possible to see sharper falls similar to those typical in BP L1 and also a good number of smoother falls similar to what is seen for Spanish L1.

³ The model consists of a f_0 normalization procedure based on measuring percent change in f_0 at the midpoint of each successive vowel along an utterance.

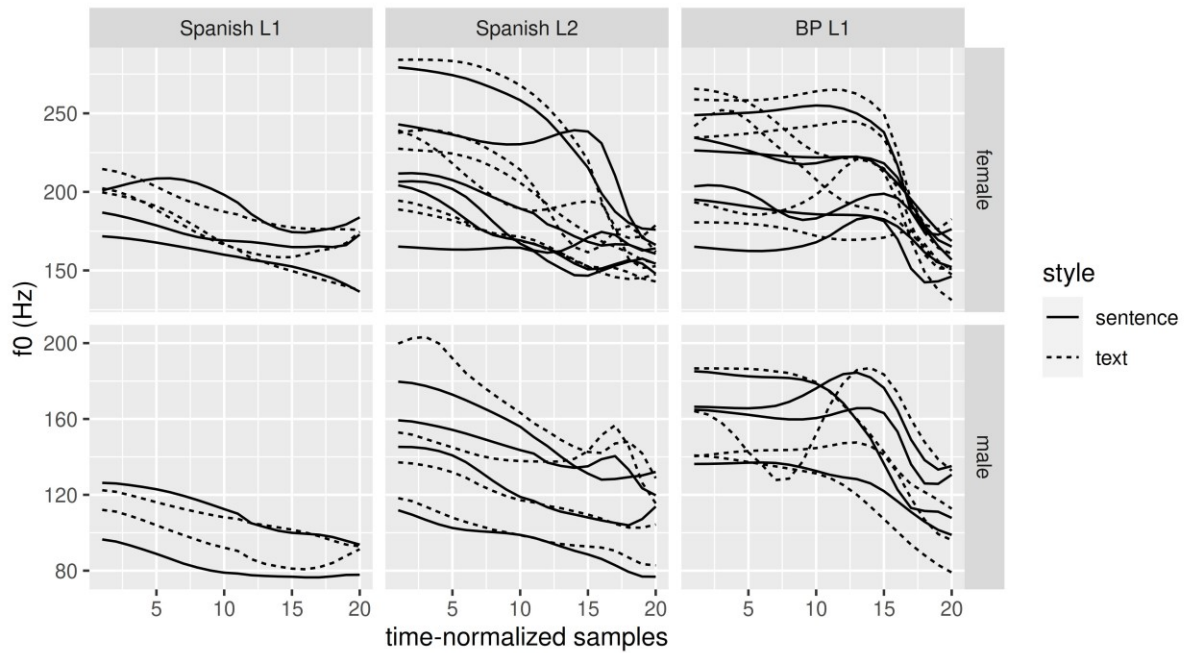


Figure 1: Time-normalized f_0 contours of the last prominent word in read sentences and running text in Spanish and BP declarative sentence *Sancho era un hombre de poca estatura y tenía un barrigón* (Spanish) - *Sancho era um homem de pouca estatura e tinha um barrigão* (BP) ‘Sancho was a short man and had a big belly’. Spanish L1 contours are shown in the left column, Spanish L2 in the middle and BP L1 in the right. Female data is shown in the top row and male data in the bottom.

Further, her results showed that in prominent phonological words in continuative contours the PENTA model parameter *height* was higher in Spanish L2, indicating that Spanish L2 contours tend, *ceteris paribus*, to be globally shifted upwards in comparison to their Spanish L1 counterparts. This effect can be seen in Figure 2 adapted from Silva (25, p. 167) that shows contours of the same Spanish wh-question produced by Spanish L1 and Spanish L2 speakers. Visual inspection shows that contours by both male and female Spanish L2 speakers tend to be produced at a slightly, although noticeable, higher register and to show greater f_0 modulation than the corresponding Spanish L1 counterparts. This contour height difference may result in L2 intonation being perceived by L1 listeners as having greater prominence than the monolingual intonation, although this hypothesis was not instrumentally tested in her work.

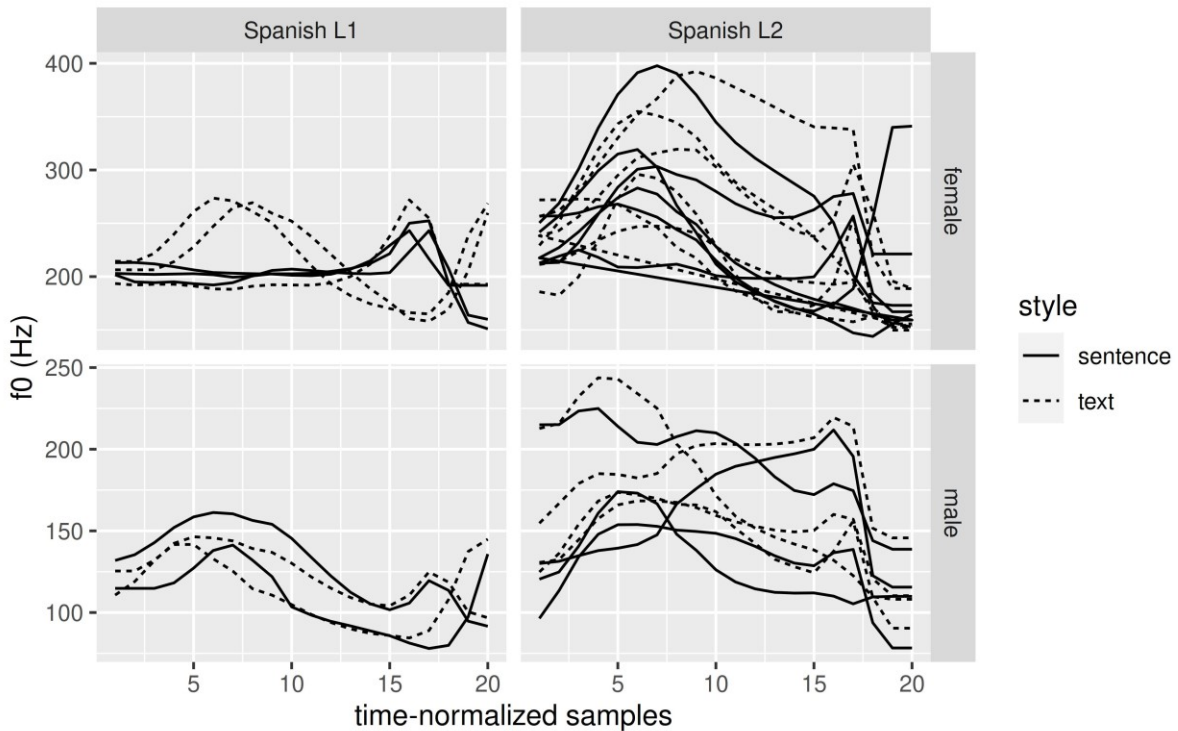


Figure 2: Time-normalized f_0 contours of the wh-question *¿Dónde están mis libros?* ‘Where are my books?’ produced by bilingual Brazilian speakers (7 female and 4 male) and Spanish speakers (3 female and 2 male).

Silva’s data present evidence of L2 learning, since some of the participants are able to reproduce aspects of Spanish L1 f_0 contours, although learning was not complete, since there is evidence of L1 transfer in the data as well. One result in Silva’s data that the author did not anticipate in her initial hypotheses was the fact that the L1 production of bilingual speakers showed more variability in the implementation of tonal patterns than what was expected by the traditional description of BP L1 intonation. It was observed that some BP L1 contours produced by BP bilinguals presented patterns similar to the ones observed in Spanish L1, hinting at the occurrence of L2 transfer to L1. Based on this result, a suggestion was made that this could be the result of language attrition. As Silva’s study did not include BP monolingual speakers that could be used as a control condition, this hypothesis could not be properly tested. One of the aims of the present study is to add this control group to the speech material studied by Silva (26) in order to further explore the possible occurrence of language attrition.

1.4 Language attrition

The phenomenon of language attrition can be defined as a “non-pathological, non-age related, structural change of an L1 within a late bilingual, assuming that the acquisition of the L1 precedes this change” (28, p. 3). Regarding the manifestation of attrition, there has been a debate in the history of the field about which particular changes in L1 could actually be considered manifestations of language attrition, as opposed to short-lived, performance-related changes⁴.

⁴ Chang (30), for instance, advances a proposal to distinguish language attrition from what he calls phonetic drift. In his view, what distinguishes drift from attrition is the fact that the first happens in the short-term and arises from recent exposure to the L2 and attrition persists past the decline of exposure to the L2. A criticism that could be levied against Chang’s distinction is that the author provides no objective definitions of “recent” and “long-term” and the author himself recognizes that “few studies [...] have actually tracked L1 learners through alternations in language environment, with the result that it is often unclear whether observed L1 changes are short-term, long-term, or medium-

Since the emergence of language attrition as a research field, the theoretical understanding of the phenomenon has evolved significantly. In current accounts, this division appears to have been replaced by a consensus view that attrition “is not [...] an ‘extreme’ form of development, experienced by a small minority of bilinguals with decades of no or very little exposure, but a form of language development experienced from the early stages of L2 development and thus, in all likelihood, common to all bilinguals” (30, p. 3), supporting the conclusion, stated elsewhere, that “all instances of L2-to-L1 transfer should therefore be considered attrition phenomena” (31, p. 5) regardless of time frame or considerations on how “deep” or “shallow” the changes in L1 are. In a recent overview presentation of the field, Kupske (33) points out that one common scenario in which changes in L1 are triggered due to the development and dominance of other languages is when there is “a change in the linguistic environment, such as migratory processes” (32, p. 101, translation is ours). This is the situation of all bilingual participants in the data sample we analyze here (see section 2.1 for sociodemographic information on the participants). Migration, however, is not the only possible scenario conducive to language attrition. It has been shown that attrition can occur in L1-dominant contexts as well (34).

In studies on attrition at the phonetic and phonological levels, there is currently a concentration on the acoustic analysis of consonants and vowels and still few studies on such influences in prosody. The studies by Flege (35); Major (36); Chang (37); Chang (38); Kupske (39), for example, analyzed the production of consonants and/or vowels in an L2-dominant context. There are also works on the production of consonants and vowels in L1-dominant contexts (34,40–45).

Regarding prosody, for example, there are few studies (29,46). Mennen (46) studied intonation in native Dutch speakers that learned Greek as L2 and looked for evidence of attrition caused by Greek in L1 Dutch. Dutch has long and short vowels, and f_0 peaks in pre-nuclear position align differently depending on vowel length. Greek, on the other hand, does not contrast vowel length. Results show that the difference between long and short vowels was reduced in the production of Dutch bilinguals. As a consequence of this quantity merger, the f_0 alignment difference is lost. Leeuw et al. (29) looked for evidence of attrition in L1 German caused by English in bilinguals. The authors looked for differences in the alignment of f_0 peaks in pre-nuclear position in both languages. Peaks align later in German in this context when compared to British English. The authors hypothesized that attrition would cause German bilinguals to align peaks earlier in pre-nuclear position influenced by the typical English pattern. Results show partial evidence for language attrition: although there is no statistically significant difference in the alignment of the peak in German bilinguals, the valley preceding the peak has a significantly earlier alignment than the German monolingual control group.

1.5 Study goals and research questions

Most of the previous literature on L2 Spanish intonation learning and cross-language influences bases its analysis on the Autosegmental-Metrical theory (47,48) and the ToBI intonational phonological notation derived from the aforementioned theory (19–22,46). However, we consider that a discrete prosodic annotation as ToBI may not be fully capable of capturing the variability in the L2 intonation (23,25,26). The studies mentioned in section 1.3 have shown that although there is transfer of f_0 contour features from L1 to L2, they do not happen systematically among all learners and do not affect all types of sentence modality in the same way (declarative sentences

term (i.e., reversible, but not quickly or easily) and, more problematically, impossible to tease apart the effect of L2 exposure from the role of L2 acquisition, L2 knowledge, or L2 use” (29, p. 202).

seem to be more easily acquired than interrogatives, for instance). In addition, a notation based on discrete tones would have problems to explain gradient variability in f_0 contours, that is, the presence of contour features that do not occur in both L1 Spanish and L1 BP but that are present in L2 Spanish (26). Also, choosing a ToBI-based system as a notation device to describe f_0 contours poses an important problem: what set of tones should be used to annotate Spanish L2 contours, those recognized by Sp-ToBI or the ones employed in ToBI-like systems developed for BP, such as Lucente (15)? Although at first it would seem obvious to answer with the first alternative, i.e. Sp-ToBI, what should be done when a contour feature from BP gets transferred to a L2 contour? Should a “BP tone” be used? That seems to be contradictory with the phonological nature of a ToBI-inspired analysis. The same conundrum arises in possible cases of cross-language influences: if a ToBI set of tones especially chosen to represent BP intonation is used to describe the L1 production of bilinguals, what to do when a Spanish contour feature intrudes in the BP L1 production? Should a Sp-ToBI tone be “borrowed”?

Besides the problem of how to best describe contour features in bilingual’s productions, another issue with the previous literature has to do with what parts of the contours are analyzed. AM-based analyses in most cases only look at pre-nuclear and nuclear pitch accents. Such a practice will fail to capture differences that modify the contour as a whole. For all these reasons, in this work we will follow a different approach. Instead of analyzing discrete tonal configurations, we apply the Dynamic Time Warping (DTW) technique as a way of finding an objective measure that will help us to quantify overall differences among f_0 contours produced under the different conditions.

DTW is a technique used to find the best possible alignment between two temporal sequences, indicating how much one of them would have to be warped in order to become identical to the other. It thus offers a measure of similarity (or distance) between these two sequences, such that the distance will have a low value when the two sequences are equal or very similar to each other and the value will grow larger when they are more dissimilar, i.e., when lots of changes would have to be applied to one of them so as to make it identical to the other (49–54). The advantage of this technique in the context of comparing pairs of f_0 contours is that it provides a measure of similarity between two time series regardless of the alignment between the sequences, that is, the analyst does not need to select in advance landmark points in the contours being compared. This is a useful feature, both because it makes the analysis less time-consuming and makes few theoretical assumptions about the contents of the contours. DTW seems to be a good alternative to labeling f_0 contours with discrete phonological tones because it can reduce to a single number the differences that spread over the whole of the two contours being compared.

This study aims at comparing the f_0 contours of declarative sentences, yes-no questions and wh-questions produced by Brazilian speakers of Spanish L2 with those produced by L1 Spanish speakers and also BP L1 contours produced by monolingual and bilingual Brazilian speakers. For this we used the DTW technique as a tool to calculate a measure of similarity between pairs of contours. For the purpose of best understanding the source of DTW distances, we performed an acoustic analysis of the f_0 contours in order to calculate three f_0 estimators: peak rate, peak range and global standard deviation. Regarding the DTW analysis of f_0 contours in this study, we have three main research questions: (1) the biggest DTW distances observed in the study will be those involving pairs of contours spoken by monolingual Spanish (MS) and monolingual BP (MP) participants, since both languages have different intonational patterns and monolingual speakers of one language have no knowledge of the intonational patterns of the other; we also predict that this result will be modulated by a modality effect, such that the DTW distance will be bigger for modalities where the intonational patterns of Spanish and BP differ more, especially interrogatives, as described earlier in sections 1.1 and 1.2; (2) if bilingual

speakers have learned the L2 patterns to a high degree, DTW distances between their contours and those produced by Spanish monolingual speakers will tend to approach the between-speaker distance variability seen in the Spanish monolingual sample; on the other hand, if bilingual speakers' L2 contours show signs of L1 transfer, i.e., contours have a mix of L1 and L2 features, DTW distances between pairs of monolingual Spanish contours and bilingual L2 contours will tend to have distance values that are lower than the ones observed for monolingual Spanish and monolingual BP pairs, but higher than the internal (between-speaker) variability of both monolingual Spanish and BP speakers; (3) if successful learning of L2 f_0 patterns by bilingual speakers do not have an impact on bilinguals' L1 performance, then DTW distances between pairs of monolingual BP contours and bilinguals' L1 contours will not differ significantly from the between-speaker distance variability seen in the BP monolingual sample; if learning L2 gives rise to systematic changes in the production of the L1, then bilinguals' L1 contours will have features of both L1 and L2 and DTW distances between pairs of monolingual BP contours and bilingual L1 contours will be higher than internal (between-speaker) variability of monolingual BP speakers, but lower than distances between pairs of monolingual Spanish and monolingual BP contours.

2 Methods

2.1 Speakers and experimental procedure

Fifteen native speakers of Brazilian Portuguese (10 women and 5 men) that speak Spanish as an L2, five native speakers of Spanish (3 women and 2 men) and five native speakers of Brazilian Portuguese (3 women and 2 men) who have never studied Spanish took part of the experiment. None of them were paid for their participation. Both Brazilian and Spanish speakers (25 subjects) declared they do not have any hearing and/or phono-articulatory problems.

The fifteen Brazilian speakers are from São Paulo State, college-educated, aged between 27, 48 years (mean age 35 years), and they learned Spanish after the age of 18. All lived in Madrid at the time of the recordings. The length of residence in Madrid ranged from six months to 16 years (mean residence time is 6 years). Ten Brazilian speakers studied Spanish in Brazil before moving to Spain. Even though all participants declared themselves proficient in Spanish, linguistic factors such as age of L2 learning, age of arrival in Spain, length of residence in Spanish-speaking countries, frequency of BP use in Madrid as well as social factors such as motivation to emigrate and personal experience with the country's culture may affect participant's performance in the L2. It was not possible to match participants on all linguistic and social factors mentioned earlier, but a sociolinguistic interview was conducted with the participants to better understand their linguistic and social experience with Spanish as L2. The interview data is summarized in Silva (25, p. 206-208). In this study, though, we do not correlate social factors with the production patterns. Because of the relative heterogeneity of participant's language experience, it is possible that learning level varies in the group. This aspect will be discussed in more detail in section 4. We consider this group to be bilingual and they provided data for conditions BL1 (Brazilian Portuguese L1 production) and BL2 (Spanish L2 production). The bilingual speaker definition can vary according to the authors or different theoretical perspectives. The studies of Conxita Lleó for example analyze L1 acquisition in bilingual contexts (55). Her studies focus especially simultaneous acquisition of two languages (the simultaneous acquisition of Spanish with another language). On the other hand, models such as the Revised Speech Learn Model proposed by Flege and Bohn (56) consider bilingual speaker a person that already has an established L1 before learning an L2 in a naturalistic context. In this case the second language learning process is not simultaneous but sequential. We consider bilingual speakers in the second

perspective, so that “bilingual” is a shorthand term for speakers that learned an L2 mostly as adults living in a foreign country where that language is spoken.

The five monolingual Spanish speakers contributed data for the MS condition (Spanish L1 production). Three speakers are from Madrid, one from Segovia and another from Ciudad Real (both cities are close to Madrid). All of them are college-educated, aged between 22 and 33 years (mean 28 years). They have never studied Portuguese as a foreign language.

Both bilingual Brazilian and monolingual Spanish speakers were recorded in three different moments. The first recording took place in a soundproof room at the phonetics laboratory of the CCHS-CSIC (*Centro de Ciencias Humanas y Sociales – Consejo Superior de Investigaciones Científicas*) in Madrid. We recorded directly on a desktop computer using the Adobe Audition 1.0 software. The audio files were sampled at a 44,100 Hz rate and saved in WAV format in one channel. For that, a headset microphone AKG C444 was used. In this first moment, we collected the data from 13 subjects. The second recording took place in each subject’s house. We collected the data from seven subjects in that way. We recorded directly on the computer using Praat software. The audio files were sampled at a 44,100 Hz rate and saved in WAV format in one channel. For that, a microphone Behringer B-2 PRO and sound card Scarlett 2i2 Focusrite were used.

The five monolingual native BP speakers are from São Paulo State in Brazil and they provided data for the MP condition (Brazilian Portuguese L1 production). All of them are college-educated, aged between 19 and 42 years (mean 25 years). They have never studied Spanish as a foreign language. The recording sessions took place in a room at the phonetics laboratory of the UFSCAR (Universidade Federal de São Carlos) in São Carlos, Brazil. We recorded directly on the computer using Praat software (57). The audio files were sampled at a 44,100 Hz rate and saved in WAV format in one channel. For that, a microphone Samson model C01 and a Behringer UMC202HD sound card was used. We collected the data from five subjects in that way.

Subjects read an excerpt from the story of *Don Quijote* adapted for children (58) and further modified by Silva (26). The excerpt chosen, which has 720 words, was the beginning of the Chapter entitled *Gigantes con aspás*, which tells the classic episode of *Don Quijote* fighting against the windmills. After reading the whole excerpt 39 isolated sentences (15 declaratives, 12 yes-no and 12 wh-questions) from the story were chosen to be read in isolation in random order.

The Portuguese translation of the Spanish excerpt has 685 words and the translation obeyed the position of the lexical stress of the final word of the sentence and, whenever possible, the same number of words per sentence, the same syntactic and prosodic structure as well as segmental content of the Spanish excerpt. Despite the structural constraints, Brazilian and Spanish participants considered the texts easy to read.

Monolingual speakers performed the experiment in one language only (Spanish for monolingual Spanish speakers and Portuguese for monolingual Brazilian speakers). Bilingual speakers performed the experiment in two languages (Spanish and Portuguese). First, they performed it in Spanish and then in Portuguese. Speakers were instructed in the language in which the task would be performed. The recording session consisted of reading aloud the text two times and reading each isolated sentence three times. First, they were instructed to read silently and comment on what they thought about the text and if it was difficult to read.

Subjects were instructed to start reading aloud the text only when they felt comfortable to do it. In addition, they were instructed to repeat any sentence that, in their opinion, had not been properly pronounced. After reading the text, they read the isolated sentences. These sentences were displayed through a Microsoft Office PowerPoint presentation. The set of 39 test sentences were presented one at a time in random order and the participant had to press a button in order to the next sentence to be shown. Finally, they were asked to tell the story read earlier in their own

words. The recording sessions took about 20 minutes per speaker in each language. A complete list of the sentences in the corpus and their translation to Brazilian Portuguese can be seen in Silva (26).

2.2 Corpus

The corpus consists of parallel speech productions by three groups: native monolingual Spanish speakers (MS), native monolingual Brazilian Portuguese (MP) speakers and native Brazilian Portuguese speakers who live in Madrid and have Spanish as L2; this former group contribute with two sets of recordings: Spanish as L2 (BL2) and Brazilian Portuguese (BL1). The four types of productions are considered levels of a variable henceforth referred to as LANGUAGE condition. For all the analysis reported here, both read text and sentences read in isolation were included but not analyzed separately due to the already complex design of the study. The three types of sentence modalities are levels of a variable we will refer to as MODALITY. As mentioned in section 2.1, there are a total of 39 sentences, 15 declarative sentences and 12 of each type of interrogative (yes-no and wh-question). Thus, in MS and MP there are 975 sentences in total (39 sentences \times 5 repetitions \times 5 speakers) per language. In BL2 and BL1 there are 2,925 sentences (39 sentences \times 5 repetitions \times 15 speakers) per language. Thus, a sum total of 7,651 were analyzed, 963 (MS) + 960 (MP) + 2876 (BL1) + 2852 (BL2).

2.3 F_0 extraction procedure

The procedure to extract the f_0 contour for each audio sample in the dataset consisted in the use of two Praat scripts. The first script (59) implements a two-pass algorithm suggested by de Looze and Hirst (60) that optimizes the selection of values for the *floor* and *ceiling* parameters used by the autocorrelation algorithm underpinning Praat's "To Pitch" function. This optimization reduces extraction errors such as octave doubling and halving. A second script (61) was run on each contour in order to look for errors not eliminated by the first script. This script goes through all voiced samples in a contour and flags as potential extraction errors two consecutive voiced samples that differ by more than a frequency threshold and are less than a duration threshold apart. Both frequency and duration thresholds are defined by the user. The values we used were 0.5 octaves for the frequency threshold and 80 ms for the duration threshold and were selected in a trial and error basis so as to maximize true positives (true errors) and minimize false positives (falsely identified errors). All f_0 samples flagged as errors by the script were later manually checked and corrected if necessary.

After the extraction and correction of the contours, a third script was used to perform the following procedures: to smooth contours using a 8 Hz bandwidth to minimize the micromelodic effects on the contours, then apply a linear interpolation over unvoiced sections in the contour and, lastly, convert the contour values from Hertz to semitones relative to the mean f_0 value in Hz of each contour processed. After these steps, the script saved each contour as time-value pairs in tab-separated raw text files for further processing.

2.4 F_0 statistical estimators

According to the description presented in Section 1, intonational differences between Brazilian Portuguese and Spanish in the three modalities studied here stem in part from the types of tones typically used to convey each modality and even peak height. Intonational differences between languages result in contours that can differ in peak height, peak density and overall contour variability. We decided to collect three parameters for each f_0 contour in our corpus: peak rate, peak range and contour standard deviation with the help of a Praat script (62). The script measures

peak rate as the number of f_0 peaks per second. Prior to peak counting, the raw f_0 contour undergoes heavy smoothing using a 2 Hz bandwidth and quadratic interpolation over unvoiced periods. All local maxima in the smoothed, interpolated f_0 contours are included when counting peaks. Calculating peak rate instead of just giving a peak count is a way to normalize the measure and compensate for the fact that the sentences to be analyzed may have different lengths. Peak range is defined as the median excursion (measured as the valley-to-peak distance) of all peaks present in the contour spanning at least 0.5 semitones. Standard deviation (SD) takes into account all the voiced samples in the contour. For the SD calculation, the contour was smoothed using a bandwidth value of 4.5 Hz and values in Hertz were converted to semitones relative to 1 Hz.

We consider that the three selected parameters can be seen as a coarse-grained way to characterize the contour features that may drive the DTW distances between pairs of f_0 contours, especially when comparing conditions where we expect to find the greatest differences (see the research questions outlined at the end of section 1).

2.5 DTW analysis

The R package DTW (53) was used in order to obtain DTW distances for pairs of different f_0 contours. In the present study, both can be repetitions of the same sentence uttered by one speaker or each can be renditions of the same sentence uttered by two different speakers.

We interpret DTW distance results as follows: if the normalized distance obtained is equal to zero this means that the two contours being compared are identical. Thus, the more the normalized distance obtained is close to zero, the more the two contours are like each other, and the more distant from zero is the value of the normalized distance, the more different the two contours are. In Figures 3 and 4, we present two examples of sentences analyzed using the algorithm and their respective normalized distance values.

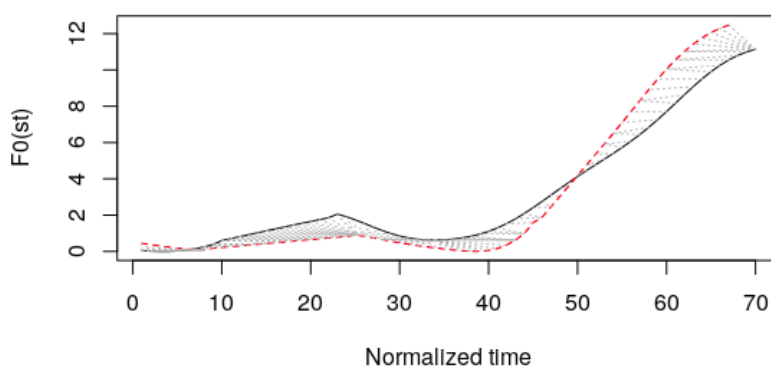


Figure 3: A pair of f_0 contours of yes-no questions ¿*Aquellos?* produced by native Spanish speakers *Sp-2* (black line) and *Sp-4* (red line). Normalized distance is 0.21 for this pair.

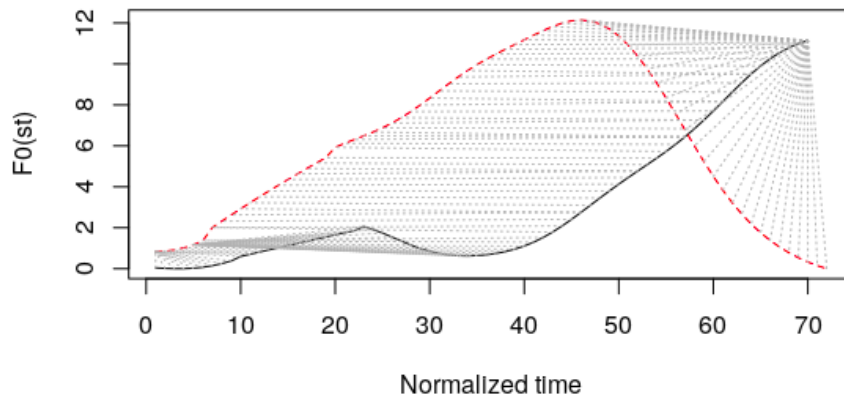


Figure 4: A pair of f_0 contours of yes-no question *¿Aquellos?*, produced by native Spanish speaker Sp-2 (black line) and Brazilian speakers Br-1 in BL2 condition (red line), Normalized distance 1.25 for this pair.

In Figure 3 we have the f_0 contours of the yes-no question *¿Aquellos?*, produced by two Spanish speakers (MS) and in Figure 4 by a Spanish and a Brazilian speaker (BL2). In Figure 3, we observe that the two f_0 contours are very similar. This similarity is reflected in a low normalized distance 0.21. In Figure 4 we have a contour produced by a monolingual Spanish speaker (the same contour represented by the black line in Figure 3) and one produced by a bilingual Brazilian speaker in Spanish L2. There is little correspondence between the two contours. The native Spanish contour presents a final rise and the bilingual speaker produces a circumflex pattern that ends in a descending movement. The cumulative alignment differences between the two contours cause a fourfold increase in the normalized distance (1.25).

Pairs of f_0 contours were analyzed in two groups:

- Within-language: both contours in the pair come from the same language condition. This group is further divided into:
 - Within-speaker: both contours in the pair come from the same speaker;
 - Between-speaker: each contour in the pair come from a different speaker.
- Between-language: contours in the pair come from different language conditions.

The within-language group allows us to estimate how variable contours are in a given language condition due to two factors: same speakers implementing contours of the same modality type in different ways (within-speaker group) and variability coming from different speakers doing different contour implementations (between-speaker group). The between-language group will allow us to compare how similar or different contours from different language conditions are.

Specifically, comparing MS to BL2 can estimate learning and comparing MP to BL1 can be a proxy for attrition. Lastly, comparing BL1 to BL2 can estimate how the bilingual speakers separate their L1 and L2 productions. The total combinations of all pairs of utterances of all speakers in the three modalities (declaratives, yes-no and wh-questions) were 567,800 according to the Table 1 and Table 2.

Table 1: Number of f_0 contour pairs comparisons for each language in the within-language group

Group	Language	Declarative	Yes-no	Wh-question
Within-speaker	MP	726	576	588
	MS	588	527	592
	BL1	1737	1576	1772
	BL2	1694	1561	1729
Between-speaker	MP	3630	2880	2940
	MS	2940	2630	2961
	BL1	30390	27476	30013
	BL2	29634	27293	30204
Total		71339	64519	70799

Table 2: Number of f_0 contour pairs comparisons for each language pair

Language	Declarative	Yes-no	Wh-question
MS-BL2	27033	21469	21902
MP-MS	9126	7200	7375
BL1-BL2	75016	59969	61260
MP-BL1	27134	21557	2210
Total	138309	110195	112639

2.7 Statistical analyses

For the experiment as a whole, there are two sets of independent variables, specified below.

- Language condition, with four levels:
 - monolingual speakers of Brazilian Portuguese (MP);
 - monolingual speakers of Spanish (MS);
 - Brazilian bilinguals speaking Brazilian Portuguese (BL1);
 - Brazilian bilinguals speaking Spanish a L2 (BL2).
- Sentence modality, with three levels:
 - declarative sentences;
 - yes-no questions;
 - wh-questions.

Dependent variables are:

- Statistical estimators extracted for each f_0 contour in the corpus: peak rate, peak range and contour standard deviation (see section 2.3.2 for a definition of each).
- DTW normalized distance, calculated for pairs of f_0 contours (see section 2.3.3 for details on the procedures).

For the statistical estimators analysis, we performed two-way ANOVA tests with language and modality as independent variables, including their interaction. Separate tests were carried for each statistical estimator (peak rate, peak range and contour standard deviation). Pairwise

independent t-tests were used to conduct post hoc comparisons among modality levels within each level of language or to compare the same modality level across two language levels. Bonferroni correction was applied to p-values in the pairwise comparisons in order to control familywise error rate. Cohen's d was used to determine effect size.

DTW data were also analyzed by means of two-way ANOVA tests with language and modality as independent variables plus their interaction followed by selected post hoc comparisons done through Bonferroni-corrected pairwise t-tests. Analysis was conducted in two groups as explained in section 2.3.3:

- Within-language: contours from the same language condition. Further divided into:
 - Within-speaker: contours from the same speaker;
 - Between-speaker: contours from different speakers.
- Between-language: contours from different language conditions.

All descriptive statistics and statistical tests were carried out within the R statistical computing environment (63). From the rstatix (64) R package we used `cohens_d` function, which provides both a d value and a verbal label to describe the effect size, ranging from “negligible” to “large”. We used both in the analysis.

3. Results

In section 3.1 we describe the effect of both LANGUAGE and SENTENCE MODALITY variables on the three f_0 dispersion estimators. Then, in section 3.2 we present the DTW analysis results for each language and modality (both within- and between-subjects) and compare the DTW results between languages pairs.

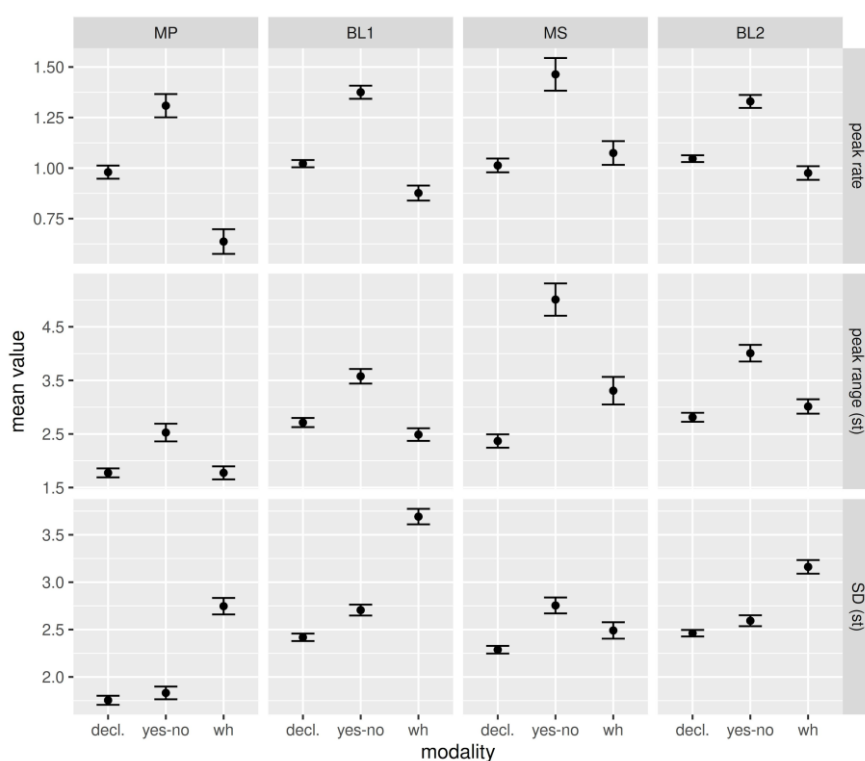


Figure 5: Mean value and 95% confidence interval for peak rate, peak range (in semitones) and standard deviation (in semitones) as a function of language (columns) and sentence modality (x-axis).

We start by presenting the results of the main ANOVA analysis run to test for main effects and interactions in the overall analysis that includes the four levels of LANGUAGE (MP, BL1, MS and BL2) and the three levels of MODALITY (declarative sentences, yes-no and wh-questions). Results are reported below by dependent variable:

- Peak rate: significant main effects of language [$F(3, 7639) = 32.7, p < 0.001, \eta^2 = 0.011$] and modality [$F(2, 7639) = 633, p < 0.001, \eta^2 = 0.14$] and a significant interaction between the two [$F(6, 7639) = 16.9, p < 0.001, \eta^2 = 0.011$].
- Peak range: significant main effects of language [$F(3, 7217) = 129, p < 0.001, \eta^2 = 0.051$] and modality [$F(2, 7217) = 340, p < 0.001, \eta^2 = 0.086$] and a significant interaction between the two [$F(6, 7217) = 24.6, p < 0.001, \eta^2 = 0.02$].
- Standard deviation: significant main effects of language [$F(3, 7639) = 246, p < 0.001, \eta^2 = 0.088$] and modality [$F(2, 7639) = 776, p < 0.001, \eta^2 = 0.17$] and a significant interaction between the two [$F(6, 7639) = 58.6, p < 0.001, \eta^2 = 0.044$].

For all three dependent variables there are significant effects of both LANGUAGE and MODALITY as well as a significant interaction between the two. Since the two independent variables yield significant effects, next we explore the results in detail. First, we look at possible effects of LANGUAGE between selected pairs of languages and then the effects of MODALITY. After that, we explore interactions between the two.

Table 3 presents comparisons between three language pairs: MS and MP, MS and BL2 and MP and BL1. In the first pair we expect to find the largest differences, since we have two different languages with different intonational patterns spoken by monolingual speakers, i.e., that have no knowledge of the other; the second pair is where we can expect to find some degree of learning, i.e., BL2 to converge towards MS in terms of the dependent variables measured in the experiment; the third pair is the one where deviations in the values of the dependent variables can constitute evidence for attrition caused by the process of learning Spanish, giving rise to deviations in the prosodic parameters from BP towards Spanish.

Looking at MS-MP comparison in Table 3, we have evidence that Spanish and BP spoken by monolinguals differ significantly in terms of the contour features captured by the three f_0 variability measures studied. Results show that f_0 contours of Spanish as L1 have higher peak rate, peaks that span a broader range and greater overall variability when compared to their BP counterparts, regardless of sentence modality. Results presented in MS-BL2 comparison show that Brazilian speakers of Spanish as L2 converge towards the values of Spanish spoken by natives; although there are statistically significant differences between mean values for both language conditions in the three acoustic parameters, the effect sizes are small or negligible as can be seen by examining Cohen's d values. As for MP-BL1 comparison, results point to the fact that speaking Spanish as L2 has a significant effect on Brazilian Portuguese speakers' performance on their L1, since there are statistically significant effects on the three acoustic parameters; notably, the drift in the three acoustic parameters is directed towards the values typical of Spanish; these results can be interpreted as language attrition, since the performance of Brazilian speakers in their L1 seems to be affected by L2.

Overall, the analysis confirms the expectation that the greatest effect sizes are found when comparing f_0 contours in BP and Spanish as L1; also, that BP speakers do change their prosodic patterns towards L2, observation that can be confirmed by the fact that the effect sizes when comparing BL2 and MS are small or negligible on the three acoustic parameters; finally, we observe that Brazilian speakers that speak Spanish as L2 do present performance differences in their L1 when compared their monolingual counterparts: their contours drift away from the patterns of BP speakers that do not speak Spanish towards the patterns of the L2. In terms of effect sizes, the attrition effects are intermediary between what is seen in the MS-MP and MS-BL2 pairs.

Table 3: Mean peak rate (peaks per second), peak range (semitones) and SD (semitones) , p -value and Cohen's d in MS-MP, MS-BL2 and MP-BL1 comparisons

MS – MP			
	Peak rate	Peak range	SD
MS	1.17	3.47	2.49
MP	0.97	2.03	2.08
p -value	< 0.001	< 0.001	< 0.001
Cohen's d	0.36	0.79	0.58
MS – BL2			
	Peak rate	Peak range	SD
MS	1.17	3.47	2.49
BL2	1.11	3.25	2.75
p -value	0.01	0.009	< 0.001
Cohen's d	0.11	0.1	0.28
MP – BL1			
	Peak rate	Peak range	SD
MP	0.97	2.03	2.08
BL1	1.08	2.93	2.40
p -value	< 0.001	< 0.001	< 0.001
Cohen's d	0.22	0.61	0.88

Table 4 shows mean value of peak rate, peak range and SD as a function of sentence modality (all languages polled). Mean values for the three modalities are significantly different for the three acoustic parameters: peak rate (all $p < 0.001$), peak range ($p < 0.02$ or less) and SD (all $p < 0.001$). Pairwise values of effect size are shown in Table 5. Peak features (rate and range) have a pattern: declarative and wh-question are similar (negligible to small effect size) and yes-no question has higher mean peak rate and peak range (effects are moderate to large). When it comes to SD, we see a ladder pattern: declarative in the bottom rung, closely followed by yes-no question and wh-question at the top. Summing up the results concerning modality comparisons, declarative sentence contours have an intermediary value of peak rate and the lowest values of peak range and SD; yes-no questions have the highest peak rate and range and intermediary SD value; wh-questions have the lowest peak rate, intermediary value of peak range and the highest SD value. As it is possible to see in Figure 5, the fact that wh-question contours have high SD in the overall mean is driven by BP, BL1 and BL2, not MS. This interaction will be further explored later in Table 6.

Table 4: Mean peak rate (peaks per second), peak range (semitones) and SD (semitones) as a function of sentence modality

Modality	Peak rate	Peak range	SD
Declarative	1.02	2.59	2.33
Yes-no question	1.36	3.79	2.56
Wh-question	0.91	2.79	3.23

Table 5: Effect size (Cohen's d) value for pairwise comparisons between sentence modalities

Comparisons	Peak rate	Peak range	SD
Declarative — Yes-no	0.79	0.64	0.3
Declarative — Wh-question	0.26	0.09	0.96
Yes-no — Wh-question	0.84	0.51	0.65

Table 6: Pairwise comparisons as a function of language pair, sentence modality and acoustic correlate. Values shown in cells are the p -value for the independent t-test and the corresponding Cohen's d

	Declarative			Yes-no question			Wh-question		
	Peak rate	Peak range	SD	Peak rate	Peak range	SD	Peak rate	Peak range	SD
MS-MP	$p = 1$ $d = 0.1$	$p < 0.001$ $d = 0.57$	$p < 0.001$ $d = 1.23$	$p = 0.002$ $d = 0.25$	$p < 0.001$ $d = 1.17$	$p < 0.001$ $d = 1.39$	$p < 0.001$ $d = 0.84$	$p < 0.001$ $d = 0.93$	$p = 0.013$ $d = 0.34$
MS-BL2	$p = 1$ $d = 0.11$	$p = 0.002$ $d = 0.34$	$p = 0.036$ $d = 0.35$	$p < 0.001$ $d = 0.22$	$p < 0.001$ $d = 0.4$	$p = 0.28$ $d = 0.2$	$p = 0.079$ $d = 0.19$	$p = 0.002$ $d = 0.14$	$p < 0.001$ $d = 0.71$
MP-BL1	$p = 1$ $d = 0.14$	$p < 0.001$ $d = 0.8$	$p < 0.001$ $d = 1.15$	$p = 1$ $d = 0.13$	$p < 0.001$ $d = 0.6$	$p < 0.001$ $d = 1.18$	$p < 0.001$ $d = 0.44$	$p < 0.001$ $d = 0.55$	$p < 0.001$ $d = 0.92$

Interactions between languages and modalities were explored by examining for different language pairs (lines in Table 6) how the three sentence modalities differ in terms of the three acoustical correlates. Since most pairwise comparisons yielded statistical significance (22 out of 27), we base our conclusions mostly on effect size as a proxy measure for distance between the language pair for a given acoustic correlate and sentence modality.

- MP-MS: declarative sentences have the least differences – small and moderate effect sizes for peak rate and peak range and a large effect for SD (MS values are greater than MP); yes-no questions have similar peak rate (small effect size) and large differences for the other two variables; wh-questions present large differences in peak rate and peak range (MS > MP) and a small difference in SD (MP > MS; wh-question is the modality for which MP has the greatest mean SD value).
- MS-BL2: three modalities are rather similar; effect sizes are small with the exception of contour SD in wh-questions, where there is a moderate difference (BL2 > MS).

- MP-BL1: negligible or small effect size for peak rate in all three modalities; large (declarative) and moderate effect sizes (yes-no and wh-question) for peak range; large size effect for SD in all three modalities. Effects are all in favor of BL1, showing that BL1 deviates from MP towards targets more compatible with MS.

Summing up, the interaction analysis shows that the largest differences are seen in the MP-MS pair, as could be expected, since both are languages spoken by monolingual speakers. Also as would be expected by the contrastive description of modality intonation presented in Sections 1.1 and 1.2, the largest differences are seen in the two interrogative modalities. The results also show that there is a great deal of learning by BP speakers, since BL2 and MS comparisons yield small effect sizes regardless of sentence modality, indicating that BL2 speakers are able to modify their contour to match to a good degree those of the target language (MS). Lastly, results also point to the fact that the performance of BP bilingual speakers in their native language (BL1) show attrition. The effect is rather uniform among the three modalities, and contour SD is the most affected acoustic parameter, followed by peak range (two moderate and one large effect size); peak rate is unaffected or shows a significant but small-sized effect (wh-question).

In order to show in visual terms the contour patterns described by the statistical analysis presented before, Figure 6 shows time-normalized f_0 contours of BP and Spanish sentences as a function of modality and language. Starting with MP and MS, we can see the intonational features described in sections 1.1 and 1.2 for both languages: i) declarative contours have a similar pattern for both languages, being comprised by a number of peaks, roughly one for each phonological word, but differing in how the final low tone aligns at the boundary - later for BP and earlier for Spanish; ii) yes-no questions in BP start with a high or extra-high peak, followed by a low section, ending in a circumflex tone; in Spanish, the recurrent features are a pre-final circumflex tone, immediately followed by a high or extra-high rising contour; iii) wh-questions in BP have a descending pattern that start at a high or extra-high point and drifts down towards the final boundary; in some cases the contour does not present peaks, just a smooth gliding towards the low final target; in Spanish, as noted in Section 1.1, there are at least two distinct patterns: the contour can start with an optional peak and it can end either with a circumflex or a rising tone; in the contours shown, there is no connection between the presence of a peak at the start and the two possible final tones.

The visual examination of the contours (Figure 6) agrees with the statistical analysis: Spanish contours tend to have more peaks, peaks that span a greater range and overall greater variability. Also, the differences seem more visible in interrogative sentences, both yes-no and wh-questions. Comparing BL2 and MS, it is immediately visible that BL2 contours show a great deal of variability, irrespective of sentence modality. This is in agreement with the statistical results that show learning, since MS contours are overall more variable than MP. Given this systematic difference, it can be said that successful learning of MS intonation by BP natives must include the magnification of contour parameters such as peak rate, peak range and overall variability (of which SD can be seen as a proxy).

Turning to the interrogative modalities, we can see that MS contours show both evidence of learning and L1 transfer. Looking at yes-no questions, we see that BL2 contours incorporate features that are crucial to MS intonational grammar: the pre-final circumflex tone followed by a final rising tone ending in extra-high values, although some of the speakers skip the final rising tone and end the contour with a circumflex tone. L1 transfer can be seen in the initial high peak that is absent in MS contours, but present in MP. Similar remarks can be made about wh-question contours: BL2 incorporates a key feature of MS contour, namely a final tone that can be either circumflex or rising (again, those are not seen in MP contours); L1 transfer is also seen, in the

form of initial peaks that tend to be high or extra-high as is typical in MP. Especially for wh-questions, there seems to be some overshooting in terms of SD in BL2 (note the moderate-sized effect size in favor of BL2 reported in Table 6).

Comparing now BL1 and MP (and keeping BL2 contours in the background), the results suggest that our data confirms the prediction that there is an interaction between learning and the occurrence of attrition. The overall (modality-independent) greater contour variability that bilingual BP speakers had to learn in order to properly speak Spanish bleeds back into their L1 production. As reported when discussing the data in Table 6, SD and peak range are the most affected parameters. Focusing on yes-no questions, we see that BL1 contours preserve MP's high or extra-high initial peak, but we also see circumflex tones that are higher and are aligned earlier than what is typical in MP and final extra-high rising tones that are totally absent in MP; for this modality in particular, BL1 and BL2 contours seem very similar. Lastly, looking at wh-questions, we also see a kind of hybrid pattern: a falling contour that starts at a high or extra-high point as in MP, but instead of smoothly going down and ending on a low final tone, we see a number of cases of final circumflex and even final rising tones that are completely absent in MP contours but are typical of MS; for this modality, BL1 and BL2 have different patterns, showing less L2 transfer to the L1.

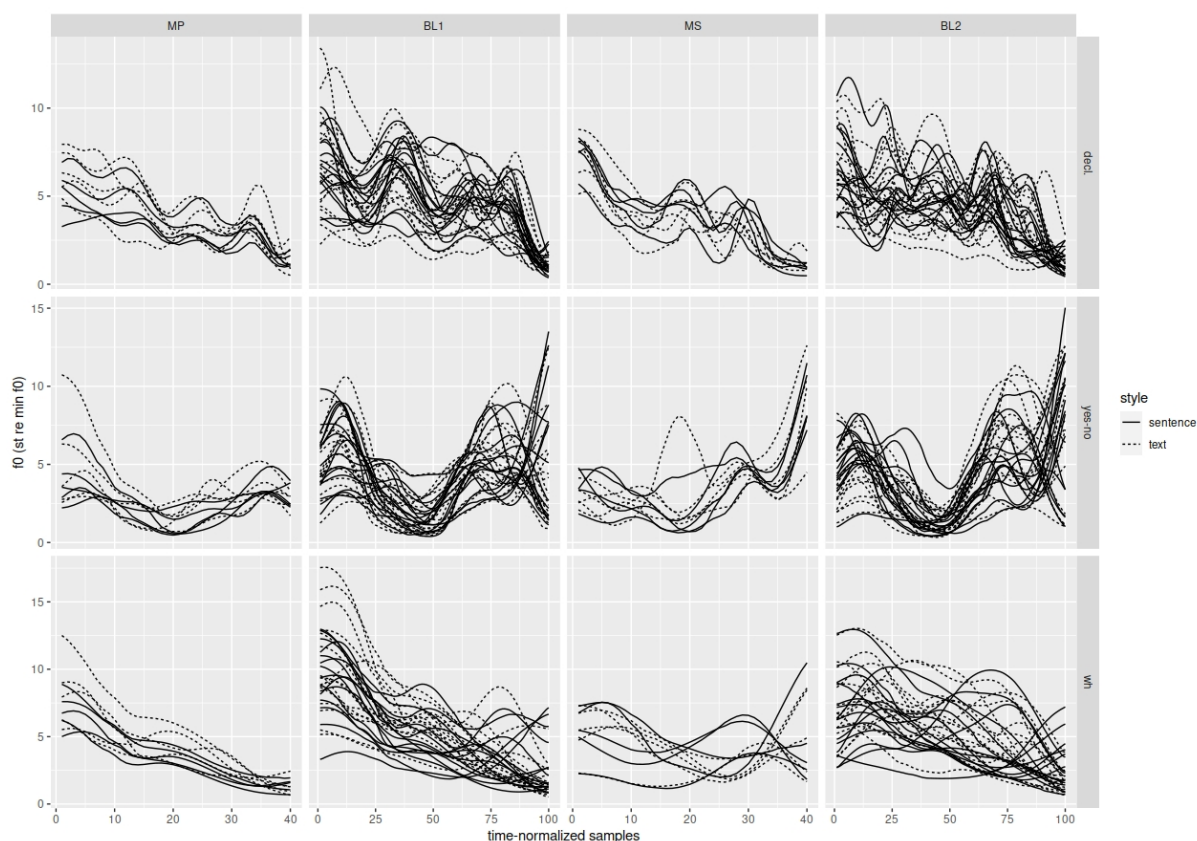


Figure 6: Time-normalized f_0 contours (semitones relative to each sentence's minimum f_0) of BP and Spanish sentences as a function of language and modality. Solid lines are read isolated sentences and dashed lines are read sentences embedded in text. MP and MS ($N = 5$ per sentence context) and BL1 and BL2 ($N = 15$ per sentence context). BP sentences are *Um encantador levou seus livros pelos ares* (A charmer blew his books – declarative), *Amigo Sancho, tens medo?* (My friend Sancho, are you afraid? - yes-no question) and *Onde estão meus livros?* (Where are my books? - wh-question). Spanish sentences are *Un encantador se ha llevado sus libros por los aires* (declarative), *Amigo Sancho, tienes miedo?* (yes-no question) and *¿Dónde están mis libros?* (wh-question).

3.2 DTW analysis

Sections 3.2.1 and 3.2.2 present the main statistical results separately by language group. In Section 3.2.1 we present the within-language analysis and in 3.2.2 the between-language analysis.

3.2.1 Within-language analysis

In this section we present the statistical analysis of effects of the independent variables LANGUAGE and MODALITY on DTW distance between contour pairs taken from the same language. Mean DTW values are presented as a function of language condition and sentence modality in Tables 7 and 8, respectively. In both cases, the results are also presented as a function of group (within- and between-speaker).

Within-speaker group: significant main effects of LANGUAGE [$F(3, 5480) = 43, p < 0.001, \eta^2 = 0.023$] and modality [$F(2, 5480) = 51, p < 0.001, \eta^2 = 0.018$] and a significant interaction between the two [$F(6, 5480) = 5.42, p < 0.001, \eta^2 = 0.006$].

Languages: all pairwise comparisons yield significant differences (all $p = 0.001$ or less) except for BL1-BL2 ($p = 0.058$); median effect size is small ($d = 0.24$). Languages spoken by monolinguals have smaller mean distances than the others.

Modalities: declarative/yes-no questions ($p = 0.51, d = 0.029$); declarative/wh-questions ($p < 0.001, d = 0.26$); yes-no questions/wh-questions ($p < 0.001, d = 0.27$). Wh-questions have a higher mean distance compared with the other two.

Between-speaker group: significant main effects of LANGUAGE [$F(3, 101225) = 872, p < 0.001, \eta^2 = 0.025$] and modality [$F(2, 101225) = 1613, p < 0.001, \eta^2 = 0.031$] and a significant interaction between the two [$F(6, 101225) = 87.9, p < 0.001, \eta^2 = 0.005$].

Languages: all pairwise comparisons yield significant differences (all $p < 0.001$); median effect size is small ($d = 0.49$). Languages spoken by monolinguals have smaller mean distances than the others.

Modalities: declarative/yes-no questions ($p < 0.001, d = 0.098$); declarative/wh-questions ($p < 0.001, d = 0.39$); yes-no questions/wh-questions ($p < 0.001, d = 0.3$). Step-like pattern: declarative > yes-no questions > wh-questions, small-size effects.

Table 7: Mean DTW distance (SD in parenthesis) as function of language condition and group

	Within	Between
MP	0.29 (0.31)	0.45 (0.37)
MS	0.35 (0.32)	0.55 (0.42)
BL1	0.42 (0.35)	0.78 (0.53)
BL2	0.44 (0.40)	0.79 (0.55)

Table 8: Mean DTW distance (SD in parenthesis) as function of sentence modality and group

	Within	Between
Declarative	0.37 (0.27)	0.67 (0.41)
Yes-no question	0.36 (0.29)	0.71 (0.46)
Wh-question	0.47 (0.48)	0.88 (0.67)

Figure 7 shows mean DTW distance as a function of both sentence modality and language condition and also group.

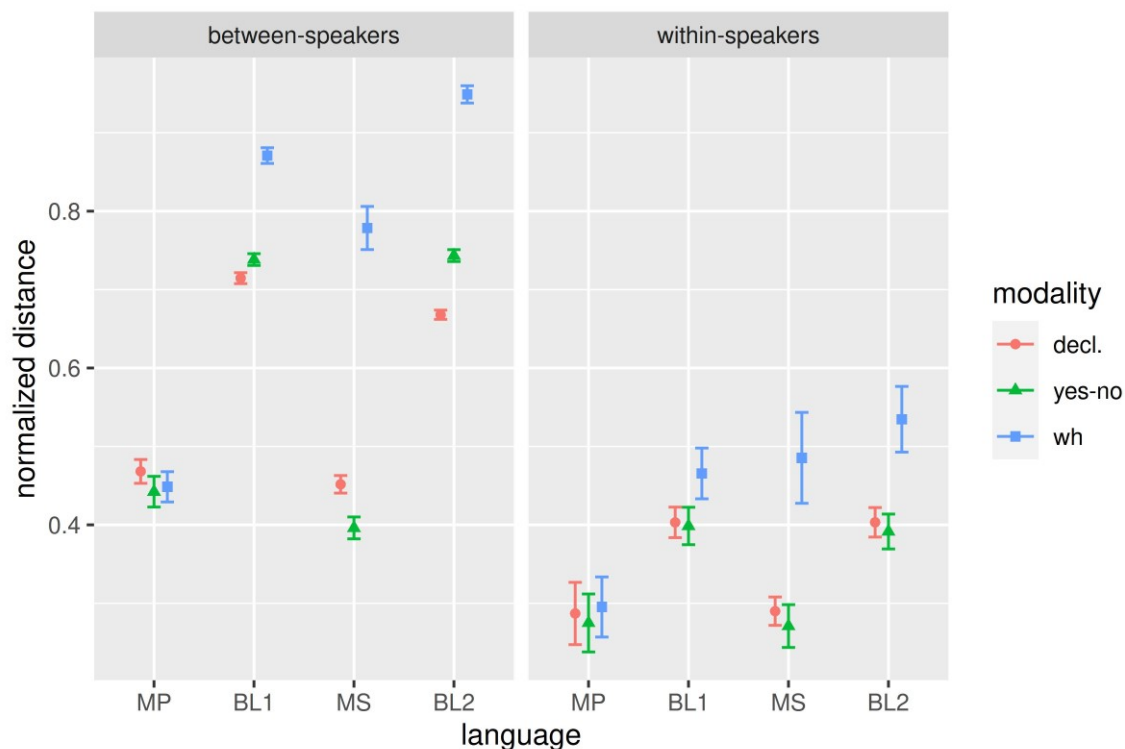


Figure 7: Mean DTW distance and 95% confidence interval as a function of language condition (markings in the horizontal axis), sentence modality (coded by both color and point shapes) and group (right and left panels).

Looking at the interaction of language and modality, the first and more obvious result is that DTW distances are smaller in within-speaker than in between-speaker group. Two other observations can be made by looking at Figure 7: (1) languages spoken by monolinguals (MP and MS) show smaller mean distances than the others (BL1 and BL2), regardless of group and (2) wh-questions have bigger distances compared with the other modalities. The latter observation does not hold for monolingual BP speakers – no significant difference in both comparison groups. For the other three languages, wh-questions have bigger mean distances than the other two modalities, regardless of group. MS is the language in which the difference between wh-questions and the other modalities is bigger – median effect size is moderate ($d = 0.58$) for within-speaker and large ($d = 0.83$) for between-speaker.

MP and MS have similar levels of mean distance in both within- and between-speaker groups, except for wh-questions. The languages spoken by monolinguals have a low level of variability at within- and between-speaker levels for declarative sentences and yes-no questions. The heightened level of DTW distance seen for MS in wh-questions is due to the variability shown in Figure 6, where it is possible to see that contours start with either a high or a level tone and end either with a final rising or a circumflex pattern. Since within-speaker distance is also higher for wh-questions than other modalities, this is evidence that the same speaker will vary in the contour implementation.

When it comes to BL1 and BL2, all three modalities hover at similar levels of mean distance in both within- and between-speaker groups. Declarative sentences and yes-no questions have lower mean distance values than wh-questions, similar to what happens with MS. The overall higher levels seen for BL1 and BL2 in comparison to MP and MS can be explained by the fact that the contours in those languages have a hybrid character, as pointed in Section 3.1: because of attrition, some BL1 contours can present tones that are typical of MS while others remain very much like the ones in MP; because there is (incomplete) learning, some BL2 contours emulate

the tones of MS but some still transfer MP tonal structure to the second language. This variability contributes to inflate DTW distances for BL1 and BL2 in general and wh-questions even more.

3.2.2 Between-language analysis

In this section we present the statistical analysis of effects of the independent variables LANGUAGE PAIR and MODALITY on DTW scores. Figure 8 shows mean values of pairwise DTW scores as a function of both variables. We start by presenting the results of the main ANOVA analysis run to test for main effects and interactions in the overall analysis that includes the four levels of the LANGUAGE (MP, BL1, MS and BL2) variable and three levels of the MODALITY variable (declarative sentences, yes-no and wh-questions). Results are reported below by dependent variable. The results show significant main effects of language [$F(3, 188257) = 371, p < 0.001, \eta^2 = 0.006$] and modality [$F(2, 188257) = 5079, p < 0.001, \eta^2 = 0.051$] and a significant interaction between the two [$F(6, 188257) = 356, p < 0.001, \eta^2 = 0.011$].

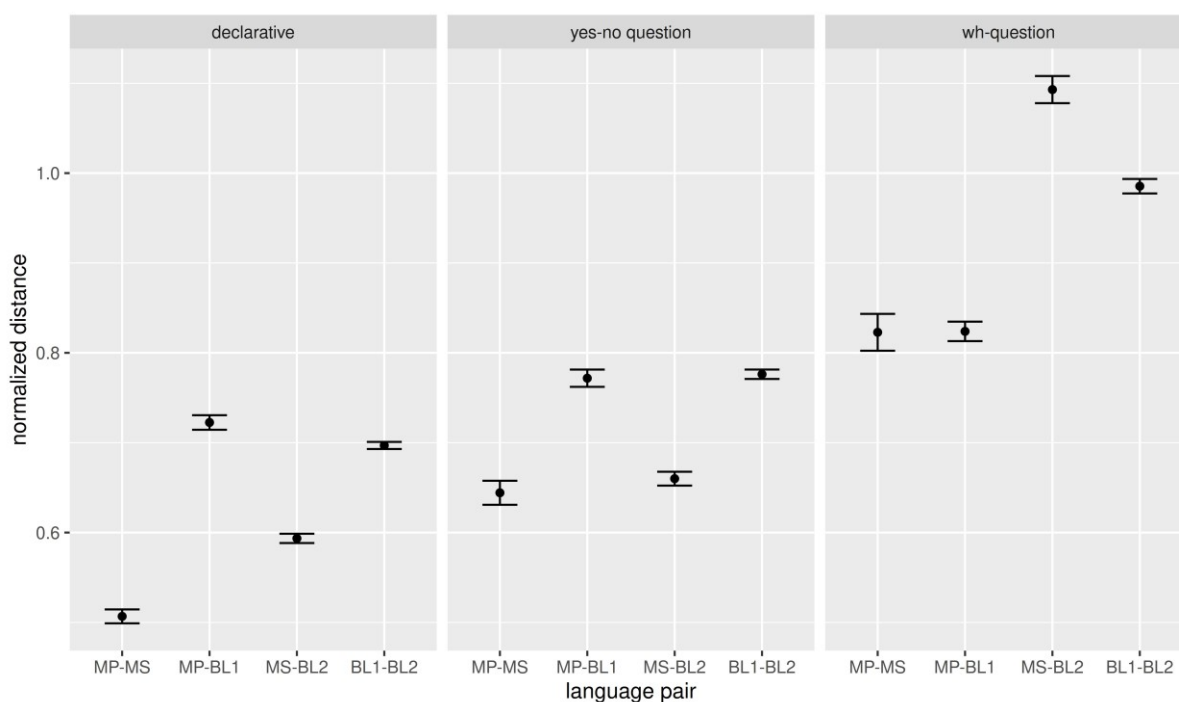


Figure 8: Mean DTW distance and 95% confidence interval as a function of language pair (markings on the horizontal axis) and sentence modality (right, central and left panels).

Since all independent variables yield significant effects, next we explore the results in detail. First, we look at possible effects of LANGUAGE between selected pairs of languages and then the effects of MODALITY. After that, we explore interactions between the two.

Mean DTW distance (SD in parentheses) for each language pair are listed below. The six pairwise comparisons yield significant differences (all $p < 0.001$) except for MP-BL1 compared to MS-BL2. The pair of languages spoken by monolingual speakers (MP-MS) has the lowest mean distance and the pair involving the BP bilinguals speaking their native language and Spanish as L2 (BL1-BL2) has the highest mean distance. The other two pairs are in-between. Effect sizes range from negligible to small (Cohen's d ranging from 0.0006 to 0.32).

- MP-MS: 0.65 (0.48)
- MP-BL1: 0.77 (0.53)
- MS-BL2: 0.77 (0.59)
- BL1-BL2: 0.81 (0.56)

Two factors may be the cause of higher DTW distances involving pairs that have BL1 or BL2 as a member: (1) higher overall variability in f_0 contours for BL1 and BL2 compared to MP and MS as can be seen in Figure 5 and (2) the hybrid intonation patterns seen in the time-normalized contours in Figure 6 for BL1 and BL2.

Mean DTW distance for each modality is listed below. Results of pairwise comparisons are: declarative/yes-no questions ($p < 0.001$, $d = 0.17$ negligible), declarative/wh-questions ($p < 0.001$, $d = 0.5$ moderate) and yes-no/wh-questions ($p < 0.001$, $d = 0.36$ small). Declarative sentence is the modality with less DTW distance, then yes-no questions and finally wh-questions. This step-like pattern can be seen in Figure 8.

- Declarative: 0.67 (0.40)
- Yes-no question: 0.74 (0.47)
- Wh-question: 0.96 (0.73)

We now investigate the interaction between both language pairs and modalities. Post-hoc comparisons are presented below as a function of modalities.

Declarative sentences: differences between all language pairs are statistically significant (all $p < 0.001$, median $d = 0.3$, ranging from 0.057 to 0.56).

Yes-no questions: non-significant difference between pairs MP-MS and MS-BL2 ($p = 0.46$) and MP-BL1 and BL1-BL2 ($p = 1$); all other comparisons yield significance (all $p < 0.001$, median $d = 0.27$).

Wh-questions: non-significant difference between pairs MP-MS and MP-BL1 ($p = 1$); all other comparisons yield significance (all $p < 0.001$, median $d = 0.24$).

Declarative sentences and yes-no questions show a similar pattern: language pairs MP-BL1 and BL1-BL2 have the greatest mean DTW distances among the four pairs; MS-BL2 and MP-MS have significant lower mean values in comparison, MP-MS being the lowest in declarative sentences. The pattern is similar, although all four pairs present a small-sized but statistically significant increase in yes-questions when compared to declarative sentences – MP-MS $p < 0.001$ $d = 0.39$; MP-BL1 $p < 0.001$ $d = 0.098$; MS-BL2 $p < 0.001$ $d = 0.18$; BL1-BL2 $p < 0.001$ $d = 0.18$.

We can explain why BL1-BL2 and MP-BL1 are the pairs that have the highest mean DTW in declarative and yes-no question bearing in mind the results described in Section 3.2.1 about the variability in contours. Starting with BL1-BL2, this is a pair where both languages are highly variable – BL1 contours show both L2 influence and faithful L1 performance; similarly, BL2 contours show both evidence for learning and L1 transfer. Since both languages in the pair present high between-speaker variability, this favors higher DTW distances. Let's consider now the MP-BL1 and MS-BL2 pairs; in both cases, one of the elements in the pair is a language with high between-speaker variability (BL1 or BL2). The fact that MP-BL1 has statistically significant higher mean distances than MS-BL2 may be considered evidence that cases of attrition in BL1 contours are more common than retention of native performance and that BL2 contours where features learned from L2 are present are more numerous than those showing transfer from L1. Lastly, the low mean values for the MP-MS pair may be explained by the fact that MP and MS both have low between-speaker variability as shown by within-language DTW distances for declarative and yes-no questions. Since monolingual speakers tend to be very consistent, f_0 contour pairs for which one is taken from MP and the other from MS will tend to generate lower DTW distances; even though there are systematic differences between contours in both languages, they tend to be constant for all speakers.

Wh-questions present an altogether different pattern. MS-BL2 is the pair with the highest mean DTW distance, followed by BL1-BL2, in turn followed by MP-MS and MP-BL1 at the same level. The same principle behind the results for the other modalities seem to explain what is

seen here. As pointed earlier, BL2 is characterized by a high level of both within and between-speaker variability; wh-questions in MS are special in the sense that this modality differs from others in presenting statistically significant higher within- and between-speaker variability in within-language DTW distance; the combination of these two factors adds up to make a good number of contours in the MS-BL2 comparison differs substantially, yielding higher DTW distances. The second highest value in the wh-question modality for the BL1-BL2 pair can be explained by the same reasons presented in the previous paragraph for this pair. The same holds for the lowest mean value associated with MP-MS; MP speakers have a low within- and between-speaker variability in wh-questions, but not MS; this simultaneously explains that MP-MS has one the lowest means in the wh-question modality but at the same time has the highest mean for the MP-MS pair considering the three modalities (all pairs $p < 0.001$). Again, the low variability of MP explains that MP-BL1 also has a low mean value within wh-questions; we can also hypothesize that BL1 has fewer cases of contours affected by attrition in this modality than the others.

4. Discussion and conclusion

In this section we review what we consider to be the main contributions brought by the study. Firstly, the use of a set of tools to describe f_0 contours that do not rely on theory-dependent transcription systems that are language-specific; we consider this point to be important, given the inherently variable nature of L2 production. Secondly, the results present an objective and quantitative account of the highly variable nature of both Spanish L2 and BP L1 production of bilingual speakers. Lastly, what we see as the most original result is the evidence of L2 influence on the L1 prosodic production of Brazilian bilinguals, a phenomenon consistent with the definition of attrition we gave in section 1.4. There is also evidence of learning in the data, although this is not entirely new in the previous literature on Spanish L2 spoken by Brazilian learners. The novelty in this respect is that our participants are bilinguals that were living in a L2-dominant context, while most previous literature study late bilinguals learning Spanish in formal settings while living in Brazil.

Three acoustic parameters - rate of f_0 peaks, mean peak range and standard deviation of the whole contour f_0 - were used as proxies for the overall variability of f_0 contours in the four language conditions studied here. A detailed statistical analysis of these parameters is reported in section 3.1 and shows that there are significant effects of both language and sentence modality. A comparison of language pairs indicates that Spanish and BP spoken by monolinguals (MS-MP) is the pair for which the differences in parameters' values are greater (using effect size as a measure). This result corroborates hypothesis 1, stated at the end of section 1, as it seems reasonable to attribute differences in the acoustic parameters to differences in intonational phonology between the two languages. The MS-BL2 pair shows the least differences, suggesting a good deal of learning, as f_0 parameters of bilinguals (BL2) come close to the values of the L2 target (MS). Results for the MP-BL1 pair show that the level of divergence in the f_0 patterns is intermediary between those of the two other pairs, MS-MP and MS-BL2. We interpret this result as evidence that the successful learning process revealed by the close approximation between BL2 and MS may have interfered with the bilingual speakers' performance of their L1, evidenced by a greater degree of divergence between the bilinguals' L1 performance and the monolingual BP performance, a situation that we interpret as an instance of language attrition. Regarding the other two language pairs, MS-BL2 presents smaller distances between the three acoustic parameters; distances are not that different between modalities, which we interpret as evidence of similar levels of learning throughout modalities. In MP-BL1 we see bigger differences compared to the

ones in MS-BL2. Considering the modalities, wh-questions yield bigger differences compared to the other two modalities.

Overall, results point to the fact that, at least in our data, successful learning in bilinguals is associated with increase in attrition, especially for the wh-question modality, the one among the three studied here where Spanish and BP differ the most. As Brazilian bilinguals strive to change their f_0 contours to match the patterns in L2, their native L1 production is warped towards the L2 patterns. We interpret this bidirectional influence as evidence in favour of Flege and Bohn's revised Speech Learning Model (SLM-r) proposition that the L1 and L2 share a common phonetic space in bilinguals and, as such, both L1 and L2 compete for the shared cognitive resources. Furthermore, the L2 Intonation Learning theory (LILt) model, proposed by Mennen (65), also seems to be relevant to explain the results presented here as we pointed out earlier in this section.

Results of contour distance as measured by the DTW technique, reported in section 3.2, seem to partially confirm hypothesis (1) stated at the end of section 1. The mean DTW distance in the within-language group for monolingual speakers (MP and MS) is around 0.45 for all modalities in MP and declarative and yes-no question in MS, the exception being wh-question in MS (mean value of about 0.8); this exception is compatible with the observation made in section 1 about the well-known variability in wh-question patterns documented in Spanish. When these figures are compared to mean DTW distance for the MP-MS pair in the between-language condition (see Figure 8), we see that the values are always higher than 0.45: mean values go up from 0.5 (declaratives) to around 0.65 (yes-no questions) and then almost 0.8 (wh-questions). Comparing the two sets of results, we can say that variability in DTW distance due to between-speaker differences within a monolingual sample (either in MP or MS) is more or less constant (with the exception of wh-question in MS for which there is a good explanation presented in section 1.1) and that variability due to between-language differences when each f_0 contour in the pair comes from a monolingual sample (MP or MS) is always at a higher mean level and varies as a function of sentence modality as well. These results suggest that DTW distances are sensitive to differences in language intonation patterns and also to modalities within each language. We said that the hypothesis was partially confirmed because the MP-MS pair is not the one yielding the highest mean DTW distances. What seems to be behind this result is not the fact that BP and Spanish intonational patterns are not so different as we thought at first, but that f_0 contours produced by speakers in our monolingual samples (MP and MS) tend to be much more homogeneous, that is, there is less between-speaker variability, than contours produced by bilinguals, either in BL1 or BL2 conditions.

Hypotheses (2a) and (2b) predicted that bilinguals' contours would be in between BP and MS: in BL2 condition, they would be closer to MS if L2 learning was predominant and closer to MP if L1 transfer was typical; in BL1 condition, contours would be close to MP in most cases, but L2 traits would appear if cases of attrition were the norm. But we tacitly assumed that speakers' behavior would be homogeneous. What the results show is that the behavior of bilinguals is significantly more variable than their monolingual counterparts, as can be seen in Figure 6. Inspection of time-normalized f_0 contours of individual speakers indicate that in BL2 condition the same speaker will produce contours that are very similar to MS and also some that are identical to the L1 pattern, constituting instances of learning; conversely, in BL1, some speakers will produce contours that are faithful to the MP pattern and others that have the MS patterns, a situation we identify as evidence of attrition. Cases of contour "hybridization", where traits of both languages are present in the same contour do happen, as can be seen in Figure 6. Inspection of the Figure 6 suggests they are not the norm, but we have not quantified how prevalent they are. Contours produced by bilinguals are more variable not only qualitatively but

also quantitatively. The most striking difference in this respect is contour standard deviation. As can be seen in Figure 5, BL1 has significantly higher SD values than MP in all modalities; BL2 has higher SD values compared to MS, especially in wh-questions.

The results generated by the DTW analysis allow us to conclude that the technique was useful as a descriptive device that helped us make sense of the f_0 data we had, although some of our initial hypotheses were not confirmed. Despite its usefulness, DTW distances alone are not enough as an analytical tool. To better make sense of our data, we also had to make use of time-normalization of the f_0 contours of specific sentences in order to visualize information about the alignment of f_0 movements to specific words which is something that the holistic nature of DTW does not allow.

One important trait of the f_0 contours produced by bilinguals is their great overall variability. This result could be explained by a bidirectional L1-L2 linkage, as proposed by the SLM-r model (56) to explain the learning of segmental features. According to the latest version of the model, there are interactions between the L1 and L2 phonetic subsystems and such interactions occur because L1 and L2 sounds exist in a common phonetic space. As a result, the theory predicts that, as a bilingual develops an L2 category alongside a previously existing L1 category, the distance between the two in the shared phonetic space tends to be magnified in order to increase the likelihood of acquisition, especially if both categories are phonetically close. The theory predicts that the contrast-enhancing strategy leads to bilinguals producing both L1 and L2 in a non-precise way. A similar prediction is made by the LILt model, that deals specifically with the learning of intonation. Our data seems to corroborate this theoretical prediction: in order to try to maintain the contrasts between the intonation in BP and Spanish, that have some contour features in common, the cost would be to produce both languages in a non-precise way. In the context of our data, this less-precise production led to increased variability in the f_0 contours. Interpreting the results within the context of the LILt model, this could be interpreted as a deviance in the *realizational* dimension.

Considering that the influence between L1-L2 is bidirectional, the SLM-r model also allows for the possibility of language attrition. When L2 sounds are very similar to those of L1, the substitution would be unnoticed by monolingual speakers of the target L2. This could be what happened in part with the intonation of declarative sentences. Our results showed smaller DTW distances between the intonation of declarative sentences compared to both types of interrogatives. On the other hand, when there is an L2 sound pattern for which a new category was not formed, the model predicts that a compromise L1-L2 category will develop based on the combined distribution of sound patterns defining the L1 and L2 categories. This is what we observed in bilingual interrogatives. In yes-no question, the bilinguals produce both final f_0 contours: rising (from Spanish) and circumflex (from BP) in both in Spanish L2 and BP L1. Similarly, in wh-question the bilinguals produce extra high pitch accent in the interrogative pronoun (from BP) and the final f_0 contour with a falling (from BP), rising (from Spanish) and circumflex (from Spanish) shape in both languages in Spanish L2 and BP L1. The fact that declaratives yield less deviance in L2 production can be explained in terms of the LILt model as arising from the fact that declaratives are phonologically more similar (called the *systematic* dimension in the model) in BP and Spanish than interrogatives, as explained in sections 1.1 and 1.2.

As mentioned in Section 1.4 there are only a couple of studies about the intonation of bilingual speakers that also discuss language attrition in their L1 intonation (29,46). As far as we know, none has published about language attrition in intonation produced by Brazilian bilingual

speakers of Spanish L2 other than a mention by Silva (26) about the possibility⁵. For this reason, this study could bring an important theoretical contribution to the studies of language attrition analyzing prosody. We consider that our results make a better case for attrition involving intonation than the ones previously reported in the literature. Compared to results in Mennen (46), the attrition we observe is not caused by L2 influence in a confounding variable (vowel length, in her case). Compared to results presented by Leeuw and colleagues (29), the influence of L2 in L1 we see in our data is more robust, affecting the choice of contour, not only the alignment of a tone.

As noted in Section 2.1, the speaker sample analyzed in this study is relatively heterogeneous in terms of time of residence in Spain, frequency of L1 usage and social setting in which their use of L2 takes place. These factors may have an impact in the observed results and explain at least part of the variability observed in the bilingual production, both in L1 and L2. Future work should look the data of individual speakers separately and explore the possible influence of speaker experience on the patterns observed here. In particular, it should try to establish correlations between individual speaker experience (length of residence, L1 and L2 frequency of use) and outcomes of learning and attrition.

REFERENCES

1. Xu Y. Speech melody as articulatorily implemented communicative functions. *Speech Commun.* 2005; 46:220–51. DOI: <https://doi.org/10.1016/j.specom.2005.02.014>.
2. Prom-on S, Xu Y, Thipakorn B. Modeling tone and intonation in Mandarin and English as a process of target approximation. *J Acoust Soc Am.* 2009 Jan;125(1):405–24. DOI: <https://doi.org/10.1121/1.3037222>.
3. Xu, Yi. Timing and coordination in tone and intonation — An articulatory-functional perspective. *Lingua.* 2009; 119:906–27.
4. Liu F, Xu Y, Prom-on S, Yu, Alan. Morpheme-like prosodic functions: Evidence from acoustic analysis and computational modeling. *J Speech Sci.* 2013;3(1):85–140.
5. Xu Y, Prom-on S. Toward invariant functional representations of variable surface fundamental frequency contours: Synthesizing speech melody via model-based stochastic learning. *Speech Commun.* 2014 Feb;5 7:181–208. DOI: <https://doi.org/10.1016/j.specom.2013.09.013>.
6. Navarro Tomás T. *Manual de pronunciación española*. 27th ed. Madrid: Editorial Consejo Superior de Investigaciones Científicas; 1999.
7. Sosa JM. *La entonación del español*. Madrid: Cátedra; 1999.
8. Quilis A, Fernández J. *Curso de Fonética y Fonología Españolas*. 18th ed. Madrid: Consejo Superior de Investigaciones Científicas; 2003.
9. Face T. F0 Peak Height and the Perception of Sentence Type in Castilian Spanish. *Rev Int Lingüíst Iberoam.* 2005;3(2(6)):49–65.
10. Estebas-Vilaplana E, Prieto P. Castilian Spanish Intonation. In: Roseano, Paolo, editor. *Transcription of Intonation of the Spanish Language*. München: Lincom; 2010. p. 17–48.
11. Henriksen N. Wh-question intonation in Peninsular Spanish: Multiple contours and the effect of task type. *J Port Linguist.* 2009;8(1):45–73.
12. Moraes JA. Intonation in Brazilian Portuguese. In: Hirst D, Di Cristo A, editors. *Intonation Systems: a Survey of Twenty Languages*. Cambridge: Cambridge University Press; 1998. p. 179–94.
13. Moraes JA. The Pitch Accents in Brazilian Portuguese: analysis by synthesis. In: Barbosa P, Madureira S, editors. *Speech Prosody*. Campinas; 2008. p. 389–97.

⁵ At the level of segments: (42,43). In the other direction (BP L2), see Capilla (66) Spanish L1 and BP L2 and Pereyron (67) Spanish L1, English L2, BP L3.

14. Tenani L. *Domínios Prosódicos no Português do Brasil: Implicações para a Prosódia e para Aplicação de Processos Fonológicos*. [Thesis]. [Campinas]: Universidade Estadual de Campinas; 2002.
15. Lucente L. *Aspectos dinâmicos da fala e da entoação do português brasileiro* [Dissertation]. [Campinas]: Universidade Estadual de Campinas; 2008.
16. Lucente L. *DaTo: um sistema de notação entoacional para o português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea* [Thesis]. [Campinas]: Universidade Estadual de Campinas; 2012.
17. Truckenbrodt H, Sandalo F, Abaurre MB. Elements of Brazilian Portuguese intonation. *J Port Linguist*. 2008;8(1):77–115.
18. Frota S, Moraes JA. Intonation in European and Brazilian Portuguese. In: Wetzels L, Costa J, Menuzzi S, editors. *The Handbook of Portuguese Linguistics*. Chichester: Wiley Blackwell; 2016. p. 141–66.
19. Sá PCF. *Análise entonacional de enunciados assertivos, continuativos e interrogativos lidos em piadas espanhol/LE e espanhol/LM* [Dissertation]. [Rio de Janeiro]: Universidade Federal do Rio de Janeiro; 2008.
20. Pinto M. *Transferências prosódicas do português do Brasil/LM na aprendizagem do espanhol/LE: enunciados assertivos e interrogativos totais* [Thesis]. [Rio de Janeiro]: Universidade Federal do Rio de Janeiro; 2009.
21. Rebollo L, Silva C, Pinho JR, Pinto M. Formulação de pedidos na aprendizagem de Espanhol/LE e transferências prosódicas por parte de falantes de português do Brasil. In: *I CIPLOM Congresso Internacional de Professores de Línguas Oficiais do MERCOSUL and I Encontro Internacional de Associações de Professores de Línguas Oficiais do MERCOSUL Línguas, sistemas escolares e integração regional*. Foz do Iguaçu; 2010. p. 580–8.
22. Silva C, Rebollo L, Pinto M. Pedidos de informação e pedidos de ação em Português do Brasil, fala carioca e em espanhol europeu, fala madrilena: variantes ou padrões entonacionais distintos? In: *III Colóquio Brasileiro de Prosódia da Fala*. Belo Horizonte; 2011.
23. Oliveira AF. *Caracterización de la entonación del español hablado por brasileños* [Thesis]. [Barcelona]: Universitat de Barcelona; 2013.
24. Dias ECO. *Declarativas e interrogativas totais no espanhol L1 e L2 falado em Bogotá: uma contribuição para estudos prosódicos* [Thesis]. [Florianópolis]: Universidade Federal de Santa Catarina; 2015.
25. Renato AC. Estudio exploratorio fonético-acústico de la transferencia prosódica del português de Brasil en el español/LE. In: Alves UK, editor. *Aquisição Fonético-Fonológica de Língua Estrangeira - Investigações Rio-Grandenses e Argentinas em Discussão*. Campinas: Pontes; 2016. p. 365–92.
26. Silva CC. *Análise fonético-experimental da entoação de declarativas e interrogativas em espanhol/LE*. [Thesis]: Universidade Estadual de Campinas; 2016.
27. Baralo M. *La adquisición del español como lengua extranjera*. 3rd ed. Madrid: Arco Libros; 2011.
28. Cantero Serena FJ. *Teoría y análisis de la entonación*. Barcelona: Edicions de la Universitat de Barcelona; 2002.
29. de Leeuw E, Mennen I, Scobbie JM. Singing a different tune in your native language: first language attrition of prosody. *Int J Biling*. 2012 Mar;16(1):101–16. DOI: <https://doi.org/10.1177/1367006911405576>.
30. Chang CB. Phonetic drift. In: Schmid MS, Köpke B, editors. *The Oxford Handbook of Language Attrition*. Oxford, UK: Oxford University Press; 2019. p. 191–203.
31. Schmid M, Köpke B. Introduction. In: Schmid M, Köpke B, editors. *The Oxford Handbook of Language Attrition*. Oxford, UK: Oxford University Press; 2019. p. 1–6.
32. Schmid M, de Leeuw E. Introduction to Linguistic Factors in Language Attrition. In: Schmid M, Köpke B, editors. *The Oxford Handbook of Language Attrition*. Oxford: Oxford University Press; 2019. p. 181–91.

33. Kupske F. Atrito Linguístico. In: Kupske F, Alves UK, Lima Jr. R, editors. *Investigando os sons de línguas não nativas: uma introdução*. Campinas: Editora da ABRALIN; 2021. p. 99–128. DOI: 10.25189/9788568990117.
34. Schereschewsky LC, Alves UK, Kupske FF. Atrito linguístico em plosivas em início de palavra: dados de bilíngues e trilíngues. *Rev Linguística*. 2019;15(2):10–29. DOI: <https://doi.org/10.31513/linguistica.2019.v15n2a21353>.
35. Flege JE. The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *J Phon*. 1987;15:47–65.
36. Major RC. Losing English as a First Language. *Mod Lang J*. 1992;76(ii):190–208.
37. Chang CB. Rapid and multifaceted effects of second-language learning on first-language speech production. *J Phon*. 2012;40:249–68. DOI: <https://doi.org/10.1016/j.wocn.2011.10.007>.
38. Chang CB. A novelty effect in phonetic drift of the native language. *J Phon*. 2013;41:520–33. DOI: <https://doi.org/10.1016/j.wocn.2013.09.006>.
39. Kupske F. *Imigração, atrito e complexidade: a produção das oclusivas surdas iniciais do inglês e do português por sul-brasileiros residentes em Londres*. [Thesis]. [Porto Alegre]: Universidade Federal do Rio de Janeiro; 2016.
40. Cohen GV. *The VOT Dimension: a bi-directional experiment with English and Brazilian-Portuguese stops*. [Dissertation]. [Florianópolis]: Universidade Federal de Santa Catarina; 2004.
41. Lord G. Second Language Acquisition and First Language Phonological Modification. In: *Selected Proceedings of the 10th Hispanic Linguistics Symposium*. 2008. p. 184–93.
42. De los Santos B da R. *A produção da vogal átona final /e/ por porto-alegrenses aprendizes de espanhol como segunda língua (L2): uma investigação sobre atrito linguístico em ambiente de L2 não-dominante*. [Dissertation]. [Porto Alegre]: Universidade Federal do Rio de Janeiro; 2017.
43. De los Santos B da R. Explorando a possibilidade de atrito linguístico: uma análise acústica da produção da vogal átona final /e/ na variedade Porto-Alegrense do Português Brasileiro. *Rev Bras Fonologia Laboratório*. 2018;3(1):14–41.
44. Schereschewsky LC, Alves UK, Kupske FF. First language attrition: the effects of English (L2) on Brazilian Portuguese VOT patterns in an L1-dominant environment. *Rev Digit Programa Pós-grad Em Let PUCRS*. 2017;10(2):700–16. DOI: <https://doi.org/10.15448/1984-4301.2017.2.26365>.
45. Osborne DM, Simonet M. Foreign-language phonetic development leads to first-language phonetic drift: plosive consonants in native Portuguese speakers learning English as a foreign language in Brazil. *Languages*. 2021;6:1–26. DOI: <https://doi.org/10.3390/languages6030112>.
46. Mennen I. Bi-directional interference in the intonation of Dutch speakers of Greek. *J Phon*. 2004;32:543–63.
47. Pierrehumbert J. *The phonology and phonetics of English intonation* [Thesis]. [Cambridge]: Massachusetts Institute of Technology; 1980.
48. Ladd, Robert. *Intonational Phonology*. Cambridge: Cambridge University Press; 2008.
49. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *Acoust Speech Signal Process*. 1978;26(1):43–9.
50. Salvador S, Chan P. Fast DTW: towards accurate dynamic time warping in linear time and space. In: *KDD workshop on mining temporal and sequential data*. Seattle; 2004. p. 70–80.
51. Müller M. *Information retrieval for music and motion*. Berlin: Springer; 2007.
52. Senin P. *Dynamic Time Warping Algorithm Review* [Internet]. University of Hawaii at Manoa; 2008 p. 1–23. Available from: http://seninp.github.io/assets/pubs/senin_dtw_litreview_2008.pdf
53. Giorgino T. Computing and Visualizing Dynamic Time Warping Alignment in R *The dtw Package*. *J Stat Softw*. 2009;31(7):1–24.
54. Morales-Cordovilla J, Cabañas-Molero P, Peinado A, Sánchez V. A Robust Pitch Extractor Based on DTW lines and CASA with Application in Noisy Speech Recognition. In: *Advances in Speech and Language Technologies for Iberian Languages* [Internet]. Madrid; 2012. p. 1–10. Available from: http://www.ugr.es/~jamc/Files/Morales12_DTWPitch.pdf

55. Lleó C. The acquisition of prosodic word structures in Spanish by monolingual and Spanish-German bilingual children. *Lang Speech*. 2006;49(2):205–29.
56. Flege JE, Bohn O-S. The Revised Speech Learning Model (SLM-r). In: Wayland R, editor. *Second Language Speech Learning* [Internet]. 1st ed. Cambridge University Press; 2021 [cited 2021 Apr 8]. p. 3–83. Available from: https://www.cambridge.org/core/product/identifier/9781108886901%23CN-bp-1/type/book_part
57. Boersma P, Weenink D. *Praat: doing phonetics by computer* [Computer program] [Internet]. 2021. Available from: <http://www.praat.org/>
58. Sánchez Aguilar A. *Don Quijote Adaptación, notas y actividades*. Barcelona: Vicens Vives; 2004.
59. Arantes P. *better_f0: A Praat script for better f0 extraction* [Internet]. Zenodo; 2019 [cited 2021 Mar 4]. Available from: <https://zenodo.org/record/3470108>
60. De Looze C, Hirst D. Integrating changes of register into automatic intonation analysis. In: *Speech Prosody*. Chicago; 2010. p. 1–4.
61. Arantes P. *parantes/f0-outliers: Version 1.0* [Internet]. Zenodo; 2021 [cited 2021 Jul 8]. Available from: <https://zenodo.org/record/5083453>
62. Arantes P. *Parantes/F0-Measures: Version 1.1.0* [Internet]. Zenodo; 2018 [cited 2021 Jan 14]. Available from: <https://zenodo.org/record/1453467>
63. RStudio Team. *RStudio: Integrated Development for R*. [Internet]. Boston; 2021. Available from: <http://www.rstudio.com/>
64. Kassambara A. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [Internet]. 2020. Available from: <https://CRAN.R-project.org/package=rstatix>
65. Mennen I. Beyond Segments: Towards a L2 Intonation Learning Theory. In: Delais-Roussarie E, Avanzi M, Herment S, editors. *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*. Berlin: Springer; 2015. p. 171–88.
66. Capilla MCC. *Espanhol e português em contato: o atrito da L1 de imigrantes espanhóis no Brasil*. [Dissertation]. [Brasília]: Universidade de Brasília; 2007.
67. Pereyron L. *A produção vocálica por falantes de espanhol (L1), inglês (L2) e português (L3): uma perspectiva dinâmica na (multi) direcionalidade da transferência linguística*. [Thesis]. [Porto Alegre]: Universidade Federal do Rio de Janeiro; 2017.