https://dx.doi.org/10.1590/1678-460X202258882

Articles

Effects of F0 movements, intensity, and duration in the perceptual identification of Brazilian Portuguese wh-questions and wh-exclamations

Efeitos de movimentos de F0, intensidade e duração na identificação perceptiva da questão parcial e da exclamação no português brasileiro

Luma da Silva Miranda[1]
João Antônio de Moraes[2,3]
Albert Rilliard[4,5]

ABSTRACT

*This study presents a perceptual analysis of the Brazilian Portuguese wh-question and wh-exclamation intonational contours to discriminate their acoustic and perceptual features. The corpus of this study is composed of the sentence "*Como você sabe*" ("How do you know" vs. "How clever you are!"), which was produced with both speech acts. Two perceptual identification experiments were designed to assess the subjects' ability to*

1. Eötvös Loránd University. Budapest – Hungary. https://orcid.org/0000-0002-5529-0338. E-mail: miranda.luma@btk.elte.hu.
2. Federal University of Rio de Janeiro. Rio de Janeiro – Brazil. https://orcid.org/0000-0002-8894-166X. E-mail: jamoraes3@gmail.com.
3. Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.
4. Université Paris-Saclay, CNRS, LISN. Orsay – France. https://orcid.org/0000-0001-6490-2386. E-mail: albert.rilliard@limsi.fr.
5. Federal University of Rio de Janeiro. Rio de Janeiro – Brazil.

identify these speech acts based on their prosodic characteristics, as well as the perceptual relevance of specific prosodic cues in the recognition of wh-questions and wh-exclamations. The results of the first perceptual test indicated that Brazilian listeners can identify these two speech acts by intonation only, whereas the second test showed that F0, duration and intensity cues contribute to the perceptual identification of the speech acts. Stimuli with a falling F0 movement in the last stressed syllable tend to be interpreted as wh-questions, whereas stimuli with a slightly rising F0 movement tend to be judged as wh-exclamations.

Keywords: *intonation; perception; wh-question; wh-exclamation; Brazilian Portuguese.*

RESUMO

Este estudo apresenta uma análise perceptiva da entoação da questão parcial e da exclamação do português brasileiro com o objetivo de discriminar suas características acústicas e perceptivas. O corpus deste estudo é constituído pela frase "Como você sabe", que foi produzida com os dois atos de fala. Aplicaram-se dois testes de percepção para avaliar a capacidade dos sujeitos em identificar esses atos de fala baseados nas suas características prosódicas, bem como a relevância perceptiva de pistas prosódicas específicas que atuam no reconhecimento da questão parcial e da exclamação. Os resultados do primeiro teste perceptivo aplicado com vinte e cinco ouvintes indicaram que eles identificam esses dois atos de fala somente pela entoação, enquanto os do segundo teste aplicado com vinte participantes mostraram que a configuração da F0 no núcleo das sentenças e as informações de intensidade e duração contribuem para o reconhecimento perceptivo desses atos de fala. Estímulos com movimento de F0 descendente na última sílaba tônica tendem a ser interpretados como questão parcial, enquanto os estímulos com movimento de F0 ascendente de baixa amplitude tendem a ser julgados como exclamação.

Palavras-chave: *entoação; percepção; questão parcial; exclamação; português brasileiro.*

1. Introduction

The illocutionary distinction (Searle, 1969) between wh-questions and wh-exclamations relies on obtaining a specific information in the former and, in the latter, on the expressive reaction of the speaker

towards either the interlocutor's message or an event. In some languages, there is a syntactic difference between these two speech acts (Nouwen & Chernilovskaya, 2015). In other languages, such as German (Repp, 2015, 2020), Hungarian (Mády et al., 2013) and Italian (Kellert et al., 2018), the syntactic structure remains the same and wh-questions and wh-exclamations are differentiated by the intonational contour, which is also the case of Brazilian Portuguese (henceforth, BP).

In terms of syntax, both sentences can be constructed in BP with the same wh-word in the initial position of the sentence; however, in the Brazilian Portuguese grammar, it is also possible to move the interrogative morpheme to another part of the clause such as the final position: "***Onde*** *você estava*?" x "*Você estava* ***onde***?" (both sentences meaning "Where have you been?"), but not in the wh-exclamative sentence. In this paper, sentences with the WH-word in the initial position of the clause were selected to be used in the perceptual analysis. Although several studies showing intonational differences between assertions and yes-no questions exist (Celeste & Reis, 2012; Frota et al., 2015; Moraes, 2008), wh-questions and wh-exclamations have seldom been investigated in BP, especially in terms of perceptual analysis. Yet, recent studies in BP have shown the relevance of the intonational contours for perceptually distinguishing these two types of speech acts in BP (Moraes, 2008; Miranda, 2015; Oliveira et al., 2014; Zendron da Cunha, 2016).

Importantly, Zendron da Cunha (2016) analyzed three types of Brazilian Portuguese wh-exclamative clauses with the wh-words "*como*" (how), "*que*" (what) and "*quanto*" (how many). The author concluded that, depending on the wh-word, there are two different intonational behaviors. On the one hand, wh-exlamatives with "*como*" (how) present a higher F0 in the wh-word in comparison with other parts of the contour and a falling nuclear F0 at the end of the utterance. On the other hand, wh-exclamatives with wh-word "*que*" and "*quanto*" have a lower F0 in the wh-word along with an intonational contour more similar to the yes-no question in BP, showing, at the end of the utterance, a rising intonation. In relation to the perceptual identification, Zendron da Cunha (2016) also shows that wh-exclamatives with "*como*" (how) are more easily recognized as exclamative uterances

than the ones with "*que*" and "*quanto*". In the present paper, only the wh-exclamations with the wh-word "como" are analyzed, since the corpus of this paper was set up with a paralelism to other types of speech acts (e.g., assertion and echo-question) with the intention of creating an analysis in a multimodal approach (Miranda, 2019). Hence, there was only one type of wh-word included in the perceptual analysis of the wh-exclamation in this paper.

Cross-linguistically, wh-questions present a falling F0 at the end of the contour. In Italian, for instance, a final F0 falling movement is the most typical pattern (Sorianello, 2011). Other phonetic cues were also described in the literature for wh-questions, when compared to assertions, such as a larger pitch range, a shallower topline decline and a less low pitch offset (Brinkmann & Benzmuller, 1999) as well as a higher nuclear accented peak (Rohloff & Michalsky, 2018) in wh-questions. Furthermore, in languages like English and Dutch, the wh-word can be accented (Chen, 2012). As for wh-exclamations, they are described, in many languages such as English (O'Connor & Gordon, 1961), German (Batliner, 1988), Italian (D'Eugenio, 1976) and French (Delattre, 1966), as having an initial extra-high pitch and, at the end of the contour, a falling intonation. However, in other languages like Hungarian (Mády et al., 2013), there is an initial low pitch. Apart from these intonation cues, other acoustic parameters may also characterize the wh-exclamations. For instance, the duration of the nuclear stressed syllable is lengthened in the Cosenza Italian dialect (Sorianello, 2011).

In Brazilian Portuguese, the wh-questions and wh-exclamations intonational contours are described with a nuclear falling F0 movement. Moraes (2008), for instance, analyzed the contours of speech acts produced by speakers from Rio de Janeiro. The author showed that, in the BP wh-question, there is a high initial F0 on the first stressed syllable followed by a falling F0 movement until the end of the utterance. It is worth mentioning that, in European Portuguese, similar results were found for the wh-questions in the studies published by Falé (2005), Frota et al. (2015) and Mata (1990). The BP wh-exclamation contour also presents a high initial F0 followed by a F0 fall; however, a slightly rising F0 movement is observed on the last stressed syllable, followed by a final falling F0 movement. In addition to the nuclear difference between the contours, Moraes (2008) also mentioned that there were

different intensity and duration cues in both contours. Miranda et al. (2020) partially corroborates this statement, since the authors did not find significant differences in relation to the duration in both speech acts, although they found a significantly different distribution of the syllables with more intensity within the utterance: the syllables "*co*" and "*mo*" were produced with more intensity in the wh-exclamation, whereas the syllable "*mo*" and "*vo*" were salient in the wh-question.

Dialectal studies available in BP intonation literature corroborate these findings. Frota et al. (2015) investigated the wh-question contour in two Portuguese varieties in the corpus "Interactive Atlas of Portuguese Prosody"[6] and found that there is a nuclear falling F0 accent in the wh-questions for both European (Lisboa, Porto, Alentejo and Algarve) and Brazilian Portuguese (Salvador, Belo Horizonte, São Paulo, and Rio Grande do Sul). Moreover, for the São Paulo dialect, Silva (2016) described the nucleus of wh-questions as predominently descendant, whereas Rosignoli (2017) reported rising, falling and monotonal F0 configurations in the nuclear region of the wh-question contour. The majority of F0 contours in Rosignoli's data (2017) was rising, which is different from other literature findings. However, Frota et al. (2015, p. 266) state that, although the nuclear F0 rising contour is typical of yes-no questions, wh-questions may present a rising nuclear F0 contour to express an added politeness to the utterance, which is in line with previous studies in other languages on wh-questions such as German (Essen, 1964; Kohler, 2004; Oppenrieder, 1988).

In addition to dialectal studies, experimental perceptual analysis with wh-questions and wh-exclamations are found in BP. Oliveira et al. (2014), besides reporting falling F0 movements for both wh-questions and wh-exclamations in the Brazilian dialect of Vitória da Conquista (Bahia State), also manipulated the F0 on the nuclear stressed syllables of the utterances. In the outcome of the experiment, the authors showed that the perceptual identification of wh-question intonational contours is favored in lower F0 regions (45 Hz till 150 Hz), while, in higher F0 ones (150 Hz till 300 Hz), the identification of wh-exclamation increases. Zendron da Cunha (2016) also reported, for Brazilian speakers from

6. For more information about the project: <http://labfon.letras.ulisboa.pt/InAPoP/presentation.html>. Accessed on 12 November 2021.

Florianópolis, a nuclear falling F0 contour for both speech acts, which are perceptually identified through their prosody. The author also mentioned that wh-questions were easier to identify perceptually than wh-exclamations, since the intonational contour of wh-exclamations varied, depending on the wh-word included in the sentence, as already mentioned in this introduction, whereas wh-questions always present the same F0 configuration.

According to the perceptual study designed by Moraes (2008), based on manipulated syllables of wh-questions and wh-exclamations, the final pre-stressed and stressed syllables were relevant for the identification of both speech acts. In the wh-exclamations, the final pre-stressed syllable was lowered, and the final stressed syllable was higher than in the wh-questions. It is worth mentioning that, in Miranda (2015), the final rising F0 movement in the nuclear stressed syllable of the wh-exclamation contour was also important for the identification of the pragmatic contrast of this speech act and the wh-question by Rio de Janeiro listeners.

Considering the findings of the previous studies, BP wh-questions and wh-exclamations present similarities throughout the intonational contour, especially in the nuclear region of the contour with a falling F0 movement. Yet, none of these studies has investigated, so far, the role of the multiple parameters of prosody, namely F0, duration and intensity, for the perceptual identification of these speech acts.

This paper presents a perceptual analysis of BP wh-questions and wh-exclamations and verifies the relative role of three acoustic features for the perceptual identification of two speech acts: the F0 configuration, intensity and duration. The present study: (a) applies an identification test of the intonational contours analyzed in this paper; (b) manipulates the F0 in the pre-nucleus and nucleus of wh-questions and wh-exclamations and (c) evaluates the role of each acoustic feature in the identification process. Based on results by Moraes (2008), Miranda (2015) and Oliveira et al. (2014), we hypothesize that (i) the speech acts will be identified by their intonational contour; and (ii) the manipulations in the nuclear rather than the prenuclear region of the contour will be perceptually relevant. Considering the intonation characteristics of the nucleus which will trigger the listeners'

identification of the contours, it is expected that (iii) a nuclear falling F0 movement favors the wh-question identification, while a slight rising F0 movement increases wh-exclamation answers. We also expect that (iv) the duration and intensity cues will improve the perceptual recognition of these intonational contours. The next sections of this paper are organized as follows: method, Experiment 1, Experiment 2, general discussion, and conclusion.

2. Method

The corpus recorded for this study is based on the sentence "*Como você sabe*" constructed with the WH-word in the initial position of the clause. This sentence was recorded while informants produced three different speech acts (Searle, 1969): wh-question (meaning: "How do you know it?"), wh-exclamation (meaning: "How clever you are!") and assertion (meaning: "As you know it"). The sentence "*Como você sabe*" when uttered as an assertion means that the speaker is giving an explanation to the interlocutor. In this study, the assertion speech act was included to serve as a basis from which the F0 manipulations could recreate the configuration of wh-question and wh-exclamative contours, without carrying other cues to the two target speech acts.

Six speakers (three female) from Rio de Janeiro recorded the corpus. All of them had an undergraduate degree and were adults (between 20 and 59 years of age). Before the recording sessions, the pragmatic context of each speech act was explained to the speakers in order to elicit the intonational contours and prepare them to produce these speech acts during the recording session. Each sentence was repeated three times by the six speakers, resulting in fifty-four recorded utterances. The recording sessions took place at the Phonetic Acoustic Laboratory of the Federal University of Rio de Janeiro (UFRJ), using microphone Shure SM 48, in digital media, using Sound Forge (version 7.0) software. The sound files were saved in wav format (mono, 16 bits and 22.050Hz). The two perception experiments were set up in the software Praat (Boersma & Weenink, 2016) with a script that was run with the function "ExperimmentMFC". The process of F0 manipulation was made in two steps, inspired by the IPO method ('t Hart et al., 1990): a close copy stylization and a standard stylization.

Detailed information about these procedures is provided in specific stylization sections of this article.

The study was conducted following the ethical requirements, which includes participants' consent either in the recording session (six speakers in the production task), or in the application of the perceptual experiment (twenty-five listeners in Experiment 1 and twenty listeners in Experiment 2). The anonymity of participants was also preserved.

3. Experiment 1: Perceptual identification and validation of speaker's production

The first perception experiment was designed to verify whether the assertions, wh-questions and wh-exclamations are identified based on their intonational contours.

Stimuli and procedure

The Experiment 1 included the fifty-four original recordings (6 speakers x 3 speech acts x 3 repetitions) in a forced-choice test. Listeners had to identify the intonational contours produced with the sentence "*Como você sabe*" given three options: assertion, wh-question and wh-exclamation. The audio files were randomized for each participant. Listeners were placed in an isolated acoustic laboratory with a computer and a headset. The tasks were taken individually, and one session lasted between 10 and 15 minutes.

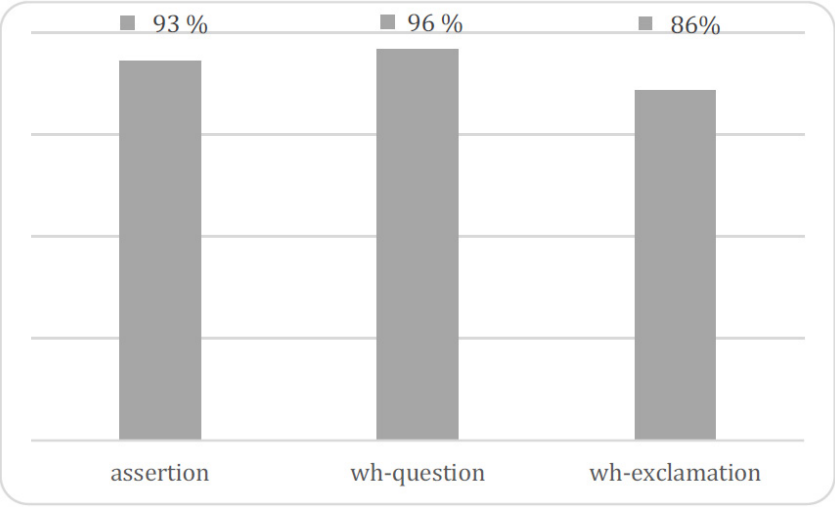
Participants

In Experiment 1, twenty-five listeners (seventeen women) participated and their mean age was 26 years old. All of them were undergraduate students of UFRJ, and none reported any auditory impairment.

Results

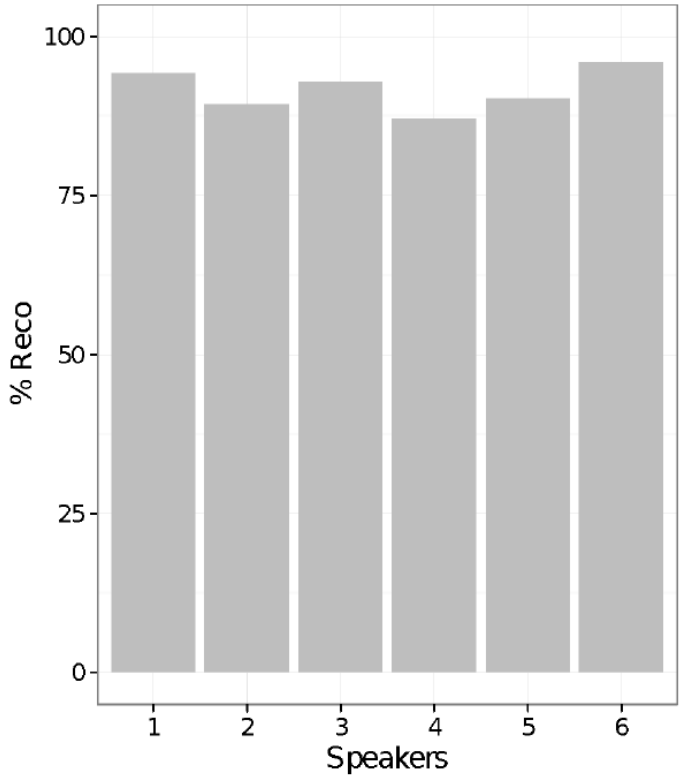
Figure 1 shows the outcome of Experiment 1. The wh-question received the highest identification rate (96%), followed by assertion (93%) and wh-exclamation (86%). The results of Experiment 1 corroborate the findings of BP intonation literature related to the identification of wh-questions and wh-exclamations, such as Moraes (2008), Oliveira et al. (2014) and Zendron da Cunha (2016).

Figure 1 – Bar plots representing the recognition percentages obtained by the 54 utterances



Regarding the speakers' performance, Figure 2 shows that the productions of the six speakers were identified above 80% by the listeners.

Figure 2 – Bar plots representing the average recognition percentages of the speech acts produced by the six speakers



In addition, since the perceptual test was a forced-choice task, it was possible to observe what types of confusion were made between these expressions. Table 1 presents the answers for the three categories used in Experiment 1 and shows the confusion between the three speech acts.

Table 1 – Confusion matrix with the results of the identification test: number of votes for each category plus the perceptual recognition percentage.

Perception / Production	Assertion	Wh-question	Wh-exclamation	Total
Assertion	**421 (93.6%)**	20 (4.4%)	9 (2.0%)	450 (100%)
Wh-question	16 (3.6%)	**430 (95.6%)**	4 (0.9%)	450 (100%)
Wh-exclamation	11 (2.7%)	52 (11.6%)	**387 (85.8%)**	450 (100%)

According to Table 1, the intonational contours uttered as assertions were mixed, in few cases, with wh-questions (4.4%) and wh-exclamations (2.0%), although assertion contours were well recognized in general (93.6%). The wh-questions were highly recognized at 95.6% and mixed more frequently with assertions (3.6%) than wh-exclamations (0.9%). The wh-exclamations presented a 85.8% identification rate, being mixed more frequently with wh-question (11.6%) than assertion (2.7%). Based on Experiment 1, it was possible to verify not only that the three modalities are distinguished through their prosodic performances, but also that the six speakers succeeded in the production task. In other words, the intended communicative purpose was correctly interpreted by the listeners through prosody, without interference from the syntactic or lexical levels. The first hypothesis of this study regarding the perceptual identification of the wh-questions and wh-exclamations was confirmed, which supports previous findings in BP (Moraes, 2008; Oliveira et al., 2014; Zendron da Cunha, 2016).

4. Experiment 2: Perceptual identification with resynthesized stimuli

Experiment 2 checks the relative role of the manipulated F0 configuration as well as the duration and intensity cues in the identification of each speech act.

Stimuli — F0 manipulation

The F0 stylizations were made with the TD-PSOLA (Moulines & Chapentier, 1990) included in the Praat software (Boersma & Weenink,

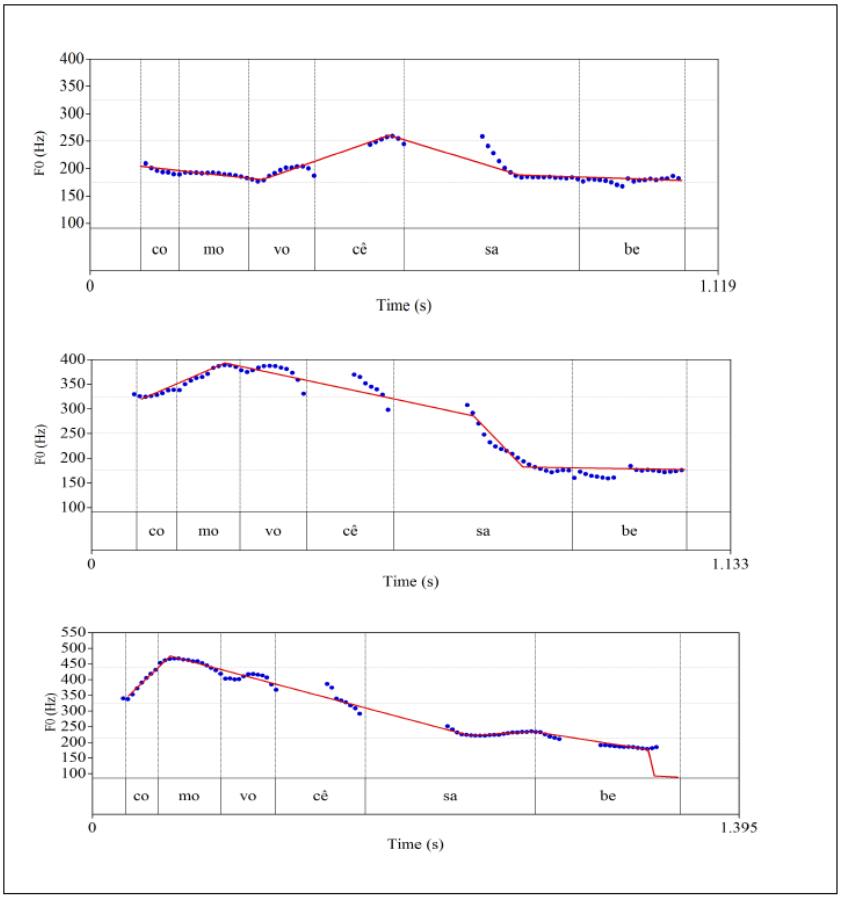
2016), which is a technique allowing modifications of the F0 from a given speech signal.

Close copy stylization

The close copy stylization proposed by the IPO (*Instituut voor Perceptie Onderzoek* - Institute for Perception Research) model developed by 't Hart et al. (1990) allows for a suppression of microprosodic variations, i.e., the elimination of influences on the melodic contour coming from the articulation of the sounds at the segmental level. Those variations that do not contribute to perception are eliminated based on the criterion of perceptual equality between the stylized and original F0 contour. By eliminating the phonetic details that do not participate in the communication of prosodic functions, the IPO model aims at unveiling the structure of the intonation which is part of the listener's linguistic knowledge. This stylization process was developed to serve as melodic configuration and was used in previous works such as Miranda (2015) and Moraes and Rilliard (2018). The creation of the stylized F0 contours themselves was made following the protocol described in Miranda (2015), which includes the following steps in software Praat (Boersma & Weenink, 2016): the function "to manipulation", "stylize pitch" and the elimination of the remaining inflection points using "CTRL+ALT+T".

For this stylization procedure, we selected one utterance of the wh-question, wh-exclamation and assertion produced by a female speaker (speaker 2), after two of the authors, who are trained linguists, have listened to the three repetitions of each speaker's speech act production. The result of the close-copy stylization of the three speech acts can be seen in Figure 3:

Figure 3 – Close-copy stylization (red line) and the original contour (blue line) of assertion (above), wh-question (middle) and wh-exclamation (below) produced by speaker 2



In Figure 3, the close-copy stylization of the assertion presents a flat F0 movement until the beginning of the syllable "*vo*", where a F0 rising movement starts and whose peak is located at the right edge of the stressed syllable "*cê*", immediately preceding the nucleus. Next, a F0 falling movement starts in the left edge of the nuclear stressed syllable "*sa*" until its middle, followed by a flat F0 movement that continues until the end of the utterance. The close-copy stylization of the wh-question starts with a high initial F0 rising until the prenuclear post-stressed syllable "*mo*" followed by a falling F0 movement. This F0 fall is steeper on the final stressed syllable "sa", down to a low level at

the end of the utterance. The wh-exclamation stylization contour starts with a high initial F0 whose peak is also located on the prenuclear post--stressed syllable "*mo*". Next, there is a falling F0 movement which is extended until the middle of the nuclear stressed syllable "*sa*". Along the nuclear stressed syllable, there is a slightly rising F0 movement in the right edge of the syllable, followed by a falling F0 movement on the final post-stressed syllable "*be*".

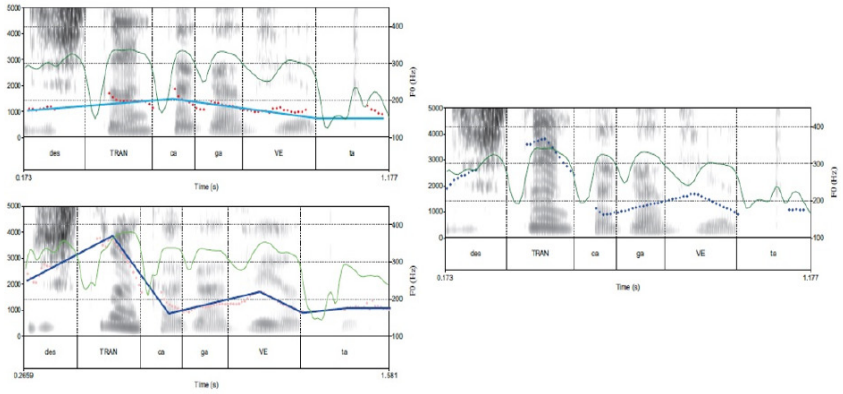
The close-copy stylization procedure showed that, in the prenuclear region, the wh-question and wh-exclamation contours presented a rising F0 movement followed by a F0 fall in the nucleus; however, the wh-exclamation contour presents a slightly rising-falling F0 movement on the nucleus instead of a steep F0 falling found in the wh-question. These phonetic characteristics of the two intonational contours were manipulated during the next stylization procedure and later were evaluated by listeners in order to verify which of these phonetic cues were perceptually relevant for the identification of the speech acts.

Standard stylization

IPO model ('t Hart et al., 1990) also includes a second process of manipulation of the melodic curve; this enables the experimenter to generalize the melodic movements that constitute an intonational contour based on perceptual equivalence. To achieve this, it is necessary to manipulate the acoustic features in the melodic contours and afterwards submit them to listeners' evaluation. The new F0 stylizations are applied onto the stylized contour obtained from the close-copy stylization procedure.

As the aim of Experiment 2 is to evaluate the influence of F0 configurations, duration and intensity on the perceptual identification of wh-question and wh-exclamation contours, six manipulations were made on the F0 stylization and applied onto the assertion, wh-question and wh-exclamation segmental base (i.e., the duration and intensity cues of each original contour were maintained). The details of the procedure can be seen in Moraes and Rilliard (2018) and the method is illustrated in Figure 4:

Figure 4 – Representation of the melodic contours transfer from the sentence "*Destranca a gaveta*" 'Unlock the drawer' to another: there are two original F0 curves on the left: assertion (above) and request (below). Original melodic curve (pitch) in red dotted line, the intensity with continuous green lines and the close--copy stylization of the intonation with blue straight lines. The graph on the right shows how the melodic curve of the request utterance is overlaid on the assertion utterance, retrieved from Moraes and Rilliard (2018, p. 246)



The variables used to manipulate the acoustic features of the wh-question and wh-exclamation intonational contours, which may contribute to their perceptual identification, were: (i) two F0 configurations in the pre-nucleus of the utterance: falling and rising; (ii) three F0 configurations on the nuclear stressed syllable: rising, falling-rising and falling, and, finally, (iii) three segmental bases, that is, the intensity and duration measurements for each of the segments for the assertion, wh-question and wh-exclamation, using the method of the F0 configuration transplantation. The combinations of the F0 in the prenucleus and nucleus of the stylized intonational contour are shown in Figure 5:

Figure 5 – Stylized F0 curves: falling (left) and rising (right) pre-nucleus combined with a rising F0 configuration on the nuclear stressed syllable

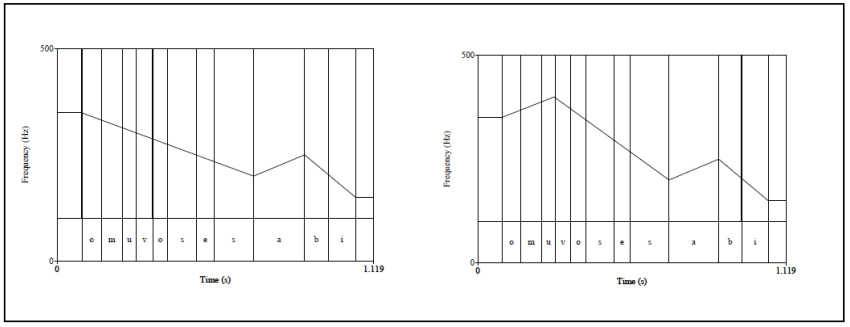
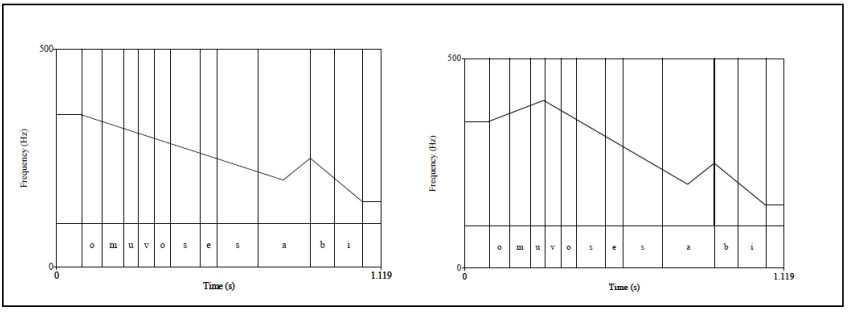


Figure 5 shows two stylized contours with two types of prenuclear F0 configurations and the same nuclear configuration. In the prenuclear region, there is a falling F0 movement with initial F0 in 400 Hz on the first stressed syllable "*co*" (left) and a rising F0 movement in 350 Hz on the same syllable followed by a F0 peak in 400 Hz on the post-stressed syllable "*mo*" (right). Along the nuclear region, there is a rising F0 movement on the stressed syllable "*sa*" whose F0 peak is in 250 Hz, followed by a F0 fall in the post-stressed syllable "*be*".

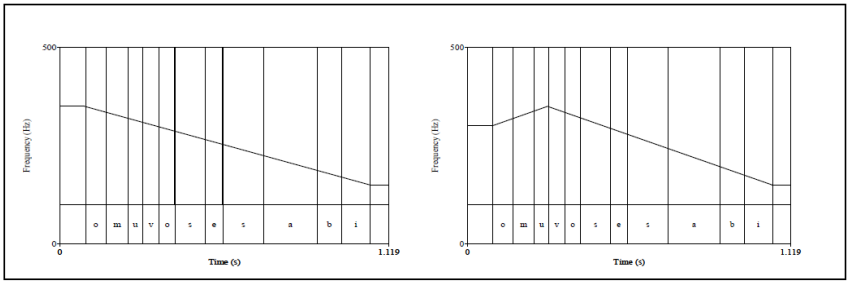
Figure 6 presents the same F0 prenuclear configurations as Figure 5, combined with a delayed F0 rise movement on the final stressed syllable *"sa"*, which creates a falling-rising shape whose F0 peak is also in 250 Hz, followed by a F0 fall over the post-stressed syllable "*be*".

Figure 6 – Stylized F0 curves: falling (left) and rising (right) pre-nucleus combined with a falling-rising F0 configuration on the nuclear stressed syllable



The last two F0 stylizations in Figure 7 are similar to F0 prenuclear configurations of Figures 5 and 6: a falling movement with initial F0 in 400 Hz on the first stressed syllable "*co*" (left) and a rising F0 movement in 350 Hz in the same syllable plus a F0 peak in 400 Hz on the post-stressed syllable "*mo*" (right). For the nuclear region, a falling F0 on the stressed syllable *"sa"* with mean F0 of 200 Hz, which continues until the post-stressed syllable "*be*", can be seen in both F0 contours.

Figure 7 – Stylized F0 curves: falling (left) and rising (right) pre-nucleus combined with a falling F0 configuration in the nuclear stressed syllable



The six F0 manipulations illustrated in Figures 5, 6 and 7 were transplanted onto the segmental bases of the three speech acts analyzed in this paper: assertion, wh-question and wh-exclamation. Therefore, in total, 18 stimuli were created with these combinations of F0 manipulations in the pre-nuclear and nuclear regions of the intonational contours as well as the three segmental bases in which the stylized contours were transplanted.

Procedure

The task of Experiment 2 consisted in identifying 18 stylized contours belonging to two speech acts: wh-question or wh-exclamation. The aim of this experiment was to verify the contribution of the type of F0 movement and the duration and intensity cues to the perceptual identification of wh-questions and wh-exclamations. The experiment was set up using a Praat script to present the stimuli randomly and collect the participants' answers. Listeners were tested in a sound-

attenuated room with a computer and a headset. The participants took the task individually and one session lasted between 10 and 15 minutes.

Participants

Twenty listeners (fifteen women) participated in the Experiment 2; none of them reported auditory impairment. All of them were either undergraduate or graduate students at UFRJ. The mean age of the participants was 26 years.

Statistical analysis

To verify whether the proportion of votes given for each speech act changes significantly as a result of the factors that were manipulated to create the stimuli, a logistic regression test was used. The results of each participant were extracted separately by the software Praat (Boersma & Weenink, 2016). Next, the results of all participants were grouped together, so that 360 responses from the experiment were analyzed. A regression analysis was applied in the proportion of "question" and "exclamation" answers received by the stimuli using the software R (2020), using the standard "stats" package. The factors included in the analysis were: (i) "prenuclear F0 configuration" with two levels: falling and rising; (ii) "nuclear F0 configuration" with three levels: rising, falling-rising and falling; and (iii) "the segmental base" with three levels: assertion, wh-question and wh-exclamation. This way, it was possible to verify the effect of the manipulated factors on the tendency of the stimuli being interpreted as "question".

Results

A logistic regression model with the three factors described in the section of standard stylization along with the interactions (doubles and triples) between them was calculated. Afterwards this was submitted to a simplification process, removing interactions and non-significant factors as well as grouping non-significant levels. The minimal adequate model (Crawley, 2013, p. 391) is described in Table 2:

Table 2 – Analysis of the regression analysis deviation of Experiment 2 (*likelihood-ratio*, *chi-square*, degree of freedom, *p*)

	LR Chisq	Df	Pr(>Chisq)
Base (Excl. – Ass./Ques.)	30.979	1	2.608e-08
Nucleus (Fall. – Rise/Fall.-Rise)	84.908	1	<2.2e-16

This model implies two factors without interactions: the "nuclear F0 configuration" and the "segmental base". The three levels of the "nuclear F0 configuration" were reduced to two levels: the rising and falling-rising movements *vs*. the falling movement ($\chi^2_{(1)} = 85, p < 0.05$). The three "segmental bases" were also reduced to two levels: assertion and wh-question bases *vs*. wh-exclamation base ($\chi^2_{(1)} = 31, p < 0.05$). The interactions and the factor "prenuclear configuration" did not have a significant effect on the distribution of responses.

As illustrated in Fig. 8, the identification of wh-question and wh-exclamation does not change significantly according to the prenuclear configuration.

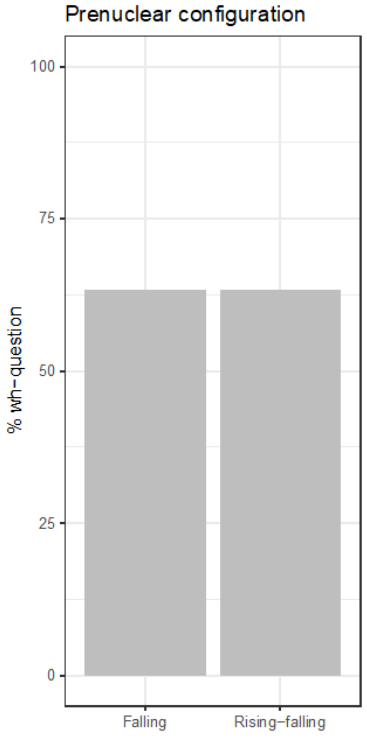


Figure 8 – Bar plots representing the votes (N = 40) obtained for wh-question, for each of the prenuclear configurations: falling (F) and rising (R); Duration and intensity patterns of original assertive sentence.

The identification of wh-question and wh-exclamation contours are very similar, regarding the pre-nuclear configuration: the falling F0 movement stands for 68% of the wh-question recognition and 32% for wh-exclamation. On the other hand, the rising configuration obtained 67% of recognition for wh-question and 33% for wh-exclamation. Note that, statistically, the types of pre-nuclear F0 configuration had no significant effect on the listener's proportion of votes. Therefore, this result shows that the prenuclear region of the wh-questions and wh-exclamations may present different shapes without affecting the identification of these speech acts. The second hypothesis of this study, which predicted that the F0 manipulation on the prenuclear region of the contour would be less important than the nuclear region, was also confirmed.

Figure 9 illustrates the effect of nuclear F0 configurations on the listeners' answers, showing only the opposition between rising and falling-rising *vs.* falling F0 movements, since there was no significant difference between rising and falling-rising configurations in the statistical analysis.

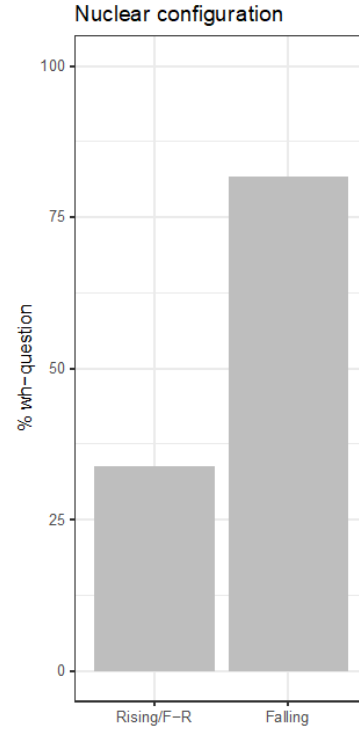


Figure 9 – Bar plots representing the votes (N = 60) obtained for wh-question, for each of the nuclear configurations: rising (R)/ falling-rising (F-R) and falling (F); Duration and intensity cues of the original assertive sentence.

With a nuclear falling F0 movement on the stressed syllable applied onto the assertive base, the stimuli receive 78% of wh-question answers, but the rising F0 movement on the same syllable receives 30% of question interpretation. The listener's interpretation is influenced by the type of F0 movement along the nuclear region of the sentence. The third hypothesis of this study was confirmed, which predicted that a falling F0 contour would favor the wh-question identification, while a slight F0 movement would increase the wh-exclamation identification, since the analysis revealed significant changes for the nuclear F0 movements.

Figure 10 shows the results of duration and intensity cues of the three segmental bases:

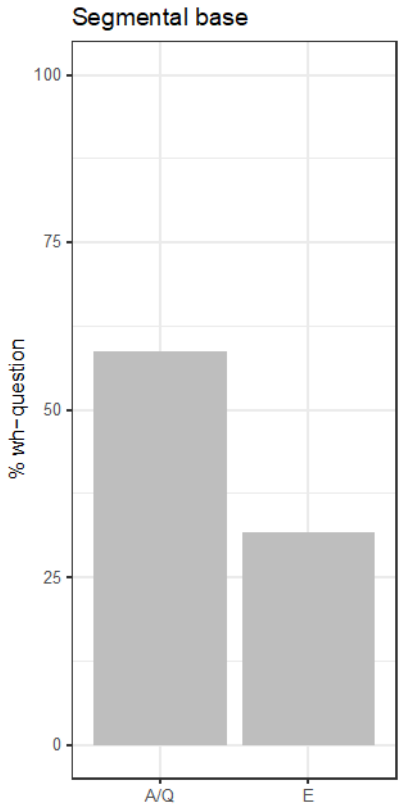


Figure 10 – Bar plots representing the votes (N = 120) obtained for wh-question based on duration and intensity of the original assertion (A)/wh-question (Q) and wh-exclamation (E).

In the statistical analysis, the difference between the levels "wh-question" and "wh-exclamation" of the factor "segmental base" was significant, while the difference between the assertion and wh-question segmental bases was not. The stimuli produced with either the assertion or the wh-question segmental base (i.e., duration and intensity cues of the assertion or wh-question original sentence) were interpreted as wh-questions, presenting 55% of the votes. When the wh-exclamation base was used, the stimuli were interpreted mostly as wh-exclamation, regardless of the F0 configuration: only 29% of question interpretation were attributed to stimuli produced with the wh-exclamation segmental base. This result confirms the fourth hypothesis of this study: the original duration and intensity cues of the wh-question and wh-exclamation contours favor the recognition of the two speech acts.

To sum up, the results of Experiment 2 showed that both the nuclear F0 movement along with the intensity and durational cues influence the perceptual recognition of wh-questions and wh-exclamations.

5. General discussion and conclusion

The present paper investigated whether Brazilian Portuguese wh-question and wh-exclamation speech acts are perceptually identified and have specific F0 configurations that influence their perceptual identification. By using an experimental design, two perceptual experiments were applied with Brazilian listeners. The outcome of Experiment 1 not only gives support to previous studies that showed that these two speech acts are identified by Brazilian listeners based on their prosody (Moraes, 2008; Oliveira et al., 2014; Zendron da Cunha, 2016), but also confirms the relevance of the intonational contour for distinguishing wh-questions and wh-exclamations in languages that present the same syntactic structure (Kellert et al., 2018; Mády et al., 2013; Repp, 2015, 2020). The results of Experiment 1 of the present study shows that BP wh-exclamative sentence with the wh-word "*como*" is highly recognized. It is worth-mentioning that Zendron da Cunha (2016) applied two identification tests with three types of wh-exclamations, including the structure with the wh-word "*como*". Zendron da Cunha's (2016) first perceptual test revealed that the identification of the wh-exclamation with the wh-word "*como*" was

higher than the others wh-words such as "*que*" and "*quanto*", while in her second perceptual test, the wh-word "*que*" was better recognized as wh-exclamation.

The results of Experiment 2 of the present study give support to the primacy of the nuclear region of the intonation contour over the prenuclear region in the identification of wh-questions and wh--exclamations, which was also confirmed by previous works in other languages. For instance, Kellert et al. (2018) applied an experiment with reaction time to find out if the prenuclear cues were used by listeners in the recognition of wh-questions and wh-exclamations in Cosenza Italian. The outcome of the experiment showed that listeners, in more than 90% of the trials, gave their responses after the end of the utterance. This paper along with the above-mentioned study (Kellert et al., 2018) give evidence to the Autosegmental-Metrical Theory (Pierrehumbert, 1980) in relation to the importance of the nuclear region for the pragmatic meaning of the intonation contour. However, there are some works claiming the relevance of the pre-nucleus in the perceptual identification of sentence mode: see, in Brazilian Portuguese, Nunes and Seara (2015) for the Sergipe dialect as well as Castelo (2016) for North Brazilian dialects, which show that the prenuclear region is important for the identification of statements and questions as well; also, see the analysis of Face (2011) for Spanish and Petrone and Niebuhr (2014) for German.

Experiment 2 also indicated that the F0 configuration within the nuclear region, specifically along the last stressed syllable, influences the identification of the illocutionary content of these speech acts in BP. Stimuli with a falling F0 movement in the nuclear stressed syllable were interpreted as wh-questions, whereas stimuli with a nuclear rising F0 movement were perceived as wh-exclamations. This is also in line with previous results in BP found in Miranda (2015), Moraes (2008), Oliveira et al. (2014), as well as Zendron da Cunha (2016), although the experimental designs of these studies are different. Regarding the role of the F0 movement in the nuclear region of the contour for the identification of wh-questions and wh-exclamations, the results of Oliveira et al. (2014)'s experiments, in which F0 values on the nuclear stressed syllable of the intonation contour were manipulated, showed that a lower F0 in the nuclear stressed syllable favors the identification

of wh-questions, whereas a higher F0 on the same syllable increases the recognition of the wh-exclamation.

In addition, as for the BP wh-exclamation, Moraes (2008) verified that a less steeper fall in the wh-exclamation contour was the most important acoustic feature in his perceptual experiment that led to the recognition of this speech act. Miranda (2015) also manipulated the wh-exclamation contour and the author found out that stimuli with a rising F0 movement on the nuclear stressed syllable with a late F0 peak are better recognized as wh-exclamation. The result of BP wh-exclamation in this paper also confirms previous findings in the literature, concerning other languages as well. Mády et al. (2013) analyzed the Hungarian wh-exclamations and showed that, in the perceptual experiment, acoustic stimuli with rising F0 movement, delayed peak and a combination of low initial boundary tones and rising accents were recognized as wh-exclamations.

Futhermore, Experiment 2 also showed that duration and intensity cues also influence the perception of the speech act. As far as we know, there is no previous study that analyzed the role of intensity and duration for the identification of these two speech acts in BP. The technique of F0 transplantation described by Moraes and Rilliard (2018) effectively analyzes the contribution of intensity and duration on the perceptual identification of Brazilian Portuguese speech acts.

In the present study, the manipulation of the fundamental frequency inspired by the IPO method (‘t Hart et al., 1990) allowed us to explore phonetically the behavior of the acoustic features in the wh-question and wh-exclamation intonational contours. Based on the second perceptual test, it was possible to investigate the weight of each acoustic parameter in the perceptual identification of these speech acts. However, the role of the prenuclear region in the identification of the analyzed speech acts remains unclear, since the first stressed syllable was kept high either with the rising (350 Hz) or the falling (400 Hz) F0 movements manipulated in this experimental study. Another way to assess the role of the prenuclear region is to manipulate the F0 level of the stressed syllable in the prenucleus. In addition, a perceptual test with a gating paradigm could be applied to verify whether Brazilian listeners identify the speech acts with their prosodic cues from the beginning of the

utterance, in a similar method applied by Falé and Faria (2006) with the European Portuguese question intonation.

This paper presented a preliminary attempt to investigate the role of prosodic cues in the perceptual identification of BP wh-questions and wh-exclamations. Based on the results of the experiments, we conclude that the prosody is an important source of information for the perceptual identification of both speech acts, which present the same syntatic structure. In addition, from the linguistic point of view, the F0 nuclear configurations as well as the duration and intensity cues are relevant for the recognition of the pragmatic meaning of Brazilian Portuguese wh-question and wh-exclamation speech acts.

Acknowledgements

This research was carried out with the support of Brazilian Federal Agency for Support and Evaluation of Graduate Education–CAPES. Funding code scholarship 88882.331896/2015-01 awarded by the first author during her Ph.D. course at Federal University of Rio de Janeiro, Brazil.

Conflict of interests

The authors declare they have no conflict of interest.

Credit Author Statement

We, Luma da Silva Miranda, João Antônio de Moraes and Albert Rilliard, hereby declare that we do not have any potential conflict of interest in this study. Luma Miranda and João Moraes participated in the study conceptualization, methodology, study design, project supervision, data collection, data generation and formal data analysis, whereas Albert Rilliard also participated in the formal data analysis as well as the statistical data analysis and data validation. We have all participated in editing the text. All authors approve the final version of the manuscript and are responsible for all aspects, including the guarantee of its veracity and integrity.

References

Batliner, A. (1988). Der Exklamativ: Mehr als Aussage oder doch nur mehr oder weniger Aussage? Experimente zur Rolle von Hohe und Position des F0-Gipfels. In H. Altmann (Ed.), *Intonationsforschungen* (pp. 243–271). Niemeyer.

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (Version 6.0.20) [Computer software]. https://www.fon.hum.uva.nl/praat/.

Brinckmann, C., & Benzmuller, R. (1999). The relationship between utterance type and F0 contour in German. In *Proceedings of the 6th European Conference on Speech Communication and Technology (Eurospeech 1999)* (pp. 21-24). https://www.isca-speech.org/archive/pdfs/eurospeech_1999/brinckmann99_eurospeech.pdf.

Castelo, J. (2016). A *entoação dos enunciados declarativos e interrogativos do Português do Brasil: uma análise fonológica em variedades ao longo da Costa Atlântica* [Unpublished doctaral dissertation]. University of Lisbon. http://hdl.handle.net/10451/26298.

Celeste, L. C., & Reis, C. (2012). Formal intonative analysis: Intsint applied to Portuguese. *Journal of Speech Sciences*, *2*(2), 3-21. https://doi.org/10.20396/joss.v2i2.15026.

Chen, A. (2012). Shaping the intonation of wh-questions: Information structure and beyond. In J. De Ruiter (Ed.), *Questions: formal, functional and interactional perspectives* (pp. 146-162). Cambridge University Press. https://doi.org/10.1017/CBO9781139045414.010.

Crawley, M. (2013). *The R Book* (2nd ed.). Wiley. https://www.perlego.com/book/1010482/the-r-book-pdf (Original work published 2012).

Delattre, P. (1966). Les 10 intonations de base du français. *The French Review*, *40*(1), 1-14.

D'Eugenio, A. (1976). The intonation systems of Italian and English. *Rassegna Italiana di Linguistica Applicata*, *8*(1), 57-85.

Essen, O. (1964). *Grundzüge der hochdeutschen satzintonation*. A. Henn Verlag.

Face, T. L. (2011). *Perception of Castilian Spanish intonation*: *Implications for intonational phonology*. LINCOM Europa.

Falé, I. (2005). *Percepção e reconhecimento da informação entoacional em português europeu* [Unpublished doctoral dissertation]. University of Lisbon.

Falé, I., & Faria, I. H. (2006). Categorical perception of intonational contrasts in European Portuguese. In R. Hoffmann & H. Mixdorff (Eds.), *Proceedings of Speech Prosody 2006*. Paper number 171.

https://www.isca-speech.org/archive/pdfs/speechprosody_2006/fale06_speechprosody.pdf.

Frota, S., Cruz, M., Svartman, F., Collischonn, G., Fonseca, A., Serra, C., Oliveira, P., & Vigário, M. (2015). Intonational variation in Portuguese: European and Brazilian varieties. In S. Frota, & P. Prieto. *Intonation in Romance* (pp 235-283). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0007.

Kellert, O., Panizza, D., & Petrone C. (2018). On the role of prosody in disambiguating wh-exclamatives and wh-interrogatives in Cosenza Italian. In I. Feldhausen, J. Fliessbach, & M. M. Vanrell (Eds.), *Methods in prosody: A Romance language perspective* (pp. 165-188). Language Science Press. https://doi.org/10.5281/zenodo.1441343.

Kohler, K. J. (2004). Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions. In G. Fant, H. Fujisaki, J. Cao, & Y. Xu (Eds.), *From traditional phonology to modern speech processing - In honour of Professor Wu Zongji's 95th birthday* (pp. 205-215). Foreign Language Teaching and Research Press.

Mády, K., Gyuris, B., & Szalontai, A. (2013, May 30). Experimental investigations on the prosody of Hungarian exclamatives [PowerPoint slides]. Docplayer. https://docplayer.net/8867107-Experimentalinvestigations-on-the-prosody-of-hungarian-exclamatives.html.

Mata, A. I. (1990). *Questões de entoação e interrogação em Português* [Unpublished master thesis]. University of Lisbon.

Miranda, L. (2015). *Análise da entoação do português do Brasil segundo o modelo IPO* [Unpublished master thesis]. Federal University of Rio de Janeiro.

Miranda, L. (2019). *Estudo fonético-perceptivo da entoação de enunciados assertivos, interrogativos e exclamativos no português do Brasil: uma análise multimodal* [Unpublished doctoral dissertation]. Federal University of Rio de Janeiro.

Miranda, L., Moraes, J., & Rilliard, A. (2020). Statistical modeling of prosodic contours of four speech acts in Brazilian Portuguese. In *Proceedings of Speech Prosody 2020* (pp. 404-408). https://doi.org/10.21437/SpeechProsody.2020-83.

Moraes, J. A. (2008). The pitch accents in Brazilian Portuguese: analysis by synthesis. In *Proceedings of Speech Prosody 2008* (pp. 389-397). https://www.isca-speech.org/archive/pdfs/speechprosody_2008/moraes08_speechprosody.pdf.

Moraes, J., & Rilliard, A. (2018). Describing the intonation of speech acts in Brazilian Portuguese: Methodological aspects. In I. Feldhausen, J. Fliessbach, & M. M. Vanrell (Eds.), *Methods in Prosody: A Romance*

language perspective (pp. 229-262). Language Science Press. https://doi.org/10.5281/zenodo.1441347.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech communication*, *9*(5-6), 453-467. https://doi.org/10.1016/0167-6393(90)90021-Z.

Nouwen, R., & Chernilovskaya, A. (2015). Two types of wh-exclamatives. *Linguistic Variation*, *15*(2), 201-224. https://doi.org/10.1075/lv.15.2.03nou.

Nunes, V. G., & Seara, I. C. (2015). Distinção de variedades dialetais e de modalidades através de contornos de regiões pré-nucleares: Análises acústicas e perceptuais. *Diadorim*, *17*(2), 34-51. https://doi.org/10.35520/diadorim.2015.v17n2a4068.

O'Connor, J. D., & Arnould, G. F. (1961). *Intonation of coloquial English*: *A practical handbook*. Longmans.

Oliveira, J. S. N., Pacheco, V., & Oliveira, M. (2014). Análise perceptual das frases exclamativas e interrogativas realizadas por falantes de Vitória da Conquista/BA. *Signum: Estudos da Linguagem*, *17*(2), 354-388. http://dx.doi.org/10.5433/2237-4876.2014v17n2p354.

Oppenrieder, W. (1988). Intonatorische Kennzeichnung van Satzmodi. In H. Altmann (Ed.), *Intonationsforschungen* (pp. 69-205). Niemeyer.

Petrone, C., & Niebuhr, O. (2014). On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech*, *57*(1), 108–146. https://doi.org/10.1177/0023830913495651.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation.* The MIT Press. https://dspace.mit.edu/handle/1721.1/16065.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Repp, S. (2015). On the acoustics of wh-exclamatives and wh-interrogatives: Effects of information structure and sex of speaker. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Paper number 0319.1-5. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0319.pdf.

Repp, S. (2020). The prosody of wh-exclamatives and wh-questions in German Speech acts differences, information structure, and sex of speaker. *Language and Speech*, *63*(2), 306-361. https://doi.org/10.1177/0023830919846147.

Rohloff, M., & Michalsky, J. (2018). Pitch scaling as a question cue in German wh-questions. In *Proceedings of Phonetics & Phonology*

in German-Speaking Countries (pp. 169–172). http://dx.doi.org/10.18452/18805.

Rosignoli, C. C. (2017). O padrão entoacional das sentenças interrogativas parciais do português brasileiro em fala manipulada. *Estudos Linguísticos (São Paulo. 1978)*, *46*(1), 41-54. (accessed 03 April, 2022). https://doi.org/10.21165/el.v46i1.1633.

Searle, J. (1969). *Speech Acts*. Cambridge University Press.

Silva, C. (2016). *Análise fonético-experimental da entoação de declarativas e interrogativas em espanhol/LE* [Unpublished doctoral dissertation]. State University of Campinas. https://doi.org/10.47749/T/UNICAMP.2016.968924.

Sorianello, P. (2011). Aspetti pragmatici e prosodici dell'atto esclamativo. *Studi Linguistici e Filologici Online (SLIFO)*, *9*(1), 287–332.

't Hart, J., Collier, R., & Cohen, A. (1990). A perceptual study of intonation: An experimental-phonetic approach to speech melody. Cambridge University Press. https://doi.org/10.1017/CBO9780511627743."

Zendron da Cunha, K. (2016). *Sentenças exclamativas em português brasileiro: Um estudo experimental de interface* [Unpublished doctoral dissertation]. Federal University of Santa Catarina. https://repositorio.ufsc.br/xmlui/handle/123456789/172264 (accessed 03 April, 2022).

Recebido em: 30/11/2021
Aprovado em: 03/03/2022