

## RESEARCH PAPER

# Listening to accented speech in Brazilian Portuguese: On the role of fricative voicing and vowel duration in the identification of /s/ – /z/ minimal pairs produced by speakers of L1 Spanish

Ubiratã Kickhöfel Alves<sup>1</sup> and Luciene Bassols Brisolara<sup>2</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, BR

<sup>2</sup> Universidade Federal do Rio Grande, BR

Corresponding author: Ubiratã Kickhöfel Alves ([ukalves@gmail.com](mailto:ukalves@gmail.com))

---

This article reports the results of two experiments investigating the combined role of vowel length and length of fricative voicing in the identification, by Brazilians, of minimal pairs such as *casa* /z/ – *caça* /s/ produced by speakers of Spanish (L1). In Experiment 1, stimuli were manipulated so that length of voicing in the fricative was tested in two levels (100% or 0% of voicing) and vowel length was tested in four levels (25%, 50%, 75% and 100% of the length of the total vowel). In Experiment 2, voicing length was tested in three levels (25%, 50% and 75% of voicing), combined with the four levels of vowel length (25%, 50%, 75% and 100% of the length of the total vowel). Both experiments were run on TP Software (Rauber et al. 2012), and forty Brazilian listeners with no experience with Spanish took part in both tasks. The results show an interaction between the two cues, especially in the stimuli with no full voicing in the fricative. These findings provide additional evidence to the gradient status of speech in production and perceptual phenomena (Albano 2001; Albano 2012; Perozzo 2017), besides shedding light on the teaching of Brazilian Portuguese as an Additional Language.

---

**Keywords:** Laboratory Phonology; Brazilian Portuguese; speech perception; L2 speech intelligibility; acoustic-articulatory gestures

---

## 1. Introduction

In Brazilian Portuguese (BP), the voiced ([z]) and voiceless ([s]) counterparts of the alveolar fricative differ phonologically, such as in the minimal pairs *casa* (‘house’) – *caça* (‘hunt-noun’) and *rasa* (‘shallow’) – *raça* (‘breed’), whose first members are produced with the voiced fricative (Cristófaros-Silva 2017; Seara, Nunes & Lazzarotto-Volcão 2015). This is a different scenario from Spanish, in which a phonological distinction between the voiceless and the voiced counterparts of the alveolar fricative does not occur (Alarcos LLorach 1965; Hualde et al. 2010; Martínez Celdrán 2003; Navarro Tomás 2004; RAE 2011). This difference between the two languages poses difficulties to L1 Spanish speakers attempting to learn BP, as they find it difficult to produce the voiced fricative in the target language (cf. Akerberg 2004; Alves 2015a; Oliveira 2016; Silveira & Souza 2011; Sobral, Nobre & Freitas 2006, among others).

Grounded on a dynamic account of phonology (Albano 2001; Albano 2012; Alves 2018a; Zimmer & Alves 2012), our observations of the fricatives produced by L1 Spanish speakers learning BP suggest that the characterization of these data should go beyond the binary description of “total presence” or “total absence” of voicing in the fricative, as the

development of the distinction between the members of these minimal pairs corresponds to a gradient process. Considering that the developmental process of a new language corresponds to learning to “orchestrate units of time” (Zimmer & Alves 2012), we have noticed that voicing in the fricative in BP is produced by these learners in a continuum, which extends from little voicing to voicing all the way through the consonant. As this gradient status challenges the so-called binary distinction formalized by the traditional [voice] distinctive feature, it was necessary to determine how much voicing in the fricative proved necessary so that L1 Brazilian Portuguese listeners could identify the members of these minimal pairs as showing /s/ or /z/.

In order to answer this question, in Alves et al. (2018) we sought to investigate the effects of voicing length in the identification of the fricative [z] produced by L1 Spanish speakers on the distinction between the categories of “voiceless” and “voiced” consonants by Brazilian listeners. Speech data were collected from six L1 Spanish speakers who had been living in Brazil for less than twelve months. From the recordings and manipulations of different degrees of voicing in the fricative (0%, 25%, 50%, 75% and 100% of the total duration of the fricative), an identification task was built on TP Software (Rauber et al. 2012). Thirty-five Brazilian participants took part in this task. The results indicated that voicing the consonant all the way through was not a necessary condition for the identification of the fricative as voiced. It was also verified that the pattern with 25% of voicing was more difficult to identify as [z].

Although the results in Alves et al. (2018) proved important to show that full voicing throughout the fricative is not mandatory for the identification of [z], since the length of the preceding vowel had been maintained the same in all stimuli (around 220 ms), it remained to be said whether vowel length also played a role in the identification of the fricative. This possibility deserves consideration since, according to the data shown in Ribeiro (2017), the voiced counterparts in minimal pairs such as *cata/cada* (‘you catch/each’) and *capo/cabo* (‘I castrate/cape’) present a significantly higher vowel length in BP. As this difference was found in production, it was possible that voicing length would play a decisive role in speech perception, especially in instances of accented speech, in which some primary cues in speech perception might not be fully produced.

The possibility described above deserves special consideration, since preceding vowel length plays an important role in the identification of consonant voicing in many languages (Chen 1970). In English, for example, vowel length plays a decisive role in the identification of minimal pairs such as “cap” vs. “cab” or “right” vs. “ride”, with the voiced counterparts of the pairs presenting higher vowel lengths (Ladefoged & Johnson 2010). Even in Dutch, in which final voicing distinctions are traditionally attested to be neutralized (Kenstowicz 1994), voicing length differences can still be found among members of minimal pairs (Ernestus & Baayen 2006; Port & O’Dell 1985). Despite these sources of evidence, Keating (1985) explains that this difference in vowel length “cannot be automatic and physiological” (Keating 1985: 123). By providing evidence from Polish and Czech, the author claims that this vowel length pattern cannot be regarded as predictable, as this phenomenon characterizes different grammars in different language systems, not constituting a physiologically intrinsic aspect. This motivated Ribeiro’s (2017) study, which suggested that this length difference is part of the grammar system of BP. In our study, therefore, we enquire if this significant difference found in production would also play a role in the perception of BP words, especially in cases in which the fricative is partially voiced.

Departing from this enquiry, in this article we report the results of two experiments that aim to provide evidence in support of an interaction between the cues of “vowel length” and “length of fricative voicing” in the identification of the fricative as [z]. In Experiment 1, stimuli were manipulated from the same original data used in Alves et al. (2018), having

“length of fricative voicing” (tested in two levels: 100% or 0% of voicing) and “vowel length” (tested in four levels: 25%, 50%, 75% and 100% of the length of the total vowel) as the main factors investigated. Vowel length was manipulated as we reduced the length of the vowels that preceded voiced fricatives, so that 100% of vowel length would correspond to the original length of the segment preceding the voiced fricative. The experiment was run as an identification task on TP Software (Rauber et al. 2012), and forty Brazilian listeners with no experience with Spanish took part in the study.

As voicing in the productions of BP /z/ by L1 Spanish speakers is developed in a gradient fashion (cf. Alves et al. 2018), it proved important to test additional experimental conditions, in which voicing in the fricative was also partial. Therefore, in Experiment 2, we added testing conditions in which voicing was not fully present or absent along the fricative. The tested conditions consisted of three different degrees of voicing in the fricative (25%, 50% and 75%), combined with the four levels of vowel length employed in the previous experiment (25%, 50%, 75% and 100%). This task was administered after two weeks to the same forty participants who had taken part in the previous experiment.

The results of these two experiments have theoretical implications for the discussion on the phonological primitives of speech perception employed in the current models of L2 speech perception. In the L2 research scenario, two perceptual models have assumed a leading role in the last twenty years: the Speech Learning Model (Flege 1995; Flege 2003, among others)<sup>1</sup> and The Perceptual Assimilation Model (Best 1995)/Perceptual Assimilation Model-L2 – (Best & Tyler 2007).<sup>2</sup> Although both models account for the dynamicity of speech sounds and their gradience in phonology, we claim that both face challenges with regard to the phonological primitive employed by each of them.

Though not explicitly put in its original texts, the Speech Learning Model (SLM) is assumed to employ a psychoacoustic approach to speech perception (Alves & Silva 2016; Nishida 2012; Nishida 2014a; Nishida 2014b; Perozzo 2017). In other words, acoustic information seems to function as the primitive unit of speech perception. This assumption, on the one hand, has contributed to regarding speech perception as a domain-general cognitive process, rather than dependent on a specific language acquisition device (Chomsky & Halle 1968). On the other hand, this has led to an ongoing discussion on The Constancy Problem (cf. Strange 1995), given the likely consequence that a unique representation for every single acoustic production should be assumed in this approach. As such a biunivocal correspondence between acoustics and phonological representations was seen as untenable, the model tended to be regarded as a phonetically-driven approach, without much to say about phonological representations.

An alternative approach was offered by Best (1995) in her Perceptual Assimilation Model (PAM), as well as in Best & Tyler (2007), in their Perceptual Assimilation Model-L2 (PAM-L2). Based on the discussion regarding L1 speech perception proposed in Fowler (1986), both the PAM and the PAM-L2 are grounded on a Direct Realist Approach to L2 Speech Perception. This proposal would prove attractive not only regarding the principles that guide the perceptual phenomena, but also in what concerns the phonological primitive adopted in the model. As for the tenets of speech perception, Fowler’s direct realist approach

<sup>1</sup> Another L2 perceptual model that has guided many studies in the last fifteen years is the L2LP Model, by Escudero (2005, 2009). However, in its original version (2005), the model formalizes different levels of representation (an auditory, a phonological and an underlying level). For this reason, this model will not be discussed in this paper.

<sup>2</sup> It should be noted that, though grounded on the same theoretical assumptions, both models differ in what regards their object of investigation. The original version of the Perceptual Assimilation Model (Best 1995) deals with assimilation patterns by naïve L2 learners. In turn, Best & Tyler (2007) propose an adaptation of the Perceptual Assimilation Model in order to deal with cases of L2 development (hence the name “Perceptual Assimilation Model-L2”). For a more complete description of these two proposals, see Alves & Silva (2016), Perozzo (2017) and Antoniou (2018).

for speech perception followed Gibson (1966, 1979), thus conceiving perception as a domain-general process. With regard to the phonological primitive employed in both PAM and PAM-L2, the authors adopted articulatory gestures (cf. Browman & Goldstein 1986; Browman & Goldstein 1992) as the basic units of speech perception, thus implementing a more dynamic approach to phonology.

In this paper, we discuss the adequacy of the two perceptual models in accounting for our data, as we consider the combined role of two cues in voicing distinctions. We then follow an alternative approach to tackle both L2 perception and the perception of accented speech (Perozzo 2017), grounded on acoustic-articulatory gestures (Albano 2001) as the primitives of speech perception. We suggest that Perozzo's (2017) model proves more epistemologically coherent and accounts for our data more appropriately, as we impose some additional challenges to be pursued by the model.

As we focus on the perception of accented Brazilian Portuguese productions, this paper also plays a role in the discussions on pronunciation teaching. Focusing on speech intelligibility as the main goal to be pursued by foreign learners (Derwing & Munro 2015a; Derwing & Munro 2015b; Levis 2018), we aim to show that native-like speech is not a pre-requisite for speech intelligibility to take place. We therefore assume that listeners can make use of other cues that might not necessarily play a role in the perception of non-accented speech. This considered, phonological oppositions may take place through a combination of different cues in action, depending on the foreign learners' L1 background and their language experience with other additional languages. This fact needs to be taken into consideration by foreign language teachers, who should be aware that having a foreign accent does not necessarily hinder intelligibility.

We therefore believe this study may contribute to discussions on both theoretical and applied linguistics, as it is our intention to contribute to (i) the description of a dynamic phenomenon that takes place in the perception of Brazilian Portuguese; (ii) the discussion on phonological primitives in speech perception and their formalization in L2 perceptual models; (iii) the debate on the necessity of an “intelligibility” (Levis 2005; Levis 2018) instead of a “nativeness” principle to L2 pronunciation teaching.

## 2. Method

### 2.1. Stimuli

The original non-manipulated data<sup>3</sup> were the same ones employed in Alves et al. (2018). In that experiment, we had manipulated different rates of voicing in the fricative. In the present study, in both experiments, we aim to test the manipulated patterns from that previous study in combination with different vowel lengths.

As explained in Alves et al. (2018), the production data that served as stimuli had been obtained from six L1 speakers of Spanish who had been living in the South of Brazil (in the city of *Rio Grande-RS*) for 3–12 months. All of them were graduate students at *Universidade Federal do Rio Grande* and had been taking Brazilian Portuguese lessons in order to improve their proficiency in the language. **Table 1** presents an overall description of these L1 Spanish speakers.

Despite their different nationalities, none of the learners voiced the fricative /s/ in syllable-initial position in their L1 variety. This is particularly important, as voicing the fricative in their L1 could affect their production patterns in BP and have an effect on the results of the perceptual task. As this study is part of a larger project, in which we carried out data collections of free speech productions in both L1 Spanish and L2 Portuguese,

<sup>3</sup> All methodological procedures were performed in accordance with the Declaration of Helsinki. This study was approved by the Ethics Committee of the institution of the first author (Approval number 1.844.916).

we analyzed the participants' L1 productions of their fricatives in our databank so as to make sure that they did not present instances of voicing. This methodological step was taken because some varieties of Colombian (RAE 2011; Vaquero de Ramírez 2003) and Mexican (Cuétara Priede 2004; Perissinotto 1975) Spanish tend to voice /s/ in syllable-initial positions.<sup>4</sup>

The six learners from **Table 1** were invited to read a set of target sentences in Brazilian Portuguese containing the /z/-/s/ minimal pairs *asa/assa* ('wing/he bakes'), *casa/caça* ('house/hunt'), *rosa/roça* ('rose/farm'), *pesa/peça* ('he weighs/piece'), *rasa/raça* ('shallow/breed') and *tese/tece* ('thesis/he weaves'). Each member of the minimal pairs was repeated five times. Further acoustic analyses of the fricative segments in these minimal pairs made it clear that voicing did not occur all the way through the fricative in these participants' productions, which motivated us to investigate the role of different voicing lengths in the fricative.

In order to build our perceptual stimuli, in view of the audio characteristics of the recorded material, in Alves et al. (2018) we chose one minimal pair for each one of the speakers, as **Table 2** shows. We selected (partially) voiced stimuli with similar lengths in the preceding vowel, making sure that their fricative segments had also been produced with very similar lengths, as shown in **Table 2**.

**Table 1:** Information on the BP Learners (L1: Spanish).

	Country of origin	Age of onset of learning of BP	Months living in Brazil
Speaker 1	Venezuela	32	3 months
Speaker 2	Peru	24	12 months
Speaker 3	Honduras	23	3 months
Speaker 4	Colombia	23	3 months
Speaker 5	Ecuador	27	9 months
Speaker 6	Mexico	23	3 months

Source: Alves et al. (2018: 447).

**Table 2:** Stimuli to be manipulated.

Participant/target word	Vowel length (ms)	Fricative length (ms)
participant1_asa	226.00	141.00
participant2_casa	222.00	146.00
participant3_pesa	218.00	143.00
participant4_rasa	219.00	143.00
participant5_rosa	218.00	145.00
participant6_tese	220.00	145.00

Source: Alves et al. (2018: 448).

<sup>4</sup> As will be shown in the Conclusion, we further investigated the effects played by this difference in the speakers' profiles, as in a follow-up verification we reran the statistics of the experiments including "speaker" as an independent variable. The results showed main effects of "speaker", as well as interactions among this variable and the two others investigated in this study ("voicing length" and "vowel length"). Since "speaker" turned out to be a confounding variable, as will be discussed in the Conclusion, we see these differences in the speakers' profiles, in terms of both nationality and length of stay in Brazil, as a limitation of the present study.

From those selected audios, we were able to manipulate stimuli so as to have voicing lengths of 0%, 25%, 50%, 75% and 100% of the total length of the fricative. All stimuli manipulations were carried out on *Audacity – version 2.3.0*. In order to manipulate different patterns of length of voicing in the fricative,<sup>5</sup> we added portions of voiced fragments of audio files that had been obtained from a production of the same lexical item that was being manipulated. This voiced portion should correspond to the length of the desired voicing pattern, besides 25% of the length of the previous vowel, so that the fricative and the vowel could be overlapped in the last 25% of the length of the vowel and then crossfaded (so that a smooth transition from the vowel to the following consonant could be created). On its right border, the voicing of the stimuli was gradually weakened, in order to simulate a gradient loss of voicing all over the fricative. Given this description of the audio portions used in the manipulation, the procedures for manipulating voicing length<sup>6</sup> adopted in Alves et al. (2018) consisted of (i) inserting some silent period on the portion of the fricative that was going to receive voicing, making sure that the adding of the new audio portion respected the zero transitions of the original audio stimuli; (ii) inserting the portion of audio with the desired voicing pattern, also making sure that the adding of the new audio portion respected the zero transitions of the original audio; (iii) crossfading the vowel-fricative transition so as to obtain a smooth change between each of the segments.

As already mentioned in the Introduction, the results in Alves et al. (2018) showed that voicing all the way through the fricative did not prove an essential condition for the identification of the segment as voiced [z]. In fact, patterns with 50%, 75% and 100% of voicing were predominantly identified as voiced, and results concerning the pattern with 25% of voicing tended to vary among the 35 listeners that took part in this study. We therefore inquired whether preceding vowel length would play a role in the participants' decisions. As we consider the stimuli in the previous study, this question was particularly important, as we had used a longer vowel to control for preceding vowel length, given that our manipulations had been carried out in the voiced counterparts of the minimal pairs. In the experiments we carry out in the present study, therefore, we not only make use of the fricative voicing manipulations already investigated in Alves et al. (2018), but we also test the effect of manipulated vowel length.

## 2.2. Experiment 1

In order to test the effects of vowel length, in Experiment 1 we considered stimuli which varied in 25%, 50%, 75% and 100% of the original vowel length. In this experiment, we tested only those stimuli from Alves et al. (2018) that were fully voiced (100% of voicing) or devoiced (0%). In order to manipulate vowel length, we extracted portions of the vowels from the original lengths shown in **Table 2**.<sup>7</sup> Therefore, 100% of voicing length corresponded to the original vowel length that preceded the (partially) voiced counterpart of the minimal pair in the original stimuli, as shown in **Table 2**. In the same fashion as in Alves et al. (2018), fricative length was also controlled in Experiment 1, exhibiting those values shown in **Table 2**.

In Experiment 1, thus, we tested the perception of eight different manipulated patterns, as shown in **Table 3**.

<sup>5</sup> As will be shown in the next subsections, in Experiment 1 we manipulated patterns with 100% of voicing in the fricative, while in Experiment 2 we tested fricatives with 25%, 50% and 75% of voicing.

<sup>6</sup> The procedures for vowel length manipulation are discussed in the next section.

<sup>7</sup> Following Motta-Ávila (2017), we reduced the length of the vowel by cutting of portions from their central region to their borders. This manual manipulation aimed to respect vowel transitions with the previous and following consonants. This was also carried out on *Audacity*.

**Table 3:** Manipulated patterns tested in Experiment 1.

<b>% voicing_% vowel length</b>
0% voicing_25% vowel length
0% voicing_50% vowel length
0% voicing_75% vowel length
0% voicing_100% vowel length
100% voicing_25% vowel length
100% voicing_50% vowel length
100% voicing_75% vowel length
100% voicing_100% vowel length

Considering the speakers/minimal pairs presented in **Table 2**, there were six different audio files/target words for each one of the patterns tested. Each one of these audio files was randomly repeated three times, which adds up to 18 tokens in each experimental condition. Experiment 1, thus, totaled 144 tokens.

The experiment was run as an identification task on TP Software (Rauber et al. 2012). As in Alves et al. (2018), participants were asked to choose whether the member of the minimal pair in the target BP sentence that had been heard (*Diga\_bem*/"Say \_\_ well") was voiceless [s] or voiced [z]. Optional pauses were provided after each 48 stimuli. Participants took from seven to 10 minutes to do the task.

Forty participants took part in the study. All of them had been taking the *Letras* (Languages and Linguistics) major at *Universidade Federal do Rio Grande do Sul (Porto Alegre-Brazil)*, but none of them were speakers of Spanish as an additional language. Participants who had some experience with accented speech in Brazilian Portuguese were also discarded from the study. Most participants spoke either French or English as an additional language.

### 2.3. Experiment 2

Although the results in our first experiment should be able to provide evidence to the combined role of vowel length and length of voicing in the fricative, it should be considered that Experiment 1 tested only those cases in which voicing in the fricative was totally present (100% of voicing) or absent (0%). Experiment 2 aimed to fill in this gap.<sup>8</sup>

For the manipulation of stimuli in Experiment 2, we took the same steps employed in Alves et al. (2018) and in Experiment 1. As shown in **Table 4**, we tested the perception of twelve different manipulated patterns, having "fricative voicing" (tested in three levels: 25%, 50% and 75% of voicing) and "vowel length" (tested in four levels: 25%, 50%, 75% and 100% of the total length of the vowel) as the main factors investigated.

<sup>8</sup> It is undeniably true that manipulating all experimental conditions in one single experiment would have allowed us to see the full interaction between the two variables in all their possible combinations. However, when Experiment 1 was run, we hypothesized that vowel length effects would have been found in both 0% and 100% of voicing in the fricative. This was not confirmed, since vowel length effects were not found in those conditions with full voicing of the fricative, as we will show in the Results section. These results motivated our second experiment (which was indeed built after the results obtained from Experiment 1), in which we tested voicing length effects in different degrees of voicing in the fricative (25%, 50% and 75%). Moreover, a single experiment would add up to 360 tokens (144 from Experiment 1 and 216 from Experiment 2), which would make the task too long and tiring, having possible undesirable effects on the results. Despite this limitation, after presenting the data of each individual experiment in the Results section, we present the results of a 5 × 4 Mixed ANOVA that considers the full interaction of all the levels of the two variables.

**Table 4:** Manipulated patterns tested in Experiment 2.

<u>% voicing_% vowel length</u>
25% voicing_25% vowel length
25% voicing_50% vowel length
25% voicing_75% vowel length
25% voicing_100% vowel length
50% voicing_25% vowel length
50% voicing_50% vowel length
50% voicing_75% vowel length
50% voicing_100% vowel length
75% voicing_25% vowel length
75% voicing_50% vowel length
75% voicing_75% vowel length
75% voicing_100% vowel length

Considering the speakers/minimal pairs presented in **Table 2**, there were also six different audio files for each one of the patterns tested. Each one of these audio files was randomly repeated three times, which adds up to 18 tokens in each experimental condition. Experiment 2, thus, totaled 216 tokens.

The same 40 participants who had taken part in Experiment 1 did the task in Experiment 2. The experiment was run two weeks after we had administered the first task. Experiment 2 was also run as an identification task on TP Software (Rauber et al. 2012). Once again, participants were asked to choose if the member of the minimal pair in the target sentence heard (*Diga\_bem*/"Say \_\_ well") was voiceless [s] or voiced [z]. Optional pauses were provided after each 58 stimuli. All participants took from 10 to 15 minutes to do the task.

### 3. Results

#### 3.1. Experiment 1

As described in the Method, in Experiment 1 "fricative voicing" was tested in two levels (100% or 0% of voicing) and "vowel length" was tested in four (25%, 50%, 75% and 100% of the total length of the vowel). **Table 5** presents the identification rates of [z] in the eight conditions investigated.

The averages described in **Table 5** show a remarkable difference in the identification of the fricatives as [z] between stimuli with and without voicing. As for the rate of duration in the preceding vowel, there seems to be a somewhat linear growth in the stimuli with 0% voicing as vowel duration becomes larger. As for the stimuli with fully voiced fricatives, the highest rate is found with 100% of the total duration of the preceding vowel. In turn, stimuli with the lowest vowel duration (25%) show the lowest rates among those with a fully voiced consonant.

In our inferential analyses (run on IBM SPSS – Version 21), we took the identification rates of [z] as the dependent variable. We ran a  $2 \times 4$  Mixed ANOVA (two levels of "fricative voicing": 0%; 100%; four levels of "vowel length": 25%, 50%, 75% and 100% of the total duration of the vowel – 220 ms). This test showed significant main effects of "fricative voicing" ( $F(1,39) = 285.26, p = 0.00$ ) and "vowel length" ( $F(3,117) = 11.23, p = 0.00$ ). As for the differences among the four levels of vowel length, post-hoc Bonferroni tests showed significant differences between 25% and 75% ( $p = 0.02$ ), 25% and 100% ( $p = 0.00$ ), 50% and 100% ( $p = 0.00$ ) and 75% and 100% ( $p = 0.02$ ). In other words,

when compared to manipulated lengths, full vowel length showed significant differences in the identification of [z].

There was also an interaction effect between “fricative voicing” and “vowel length” ( $F(3,117) = 5.51, p = 0.00$ ). These results are suggestive that both the length of the preceding vowel and the length of voicing in the fricative play a role in the identification of the consonant as voiced ([z]). As for the stimuli with no voicing in the fricative, we ran a Repeated-Measures ANOVA ( $F(3,117) = 11.55, p = 0.00$ ). The results of the post-hoc Bonferroni tests are shown in **Table 6**.

As seen in **Table 6**, in those conditions with 0% of voicing, the shortest length of the vowel (25%) allowed for significantly lower results in the identification of the consonant as [z], when compared to all the other vowel lengths. A comparison between 50% and 100% of vowel length also showed significance. The only comparisons in which significant differences had not been found were in those conditions in which vowel length differed between 50% and 75% or 75% and 100%. These results suggest that, mainly in those cases in which voicing does not seem to provide enough acoustic-articulatory information to allow for the identification of the consonant as voiced, longer vowel durations seem to contribute to an identification of the consonant as voiced [z].

As for the differences among the tested conditions with full voicing in the fricative, we ran a Repeated-Measures ANOVA ( $F(3,117) = 2.32, p = 0.08$ ) which showed marginally significant results. Given these results, we decided to run post-hoc Bonferroni tests that revealed significant differences between “100% voicing\_75% vowel length” and “100%

**Table 5:** Identification rates of [z] in each of the patterns tested (number of tokens per pattern = 18; number of participants = 40).

% voicing in [z]_% vowel length	Mean	Standard Deviation
0% voicing_25% vowel length	1.25	2.81
0% voicing_50% vowel length	1.68	2.78
0% voicing_75% vowel length	2.10	3.17
0% voicing_100% vowel length	2.55	3.00
100% voicing_25% vowel length	16.60	2.96
100% voicing_50% vowel length	16.73	3.17
100% voicing_75% vowel length	16.65	3.08
100% voicing_100% vowel length	16.98	2.97

**Table 6:** Post-hoc Bonferroni test results (significant results only) among the test conditions that presented 0% of voicing.

Post-hoc comparisons	p value
“0% voicing_25% vowel length” and “0% voicing_50% vowel length”	<b>0.02</b>
“0% voicing_25% vowel length” and “0% voicing_75% vowel length”	<b>0.01</b>
“0% voicing_25% vowel length” and “0% voicing_100% vowel length”	<b>0.00</b>
“0% voicing_50% vowel length” and “0% voicing_100% vowel length”	<b>0.01</b>

voicing\_100% vowel length” ( $p = 0.031$ )<sup>9</sup> only. These results suggest that, when voicing extends all the way through the consonant, the role played by vowel length does not seem to be as pivotal as in the case when voicing is not long enough to account for the difference between [s] and [z]. It is also important to mention that the stimuli with full voicing and full vowel length showed the highest rates of identification of [z] in the data presented in **Table 5** (with an average of 16.98 out of 18), providing additional evidence to the important role of both cues when acting in combination.

In summary, the data in Experiment 1 suggest that the length of the preceding vowel also plays a role in the identification of syllable-initial [s] or [z], especially in those cases in which voicing does not go all the way through the fricative, which tends to be the case of learners of BP whose L1 is Spanish (cf. Akerberg 2004; Alves 2015a; Oliveira 2016; Silveira & Souza 2011; Sobral, Nobre & Freitas 2006, among others). Though these results provide important information on the combined role of these cues, the data in Experiment 1, however, only account for those cases in which voicing is either fully present or totally absent in the fricative. Following a more dynamic account, in Experiment 2 we present the identification rates of [z] in those conditions in which voicing does not go all the way through the consonant.

### 3.2. Experiment 2

As explained in the Method, in Experiment 2 the tested conditions with three different degrees of voicing in the fricative (25%, 50% and 75%) were combined with the four levels of vowel length employed in the previous experiment (25%, 50%, 75%, 100%). **Table 7** presents the descriptive data regarding the identification of the consonant as [z].

As previously seen in **Table 5**, the identification rates of [z] increase as the rate of fricative voicing becomes larger. The descriptive data in **Table 7** also show that there is a gradient increase in the identification rates in view of vowel duration, when the length of fricative voicing is the same.

We ran a  $3 \times 4$  Mixed ANOVA (three levels of “fricative voicing”: 25%, 50% and 75% and four levels of “vowel length”: 25%, 50%, 75% and 100% of the total duration of the vowel – 220 ms). Once again, we found significant main effects of “fricative voicing” ( $F(2,78) = 201.76, p = 0.00$ ). It is important to mention that post-hoc Bonferroni tests showed significant differences among all conditions (“25%–50%”,  $p = 0.00$ , “25%–75%”,  $p = 0.00$ , “25%–100%”,  $p = 0.00$ , “50%–75%”,  $p = 0.00$ , “50%–100%”,  $p = 0.00$ , “75%–100%”,  $p = 0.00$ ). This makes it clear that voicing in the fricative is a decisive cue in the identification of the consonants as voiceless or voiced, as a longer voicing length entails more identifications of the consonant as [z].

We also found main effects of “vowel duration” ( $F(3,117) = 93.35, p = 0.00$ ). Post-hoc Bonferroni tests also showed significant differences among all four levels of vowel length (“25%–50%”, “25%–75%”, “25%–100%”, “50%–75%”, “50%–100%”, “75%–100%”), all of them with a  $p$  value of 0.00. In other words, when its effects are considered in isolation, vowel length is also a decisive cue in identification, with longer vowels accounting for higher rates of [z].

Finally, as for the interaction between “fricative voicing” and “vowel length”, our ANOVA also revealed significant results ( $F(6,234) = 11.91, p = 0.00$ ). These results

<sup>9</sup> It should also be mentioned that, in all post-hoc tests in which patterns with full voicing were compared to conditions exhibiting no voicing, there were significant differences among the tested conditions. This suggests that, all else being equal, full voicing may be solely responsible for the identification of the fricative as [z] (especially when contrasted with no voicing at all, as in Experiment 1). However, as already mentioned, full voicing does not seem to be the main production pattern produced by L1 Spanish learners of BP. This empirical fact motivated the proposal of the two experiments in this article.

once again reinforce the combined role of preceding vowel length and fricative voicing in the identifications of [z]. Given this interaction, we investigated possible differences among those test conditions with the same degree of voicing, in order to further discuss the role of vowel length. As for the differences among the tested conditions with 25% of voicing in the fricative (25% voicing\_25% vowel length, 25% voicing\_50% vowel length, 25% voicing\_75% vowel length, 25% voicing\_100% vowel length), a Repeated-Measures ANOVA ( $F(3,117) = 44.54, p = 0.00$ ) followed by Bonferroni tests revealed significant differences among all pairs of conditions (all of them showing a  $p$  value of 0.00). As for the differences among the tested conditions with 50% of voicing in the fricative (50% voicing\_25% vowel length, 50% voicing\_50% vowel length, 50% voicing\_75% vowel length, 50% voicing\_100% vowel length), a Repeated-Measures ANOVA ( $F(3,117) = 58.44, p = 0.00$ ) followed by Bonferroni tests also revealed significant differences among all pairs of conditions (once again, all of them showing a  $p$  value of  $p = 0.00$ ).

As for the differences among the tested conditions with 75% of voicing in the fricative, a Repeated-Measures ANOVA ( $F(3,117) = 8.31, p = 0.00$ ) also showed significant results. However, the post-hoc Bonferroni tests did not reveal significant differences among all conditions. **Table 8** presents the results of those comparisons showing significant results.

As can be seen in **Table 8**, from the six pairwise comparisons carried out, only two of them showed significant results. These results are different from the comparisons with the conditions that presented 25% or 50% of voicing, in which all post-hoc comparisons were significant. These results show the decisive role of fricative voicing and once again suggest that vowel length plays a significant role especially when the length of voicing

**Table 7:** Identification rates of [z] in each of the patterns tested (number of tokens per pattern = 18; number of participants = 40).

% voicing in [z]_ % vowel length	Average	Standard Deviation
25% voicing_25% vowel length	2.03	2.81
25% voicing_50% vowel length	3.50	3.33
25% voicing_75% vowel length	4.78	3.59
25% voicing_100% vowel length	6.93	3.91
50% voicing_25% vowel length	8.65	3.25
50% voicing_50% vowel length	10.28	3.45
50% voicing_75% vowel length	11.68	3.50
50% voicing_100% vowel length	13.68	3.02
75% voicing_25% vowel length	15.38	2.78
75% voicing_50% vowel length	15.95	2.77
75% voicing_75% vowel length	16.05	2.83
75% voicing_100% vowel length	16.77	2.96

**Table 8:** Post-hoc Bonferroni test results (significant results only) among the test conditions that presented 75% of voicing.

Post-hoc comparisons	p value
“75% voicing_25% vowel length” and “75% voicing_100% vowel length”	<b>0.00</b>
“75% voicing_75% vowel length” and “75% voicing_100% vowel length”	<b>0.00</b>

in the fricative is shorter than the consonant itself (which is particularly the case of L1 Spanish learners of BP, who find it difficult to voice the fricative all the way through).

As a final comparison, in order to highlight the paramount role of voicing in the participants' identification process, we also tested differences among all the conditions with the same degree of vowel length, with all of them showing significant results (25% ( $F(2,78) = 218.07, p = 0.00$ ), 50% ( $F(2,78) = 190.23, p = 0.00$ ), 75% ( $F(2,78) = 135.45, p = 0.00$ ) and 100% ( $F(2,78) = 1014.63, p = 0.00$ ). Post-hoc Bonferroni tests showed significant differences between all the tested pairs (with a  $p$  value of 0.00 for all pairwise comparisons). These results reinforce the fundamental role of fricative voicing: when two different degrees of voicing in the fricative are tested, there is always a significant difference due to the higher identification rates in the pattern with longer fricative voicing, regardless of vowel length.

This result should not be seen as surprising, as syllable-initial [z] in BP tends to be voiced all the way through by native speakers; vowel length, then, should play the role of a secondary cue in perception. However, when there is no full voicing in the fricative (which is the case of the manipulated stimuli in this study and in real productions by L1 Spanish learners of BP), vowel length may also be used as a cue by native listeners of BP.

### 3.3. Data from Experiments 1 and 2 taken together

Finally, in order to see the full interaction between the two variables in all the possible combinations, we ran a  $5 \times 4$  Mixed ANOVA (five levels of "fricative voicing": 0%, 25%, 50%, 75% and 100%; four levels of "vowel length": 25%, 50%, 75% and 100% of the total duration of the vowel – 220 ms), with the results obtained from the two experiments together. This follow-up test was possible as the participants in the two experiments were the same, and the time lag between the two tests was not longer than two weeks. As seen in the previous tests, we found significant main effects of "fricative voicing" ( $F(4,156) = 226.00, p = 0.00$ ). Once again, post-hoc Bonferroni tests showed significant differences among all conditions ("0%–25%",  $p = 0.00$ , "0%–50%",  $p = 0.00$ , "0%–75%", "0%–100%",  $p = 0.00$ , "25%–50%",  $p = 0.00$ , "25%–75%",  $p = 0.00$ , "25%–100%",  $p = 0.00$ , "50%–75%",  $p = 0.00$ , "50%–100%",  $p = 0.00$ , "75%–100%",  $p = 0.01$ ).

We also found a main effect of "vowel length" ( $F(3,117) = 97.14, p = 0.00$ ). In the same fashion as in the previous experiments, post-hoc Bonferroni tests also showed significant differences among all four levels of vowel length ("25%–50%", "25%–75%", "25%–100%", "50%–75%", "50%–100%", "75%–100%"), all of them with a  $p$  value of 0.00. This once again confirms the decisive role played by vowel length in the identification of fricative voicing.

Our  $5 \times 4$  Mixed ANOVA also revealed an interaction between "fricative voicing" and "vowel length" ( $F(12,468) = 18.46, p = 0.00$ ). As for the post-hoc tests, the results of the comparisons among the test conditions with the same degree of voicing have already been shown in Experiments 1 and 2. We then investigated possible differences among each one of the five test conditions with the same vowel length: the results of all Repeated-Measures ANOVAS (25% of vowel length:  $F(4,156) = 236.91, p = 0.00$ ; 50%:  $F(4,156) = 208.89, p = 0.00$ ; 75%:  $F(4,156) = 159.42, p = 0.00$ , 100%:  $F(4,156) = 158.88, p = 0.00$ ), followed by Bonferroni tests, showed significant differences among all pairs of conditions (all of them with a  $p$  value of 0.00), except between 75% and 100% of voicing in the fricative when vowel length was 50%, 75% or 100% of the total duration of 220 ms. It therefore seems that a voicing length of either 75% or 100% does not play a decisive role when vowel length is at least 50%, but it does play a role when the duration of the preceding vowel is shorter. Taken isolatedly or in combination, the data from both experiments therefore suggest that one of the cues ("vowel length" or "fricative voicing") seems to compensate the other when one of them does not seem to be long enough.

In sum, running all the data from the two experiments together seems to confirm what had previously been shown in the individual analyses of Experiments 1 and 2: both the duration of voicing in the fricative and the length of the preceding vowel interact and play a role in the distinction of /s/ – /z/ by Brazilian Portuguese hearers. These results are especially important with regard to the perception of accented L2 speech, especially among L1 Spanish speakers, who might find it difficult to fully voice the fricatives in Brazilian Portuguese.

#### 4. Discussion

The results in the previous section have shown an interaction between “fricative voicing” and “vowel length” in the identification of BP [z] by Brazilian listeners. Our analyses also suggest that an important role of vowel length may be found when voicing in the fricative is partial. These findings have both theoretical and pedagogical implications, as we will discuss in what follows.

##### 4.1. Implications for Laboratory Phonology

As for the theoretical implications for the field of Laboratory Phonology, besides providing additional evidence to the gradient status of speech, be it in the L1 (Albano 2001; Albano 2012; Albano 2017; Berti & Ferreira-Gonçalves 2012; Meneses & Albano 2015) or in the L2 (Alves 2018a; Alves et al. 2018; Silva 2014; Zimmer & Alves 2012), the study reinforces the combined action of different cues in the identification of phonological distinctions. Given this gradient status, it is important to discuss whether the current perceptual models prove adequate to account for our data. In order to do so, it is vital to discuss the perceptual primitives assumed by each one of these models.

As mentioned in the Introduction, the two perceptual models that have explained L2 perceptual phenomena in the last two decades are Flege’s (1995, 2003) SLM and Best & Tyler’s (2007) PAM-L2. Whereas the former assumes a psychoacoustic approach (and is therefore subject to “The Constancy Problem”, cf. Strange 1995), the latter assumes a Direct Realist approach so as to adopt phonological gestures as the perceptual units.

As an alternative account, Perozzo (2017) revisits some of the tenets proposed in Best & Tyler (2007). Among the main differences between the PAM-L2 and his proposal, Perozzo (2017) defends an Indirect Realist (Jackson 1977; Jackson 2010) approach to perception. According to the author, such an indirect approach proves more adequate when accounting for abstract, phonological entities in speech perception; at the same time, the realist account also allows for a dynamic primitive, without the need to resort to a distinction between underlying and surface forms.

As for the primitive involved in speech perception, Perozzo (2017) defends the adoption of acoustic-articulatory gestures (Albano 2001). According to Perozzo (2017), the adoption of this primitive would prove more epistemologically coherent in accommodating speech perception data, which, given their multimodality, needs a representational unit that accounts for both acoustic and articulatory information.<sup>10</sup>

Following Perozzo (2017), in this paper we state that an acoustic-articulatory gesture (Albano 2001) proves a more appropriate primitive in accounting for the process of sound assimilation<sup>11</sup> in our data. Firstly, assuming acoustic-articulatory gestures proves more

<sup>10</sup>For additional reasons that lead the author to adopt acoustic-articulatory gestures as the primitive unit of L2 perception, see Perozzo (2017).

<sup>11</sup>It should be said that, in Perozzo’s (2017) proposal, the author is dealing with the assimilation of target L2 sounds by foreign learners. In this paper, we extend Perozzo’s proposal to account for cases of perception of L2 accented speech by native speakers of Brazilian Portuguese. We argue that the same premises of sound assimilation portrayed in L2 assimilation models can be used to account for the perception of L2-accented speech as well as for the speech perception/assimilation processes of other dialects of the same language, as previously proposed in studies by Shaw et al. (2018).

appropriate when considering the acoustic stimuli provided in our study; indeed, the adoption of an acoustic-articulatory primitive seems to bridge the gap between the two aforementioned L2 perceptual models (the acoustic primitive that guides the SLM and the articulatory gesture adopted in the PAM-L2). Secondly, Albano's proposal proves more adequate to connect the processes of language development and language attainment. Devoting a considerable part of her work to the process of language development, Albano (2001) states that gestures are pre-linguistic units, and gestural constellations are formed as individuals learn to orchestrate time by engaging in such orchestrations ("learning by doing it", as stated by the author).<sup>12</sup> In this sense, the discussion on gestural borders promoted in Albano (2001) seems to be of paramount importance in understanding our data, as we have shown that the identification of the voiced or voiceless counterpart of the tested minimal pairs depends on the timing relations established between the acoustic-articulatory cues of "fricative voicing" and "vowel length", which interact with one another. Finally, the adoption of an acoustic-articulatory gesture also proves relevant due to its intrinsic connection to not only a gestural grammar itself, but also other modules of grammar, as proposed in Albano (2001).

It is regarding this latter aspect that our data pose additional research challenges to the perceptual model proposed in Perozzo (2017). As seen in the previous section, the effects of vowel length are particularly felt when voicing throughout the fricative is not long enough to account for the identification of [z]; however, in cases in which the fricative is voiced all the way through, vowel length does not assume a priority role in the voicing distinctions of the following consonant. In other words, fricative voicing works as a more weighed cue in this distinction; should it not be enough, less-weighed vowel length may play its complementary role in an interaction between these two cues.<sup>13</sup>

The role of cue weighting on speech perception has been long recognized in the L2 field (Flege 1995; Flege 2003), as theoretical models aiming at this phenomenon have played an important role in the last few years (Holt & Lotto 2006; Lehet & Holt 2017). It should be mentioned, however, that these proposals tend to assume acoustic features as main primitives, once again bringing the discussion on The Constancy Problem (cf. Strange 1995) into light. In accordance with Perozzo (2017), we believe the adoption of an acoustic-articulatory primitive for speech perception might be an epistemologically coherent means of dealing with cue competition.<sup>14</sup> In an attempt to suggest possible paths for future research, in this paper we suggest that acoustic-articulatory cue weighting in perception can be modelled through a theory of grammar. This proves coherent as long as we adopt Albano's acoustic-articulatory gestures as phonological primitives, not only due to the model's commitment to the process of language development, but also because Albano's proposal attempts to connect such primitives to a larger theory of grammar. In this sense, some contributions to the debate on the connection between acoustic-articulatory gestures and grammar theory have already been proposed in the Brazilian scenario (cf. Ferreira-Gonçalves 2017; Ferreira-Gonçalves & Alves 2013; Schmitt & Alves 2014). Even though we consider that such a formalization task still imposes several challenges (mainly concerning our aim not to fall back on a transformational grammar, which would "convert" an input into an output form), we believe the adoption of articulatory-gestures, besides

<sup>12</sup> This conception also has pedagogic implications for L2 teaching, as will be seen later in this section.

<sup>13</sup> As shown in the previous section, whenever fully voiced fricatives were compared to other manipulated patterns, there was a significant superiority in the identification of [z] in these consonants regardless of vowel length. Effects of vowel length tended to be found when the fricative was partially voiced.

<sup>14</sup> Indeed, this issue has been mentioned in Perozzo (2017) as one of the main theoretical challenges which remain to be faced by his model.

bridging the gap between “development” and “time orchestration”, may still allow for future discussions on what grammar is and how (or whether) it should be formalized.

#### **4.2. Implications for L2 Pedagogy**

The results presented in this paper, in conjunction with our dynamic account of speech perception, have pedagogical implications to the teaching of BP to foreign learners. These implications involve not only the definition of the syllabus, but also the methodological steps to be adopted in the teaching of L2 pronunciation.

As for the definition of the teaching syllabus, our findings provide relevant information to the teaching of minimal pairs such as *casa* /z/ – *caça* /s/ by showing that vowel length also plays a role in the intelligibility<sup>15</sup> of accented speech by BP listeners. As shown in the previous section, this cue may assume an important role especially when voicing in the fricative does not go all the way through the consonant. Our results suggest that, besides the teaching of syllable-initial fricative voicing itself, explicit teaching of vowel length should also be pertinent in the L2 BP classroom. Especially with those learners who find it more difficult to voice the fricative, teachers might find it useful to provide explicit instruction, as well as a more extensive practice, on the acoustic-articulatory cue of vowel length. Based on our findings, we therefore suggest that both preceding vowel length and syllable-initial fricative voicing be considered in an integrated fashion in the teaching of these minimal pairs. We consider this recommendation to be innovative, as, to our knowledge, pronunciation manuals on L2 Brazilian Portuguese (Almeida, Marcato & Roos 2013; Alves, Brisolara & Perozzo 2017; Roos 2010, among others) tend to focus on consonant voicing only.

When referring to the addition of vowel length to the syllabus, we advocate that the teaching of L2 pronunciation should not aim at a native pattern, but at intelligibility (Derwing & Munro 2015a; Derwing & Munro 2015b; Munro & Derwing 2015, among many others). According to this “intelligibility principle” (Levis 2005; Levis 2018), a foreign pronunciation does not need to fully reproduce a native pattern. This considered, in the case of the minimal pairs investigated in this paper, not voicing the fricative all the way through should not be regarded as a problem. This is particularly true as long as other cues, such as vowel length, can be integrated in order to account for the functional distinctions found in the additional language.

We consider this functional approach to L2 pronunciation and teaching to be in conformity with the dynamic view on language development that grounds the present paper. As already discussed, under this account, building an L2 phonology implies learning to orchestrate time (Zimmer & Alves 2012). This is possible as long as we adopt a phonological primitive with intrinsic timing (cf. Fowler 1980), such as the acoustic-articulatory gesture (Albano 2001). The adoption of such a primitive (and hence the conception of language development it is grounded on) also imposes several challenges to the language classroom, as we consider that the teaching of foreign languages must be closely connected to a view on language development.

The adoption of an approach to language development that states that we “learn how to do by doing it” may bring improvements to the L2 pronunciation classroom, traditionally

<sup>15</sup>As highlighted in Alves (2015b), a clear-cut definition to the construct of “intelligibility” is still elusive nowadays, as it may vary according to the authors that propose it. We believe that the sound identification task we carried out in this experiment may provide insights to what Munro & Derwing (2015) call “local intelligibility”, related to the understanding of shorter discourse units such as lexical items, for example. We believe, however, that a more appropriate definition of intelligibility, following a more dynamic account of language, is still under construction. As for an ecological and a dynamic characterization of intelligibility, see Alves (2018b) and Albuquerque (2019), respectively.

characterized by a more mechanical approach. In a dynamic approach to language teaching, in order to learn how to both perceive (others and themselves) and produce language, learners have to engage in communicative challenges (Kupske & Alves 2017). This contrasts with the so-called “traditional approach” to L2 pronunciation teaching, characterized by mechanical practices and drills, with no engagement between form and function. In other words, as we consider that we learn how to orchestrate gestures based on a communicative need and within a communicative environment, a dynamic approach to speech perception and production (and broadly speaking, to grammar itself) may be beneficial not only to Theoretical Linguistics, but also to the L2 language classroom.

## 5. Conclusion

In this paper, we investigated the combined role of fricative voicing and length of the preceding vowel in the identification of members of minimal pairs such as *rasa* (‘shallow’) – *raça* (‘breed’). This investigation proves relevant since learners of BP Portuguese whose L1 is Spanish find it difficult to produce fricative voicing, which might lead to intelligibility problems concerning these minimal pairs. By highlighting the importance of vowel length among Brazilian listeners, we hope that the results of this paper might provide the L2 classroom with suggestions on syllabus design in order to make BP accented speech more intelligible.

By conceiving a dynamic view on language (Albano 2001; Alves 2018a; Fowler 1980; Zimmer & Alves 2012) and on L2 perceptual phenomena (Perozzo 2017), we believe our findings have also posed theoretical challenges to a gestural account of speech (Albano 2001; Browman & Goldstein 1986; Browman & Goldstein 1992; Perozzo 2017), especially in what concerns the formalization of (or at least a theoretical explanation on) the role of primary and secondary acoustic-articulatory cues in speech perception. Although the phenomenon of cue weighting concerning acoustic primitives has been dealt with in many studies (Holt & Lotto 2006; Lehet & Holt 2017; Schultz, Francis & Llanos 2012), we believe it is important to investigate how a gestural account can account for this fact. By highlighting the role of vowel length as a secondary cue, whose effects can be felt especially in those cases in which voicing does not extend all the way through the fricative, our data might pave the way for future research on the formalization of this phenomenon.

It should also be mentioned that the results presented in this paper open new avenues for the discussion on vowel length as a secondary cue in voicing distinctions in Brazilian Portuguese. To our knowledge, this is the first study in which vowel length effects on the perception of voice distinctions have been reported in BP. In future studies, we aim to collect data on the relevance of vowel length in the perception of fricative voicing as produced by L1 BP speakers. As native speakers of BP tend to voice the fricative all the way through (Alves et al. 2018), possible effects of vowel length in the perception of native speech production may provide additional evidence in favor of this cue.

As in most experimental research, this study had to face several challenges and limitations. Firstly, one of its main limitations concerns the treatment of the L1 Spanish speakers who recorded the stimuli. As each one of the six speakers not only had come from a different country, but also exhibited a different length of stay/experience in Brazil, despite all our attempts to make sure that these differences did not affect the results (as described in 2.1), such differences seem to have played a role in the answers provided by the Brazilian participants. Given this possibility, as indicated by an anonymous reviewer, we reran the statistical analyses including “speaker” as an independent variable in the results of Experiments 1 and 2 taken individually, as well as in the results of the two experiments taken together. These tests revealed a main effect of “speaker” in each one of these scenarios,

as stimuli produced by Speaker 6 (see **Table 1**) seemed to show much greater averages than the other L1 Spanish speakers, while the stimuli produced by Speaker 2 tended to show much lower identification rates.<sup>16</sup> Interactions between “speaker” and “fricative voicing”, “speaker” and “vowel length” and also among the three independent variables were also found.<sup>17</sup> As in these new verifications an interaction between “fricative voicing” and “vowel length” was always found, although we can assure that the interaction among these two variables in voicing distinctions seems to be preserved, it is undeniable that “speaker” has played the role of a confounding variable in our data. By recognizing such a limitation, we once again highlight our difficulty in finding L1 Spanish participants who resided in Southern Brazil and presented a similar profile. At the same time, these verifications highlight the importance of a closer and more detailed inspection of individual participant differences in future studies, since their unique developmental trajectories may play a role in the results obtained.

Secondly, despite our attempt to reproduce patterns with different cue lengths, it is well true that we had to rely on very clear-cut portions of vowel and voicing lengths in our stimuli (by employing the percentages of 25%, 50%, 75% and 100%). Although we recognize that these pre-established patterns do not fully represent the dynamicity of language, we as well believe they have sufficed to account for the important roles played by the acoustic-articulatory cues investigated. Further studies with different rates of voicing in the fricative and a larger variety of vowel lengths are therefore necessary, in order for us to have a clearer picture of the time relations at stake in the perception of fricative voicing.

Other operational drawbacks might be associated with the length of the experiments and the group of participants who took part in the study. As for the length of the perceptual

<sup>16</sup> With regard to the data from Experiment 1, we ran a  $2 \times 4 \times 6$  Mixed ANOVA (two levels of “fricative voicing”: 0% and 100%; four levels of “vowel length”: 25%, 50%, 75% and 100% of the total vowel duration – 220 ms; and six levels concerning each one of the “speakers”). The results showed not only main effects of “fricative voicing” ( $F(1,39) = 291.32, p = 0.00$ ) and “vowel length” ( $F(3,117) = 11.29, p = 0.00$ ), but also a main effect of “speaker” ( $F(5,195) = 12.73, p = 0.00$ ). With regard to this latter variable, pairwise Bonferroni tests showed a difference between Speaker 6 and all other speakers ( $p = 0.00$  in all of the pairwise comparisons involving Speaker 6). As for the data from Experiment 2, our  $3 \times 4 \times 6$  Mixed ANOVA (three levels of “fricative voicing”: 25%, 50% and 75%; four levels of “vowel length”: 25%, 50%, 75% and 100% of the total vowel duration – 220 ms; and six levels concerning each one of the “speakers”) also revealed main effects of “fricative voicing” ( $F(2,78) = 206.41, p = 0.00$ ), “vowel length” ( $F(3,117) = 100.77, p = 0.00$ ) and “speaker” ( $F(5,195) = 25.14, p = 0.00$ ). As for “speaker”, pairwise Bonferroni tests showed significant differences between Speaker 2 (whose stimuli presented a lower identification average) and all other speakers ( $p = 0.00$  in all pairwise comparisons involving Speaker 2), between Speaker 6 (whose stimuli presented a lower identification range) and all other speakers ( $p = 0.00$  in all pairwise comparisons) and all other participants, and between “Speaker 1” and “Speaker 5” ( $p = 0.02$ ). Finally, when data from the two experiments were run together, our  $5 \times 4 \times 6$  Mixed ANOVA once again revealed main effects of “fricative voicing” ( $F(4,156) = 230.29, p = 0.00$ ), “vowel length” ( $F(3,117) = 103.49, p = 0.00$ ) and “speaker” ( $F(5,195) = 25.06, p = 0.00$ ). As for “speaker”, in a similar fashion to the results from Experiment 2, pairwise Bonferroni tests showed significant differences between Speaker 2 and all other speakers ( $p = 0.00$  in all pairwise comparisons involving Speaker 2, except for the comparison between “Speaker 2” and “Speaker 5”, in which  $p = 0.01$ ), between “Speaker 6” and all other speakers ( $p = 0.00$  in all pairwise comparisons), and also between “Speaker 1” and “Speaker 5” ( $p = 0.01$ ).

<sup>17</sup> In all the three follow-up Mixed ANOVAs that we ran (see the previous note for a detailed description of each of the tests), we found an interaction between “fricative voicing” and “speaker” (Experiment 1:  $F(5,195) = 15.76, p = 0.00$ ; Experiment 2:  $F(10,585) = 3.30, p = 0.00$ ; data from both experiments:  $F(20,780) = 18.54, p = 0.00$ ), an interaction between “vowel length” and “speaker” (Experiment 1:  $F(15,585) = 2.34, p = 0.00$ ; Experiment 2:  $F(15,585) = 3.30, p = 0.00$ ; data from both experiments:  $F(15,585) = 3.00, p = 0.00$ ), and an interaction among “fricative voicing”, “vowel length” and “speaker” (Experiment 1:  $F(15,585) = 2.44, p = 0.00$ ; Experiment 2:  $F(30,1170) = 2.94, p = 0.00$ ; data from both experiments:  $F(60,2340) = 2.0, p = 0.00$ ). These data confirm the status of “speaker” as a confounding variable. It should be said, however, that an interaction between “fricative voicing” and “vowel length” was also found in all of the tests (Experiment 1:  $F(3,117) = 5.191, p = 0.00$ ; Experiment 2:  $F(6,234) = 11.95, p = 0.00$ ; data from both experiments:  $F(12,468) = 18.54, p = 0.00$ ). In other words, despite the confounding role of “speaker”, our original hypothesis that voicing distinctions depend on both fricative voicing and vowel length is confirmed.

experiments, the large number of tokens in each task (especially in the second one) might have made them a bit tiring. Even though stimuli were presented in a random order and participants were advised to take pauses after blocks of stimuli, we acknowledge that listeners might have gotten tired or bored when solving the tasks. With regard to the group of participants who took the study, it is important to mention that, even though they had no previous experience with Spanish or with accented speech in BP, all of them were taking the *Letras* (Linguistics and Literature) major. Therefore, the participants in our study are likely to show more awareness of consonant voicing. Future studies with students from other majors are therefore necessary in order to determine whether this has had an effect on our data.

Finally, given the growth of Portuguese as a Lingua Franca, we highlight the importance of investigating listeners with different L1 systems in future studies. This is in line with the assumption that intelligibility is not solely dependent on the speaker's role, but it is rather established by the interaction among both speakers and listeners (Albuquerque 2019; Derwing & Munro 2015a; Derwing & Munro 2015b; Munro & Derwing 2015). Therefore, the language systems of both, as well as their experiences with other languages, play a role in characterizing intelligible speech. This study, thus, is the starting point for a long-term project aiming to investigate the role of perceptual cues among listeners with different language backgrounds. We believe the empirical data gathered in these future experiments will not only provide relevant insights to the teaching of BP as an L2, but they might also contribute to a better understanding of different language systems and their perceptual responses, fostering further theoretical advances in a dynamic approach to language. That said, the results in this study, as well as the future investigations derived from this article, highlight the importance of carrying out experimental research to both the fields of Formal and Applied Linguistics.

### Competing Interests

The authors have no competing interests to declare.

### References

- Akerberg, M. A.** (2004). A importância da palavra escrita para a pronúncia [The importance of the written word to pronunciation]. In A. R. M. Simões, A. M. Carvalho & L. Weidemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 115–124). Campinas: Pontes Editores.
- Alarcos Llorach, E.** (1965). *Fonología española [Spanish Phonology]*. (4th ed). Madrid: Editorial Gredos.
- Albano, E. C.** (2001). *O gesto e suas bordas: Esboço da Fonologia Acústico-Articulatória do Português Brasileiro [The gesture and its borders: An outline of the acoustic-articulatory phonology of Brazilian Portuguese]*. Campinas: Mercado de Letras/São Paulo – Fapesp.
- Albano, E. C.** (2012). Uma introdução à dinâmica em fonologia [An introduction to the dynamics in phonology]. *Revista da ABRALIN*, 11(1), 1–30. DOI: <https://doi.org/10.5380/rabl.v11i1.32462>
- Albano, E. C.** (2017). Fonologia de Laboratório [Laboratory Phonology]. In D. Hora & C. L. B. Matzenauer (Eds.). *Fonologia, Fonologias: Uma introdução* (pp. 169–181). São Paulo: Contexto.
- Albuquerque, J. I. A.** (2019). *Caminhos dinâmicos em Inteligibilidade e Compreensibilidade de Línguas Adicionais: Um estudo longitudinal com dados de fala de haitianos aprendizes de Português Brasileiro [Dynamic paths of Intelligibility and Comprehensibility in additional languages: A longitudinal study on the speech of Haitian learners of Brazilian Portuguese]*. (Unpublished doctoral dissertation). Universidade Federal do Rio Grande do Sul, Brazil.

- Almeida, A., Marcato, L., & Roos, L.** (2013). *Fonética com música – Português Brasileiro para Hispanofalantes [Phonetics with songs – Brazilian Portuguese for Spanish speakers]*. Buenos Aires: LiBreAr.
- Alves, U. K.** (2015a). O ensino das fricativas coronais do Português (LE) para falantes do Espanhol (L1): Análise de livro didático [The teaching of the coronal fricatives of Portuguese (FL) to speakers of Spanish (L1): Analysis of a coursebook]. In R. Silveira & I. Emmel (Eds.), *Um retrato do Português como segunda língua: Ensino, aprendizagem e avaliação* (pp. 75–96). Campinas: Pontes Editores.
- Alves, U. K.** (2015b). Ensino de pronúncia na sala de aula de língua estrangeira: Questões de discussão a partir de uma concepção de língua como sistema adaptativo e complexo [Pronunciation teaching in the foreign language classroom: Discussion issues from a view of language as a complex, adaptive system]. *Versalete*, 3(5), 374–396.
- Alves, U. K.** (2018a). Teoria dos Sistemas Dinâmicos e desenvolvimento fonético-fonológico em uma nova língua [Dynamic Systems Theory and phonetic-phonological development in a new language]. In E. Ortiz-Preuss & I. Finger (Eds.), *A dinâmica do processamento bilíngue* (pp. 117–161). Campinas: Pontes Editores.
- Alves, U. K.** (2018b). Desafios e implicações de uma caracterização ecológica para o construto ‘Inteligibilidade da fala em Língua Estrangeira’: Reflexões preliminares [Challenges and implications from an ecological characterization to the construct of ‘Foreign Language Intelligibility’: Preliminary discussions]. *Revista de Letras*, 37(2), 58–70.
- Alves, U. K., Brisolara, L. B., & Perozzo, R. V.** (2017). *Curtindo os sons do Brasil – Fonética do Português do Brasil para Hispanofalantes [Enjoying the sounds of Brazil – Brazilian Portuguese phonetics to Spanish speakers]*. Lisboa: Lidel Edições Técnicas.
- Alves, U. K., Brisolara, L. B., Rosa, L. C., & Buske, A. C. S.** (2018). Efeitos da duração do vozeamento da fricativa [z] na identificação, por brasileiros, de pares mínimos produzidos por hispânicos: Insumos para uma discussão sobre inteligibilidade de fala estrangeira [Effects of voicing length in the production of [z] on the perceptual identification of minimal pairs by Brazilian listeners: A discussion on foreign language intelligibility]. *Diacrítica*, 32(2), 437–465. DOI: <https://doi.org/10.21814/diacritica.449>
- Alves, U. K., & Silva, A. H. P.** (2016). Implicações de uma perspectiva Realista Direta para o PAM-L2: Desafios teórico-metodológicos [Implications of a Direct Realist approach to PAM-L2: Theoretical-methodological challenges]. *Revista do Gel*, 13(1), 107–131. DOI: <https://doi.org/10.21165/gel.v13i1.837>
- Antoniou, M.** (2018). Speech perception. In F. Grosjean & K. Byers-Heinlen (Eds.), *The listening bilingual – Speech perception, comprehension, and bilingualism* (pp. 43–64). Oxford: WILEY Blackwell. DOI: <https://doi.org/10.1002/9781118835722.ch3>
- Berti, L., & Ferreira-Gonçalves, G.** (2012). A aquisição do contraste /t/ e /k/ sob a ótica dinâmica [The acquisition of the /t/ – /k/ contrast from a dynamic account]. *Revista da ABRALIN*, 11(1), 139–136. DOI: <https://doi.org/10.5380/rabl.v11i1.32465>
- Best, C. T.** (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonuim, MD: York Press.
- Best, C., & Tyler, M. D.** (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O. Bohn & M. J. Munro (Eds.), *Language experience in second language learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/llt.17.07bes>
- Browman, C., & Goldstein, L.** (1986). Towards an Articulatory Phonology. *Phonology Yearbook*, 3, 219–252. DOI: <https://doi.org/10.1017/S0952675700000658>

- Browman, C., & Goldstein, L.** (1992). Articulatory Phonology: an overview. *Phonetica*, 49, 155–180. DOI: <https://doi.org/10.1159/000261913>
- Chen, M.** (1970). Vowel length variation as a function of the voicing of consonant environment. *Phonetica*, 22, 129–159. DOI: <https://doi.org/10.1159/000259312>
- Chomsky, N., & Halle, M.** (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Cristófaros-Silva, T.** (2017). *Fonética e Fonologia do Português: Roteiro de estudos e guia de exercícios [Portuguese Phonetics and Phonology: study guide and exercise booklet]*. (11<sup>a</sup> edição). São Paulo: Contexto.
- Cuértara Priede, J.** (2004). *Fonética de la ciudad de México – Aportaciones desde las tecnologías del habla [Mexico City Phonetics – contributions from speech technology]*. (Unpublished master's thesis). Universidad Nacional Autónoma de México, Mexico.
- Derwing, T., & Munro, M.** (2015a). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/llt.42>
- Derwing, T., & Munro, M.** (2015b). The interface of teaching and research: What type of pronunciation instruction should L2 learners expect? In P. L. Luchini, M. A. García Jurado & U. K. Alves (Eds.), *Fonética y Fonología: Articulación entre enseñanza e investigación* (pp. 14–26). Mar del Plata: Editorial de la Universidad Nacional de Mar del Plata.
- Ernestus, M., & Baayen, H.** (2006). The functionality of incomplete neutralization in Dutch: the case of past-tense formation. *Laboratory Phonology*, 8, 27–49. DOI: <https://doi.org/10.1515/9783110197211.1.27>
- Escudero, P.** (2005). *Linguistic perception and second language acquisition*. (Unpublished doctoral dissertation). Utrecht University, The Netherlands.
- Escudero, P.** (2009). The linguistic perception of similar L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151–190). Berlin: Mouton de Gruyter.
- Ferreira-Gonçalves, G.** (2017). Fonologia Gestual e Teoria da Otimidade [Gestural Phonology and Optimality Theory]. *Gradus – Revista Brasileira de Fonologia de Laboratório*, 2(1), 73–82.
- Ferreira-Gonçalves, G., & Alves, U. K.** (2013). Os gestos em restrições: Fonologia Gestual e Teoria da Otimidade [Gestures as constraints: Gestural Phonology and Optimality Theory]. In G. Ferreira-Gonçalves & M. R. Brum-de-Paula (Eds.), *Dinâmica dos movimentos articulatorios: Sons, gestos, imagens* (pp. 37–65). Pelotas: Editora da UFPel.
- Flege, J. E.** (1995). Second Language Speech Learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E.** (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.
- Fowler, C.** (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113–133. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31446-9](https://doi.org/10.1016/S0095-4470(19)31446-9)
- Fowler, C.** (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3–28. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30607-2](https://doi.org/10.1016/S0095-4470(19)30607-2)
- Gibson, J. J.** (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

- Gibson, J. J.** (1979). *The ecological approach to visual perception*. London: Psychology Press.
- Holt, L. L., & Lotto, A. J.** (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119(5), 3059–3071. DOI: <https://doi.org/10.1121/1.2188377>
- Hualde, J. I., Olarrea, A., Escobar, A. M., & Trevis, C. E.** (2010). *Introducción a la lingüística hispánica [Introduction to Spanish linguistics]*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511808821>
- Jackson, F.** (1977). *Perception*. New York: Cambridge University Press.
- Jackson, F.** (2010). Representative realism. In J. Dancy, E. Sosa, & M. Steup. (Eds.), *A companion to epistemology* (2<sup>nd</sup> ed.) (pp. 702–705). Malden: Blackwell.
- Keating, P. A.** (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (Ed.), *Phonetic Linguistics – Essays in honor of Peter Ladefoged* (pp. 115–132). Orlando: Academic Press.
- Kenstowicz, M.** (1994). *Phonology in Generative Grammar*. Cambridge, MA: Blackwell.
- Kupske, F. F., & Alves, U. K.** (2017). Orquestrando o caos: O ensino de pronúncia de língua estrangeira à luz do paradigma da complexidade [Orchestrating chaos: Teaching foreign language pronunciation in the Complexity paradigm]. *Fórum Linguístico*, 14(4), 2771–2784. DOI: <https://doi.org/10.5007/1984-8412.2017v14n4p2771>
- Ladefoged, P., & Johson, K.** (2010). *A Course in Phonetics* (6<sup>th</sup> edition). Boston: Cengage Learning.
- Lehet, M., & Holt, L. L.** (2017). Dimension-based statistical learning affects both speech perception and production. *Cognitive Science*, 41(S4), 885–912. DOI: <https://doi.org/10.1111/cogs.12413>
- Levis, J. M.** (2005). Changing concepts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377. DOI: <https://doi.org/10.2307/3588485>
- Levis, J. M.** (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781108241564>
- Martínez Celdrán, E.** (2003). *El sonido en la comunicación humana: Introducción a la fonética [The sound in human communication: An introduction to Phonetics]*. (2a ed.). Barcelona: Ediciones Octaedro.
- Meneses, F., & Albano, E. C.** (2015). From reduction to apocope: Final poststressed vowel devoicing in Brazilian Portuguese. *Phonetica*, 72, 121–137. DOI: <https://doi.org/10.1159/000439599>
- Motta-Ávila, C.** (2017). *A formalização fonético-fonológica da percepção de plosivas surdas sob múltiplas manipulações de Voice Onset Time (VOT) por brasileiros e americanos à luz do modelo Biphon [Phonetic-phonological formalization of the perception, by Brazilian and North-American listeners, of voiceless stops with multiple manipulations of Voice Onset Time (VOT): A Bidirectional Phonology account]*. (Unpublished master's thesis). Universidade Católica de Pelotas, Brazil.
- Munro, M., & Derwing, T.** (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 377–396). Malden, MA: Wiley Blackwell. DOI: <https://doi.org/10.1002/9781118346952.ch21>
- Navarro Tomás, T.** (2004). *Manual de pronunciación española [Manual of Spanish pronunciation]*. Madrid: Consejo Superior de Investigaciones Científicas.
- Nishida, G.** (2012). *Sobre teorias de percepção de fala [On theories of speech perception]*. (Unpublished doctoral dissertation). Universidade Federal do Paraná, Brazil.

- Nishida, G.** (2014a). As bases acústica e articulatória das teorias de percepção da fala. [Acoustic and articulatory bases of the theories of speech perception]. *Revista do GEL*, 11(1), 142–167.
- Nishida, G.** (2014b). Primitivos de análise na percepção da fala [Primitives of analysis in speech perception]. In A. Brawerman-Albini & M. L. C. Gomes (Eds.), *O jeitinho brasileiro de falar inglês: Pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 49–63). Campinas: Pontes Editores.
- Oliveira, R. A.** (2016). *A influência da instrução fonética explícita na produção e na percepção dos fricativos sonoros por hispanofalantes aprendizes de Português como Língua Estrangeira – PLE [The influence of explicit phonetic instruction on the production and perception of voiced fricatives by L1 Spanish learners of Portuguese as a Foreign Language – PFL]*. (Unpublished master's thesis). Universidade Federal de Pernambuco, Brazil.
- Perissinotto, G.** (1975). *Fonología del español hablado en la Ciudad de México: Ensayo de un método sociolingüístico*. [Phonology of the Spanish spoken in Mexico City: An essay on a sociolinguistic method]. México: El Colegio de México.
- Perozzo, R. V.** (2017). *Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não-nativa: Para além do PAM-L2 [On the cognitive, acoustic-articulatory and indirect realist accounts in non-native phonic perception: Moving beyond PAM-L2]*. (Unpublished doctoral dissertation). Brazil: Universidade Federal do Rio Grande do Sul.
- Port, R., & O'Dell, M. L.** (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30797-1](https://doi.org/10.1016/S0095-4470(19)30797-1)
- RAE – Real Academia Española & Asociación de Academias de la Lengua Española.** (2011). *Nueva gramática de la lengua española: Fonética y Fonología*. [New Grammar of the Spanish language: Phonetics and Phonology]. Barcelona: Espasa Libros.
- Rauber, A. S., Rato, A., Kluge, D. C., & Santos, G. R.** (2012). *TP: perception tests and perceptual training with immediate feedback*. Retrieved March 15, 2019, from [http://www.worken.com.br/tp\\_regfree.php](http://www.worken.com.br/tp_regfree.php)
- Ribeiro, R. S.** (2017). *Duração de vogais tônicas antecedentes a consoantes plosivas no Português Brasileiro [Duration of stressed vowels preceding plosive consonants in Brazilian Portuguese]*. (Unpublished undergraduate dissertation). Universidade Federal do Rio Grande do Sul, Brazil.
- Roos, L.** (2010). *Fonética Lúdica – Brincando com os sons [Ludic Phonetics – Playing with sounds]*. Buenos Aires: Editora Sotaque.
- Schmitt, B. K., & Alves, U. K.** (2014). The acquisition of /p/ and /k/ word-mid codas of English (L2) by learners from Southern Brazil (L1): a gestural analysis in Stochastic Optimality Theory. *Letrônica*, 7(2), 765–794. DOI: <https://doi.org/10.15448/1984-4301.2014.2.17879>
- Schultz, A., Francis, A., & Llanos, F.** (2012). Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America*, 13, 95–101. DOI: <https://doi.org/10.1121/1.4736711>
- Seara, I. C., Nunes, V. G., & Lazzarotto-Volcão, C.** (2015). *Para conhecer Fonética e Fonologia do Português Brasileiro [Getting to know Brazilian Portuguese Phonetics and Phonology]*. São Paulo: Contexto.
- Shaw, J. A., Best, C. T., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., & Mulak, K. E.** (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology*, 9(1), 11–36. DOI: <https://doi.org/10.5334/labphon.87>
- Silva, A. H. P.** (2014). A variável tempo nos estudos em aquisição [The time variable in acquisition studies]. In A. Brawerman-Albini & M. L. C. Gomes (Eds.), *O jeitinho*

- brasileiro de falar inglês: Pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 33–48). Campinas: Pontes Editores.
- Silveira, R., & Souza, T. T.** (2011). A percepção e a produção das fricativas alveolares da língua portuguesa por hispano-falantes [The perception and production of alveolar fricatives in Brazilian Portuguese by Spanish speakers]. *Revista de Estudos Linguísticos*, 19(2), 167–184. DOI: <https://doi.org/10.17851/2237-2083.19.2.167-184>
- Sobral, C. S., Nobre, M. M. R., & Freitas, M. A.** (2006). Relação fone-fonema-grafema na produção oral de aprendizes de PLE [Phone-phoneme-grapheme relationships in the oral production of learners of Portuguese as a Foreign Language]. *Portuguese Language Journal*, 1, 1–18.
- Strange, W.** (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 3–45). Baltimore: York Press.
- Vaquero de Ramírez, M.** (2003). *El español de América I: pronunciación [Spanish in America I: Pronunciation]*. Madrid: Arco Libros.
- Zimmer, M. C., & Alves, U. K.** (2012). Uma visão dinâmica da produção da fala em L2: o caso da Dessonorização Terminal [A dynamic view on L2 speech production: The case of Terminal Devoicing]. *Revista da ABRALIN*, 11(1), 221–272. DOI: <https://doi.org/10.5380/rabl.v11i1.32467>

**How to cite this article:** Alves, U. K., & Brisolara, L. B. (2020). Listening to accented speech in Brazilian Portuguese: On the role of fricative voicing and vowel duration in the identification of /s/ – /z/ minimal pairs produced by speakers of L1 Spanish. *Journal of Portuguese Linguistics*, 19: 6, pp. 1–23. DOI: <https://doi.org/10.5334/jpl.237>

**Submitted:** 15 September 2019

**Accepted:** 31 March 2020

**Published:** 13 May 2020

**Copyright:** © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Journal of Portuguese Linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 