Word-final devoicing in Brazilian Portuguese and L2 English

Wellington Mendes Junior, Thaïs Cristófaro Silva

Federal Center for Technological Education of Minas Gerais, Federal University of Minas Gerais
wellington@cefetmg.br, thaiscristofaro@gmail.com

ABSTRACT

This study examines word-final devoicing in Brazilian Portuguese (BP) and L2 English. It aims at assessing whether an ongoing sound change related to devoicing in the L1 affects the production of L2 forms. We considered the production of plural suffixes comprised of a (stop + sibilant) cluster (e.g. *ringues* [hĩgs] ~ [hĩks] 'rings' in BP and *bags* [bægz] ~ [bæks] in English). Harmonics-to-noise ratio was used to measure the degree of sibilant voicing. Results showed that devoicing was significantly attested in both L1 and L2. This provides evidence that word-final devoicing, which is an ongoing sound change in the L1, affects the production of L2 forms. Moreover, L2 English sibilant voicing is influenced by the following phonological context, a pattern which also seems to be imported from the L1. We show that not only segments, but also fine phonetic detail is transferred to the L2 as suggested by Exemplars Model.

Keywords: Word-final devoicing; Brazilian Portuguese; L2 English; Exemplars Model.

1. INTRODUCTION

This paper examines word-final devoicing (WFD) in Brazilian Portuguese (e.g. *ringues* [hĩgs] ~ [hĩks] 'rings') and in English as a Second Language (e.g. *bags* [bægz] ~ [bæks]). It intends to contribute to the understanding of L2 phonology by examining how emerging sound patterns from the L1 are adopted in the L2. The analysis is grounded on the premises of Exemplars Model which propose that fine phonetic detail is part of phonological representations (c.f. [3], [4], [10]).

In BP, word-final devoicing has been associated with an ongoing sound change that favors the production of voiceless consonants word-finally (e.g. *ringues* [hĩgs] ~ [hĩks] 'rings'). This paper intends to be a contribution to examining the earlier stages of WFD in BP phonology, a phenomenon which occurred in several Indo-European languages (c.f. [2], [8], [9], [11]).

It has been noticed that unstressed high front vowels in BP are weakened and lost when flanked between a consonant and a word-final sibilant. For example, *lápis* ['la.pɪs] > [laps] 'pencil' (c.f. [12]). This pattern applies to plural forms, as in *ringues* ['hĩ.gɪs] > [hĩgs] ~ [hĩks] 'rings' and *cheques* ['ʃɛ.kɪs] > [ʃɛks] 'cheques'. The word-final sibilant is always voiceless in BP, regardless of the voice quality of the preceding consonant (c.f. [5]). Across morphemes or word-boundaries, a regressive assimilation rule leads the sibilant to be voiced when it is followed by a vowel or a voiced consonant: [hĩgs] *ringues* 'rings' > [hĩgz#a.ma.'ɾɛ.lʊs] *ringues amarelos* 'yellow rings' and [ʃɛks] *cheques* 'cheques' > [ʃɛkz#a.ma.'ɾɛ.lʊs] *cheques amarelos* 'yellow cheques'.

In this paper, we examine the ongoing sound change in BP which is related to the alternation between voiced and voiceless consonants (e.g. [hĩgz#a.ma.'ɾɛ.lʊs] ~ [hĩks#a.ma.'ɾɛ.lʊs]) and we also consider whether this sound change influences the production of L2 English plural forms (e.g. *bags* [bægz] ~ [bæks], *bags are* [bægz#ar] ~ [bæks#ar]).

In English, the regular plural suffix and 3rd person singular present is posited to be /z/ (c.f. [7]). A progressive assimilation rule predicts that if a vowel or a voiced consonant precedes /z/, the output is [z], as in *dogs* [dɔgz], *trees* [triːz] and *pies* [paɪz]. On the other hand, if a voiceless consonant precedes /z/, it surfaces as [s], as in *cups* [kʌps], *cats* [kæts] and *ducks* [dʌks]. Finally, if a sibilant or an affricate precedes the sibilant, the outcome is [ɪz], as in *buses* [bʌsɪz], *quizzes* ['kwɪz.ɪz] and *watches* ['wɔtʃ.ɪz]. Whereas in English sibilants are prone to progressive assimilation, in Brazilian Portuguese sibilants undergo regressive assimilation.

The main question we posit is whether WFD in PB influences the production of L2 English plural forms. If that is the case, we expect that English plural forms produced by BP speakers will favor voiceless consonants word-finally, as it is the more robust pattern in L1, inhibiting the English progressive assimilation rule that applies to plural forms (e.g. [bæks] instead of [bægz] *bags*). We also posit that voiced sibilants will be favored when followed by a word-initial vowel (e.g. [bægz#ar] *bags are*). This follows from the BP regressive assimilation rule which leads sibilants to be voiced when followed by a vowel.

In sum, we will investigate whether word-final devoicing takes place in both languages and

whether sibilant voicing can be influenced by the following phonological environment. An experiment was designed to test the production of (stop + sibilant) clusters in English spoken by Brazilian speakers. Harmonics-to-noise ratio (HNR) was used to measure the degree of sibilant voicing, aiming at offering an accurate evaluation of the experimental data carried out by acoustic analysis.

2. METHODOLOGY

A group of twenty Brazilians studying at the Federal Center for Technological Education of Minas Gerais took part in this study. All participants were high school students who had been taking English classes as part of the school's curriculum for at least one year. Half of the participants displayed basic levels of L2 English (A1 or A2) whereas the other half displayed pre-advanced (B2) to advanced (C1) levels, according to the Common European Framework of Reference for Languages.

A set of 36 plural nouns ending in a (stop + sibilant) cluster were considered in BP, e.g. *cheques* [ʃɛks] 'cheques' and *botes* [bɔts] 'boats'. For the L2 English case study, another set of 36 words were selected, e.g. *bags* [bægz] and *boats* [bəʊts].

The experiment was comprised of two trials. The first one involved a picture-counting task in which participants were asked to count and name the elements shown in the pictures. This would lead to the production of the target words in their plural forms (e.g. *two cakes*). In a second application of the picture-counting task, participants were asked to pronounce a short carrier-sentence following the sibilant (e.g. *two cakes are seen*). The goal was to assess whether they would voice the word-final sibilant due to the presence of a word-initial vowel in sequence. The second trial involved sentence reading. The number of sentence syllables, intonation patterns and words' parts of speech were controlled. The BP study took place after the L2 English study and considered the same phonological environments and task types.

Due to the COVID-19 pandemic, data collection was performed remotely. Experiments were recorded with the Open Broadcaster Software Studio at 48 kHz sampling rate. The obtained recordings were converted into WAVEform audio format by the software Adobe Premiere 2020, which was able to maintain the same sampling rate as the original files. The average time to complete the experiment was 1 hour and 20 minutes. A total of 2833 tokens were considered for the L2 English study. For the BP study, 2831 tokens were considered. Samples were edited and manually annotated using Praat TextGrids (c.f. [1]).

This research mainly considered the voice quality of word-final sibilants. In BP, only voiceless sibilants occur word-finally, unless a vowel follows it, to which a voiced sibilant emerges. In English, voiced and voiceless sibilants occur word-finally. When a vowel follows the sibilant, the voice quality remains as it formerly was (rather than changing as it occurs in BP). Thus, we posited that word-final voiceless sibilants would be favored in L2 English, as it is the more robust pattern in L1. We also posited that voiced sibilants occur at higher rates when followed by a word-initial vowel.

Voicing was measured under Harmonics-to-noise ratio, which represents the degree of acoustic periodicity. Each token was extracted to a separate sound object and a harmonicity object was created, from which the mean harmonicity was calculated, hereafter the HNR. The details of its calculation can be found in [1]. Based on the discussion from Praat's manual [1], higher values of HNR should correspond to higher voicing rates. This is consistent with prior research which has shown that voiceless intervocalic sibilants in English ([s] and [ʃ]) typically have HNR values below 5 dB, whereas voiced sibilants ([z] and [ʒ]) fall within the 5-10 dB range [13].

3. RESULTS

Results showed that Brazilian Portuguese speakers tend to present a voiceless sibilant word-finally *cheques* [ʃɛks#], even in contexts where a voiced counterpart would be expected. i.e., *cheques amarelos* [ʃɛkz#amaɾɛlʊs]. Consider Figure 1.

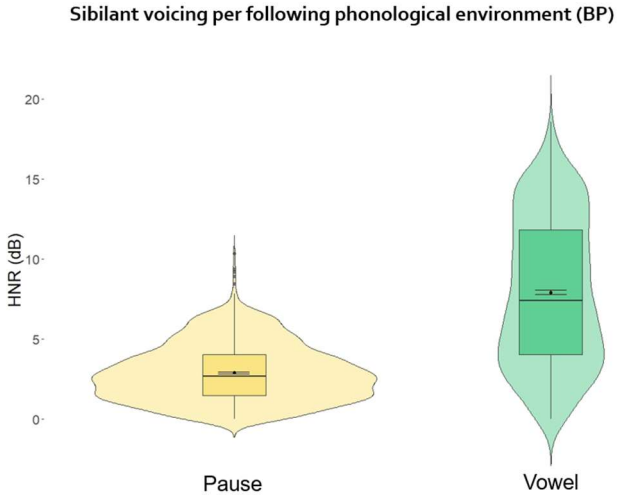

Figure 1: Sibilant voicing per following phonological environment in Brazilian Portuguese.

Figure 1 displays the rates of sibilant voicing per following phonological environment. Sibilants followed by a pause (left violin plot) were expected

to display low rates of HNR and sibilants followed by a vowel (right violin plot) were expected to be voiced, thus displaying higher rates of HNR.

As expected, lower voicing rates were attested when the sibilant was followed by a pause (e.g. *cheques* [ʃɛks#], mean = 2.8 dB), whereas higher voicing rates were attested when the sibilant was followed by a word-initial vowel (e.g. *cheques amarelos* [ʃɛkzamaɾɛlʊs], mean = 7,9 dB). Wilcoxon test results revealed that this factor is statistically significant (W = 340138, p-value < 0.01). We highlight the fact that sibilants followed by vowels showed considerable variability, with a standard deviation of 4.5. This indicates that such sibilants were produced with variable degrees of voicing, some of them with unexpectedly low rates of HNR, which indicates a tendency towards the production of [s]. Spectrographic data analysis corroborates this observation. Consider Figure 2.

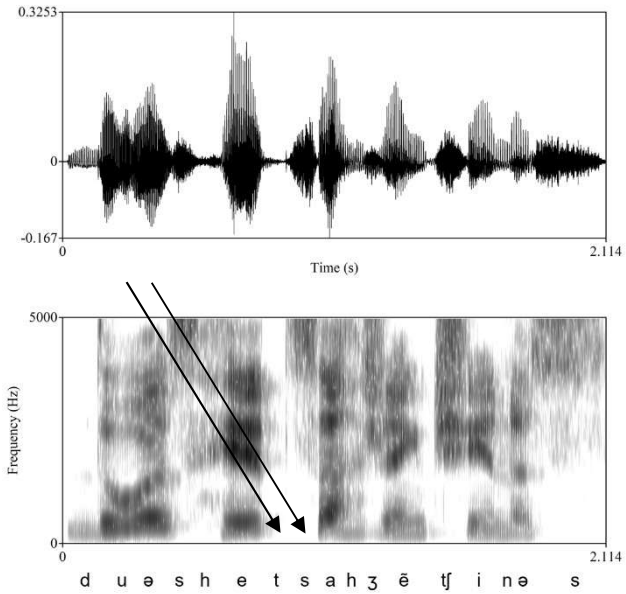

Figure 2: Spectrogram analysis of word-final devoicing in BP.

Figure 2 shows the spectrogram of the sentence *duas redes argentinas* 'two argentine hammocks', produced by a 17 year-old male participant. A voiced sibilant was expected as the following context is a vowel. However, two arrows indicate the absence of voicing during the production of the (stop + sibilant) sequence in the target word. If there was voicing, the bottommost part of the spectrogram would be dark instead of being white.

Additionally, Figure 2 shows that not only did the sibilant remain voiceless, but the preceding stop – which was formerly voiced - also shared the same property, being voiceless. Thus, [hets ah.ʒẽ.'ʧi.nəs] instead of [hedz ah.ʒẽ.'ʧi.nəs] was attested. Our data showed that word-final devoicing occurred with all cluster types followed by vowels assessed in this study ([ps, ts, ks, bs, ds, gs]), pointing out to an ongoing sound change in BP which triggers only voiceless consonants word-finally. This tendency was attested in BP even with words to which word-final devoicing leads to contrast loss: *grades* 'grids' and *grátis* 'free' [gɾats], *sedes* 'headquarters' and *setes* 'seven' [sɛts], *ringues* 'rings' e *rinques* [hĩks] 'rinks'.

It appears that word-final devoicing in BP might be associated with an Indo-European historical trend, as observed in languages such as Dutch [11], German [2], French [9] and Romanian [8]. This study thus contributes to examining the earlier stages of word-final devoicing in BP.

Let us now examine WFD in L2 English. An important question that arises is whether the ongoing sound change in BP related to word-final devoicing influences the production of L2 English plural forms. Consider Figure 3.

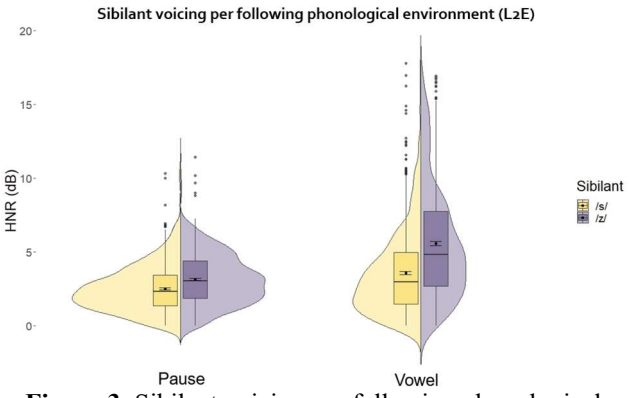

Figure 3: Sibilant voicing per following phonological environment in L2 English.

Figure 3 displays sibilant voicing rates per following phonological environment in L2 English by BP speakers. Sibilants followed by a pause (left violin plots) were expected to display low rates of HNR and sibilants followed by a vowel (right violin plots) were expected to be voiced, thus displaying higher rates of HNR. Yellow plots represent tokens whose voiceless sibilant [s] was expected (e.g. *cups* [kʌps] and *cats* [kæts]), whereas purple plots represent tokens whose voiced sibilant [z] (e.g. *jobs* [dʒɔbz], *beds* [bɛdz]) was expected. Results show that, as we hypothesized, lower voicing rates were attested when the sibilant was followed by a pause (e.g. *cups*, *jobs*) – with an average of 3.0 dB - and that occurred regardless of the consonant that precedes the sibilant. On the other hand, higher rates of voicing were attested when there was a vowel following the sibilant (e.g. *cups are*, *jobs are*) – with an average of 4.3 dB. Wilcoxon test results revealed that the following phonological

environment is statistically significant in L2 English, both within the group whose sibilant [s] is expected (W = 203088, p-value < 0, 01) and within the group whose sibilant [z] is expected (W = 155599, p-value < 0.01).

Notice that in L2 English voiceless consonants occur regardless of the preceding phonological context - which is actually what triggers voicing in L1 English due to progressive assimilation. The predominance of voiceless clusters in L2 English points out to the influence of L1 sound patterns for the construction of L2 phonological representations (c.f. [6]).

Moreover, when comparing Figures 1 and 3, we observe that learners transfer an allophonic pattern from L1 BP (i.e. regressive assimilation) to L2 English. Our results additionally suggest that L2 learners not only import allophonic sound patterns from the L1 – i.e. regressive assimilation -, but also emerging sound patterns – i.e. word-final devoicing. Once again, consider the data represented in purple in Fig. 3. These data involve words whose sibilant [z] was expected (e.g. [bægzɑ:r] *bags are*), not only due to progressive assimilation attested in English but also due to BP's regressive assimilation. Note that even though most sibilants followed by a vowel had higher rates of HNR, some tokens were associated with low rates (standard deviation = 3.5). The L2 English data in purple correspond to the BP data reported in Figure 1: in both languages the production of voiced clusters occurs only to some extent when there is a vowel following the sequence. This offers evidence that word-final devoicing, a phenomenon that still emerging in L1 BP, is also aggregated in L2 English. Consider Figure 4.

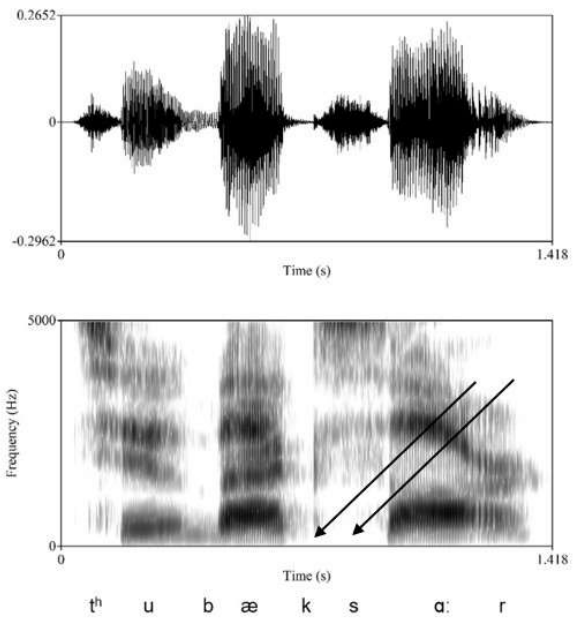

Figure 4: Spectrogram analysis of word-final devoicing in L2 English.

Figure 4 shows the spectrogram of the sentence "two bags are [...]", produced by the same participant depicted in Figure 2. The arrows indicate the absence of voicing in the final consonant cluster. It is clear that there was the production of [bæksɑ:r] instead of [bægzɑ:r], in which both the stop and the word-final sibilant were devoiced. Notice that WFD occurred with all voiced cluster types assessed in this study (i.e. [bz, dz, gz]), not only when the sibilant was followed by a pause (e.g. *two labs* [tu: læps]) but also when the sibilant was followed by a vowel (e.g. *two labs are* [tu: læps ɑ:r]). Alike our BP data, this was attested even in words which word-final devoicing leads to contrast loss: *coats* and *codes* [kəʊts], *sites* and *sides* [saɪts], *laps* and *labs* [læps] *kits* and *kids* [kɪts], *pigs* and *pix* [pɪks]. This suggests that not only segments, but also fine phonetic detail – associated with a phonological pattern that is still emerging in the L1 – is transferred to the L2.

4. CONCLUSIONS

Our results show that word-final devoicing is widely attested in both L1 BP and L2 English. This provides evidence that an ongoing sound change in the L1 affects the production of L2 forms. We also found that not only do the final sibilants devoice but also the preceding stops are also devoiced. Moreover, L2 English sibilant voicing is influenced by the following phonological context, a pattern which also seems to be imported from the L1. We argue that exemplars associated with an allophonic sound pattern (production of voiced clusters followed by vowels) are in competition with an emerging sound pattern (devoicing of word-final clusters). This offers evidence that fine phonetic detail is part of phonological representations, as posed by the Exemplars Model (c.f. [3, 4, 10]). Finally, this paper shows that word-final devoicing, an emergent phonological pattern in L1 BP, is also aggregated in English L2.

5. REFERENCES

[1] Boersma, P., & Weenink, D. (2020). Praat manual. *Amsterdam: University of Amsterdam, Phonetic Sciences Department*.

[2] Brockhaus, W. (2012). Final devoicing in the phonology of German. In *Final Devoicing in the Phonology of German*. Max Niemeyer Verlag.

[3] Bybee, J. (2001). Frequency effects on French liaison. *Typological studies in language*, *45*, 337-360.

[4] Bybee, J. (2008). Usage-based grammar and second language acquisition. In *Handbook of cognitive linguistics and second language acquisition* (pp. 226-246). Routledge.

[5] Cristófaro-Silva, T, Mendes-Jr., W. (2022). Plural formation in English: a Brazilian Portuguese case study. In: *Second Language Pronunciation: Different Approaches to Teaching and Training*. De Gruyter Mouton.

[6] Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *Second language speech learning: Theoretical and empirical progress*, 3-83.

[7] Hayes, B. (2011). *Introductory phonology*. John Wiley & Sons.

[8] Hutin, M., Niculescu, O., Vasilescu, I., Lamel, L., & Adda-Decker, M. (2020, maio). Lenição e fortificação de stop codas em romeno. In *SLTU-CCURL*.

[9] Jatteau, A., Vasilescu, I., Lamel, L., Adda-Decker, M., & Audibert, N. (2019, September). " Gra [f] e!" Word-Final Devoicing of Obstruents in Standard French: An Acoustic Study Based on Large Corpora. In *INTERSPEECH* (pp. 1726-1730).

[10] Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. *Experimental approaches to phonology*, 25-40.

[11] Simon, E. (2010). Phonological transfer of voicing and devoicing rules: evidence from L1 Dutch and L2 English conversational speech. *Language Sciences*, *32*(1), 63-86.

[12] SOARES, V. H. M. (2016). *Encontros consonantais em final de palavra no português brasileiro* (Master's dissertation, Universidade Federal de Minas Gerais.

[13] MANIWA, K., JONGMAN, A., & WADE, T. (2009). Acoustic characteristics of clearly spoken English fricatives. In *The Journal of the Acoustical Society of America*, 125(6), 3962-3973.