revistadelta

https://dx.doi.org/10.1590/1678-460X202642162587

Articles

From language-specific perceptual strategies to phonetic drift: Perception of stops in L2 English and L1 Brazilian Portuguese

De estratégias perceptuais específicas de língua a desvio fonético: Percepção de oclusivas em inglês como L2 e em português brasileiro como L1

Denise Maria Osborne*

ABSTRACT

This study investigates interactions between L1-L2 phonetic systems at the perceptual level. Brazilian Portuguese speakers with English as their foreign language completed two listening tests in which they heard the same speech stimuli but were primed to perceive them as either English or Portuguese. The study examined whether listeners could shift perceptual boundaries along a voice onset time continuum from /b/ to /p/ as a function of perceived language context. The results indicate that phonetic boundaries (crossover points) and categorization sharpness (slope steepness) constitute distinct perceptual processes. For instance, evidence of phonetic drift was observed in L2 learners' boundary locations in the perception of Portuguese stops, whereas no corresponding effects emerged for slope steepness. On the other hand, L2 learners exhibited language-specific slope patterns across Portuguese and English contexts, despite showing stable crossover points across modes. This asymmetry between boundary location and slope suggests that L2 learners attend to specific acoustic-phonetic cues that best support the differentiation of target-language categories. It also supports a view of speech perception in which L1 and L2 representations interact dynamically.

Keywords: *perception; second language; VOT; phonetic drift*

* Cornell University. Ithaca, New York – United States. https://orcid.org/0000-0002-9218-3101. E-mail: dmo67@cornell.edu

RESUMO

Este estudo investiga as interações entre os sistemas fonéticos de L1 e L2 no nível perceptual. Falantes do português brasileiro que têm o inglês como língua estrangeira participaram de dois testes de percepção auditiva, nos quais ouviram o mesmo conjunto de estímulos de fala, mas foram induzidos a percebê-los como pertencentes ao inglês ou ao português. O estudo examinou se os ouvintes seriam capazes de deslocar seus limites perceptuais ao longo de um contínuo de voice onset time de /b/ a /p/ em função do contexto linguístico percebido. Os resultados indicam que os limites fonéticos (pontos de cruzamento) e a nitidez da categorização (inclinação da curva) constituem processos perceptuais distintos. Por exemplo, foram observadas evidências da deriva fonética na localização dos limites perceptuais de aprendizes de L2 na percepção de oclusivas do português, enquanto não se verificaram efeitos correspondentes na inclinação das curvas. Por outro lado, os aprendizes de L2 apresentaram padrões de inclinação específicos para o português e para o inglês, apesar de manterem pontos de cruzamento estáveis entre os diferentes modos de percepção. Essa assimetria entre localização do limite e inclinação da curva sugere que aprendizes de L2 atentam para pistas acústicas-fonéticas específicas que melhor sustentam a diferenciação entre categoria da língua-alvo. Além disso, os resultados sustentam uma visão da percepção da fala na qual as representações de L1 e L2 interagem de forma dinâmica.

Palavras-chaves: *percepção; segunda língua; VOT; desvio fonético*

1. Introduction

This study examines the perception of bilabial stop consonants /p/ and /b/ in English and in Brazilian Portuguese (BP) by adult native speakers of BP who had acquired English as a foreign language (L2). Grounded in theoretical frameworks that highlight the dynamic interplay between the first language (L1) and L2 phonetic systems (e.g., Gonzales & Lotto, 2013; Kellogg & Chang, 2023; Kupske & Lima Jr., 2022; López, 2012; Mack, 2003; Osborne & Simonet, 2021), this study investigates the development of L2 phonetic categories in perception, with particular attention to language-specific perceptual strategies or knowledge. Additionally, it explores the potential influence of L2 exposure on the perception of L1 categories, a phenomenon known as phonetic drift (e.g., Chang, 2012; Lev-Ari & Peperkamp, 2013, Sancier & Fowler, 1996). By focusing on learners who had acquired English primarily through formal, classroom-based instruction with

revistadelta
42.1
2026

limited exposure to native speaker input, this study also aims to advance our understanding of how L1 and L2 phonetic systems interact under conditions of restricted L2 exposure.

Voice onset time (VOT), one of the most extensively studied acoustic correlates of stop consonants, refers to the temporal interval (measured in milliseconds) between the release of a stop closure and the onset of vocal fold vibration of the following vowel. VOT serves as a critical acoustic and perceptual cue for distinguishing between voiced and voiceless, as well as aspirated and unaspirated, stop consonants (e.g., Abramson & Lisker, 1970). Although both Portuguese and English contrast stops phonologically as voiced and voiceless, speakers of each language rely on distinctive phonetic cues to realize these contrasts. In word-initial position, for instance, Portuguese /b/ is typically produced with prevoicing (also known as voicing lead), resulting in negative VOT values. In contrast, English /b/ is generally realized with short-lag VOT values. Chodroff and Wilson (2017) reported a mean VOT of 13 ms for English /b/, with observed values ranging from 11 to 20 ms. While English /b/ is phonemically categorized as voiced, it is often phonetically voiceless in word-initial position due to the absence of prevoicing. Nevertheless, voicing in English stops can occur in certain contexts, such as intervocalic environments.

The production of L2 stops is closely tied to the dynamic interaction between the L1 and L2 phonetic systems. L2 learners must navigate and manage two systems, which exert mutual influence on each other (e.g., Kellogg & Chang, 2023; Kupske & Lima Jr., 2022; Mack, 2003; Osborne & Simonet, 2021). Research in speech perception has similarly highlighted the prevalence of L1-L2 phonetic interplay in perceptual processes (e.g., Gonzales & Lotto, 2013; Kellogg & Chang, 2023). However, the mechanisms underlying this interaction, as well as the degree to which L1 and L2 systems influence one another, remain incompletely understood. Some evidence suggests that only highly proficient or fluent bilinguals are able to establish distinct phonetic categories for each language in perception (e.g., García-Sierra et al., 2009; Gonzales & Lotto, 2013). These individuals demonstrate the ability to shift their perceptual boundary along a VOT continuum (e.g., from /b/ to /p/) depending on the activated language mode, indicating the development of language-specific perceptual strategies. More recent findings by Kellogg and Chang (2023) provide further evidence for phonetic drift, i.e., the influence of L1 on L2 phonetic representations, underscoring the plasticity of the L2 system. The present study seeks to advance this line of inquiry by examining whether L2 proficiency levels modulate stop perception in both the L1 and L2, thereby contributing to a deeper understanding of cross-linguistic phonetic interaction in bilingual speakers.

2. Literature review

A growing body of research on L2 phonetic boundaries has highlighted the dynamic influence that one language can exert on another in L2 learners (e.g., Caramazza et al., 1973; García-Sierra et al., 2009; Gonzales & Lotto, 2013; Gorba & Cebrian, 2021; Major, 1992). While most studies support the view that cross-linguistic interaction shapes phonetic categorization, findings across time reflect shifting theoretical frameworks (Caramazza et al., 1973; Flege & Eefting, 1987). For instance, Caramazza et al. (1973) early work, published only one after after Selinker's (1972) introduction of the concept of interlanguage, reported no language-specific adjustment in the phonemic boundaries of English-French bilinguals. Their conclusion, which at the time suggested a lack of language-specific perceptual modulation, must be understood in the context of the theoretical landscape of the early 1970s, before the development of later models such as the SLM (Flege, 1995) and SLAM-r (Flege & Bohn, 2021).

In contrast, more recent research conducted within contemporary theoretical frameworks has shown evidence of context-sensitive phonetic behavior. Gorba and Cebrian (2021), for instance, demonstrated that L2 English learners of Spanish modulated their VOT production according to the language context (e.g., producing longer VOT in English than in Spanish), indicating sensitivity to language modes. However, their work did not reveal a parallel influence of L2 experience on the perception of stop consonants. Taken together, when the temporal and theoretical contexts of these studies are considered, the literature suggests that production of these processes has evolved substantially since early bilingualism research in the 1970s.

A substantial body of research has underscored the role of language mode in activating language-specific processing mechanism that enable bilinguals to employ distinct perceptual strategies for their L1 and L2 (e.g., Flege & Eefting, 1987; García-Sierra et al., 2009; Gonzales & Lotto, 2013; López, 2012). Such studies suggest that L2 learners can develop novel perceptual categories for L2 stop consonants; however, these categories

revistadelta
42.1
2026

may not fully align with those of native monolingual speakers of the target language. This divergence shows the nuanced and dynamic nature of phonetic category formation in L2 perception.

Early bilinguals have shown to develop language-specific perceptual strategies, evidenced by their ability to shift phonetic boundaries according to language context, thereby demonstrating sensitivity to fine-grained phonetic differences in stop consonants (e.g., Gonzales & Lotto, 2013). In a study by Gonzales and Lotto, Spanish- English bilinguals who had learned both languages before the age of 8, along with English monolinguals, were exposed to a VOT continuum in the two language contexts to assess the presence of dual phonetic boundaries. Participants were presented with the same set of VOT stimuli in both Spanish and English contexts and asked to categorize the stimuli as either *pafri* or *bafri*. The results revealed that only bilingual participants showed context-dependent shifts in perception: they identified a greater number of voiceless responses (*pafri*) in Spanish mode than in the English mode. On the other hand, monolinguals did not demonstrate such boundary shifts, indicating a lack of context-sensitivity phonetic adjustment. These findings support prior claims that "highly confident L2 bilinguals [...] are more likely to possess a double phonemic representation" (García-Sierra et al., 2009, p. 378), and that "presumably, more competent bilinguals overcome this clash of phonetic systems" (Gonzales & Lotto, 2013, p. 2135). However, it remains uncertain whether late L2 learners with limited L2 exposure and lower proficiency levels are capable of developing similar language-specific perceptual strategies.

The present study seeks to investigate whether L2 learners who have acquired their L2 in a formal instructional context within their L1 environment are capable of shifting their phonemic boundaries in response to language-specific processing models (L1 vs. L2). If such learners demonstrate this perceptual flexibility, the findings will provide evidence that language-specific perceptual strategies can be acquired even in the absence of immersive L2 exposure. Additionally, the study investigates the potential for phonetic drift in L1 VOT as a function of L2 proficiency, thereby exploring the bidirectional nature of cross-linguistic phonetic influence. More broadly, this research aims to enhance our understanding of how L2 learners process and retrieve linguistic representations, particularly in cases where L1 and L2 share overlapping phonetic features that are categorized differently at the phonemic level across the two languages.

3. The present study

This study investigated the perception of word-initial bilabial stop consonant in both L2 English and L1 Portuguese by Brazilian Portuguese (BP) learners of English residing in Brazil. The primary objective was to determine whether these learners demonstrate evidence of phonetic drift in their L1 perceptual categories as a result of L2 acquisition, as well as whether they demonstrate language-specific perceptual strategies when processing stop consonants in each language. To elicit perceptual data, participants completed two two-alternative forced-choice (2AFC) identification tasks, administered in separate language modes: one in Portuguese (L1) and the other in English (L2).

There are three research questions:

RQ1. *To what extent does L2 English proficiency influence the perception of English stops?*

RQ2. *Does the acquisition of English as L2 induce phonetic drift in the perception of stop consonants in L1 Portuguese?*

RQ3. *Do L2 learners demonstrate language-specific perceptual categorization of VOT when processing stop consonants in L1 and L2 contexts?*

It was hypothesized that learners with higher proficiency in English would be more likely to demonstrate language-specific perceptual strategies, characterized by context-dependent shifts in phonemic boundary perception (e.g., García-Sierra et al., 2009). In contrast, lower-proficiency learners were not expected to demonstrate significant perceptual adjustments across language contexts; their responses in both language modes were anticipated to closely align with those of monolingual BP speakers. With respect to phonetic drift, it is hypothesized that L2 learners with higher levels of L2 experience may show L2 learning effects on the perception of the L1 Portuguese stop consonant system, reflecting cross-linguistic interaction and the plasticity of phonetic representations in L1 (e.g., Chang, 2012; Flege, 1987; Kellogg & Chang, 2023).

42.1
2026

4. Method

Participants: L2 English Learners

A total of 36 Brazilian Portuguese (BP) learners of English participated in this study (26 female, 10 male), with ages ranging from 18 to 50 (*mean* = 29.55 years). All participants had completed at least a high school education. The majority were born in the same small town in the state of Minas Gerais, with the exception of two individuals, who were born in São Paulo and Rio de Janeiro, respectively. At the time of data collection, all participants resided in Minas Gerais. Participants self-reported adequate hearing and vision, either uncorrected or corrected with glasses or contact lenses.

After signing their consent form, participants took part in the St. George International Online English Test, a standardized assessment used to evaluate their proficiency in L2 English. Their scores ranged from 10 to the maximum of 40, with a mean of 26.94 (*SD* = 8.82, *median* = 27.5, *skew* = -0.17, *SE* = 1.47). The distribution of scores was approximately symmetrical, with little skew and spread across the proficiency range. This indicated that L2 learners varied in their L2 proficiency, ranging from beginner to advanced levels, and were fairly evenly distributed across the continuum. Such variability provides an ideal condition for analyzing how proficiency influences phonetic categorization. Participants also answered a language background questionnaire, available in online Supplementary Material. Table 1 shows the summary of their responses.

Table 1. Background information of Brazilian Portuguese learners of English as an L2 (*n* = 36).

Gender	26 (female), 10 (male)
Age	*mean* = 29.83, *SD* = 9.21, *range* = 18-50
Age of Acquisition	*mean* = 16.22, *SD* = 8.41, *range* = 6-44
Length of Learning	*mean* = 6.48, *SD* = 6.28, *range* = 0.41-21 years
Contact with English native speakers	No = 0 (69.45%), Yes = 11 (30.55%)
Education	High school = 7 (19.45%) College = 25 (69.45%) Graduate school = 4 (11.10%)

Note. Length of Learning includes continuous and non-continuous study of English.

Muñoz (2010) argues that, in research on L2 language acquisition in foreign language contexts, it is important to distinguish between the age of first exposure and the age of acquisition. According to her, the age of first exposure may involve “only insignificant exposure” (p. 44), whereas the age of acquisition refers to the onset of meaningful contact with the L2. In this study, age of acquisition is defined as the age at which learners began to experience substantial and meaningful exposure to the L2, as self-reported by L2 English learners.

We were also interested in learners’ use of English in their daily lives. L2 learners reported using English most frequently at their English school (*mean* = 80.3%, *SD* = 23.72), with high average use and low variability, a consistent result among L2 learners. The least use of English was with friends (*mean* = 18.3%, *SD* = 20.35). In terms of variability in English use, the largest was at work (*mean* = 39.72%, *SD* = 39.39), where some learners reported using English 100% of the time (*n*=3; they we all English instructors), and others 0% of the time (*n*=13). A plot showing the distribution of L2 English use across different environments, along with the respective means, is provided in the online Supplementary Material. Overall, these findings suggest that while formal learning environment provides structure exposure and serves as the primary context for English use, L2 learners’ real-world use of English (outside the classroom) varied considerably. English is frequently used in contexts such as the Internet and films, whereas in real-life social interactions, such as with friends and at home, English use is limited or often non-existent.

Participants: Monolingual speakers of Brazilian Portuguese

A total of 36 monolingual Brazilian Portuguese speakers participated in the study (26 females, 10 males). On average, this group was older than the L2 learner group, with a mean age of 33.08 years (*range* = 18-65). All participants were born in the state of Minas Gerais and were residing in the same city where the data collection took place. Most were either college students or had already completed a college degree (15 participants); others were enrolled in graduate programs (13 participants), and some had completed high school (8 participants). All participants self-reported adequate hearing and vision, either uncorrected or corrected with glasses or contact lenses.

42.1
2026

A small number of monolingual BP speakers (4 participants) had taken some English classes in the past. Following the criteria outlined by Chang (2012), careful consideration was given to defining L1 speakers as a monolingual group. To ensure that these individuals could still be classified as monolinguals despite limited exposure to English, the researcher asked them to produce a sentence in English. As they were unable to do so, they were included in the monolingual group.

Stimuli

The stimuli used in this study were the same as those from Gozales and Lotto's (2013) study and were generously provided by Dr. Kalim Gonzales and Dr. Andrew J. Lotto. They were created as follows: a Spanish-English bilingual speaker produced the nonwords *bafri* and *pafri* in both Spanish and English accents. The initial *paf-* from the Spanish *pafri* was extracted and used to create a 14-step VOT continuum, ranging from -35 ms to +35 ms increments, excluding 0 ms. The final *-ri* portions from both Spanish and English productions were resynthesized to control for duration, F0 onset, and contour, while preserving the key phonetic features, namely, the retroflex /ɹ/ for English and the flap /ɾ/ for Spanish. To create the final stimuli, each of the 14 VOT steps was paired with the English *-ri* to produce the English set, and with the Spanish *-ri* to produce the Spanish set. Thus, both sets shared the same VOT continuum, differing only in the final segment.

Although the stimuli from Gonzales and Lotto (2013) were produced by a Spanish-English bilingual speaker, they were appropriate for use in this study, as the Spanish and Portuguese alveolar flaps are phonetically similar in several respects. In both languages, the /r/ sound is described as an alveolar flap that does not occur in word-initial position (e.g., Mateus & d'Andrade, 2000). Flaps can appear in consonant clusters in onset position, such as after an initial consonant, as in the stimuli *bafri* and *pafri.* Moreover, *bafri* and *pafri* are nonwords not only in English and Spanish, but also in Portuguese; however, they conform to the phonotactic rules of all three languages and are thus phonologically possible. Additionally, Spanish and Portuguese speakers typically prevoice initial /b/ and produce initial /p/ with short-lag VOT values.

Experimental Procedures

This study followed standard procedures commonly used to investigate whether L2 learners exhibit distinct phonetic boundaries for their L1 and L2. In such paradigms, participants are exposed to the same speech stimuli in two separate language contexts (L1 and L2 modes) and asked to categorize sounds along a continuum, e.g., from /b/ to /p/. The goal is to determine whether L2 learners shift their phonetic boundary based on language mode alone. A shift in categorization depending on context suggests that learners employ language-specific perceptual strategies for their L1 and L2. Crucially, however, establishing a strong and appropriate language context is essential to activate the intended language mode (e.g., Grosjean, 2001, 2013). Without sufficient activation of the target language, experimental results may be compromised, leading to ambiguous or inconclusive findings.

Data from the L2 learners were collected in quiet rooms at two English schools in a city in Minas Gerais. This setting was chosen to help control the language context and encourage participants to process the sounds in L2 English (e.g., Hay & Drager, 2010). Data from the monolingual Brazilian Portuguese speakers were collected in a quiet room at a local college in the same city.

L2 learners participated in two language sessions: one in English and one in Portuguese. The Portuguese session took place either on the same day as the English session or on a different day, depending on participants' availability. The two sessions were designed as mirror images of each other, that is, the English and Portuguese perception experiments followed the same structure and procedures. Monolingual Brazilian Portuguese speakers participated only in the Portuguese session.

A native English speaker (the research assistant), born and raised in New York City and trained by the researcher, conducted the English sessions. The Portuguese sessions were conducted by the researcher, a native speaker of Brazilian Portuguese. Because eliciting an English-language context is more challenging in a non-English-speaking country, all participants began with the English session. During this session, participants interacted only with the English-speaking assistant, while the Portuguese-speaking researcher waited in a separate room. Once the English session ended, the assistant called the researcher to begin the Portuguese session and then left the room. Eliciting a Portuguese language mode was presumably straightforward, as participants were likely to feel at ease speaking with the researcher, who shared both their native language and regional variety. Nonetheless, to ensure that the Portuguese mode was fully established, the researcher always initiated the session with a brief conversation in Portuguese. The tests were conducted using PsychoPy software (Peirce, 2014) on the researcher's computer (MacBook Air, 2011) with AKG K77 headphones. The volume was adjusted for each participant's comfort.

In the English session, participants were informed that they would hear two English words repeated multiple times. Neither the researcher nor the research assistant pronounced the words *pafri* and *barfi,* and no indication was given as to whether the words were real or imaginary. Participants both heard and read the instructions on the screen. The English instructions were as follows:

You will hear two words in English: BAFRI and PAFRI.
If the word starts with B, click the left arrow.
If the word starts with P, click the right arrow.
Click SPACE to start.

To help prime the intended language context, BAFRI and PAFRI were color-coded: in the English session, BAFRI appeared in red and PAFRI in blue, while in the Portuguese session, BAFRI appeared in green and PAFRI in yellow, colors reflecting the American and Brazilian flags, respectively. Additionally, the words *English* or *Português* was displayed above the corresponding stimuli during each session (a screenshot of the computer interface is provided in the online Supplementary Material).

The words were presented to participants in a randomized order. If a participant did not respond within 4.1 seconds, the computer automatically played the next stimulus; however, the screen remained unchanged during this time. The experiment consisted of a single block of 80 randomly ordered trials.

Statistical Procedures

To investigate the perception categorical boundaries, this study focuses on analyzing crossover points and slopes. The crossover point represents the phonetic boundary between the categories of /b/ and /p/ sounds, that is, the VOT value at which the listener is equally likely to perceive the sound /b/ or /p/, marking the transition where the probability of categorizing the sound as /p/ reaches 0.50. The VOT analyses examine the sharpness of the categorization, which provides insight, for example, into whether L2 English learners identify stops in English and Portuguese modes with the same degree of categorical confidence.

To assess statistical differences in crossover point means (one continuous value per subject), simple linear regression model were fitted using *lm( )*. For analyzing slope sharpness with response as binary variables, generalized linear mixed-effect models were employed using *glmer( )* with a binomial family and logit link. Random intercepts and random slopes for VOT were included to account for subject-level variation and the repeated-measures structure of the data. The best-fitting models were selected using the *anova( )* function to compare nested models by Chi-square tests and by identifying the lowest Akaike Information Criterion (AIC). The *lm( )* function is part of base R, while *glmer( )* is provided by the *lme4* package in R (R Development Core Team, 2021). In the event of convergence issues, continuous predictors were standardized (z-score: *mean* = 0, *SD* = 1) to enhance model interpretability, following the recommendations of Schielzeth (2010). The model was then re-run using these standardized predictors.

In this study, L2 English proficiency, measured by participants' score on the St. George International Online English Test, is treated as a continuous, gradient variable[1]. This approach enables the analyses to explore whether gradual changes in proficiency affect L2 learners' categorization, therefore avoiding arbitrary cutoff thresholds.

To identify listeners processing the /b/-/p/ contrast categorically rather than responding randomly, a categorical perception index was developed based on the following criterion: individual crossover values falling outside the plausible range of the experimental stimuli were considered as outliers. More precisely, crossover points were expected to lie within the experimental stimuli VOT range of -35 ms to + 35 ms. Outliers were defined as crossover points below -35 ms or above + 35 ms. Such extreme values may arise when participants' slopes are flap (slope near zero), indicating a non-categorical phonetic response, potentially due to inattention or measurement error. After calculating the crossover points and slopes of all participants, comprising originally of 36 L2 English learners and 36 monolingual BP speaker (a

1. I would like to acknowledge contribution of one of the anonymous reviewers to this decision.

revistadelta
42.1
2026

total of 72 participants), 8 of them were identified as outliers (4 L2 English learners and 3 L1 BP speaker). Consequently, the final sample consisted of 32 L2 English learners and 33 L1 BP speakers, totaling 65 participants.

5. Results

In the English mode, the test yielded a total of 4,345 responses by L2 learners (14 VOT steps x 10 repetitions x 32 participants, minus 135 responses that were not captured by the computer). In the Portuguese mode, the test yielded 8,993 responses from both L2 learners and BP monolingual speakers (14 VOT steps x 10 repetitions x 65 participants, minus 107 responses missing responses). As presented in Tabel 2, the highest mean rate of /p/ responses is observed among L1 BP speakers in the Portuguese mode. L2 English learners perceived more /p/ stops in the Portuguese mode than in the English mode, but fewer than L1 BP speakers in the Portuguese mode.

Table 2. Overall /p/ (*pafri*) identification across language groups and language modes.

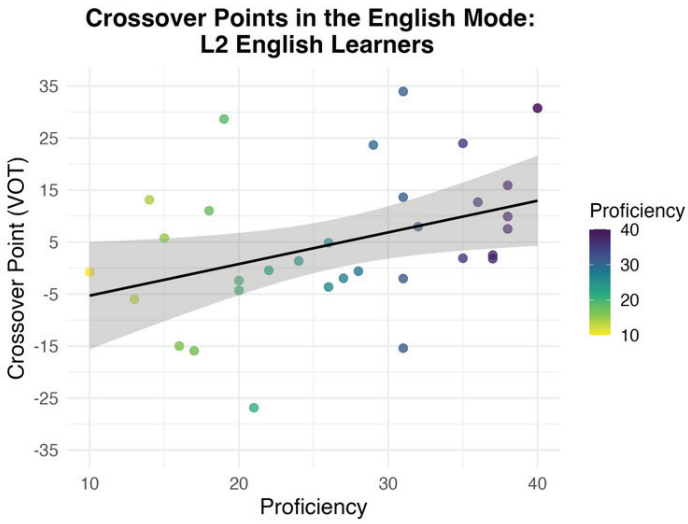
Language Mode	Language Group	/p/ Responses (out of total tokens)	Percentage of Responses (95% CI)
Portuguese	L1 BP	2,595 (4,562)	56.88% (55.44-58.32)
	L2 English	2,188 (4,431)	49.37% (47.90-50.85)
English	L2 English	1,952 (4345)	44.92% (43.44-46.40)

Perception of bilabial stops by L2 English Learners in the English Mode

RQ1. *To what extent does L2 English proficiency influence the perception of English stops?*

To investigate whether L2 learners' proficiency in English influences their categorization of English stops, analyses were conducted exclusively on the L2 learner group in the English mode. In Figure 1, the fitted regression line reveals a positive linear relationship: as proficiency increases, crossover points tend to shift later VOT values, moving toward higher values. Lower proficiency L2 English learners tend to shift phonetic boundaries earlier than higher proficiency counterparts. The shaded area indicates considerable variability among L2 learners, with a narrower range of crossover points centered around +5 ms VOT.

Figure 1. Boundary placement for /p/ by L2 English learners in the English mode as a function of proficiency score. Each dot represents an individual learner's mean crossover point, with color indicating their proficiency level, ranging from lower (lighter colors) to higher scores (darker colors). The solid line represents the fitted repression line, capturing the relationship between proficiency level and crossover point. The surrounding shaded area indicates the 95% confidence interval of the estimate.



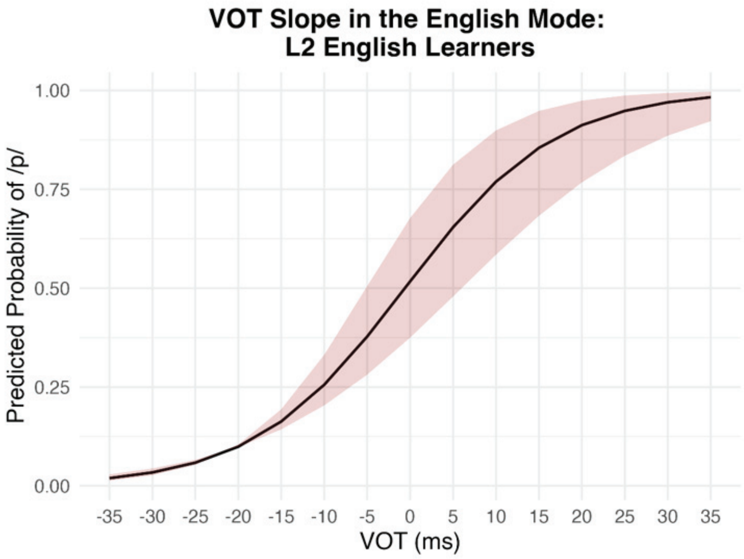
Source: Prepared by the author.

The mean crossover point by L2 English learners in the English mode was 4.87 ms (*SD* = 13.9, 95% CI: -0.15, 9.88); the minimum crossover point was -26.9 ms and the maximum was 34 ms. The linear regression model, examining the relationship between crossover (the dependent variable) and proficiency (the independent variable), revealed a significant positive effect of proficiency score on crossover point ($\beta = 0.61$, $SE = 0.27$, $t(30) = 2.27$, $p = 0.031$). This indicates that, in a within-by subject comparison, higher

revistadelta
42.1
2026

proficiency scores are statistically associated with later crossover points (higher VOT values), as the positive coefficient indicates. This finding supports the view that increased proficiency leads L2 learners' phonetic categorization to more closely resemble that of native English speakers in perceiving the /b/-/p/ boundary, reflecting a more native-like sensitivity to VOT distinctions in English stops.

Figure 2. Predicted probabilities of categorization a sound as /p/ across a range of VOT values for L2 learners in the English mode. The black solid line represents the mean logistic curve, derived from the mean intercept and slope of individual logistic regression models. The shaded brown area indicates 95% confidence interval around the mean curve, calculated via bootstrap, reflecting variability across participants.



Source: Prepared by the author.

The subsequent analyses focused on the sharpness of slopes, which reflect L2 learners' sensitivity to phonetic boundaries. As shown in Figure 2, the overall moderate slope indicates a gradual rather than abrupt shift from /b/ to /p/ responses. The wide standard error band, extending not only through the ambiguous region (approximately -5 to +5 ms), but also remaining broad beyond +5 ms, suggests considerable variability across participants, particularly at mid-range VOT values. This variability tends to diminish as VOT increases, indicating greater agreement among L2 learners in categorizing sounds as /p/ at higher VOT values in the English mode.

To verify whether proficiency in L2 English affected differences in slope, a generalized linear mixed-effects model was fitted with /p/ responses as the dependent variable, the interaction of VOT and proficiency as the fixed effects, and random intercepts and random slopes for VOT by subjects. Since the initial model failed to converge, proficiency values were centered and scaled to improve model stability. The model was then re-run with the standardized variables. The interaction was not significant (β = -0.004, *SE* = 0.01, *z value* = -0.32, p = 0.7511), indicating that the slopes of the categorization curves did not vary significantly as a function of proficiency. Both main effects were significant: proficiency (β = -0.51, *SE* = 0.21, *z value* = -2.36, p = 0.0183) and VOT (β = 0.10, *SE* = 0.01, *z value* = 8.85, $p < 0.001$). These results suggest that while VOT and proficiency each independently influence categorization responses, proficiency does not modulate the effect of VOT, that is, the rate of change in /p/ responses across VOT continuum (i.e., the slope) remains stable across proficiency levels. The significant negative effect of proficiency (main effect) indicates that, across all VOT values, higher proficiency is associated with lower baseline log-odds of /p/ responses, meaning, a reduction in the likelihood of categorizing a sound as /p/ near the category boundary (at VOT = 0). Regardless of proficiency, L2 English learners showed robust sensitivity to VOT, with categorization following an S-shaped curve, as illustrated in Figure 2. In summary, while proficiency shifts the overall response tendency, it does not affect the sharpness of the categorization boundary, as reflected in the slope of the VOT curve.

In response to RQ1 (*To what extent does L2 English proficiency influence the perception of English stops?*), the findings indicate that proficiency does not influence the steepness of the VOT slope, that is, the sharpness of phonetic categorization remains stable across proficiency levels. In contrast, the crossover point analysis revealed a significant relationship between proficiency and boundary placement, such that higher proficiency was associated with later crossover points along the VOT continuum, reflecting a shift toward more native-like English perception. In sum, these results suggest that L2 proficiency influences the location of the phonetic boundary but does not affect the rate of change (slope) in categorization across VOT values. This selective pattern echoes findings from Gorba & Cebrian (2021), who showed that certain dimensions of phonetic categories (e.g., boundary placement) may remain malleable across levels of L2 experience while others (e.g., slope steepness) may remain stable, highlighting that phonetic categories need not shift uniformly across all dimensions.

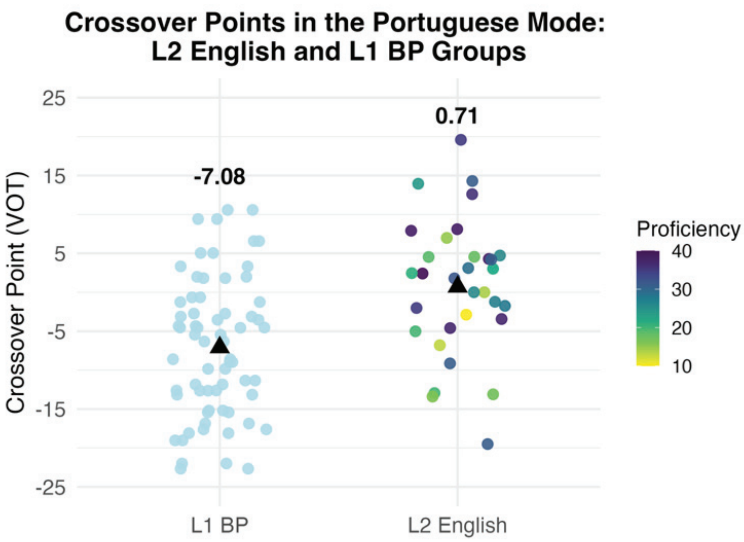
revistadelta
42.1
2026

Perception of bilabial stops by Monolingual Brazilian Portuguese speakers and L2 English Learners in the Portuguese Mode

RQ2. *Does the acquisition of English as L2 induce phonetic drift in the perception of stop consonants in L1 Portuguese?*

To examine whether L2 English listeners' perception of L1 Portuguese exhibits phonetic drift, their responses were compared to those of monolingual Brazilian Portuguese (L1 BP) speakers. Crossover points for L1 BP ranged from -22.7 to +10.6 ms, with a mean of -7.08 ms (*SD* = 8.94, 95% CI [-10.25, -3.91]). For the L2 English group, the crossover range is from -19.5 to +19.6 ms, with a mean of 0.71 ms (*SD* = 8.69, 95% CI [-2.42, 3.84]), indicating that L2 English group's crossover is higher than that of L1 BP in the Portuguese mode. The difference in mean crossover between the groups is 7.79 ms. Figure 3 shows a general trend for L2 learners to shift perception from Portuguese /b/ to /p/ at higher VOT values than their L1 BP counterparts, an indication of possible phonetic drift in the direction of English.

Figure 3. Bound placement for L2 English learners and L1 BP group in the Portuguese mode. Each point represents a subject's estimated crossover point. For L2 English participants, the color gradient from lighter to darker colors yellow indicates increasing proficiency scores. Group means are highlighted with black triangles and labeled with their respective values.

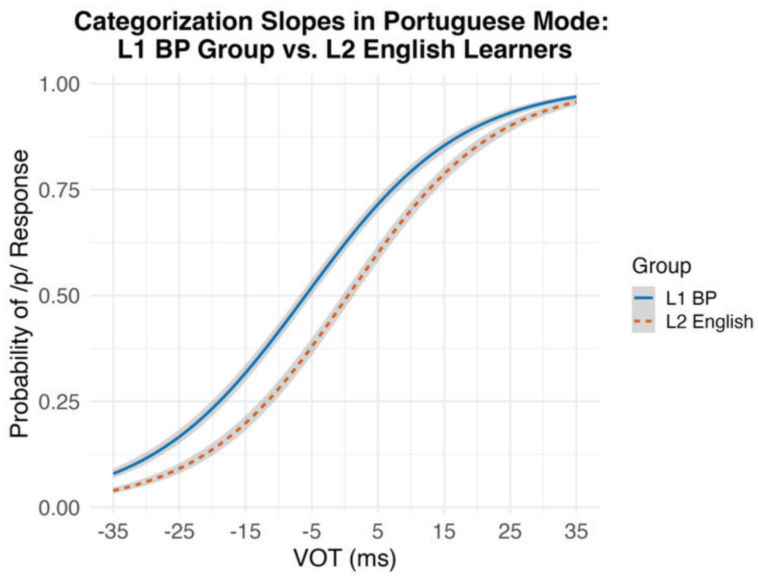


Source: Prepared by the author.

A linear model with crossover points as the dependent variable and language group as the predictor was fit to examine whether crossover points differ statistically between groups. The analysis revealed a significant effect of group membership on the crossover point ($\beta = 7.79$, $SE = 2.19$, $t(63) = 3.251$, $p < 0.001$), indicating that L1 BP participants had significantly earlier crossover points than L2 English participants. On average, L2 English learners showed crossover points that were 7.79 ms higher than those of L1 BP participants. These higher VOT values at the categorical boundary in the Portuguese mode may reflect the influence of L2 English experience, suggesting that their crossover points are shifting toward a more English-like perceptual pattern.

As a follow-up analysis, the potential influence of proficiency on crossover points within the L2 English group was examined. A linear model was fitted with crossover points as the dependent variable and proficiency as the predictor. The analysis revealed a non-significant relationship between proficiency and crossover point location ($\beta = 0.30$, $SE = 0.17$, $t(30)$ *value* $= 1.75$, $p = 0.0894$), showing that, within the L2 English group, proficiency alone was not a strong predictor of perceptual boundary shift.

Figure 4. Predicted probability of responding /p/ in Portuguese mode as a function of VOT for L1 BP speakers and L2 English learners. The lines represent predictions from a logistic regression model. The shaded ribbons show 95% confidence intervals.



Source: Prepared by the author.

revistadelta
42.1
2026

Figure 4 shows that the slopes of L1 BP and L2 English learners in the Portuguese mode follow similar trajectories, indicating that both groups exhibit a comparable overall pattern of VOT-based categorization. In other words, the slopes display similar steepness, suggesting a similar rate of categorical change. However, the slope of the L2 English learners is consistently lower across the VOT continuum and does not overlap with that of the L1 BP group (i.e., it has a lower intercept). This pattern suggests that L2 English learners have a systematically reduced likelihood of categorizing stimuli as /p/ in the Portuguese mode compared to monolingual Portuguese speakers. While L1 BP listeners show an earlier and stronger bias toward identifying sounds as /p/, the slope of L2 English learners reflects greater reluctance in doing so across VOT values.

To test whether the sharpness of the categorization slope differed between groups in the Portuguese mode, a generalized linear mixed-effects model was fitted with /p/ responses as the dependent variable, the interaction of VOT and language group as the fixed effects, and random intercepts and random slopes for VOT by subject. The interaction was not significant (β = 0.006, *SE* = 0.02, *z value* = 0.38, *p* = 0.7055), indicating that both groups showed similar sensitivity to changes in VOT in Portuguese, that is, the rate of transition from /b/ to /p/ as VOT increases was comparable. As expected, the main effect of VOT was significant (β = 0.12, *SE* = 0.01, *z value* = 10.38, $p < 0.001$), reflecting the overall influence of VOT on categorization. The main effect of language group was also significant (β = -0.62, *SE* = 0.31, *z value* = -1.99, *p* = 0.0457), showing a group difference in the intercepts: L2 English learners were significantly less likely than L1 BP speakers to categorize a stimulus as /p/ at the baseline VOT (VOT = 0). Since the interaction term, which captures group differences in slope, was not - significant, we conclude that there was no evidence of slope differences between groups in their categorization of Portuguese stops.

To assess whether proficiency influenced L2 English learners' categorization of Portuguese stops, a follow-up analysis was conducted using data exclusively from the L2 English group. This analysis incorporated a generalized linear mixed-effects model with the interaction of VOT and proficient level, along with random intercepts and random slopes for VOT by subject. Due to convergence issues, proficiency values were centered and scaled. The interaction between VOT and proficiency was not significant (β = 0.003, *SE* = 0.01, *z value* = 0.25, *p* = 0.805); the main effect of proficiency was also not significant (β = -0.30, *SE* = 0.24, *z value* = -1.25, *p* = 0.211).

However, the main effect of VOT was significant, as anticipated ($\beta = 0.13$, *SE* = 0.01, *z value* = 10.16, $p < 0.001$). The non-significant interaction indicates that proficiency level of L2 English does not affect slope (i.e., the steepness of the transition) of the logistic curve when perceiving stops in the Portuguese mode.

Addressing RQ2 (*Does the acquisition of English as L2 induce phonetic drift in the perception of stop consonants in L1 Portuguese?*), the results provide evidence of phonetic drift for L2 English group in the perception of Portuguese stops, particularly in terms of crossover points, but not in the sharpness of the slopes. While crossover points differed significantly between groups, with L2 English listeners showing higher crossover points than the L1 BP listeners, the rate of transition (VOT slopes) showed no significant difference. Furthermore, within the L2 English group, degree of L2 experience did not significantly predict either crossover location or slope steepness, a pattern consistent with cue-centric accounts suggesting that L1-attuned perceptual routines often remain stable and are not necessarily reshaped as proficiency increases (Chang, 2018; Gorba & Cebrian, 2021). Taken together, the results support the view that L2 acquisition can influence L1 perception, but such effects may be selective, such as affecting boundary placement more readily than slope sharpness, and may not scale straightforwardly with L2 proficiency.

Perception of bilabial stops by L2 English Learners across Portuguese and English Modes

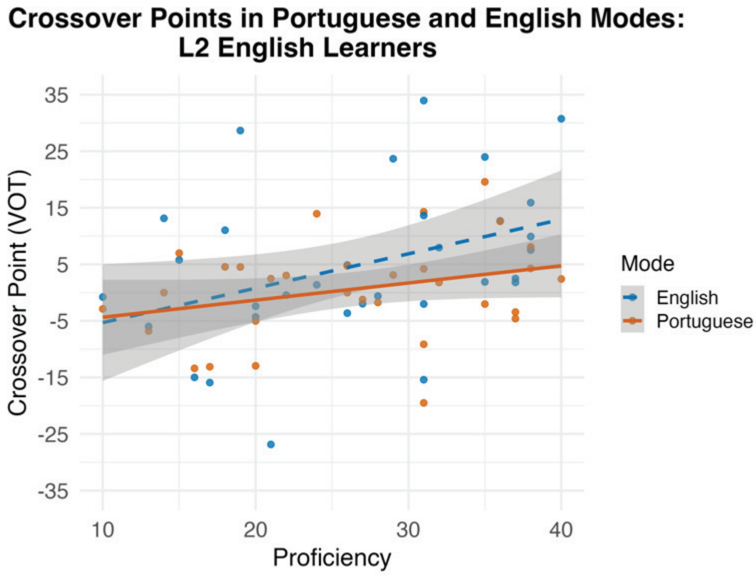
RQ3. *Do L2 learners demonstrate language-specific perceptual categorization of VOT when processing stop consonants in L1 and L2 contexts?*

The following analyses investigate whether L2 learners, with limited exposure to the target L2 English, developed language-specific perceptual strategies depending on language mode. These analyses focus on the L2 English learners' responses in both Portuguese and English modes. The analyses of the crossover points indicate that, in the English mode, the mean crossover point was 4.86 ms (*SD* = 13.91, 95% CI: -0.15, 9.88), while in the Portuguese mode was 0.71 ms (*SD* = 8.69, 95% CI: -2.42, 3.84). Figure 5 shows that both lines are positive, suggesting that higher proficiency is associated with increased crossover VOT values in both language modes,

revistadelta
42.1
2026

with a steeper increased observed in the English mode. The 95% confidence intervals show partial categorization overlap between language modes across VOT values at all proficiency levels, with greater overlaps for the lower levels.

Figure 5. Visualization of boundary placement for L2 English learners across English and Portuguese modes. The dots represent individual observed crossover points, color--coded by language mode. Smooth lines represent the predicted relationship between proficiency and crossover VOT for each mode, with shaded 95% confidence intervals indicating uncertainty around the model estimates.

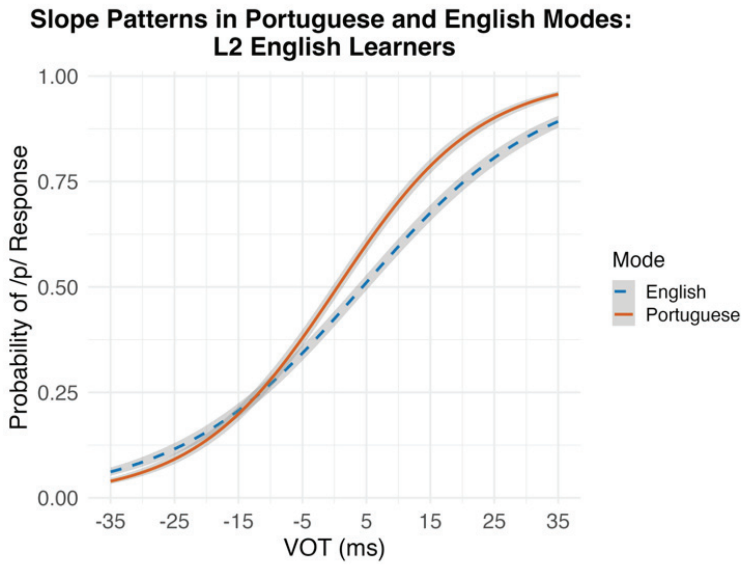


Source: Prepared by the author.

To investigate whether crossover points differ statistically across language modes, a linear simple model was fitted with crossover points as the dependent variable, and the interaction between language mode and proficiency. The interaction was not significant (β = -0.30, *SE* = 0.32, *t*(60) value = -0.95, *p* = 0.3443), indicating that the effect of proficiency on crossover points does not differ across language modes. The main effect of language mode was also not significant (β = 3.98, *SE* = 8.97, *t*(60) value = 0.44, *p* = 0.6588), suggesting no overall shift in boundary placement between English and Portuguese contexts. However, proficiency had a significant positive effect on crossover point location (β = 0.61, *SE* = 0.22, *t*(60) value = 2.70, *p* = 0.0092), indicating that more proficient L2 learners tend to place the crossover point between /b/ and /p/ at higher VOT values, regardless of language mode. Together, these results show that L2 proficiency influences phonetic boundary placement in both L2 English and L1 Portuguese, with no evidence of mode-specific effects. In other words, increased proficiency leads to higher crossover points across both modes, suggesting a general influence of L2 experience on perceptual boundaries, consistent with phonetic drift.

The next set of analyses investigated the gradient of the categorization slopes for /p/ responses in both Portuguese and English modes among L2 English learners, and whether proficiency moderated these differences. As shown in Figure 6, the slopes for Portuguese and English modes largely overlap in the negative VOT range. However, as VOT increases, the slopes begin to diverge. In English mode, learners exhibit a flatter slope compared to Portuguese mode, suggesting less abrupt categorization of /b/-/p/ sounds. This indicates that learners show more categorical perception in the Portuguese mode. Additionally, the rightward shift of the boundary in English mode, relative to the more leftward position in Portuguese, points to potential mode-dependent perceptual differences.

Figure 6. Probability of /p/ responses by L2 English learners in Portuguese and English modes. The y-axis represents the proportion of /p/ responses, ranging from 0 (/b/) to 1 (/p/). Smoothed logistic curves are plotted for each language mode, with confidence bands representing the uncertainty in the model's estimates.



Source: Prepared by the author.

revistadelta
42.1
2026

A generalized linear mixed-effects model was fitted to examine the effects of VOT, language mode, and proficiency level on L2 English learners' perception of /p/ versus /b/. The model included /p/ responses as the dependent variable, two 2-way interactions (VOT x mode, VOT x proficiency) as the fixed effects, and random intercepts and slopes for VOT by subject. A model with three-way interaction (VOT x mode x proficiency) showed a worse fit and was not pursued. Initial convergence issues were resolved by centering and scaling proficiency values. The VOT x proficiency interaction (β = -0.002, *SE* = 0.01, *z value* = -0.18, p = 0.8535) and the main effect of proficiency (β = -0.41, *SE* = 0.23, *z value* = -1.82, p = 0.0683) were not significant, indicating no evidence that proficiency affects the VOT slope or baseline perception in either mode. On the other hand, VOT x mode interaction (β = 0.02, *SE* = 0.003, *z value* = 7.09, $p < 0.001$), main effects of language mode (β = 0.29, *SE* = 0.06, *z value* = 4.81, $p < 0.001$), and main effect of VOT (β = 0.10, *SE* = 0.01, *z value* = 8.77, $p < 0.001$) were significant. The VOT slope was steeper in Portuguese mode than in English mode, indicating that the probability of perceiving /p/ increases more rapidly with VOT in Portuguese than in English mode. The shallower slope in English mode suggests L2 learners are adapting toward native-like English perception, requiring higher VOTs for /p/ in English, consistent with an effect of L2 learning.

In response to RQ3 (*Do L2 learners demonstrate language-specific perceptual categorization of VOT when processing stop consonants in L1 and L2 contexts?*), the findings suggest partial language-specific perceptual strategies among L2 learners. While no significant difference in crossover points was observed between Portuguese and English modes, the slope of the categorization function differed significantly. In the Portuguese mode, the slope was steeper, indicating a more abrupt shift in categorization compared to the flatter slope observed in English mode. This suggests that L2 English learners require higher VOT values to identify a /p/ in English than in Portuguese, pointing to system-wide shift that is not a proficiency-dependent effect, given that proficiency did not significantly affect the crossover points or slope sharpness. This pattern aligns with previous work showing that L2-driven phonetic adjustment do not necessarily depend on high proficiency; as Chang (2012) notes, "a higher level of L2 proficiency is not necessary for L2 input to influence L1 representations" (p. 264).

6. Discussion and conclusion

This study investigates whether Brazilian Portuguese speakers learning English as their L2 in Brazil, most of whom had limited exposure to English, typically restricted to 50-minute classes twice a week, exhibit perceptual learning effects in their L2 and potential phonetic drift in their L1. Specifically, it examines whether L2 English learners develop language-specific perceptual strategies and whether these strategies influence their native Portuguese phonetic boundaries. To assess this, participants were exposed to identical stimuli in two language contexts: English and Portuguese. The experiment tested whether listeners would shift their perceptual boundaries when processing bilabial stops, depending on the activated language mode. Analyses focused on two key measures: crossover points (the VOT at which /p/ is perceived 50% of the time), reflecting the categorical location, and slopes (the steepness of the logistic identification function), indicating the sharpness of the perceptual transition between phoneme categories.

In response to RQ1 (*To what extent does L2 English proficiency influence the perception of English stops?*), the results indicate that proficiency affects only certain aspects of perceptual category. Significant effects of L2 proficiency were observed, but only for crossover points in the English mode, based on a within-subject comparison. Specifically, higher proficiency was associated with crossover points at higher VOT values, indicating a shift toward native-like English phonetic boundaries. In other words, as learners became more proficient, their perception of voiceless bilabial stops increasingly resembled English norms, which involve longer VOTs than those of Portuguese. However, L2 proficiency did not significantly influence steepness of the categorization slope, indicating that while the location of the perceptual boundary becomes more target-like with increased proficiency, the sharpness of the category transitions remains stable.

The proficiency-related shift in crossover points, but not in slope steepness, suggests that listeners adjust their cue weighting in their perception of L2 English, a process that may be understood in terms of the Automatic Selective Perception (ASP) framework (Strange, 2011). In the present study, the patterns observed for L2 English listeners indicate that higher-proficiency learners have begun to develop L2 cue-specific routine: they attend more closely to the long-lag VOT cues that are critical for distinguishing English voiceless stop categories, thereby shifting their

revistadelta
42.1
2026

perceptual boundary toward more native-like values. In Strange's terms, this reflects the emergence of L2 cue-specific perceptual routines, indicating that learners are gradually learning to detect and rely on certain acoustic-phonetic cues that best differentiate target-language categories. However, the absence of proficiency effects on the sharpness of category boundary (slope steepness) indicates that these routines have not yet reached the level of automaticity associated with native-like categorical precision.

For RQ2 (*Does the acquisition of English as L2 induce phonetic drift in the perception of stop consonants in L1 Portuguese?*), the results indicate that it does. In the Portuguese mode, the crossover points of L2 English learners differed significantly from those of monolingual BP speakers, indicating that the two groups relied on distinct VOT boundaries when processing speech in their L1 Portuguese. More specifically, L2 learners demonstrated crossover points shifted further along the VOT continuum and closer to English norms than those of monolingual BP listeners. Although this type of perceptual flexibility is often associated with highly proficient bilinguals (García-Sierra et al., 2009; Gonzales & Lotto, 2013), this study findings show that even L2 learners in foreign-language learning contexts, with limit and predominantly formal input, can demonstrate sensitivity to cross-linguistic VOT differences, at least with respect to category boundary placement.

Despite significant differences in crossover points between the L2 English and monolingual BP groups, the two groups did not differ significantly in the steepness of the VOT category slope. That is, categorization sharpness remained stable across language groups demonstrating that both L2 English learners and monolingual BP speakers showed similar robust categorical patterns when processing sounds in the Portuguese mode. This dissociation between boundary location and slope broadly resonates with accounts of selective phonetic drift, according to which L2 experience may shift certain properties of L1 categories without uniformly altering category robustness or precision. For example, Chang (2018) showed that, in speech production, L1 phonetic drift can affect the weighting of specific acoustic cues, such as increased reliance to vowel-to-consonant transition information, while leaving other aspects of phonetic categories largely unchanged. Although Chang's study focused on production rather than perception, the selective nature of the observed L1 adjustments parallels the present findings, suggesting that category location may drift even when categorical sharpness remains intact.

The result for RQ3 (*Do L2 learners demonstrate language-specific perceptual categorization of VOT when processing stop consonants in L1 and L2 contexts?*) point to partial language-specific perceptual category in L2 learners. Although the crossover point did not differ across Portuguese and English processing modes, indicating a shared category boundary, listeners demonstrated significantly different slope values across the two languages. The steeper slope in Portuguese mode suggests more categorical, native-like L1 perception, whereas the flatter slope in English mode indicates greater uncertainty and a wider perceptual transition region.

Although crossover points remained stable across Portuguese and English processing modes, suggesting a shared or partially merged boundary, slop steepness varied significantly, with steeper slopes in Portuguese mode and flatter slopes in English mode. This pattern aligns with Grosjean's (2001, 2013) language-mode framework, which predicts that bilingual speech processing varies as a function of contextual language activation. In Portuguese mode, greater activation of the L1 facilitates the deployment of highly automatized L1 perception routines, resulting in more categorical, native-like responses. In contrast, English mode induces greater co-activation and competition between languages, leading to reduced automaticity and increased perceptual uncertainty. From the perspective of the Automatic Selective Perception model (Strange, 2011), this dissociation reflects differential engagement of perceptual routines rather than separate phonological representations, revealing language-specific processing strategies without fully language-specific category restructuring.

Across analyses, an asymmetry emerged between crossover points and VOT slopes, which capture distinct dimensions of phonetic categorization. While crossover points reflect the placement of category boundaries along the VOT continuum, slope steepness indexes how sharply listeners transition between categories. The results indicate that L2 phonetic perception involves both boundary adaptation and transition modulation. Crossover points shifted in two instances: as a function of L2 English proficiency in the English mode and in comparison with monolingual BP speakers in the Portuguese mode, providing evidence of phonetic drift in L1 perception. However, across L2-L1 language modes, learners maintained similar crossover points while showing systematic differences in VOT slope steepness, indicating that language context primarily modulates the deployment of phonetic categories rather than boundary location. These findings. Suggest that slope modulation may function as a strategic mechanism in L2 speech perception.

Finally, it is important to note that the stimuli in this study extended up to +35 ms, which may not fully capture the prototypical /p/ in native English. While this VOT range is likely sufficient to elicit clear categorical boundaries in the Portuguese mode, it may be insufficient to produce a fully realized categorical response in English. However, this study does not aim to model the complete native VOT range. Rather, its focus lies in examining whether language mode or L2 proficiency influences L2 learners' categorization of stops within a constrained VOT window. In this respect, the experimental design aligns with the study's objectives.

Takin together, the present findings challenge rigid distinctions between L1 and L2 phonetic systems, instead supporting an integrated view in which bilingual perceptual systems are dynamically organized. These effects emerge even in foreign-language learning environments characterized by limited and predominantly formal exposure, underscoring the extent to which L2 experience can reshape the perceptual architecture of speech.

Acknowledgments

I would like to thank Dr. Kalim Gonzales and Dr. Andrew J. Lotto for generously providing the stimuli used in this study, the reviewers for their invaluable feedback, and the participants for their enthusiastic involvement with this research.

Conflicts of interest

The author declare they have no conflict of interest.

Data available

The data used in this study are available in the article.

References

Abramson, A. S., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross language tests. In B. Hália, M. Romportl, & P. Janota (Eds.), *Proceedings of the 6th international congress of phonetic sciences* (pp. 569–573). Academia.

Caramazza, A., Yeni-Komshian, G., Zurif, E., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in

French-English bilinguals. *Journal of the Acoustical Society of America*, *54*(2), 421–428. https://doi.org/10.1121/1.1913594

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007

Chang, C. B. (2018). Perceptual attention as the locus of transfer to nonnative speech perception. *Journal of Phonetics*, *68*, 85–102. https://doi.org/10.1016/j.wocn.2018.03.003

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, *61*, 30–47. https://doi.org/10.1016/j.wocn.2017.01.001

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of bilingual classification. *Journal of Phonetics*, *15*(1), 47–65. https://doi.org/10.1016/S0095-4470(19)30537-6

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. 229–273). York Press.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*(3), 185–202. https://doi.org/10.1016/0167-6393(87)90025-2

García-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, *51*(4), 369–378. https://doi.org/10.1016/j.specom.2008.11.005

Gonzales, K., & Lotto, A. J. (2013). A Bafri, un Pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, *24*(11), 2135–2142. https://doi.org/10.1177/0956797613486485

Gorba, C., & Cebrian, J. (2021). The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. *Journal of Phonetics*, *88*, Article 101094. https://doi.org/10.1016/j.wocn.2021.101094

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages* (pp. 1–22). Blackwell.

Grosjean, F. (2013). Bilingual and monolingual language modes. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Blackwell Publishing.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892. https://doi.org/10.1515/LING.2010.027

Kellogg, J., & Chang, C. B. (2023). Exploring the onset of phonetic drift in voice onset time perception. *Languages*, *8*(1), Article 78. https://doi.org/10.3390/languages8010078

Kupske, F. F., & Lima Jr., R. M. (2022). Integrating into L2-dominant contexts and L1 phonological adaptability: An analysis of the production of Brazilian Portuguese voiceless stops. *Organon*, *37*(73), 173–198.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*(5), 320–331. https://doi.org/10.1016/j.wocn.2013.06.002

López, V. G. (2012). Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research*, *28*(2), 243–263. https://doi.org/10.1177/0267658312439821

Mack, M. (2003). The phonetic systems of bilinguals. In M. T. Banich & M. Mack (Eds.), *Mind, brain, and language: Multidisciplinary perspectives* (pp. 309–350). Lawrence Erlbaum Associates.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, *76*(2), 190–208. https://doi.org/10.2307/329772

Mateus, M. H., & d'Andrade, E. (2000). *The phonology of Portuguese*. Oxford University Press.

Muñoz, C. (2010). On how age affects foreign language learning. *Advances in Research on Language Acquisition and Teaching*, 39–49.

Osborne, D. M., & Simonet, M. (2021). Foreign-language phonetic development leads to first-language phonetic drift: Plosive consonants in native Portuguese speakers learning English as a foreign language in Brazil. *Languages*, *6*(4), Article 177. https://doi.org/10.3390/languages6030112

Peirce, J. (2014). PsychoPy (Version 1.80.01) [Computer software]. http://www.psychopy.org/

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *25*(4), 421–436. https://doi.org/10.1006/jpho.1997.0051

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209–231. https://doi.org/10.1515/iral.1972.10.1-4.209

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, *39*(4), 456–466. https://doi.org/10.1016/j.wocn.2010.09.001

OnLine Supplementary Materials

Supplementary materials can be accessed at Open Science Framework (OSF) via the following anonymous link: https://osf.io/g4xqd/files/osfstorage?view_only=6b233fc6414d4c339f76530cd0232d90. All figures in this article use color palettes that are color-blind friendly.

Recebido em: 01.07.2023
Aprovado em: 27.01.2026

EDITORA GERAL: Marcia Veirano Pinto