



OPEN ACCESS

EDITED BY

Paola Escudero,
Western Sydney University, Australia

REVIEWED BY

Silke Hamann,
University of Amsterdam, Netherlands
Anne-Michelle Tessier,
University of British Columbia, Canada

*CORRESPONDENCE

Jennifer Cabrelli
✉ cabrelli@uic.edu

RECEIVED 31 March 2025

REVISED 05 December 2025

ACCEPTED 08 December 2025

PUBLISHED 12 February 2026

CITATION

Cabrelli J, Cruz J, Escalante Martínez J, Finestrat I and Luque A (2026) From L2 acquisition to L1 restructuring: phonotactics in perception and production. *Front. Lang. Sci.* 4:1603764. doi: 10.3389/flang.2025.1603764

COPYRIGHT

© 2026 Cabrelli, Cruz, Escalante Martínez, Finestrat and Luque. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

From L2 acquisition to L1 restructuring: phonotactics in perception and production

Jennifer Cabrelli^{1*}, Jason Cruz¹, John Escalante Martínez¹, Irene Finestrat² and Alicia Luque^{3,4}

¹Multilingual Phonology Laboratory, Department of Hispanic and Italian Studies, University of Illinois Chicago, Chicago, IL, United States, ²Department of Spanish and Portuguese, Northwestern University, Evanston, IL, United States, ³Department of Applied Language Studies, Nebrjia University, Madrid, Spain, ⁴Nebrjia Research Center in Cognition, Nebrjia University, Madrid, Spain

Introduction: Research on first-language (L1) perceptual attrition indicates that second-language (L2) learners can acquire syllabic structures that are dispreferred in the L1 and that such acquisition can yield L1 phonotactic restructuring in phonetic and phonological processing and production. In the current study, we examine the production of coda stops in the L1, Brazilian Portuguese, and the adult L2, English, of a bilingual sample of speakers in an L2 immersion environment. Our first objective is to determine how these bilinguals' languages interact across perception and production. Specifically, we address the following questions within and between languages, respectively: (1) to what degree of accuracy do these speakers produce coda stops in the L2, and does L2 perception accuracy predict L2 production accuracy?; (2) to what degree of accuracy do these speakers produce coda stops in the L1, and does L1 perception accuracy predict L1 production patterns?; (3) does L2 production accuracy predict L1 production patterns? Our second objective is to model potentially asymmetric perception and production relationships in the L1 and L2 after extensive L2 exposure while accommodating variable production patterns within and across speakers.

Methods: Fifteen adult bilinguals completed a syllable concatenation task in both languages in which they concatenated disyllabic forms from monosyllabic nonce-word pairs. Productions were coded for coda stop realization and repair strategies. Production data were analyzed using mixed-effects logistic regression models, and previously published ABX perception data were used to examine perception-production relationships within and across languages.

Results: Participants reliably produced a syllabic target free from epenthesis in English and did so 66 percent of the time in Portuguese. However, they avoided coda stops in 19% of L2 productions and 46% of L1 productions. In cases of coda avoidance, speakers largely favored epenthesis of the coda stop, followed by palatalization and deletion.

Discussion: Perception accuracy did not predict production accuracy in either language. In contrast, second-language production accuracy predicted first-language production patterns. To model the speakers' asymmetric comprehension and production grammars, variable coda repair strategies, and the variable relationship between the grammars over time, we adopt the Bidirectional Phonetics and Phonology framework.

KEYWORDS

crosslinguistic influence, formal phonology, L1 restructuring, language attrition, perception-production link, perceptual epenthesis, phonotactics, second language acquisition

1 Introduction

Bilingualism research has increasingly shown that acquiring a second language (L2) can reshape speakers' first language (L1) (see Schmid and Köpke, 2019). This influence arises from the dynamic interaction between a bilingual's languages, affecting both linguistic representations and processing (e.g., de Leeuw and Chang, 2023). In phonological acquisition, this interaction is particularly evident when the L2 permits phonotactic structures that are restricted or illicit in the L1, potentially leading to restructuring of the L1 phonological system (Celata, 2019).

For example, native speakers of Brazilian Portuguese (BP) acquiring English as an L2 must navigate phonotactic differences between the two languages, such as the status of coda stops. While coda stops are permitted and frequent in English, they are dispreferred in BP and often subject to repair. The most common repair strategy is vowel epenthesis (see e.g., Quintanilha-Azevedo, 2016), which syllabifies the stop in onset position, such as coda /p/ in *optar* "to opt" (/o.p.tar/ vs. /o.pi.tar/). Prolonged exposure to a more permissive phonotactic system such as that of L2 English may induce restructuring of the L1 grammar, raising key questions about how bilinguals (variably) reconcile these differences in perception and production, and what broader effects crosslinguistic interaction may have on their L1 (Alcorn, 2018; Cabrelli et al., 2019).

These questions intersect with a broader body of research on the perception–production relationship in bilingual speech (see Nagle and Baese-Berk, 2022, for an overview). A central issue is whether perception precedes production, and how this relationship unfolds across languages. In this study, we examine coda stop production and perception in both L1 BP and L2 English among bilinguals immersed in an L2-dominant environment. Our specific objectives are to assess whether L2 perception predicts L2 production, whether these relationships extend to the L1, and whether the bilinguals' grammars show parallelism or asymmetry across the two languages. We further aim to determine how individual variability modulates these trajectories, conditioning the degree of convergence or divergence observed within and across languages.

We present data from a concatenation task that captures bilinguals' production strategies and evaluate individual variability in perception and production relationships. The perception data come from the same participants and were previously reported in Cabrelli et al. (2019); here, we reanalyze those data alongside the new production data to examine perception–production links. Our results indicate that while perception did not predict production in either language, L2 production accuracy influenced L1 production. Participants employed diverse repair strategies, highlighting substantial individual variability mismatches between perception and production. To further examine these outcomes and formally illustrate the interlanguage grammars that could give rise to the attested mappings, we use Parallel Bidirectional Phonology and Phonetics (BiPhon; Boersma, 2011), a stochastic Optimality Theory framework that allows us to model L1 and L2 perception and production within a single phonological system. Simulations trained on participants' input-output distributions via the Gradual Learning Algorithm (GLA; Boersma and Hayes, 2001) conceptually demonstrate how probabilistic constraint reranking can yield both categorical and variable mappings across modalities. While not

intended to generate new empirical predictions, these simulations clarify how a unified grammar can give rise to asymmetrical but systematically related perceptual and productive patterns.

This article is structured as follows: we first review relevant literature on bilingual phonological adaptation and theoretical models of perception–production links. We then describe our methodology, followed by an analysis of production data in both languages and their relationship to perception. Finally, we discuss implications for bilingual phonological models and propose directions for future research.

2 Cross-linguistic influence in perception and production

The relationship between perception and production in bilingualism has long been debated, with recent work extending this discussion to L1 attrition. We adopt Schmid and Köpke's (2019) inclusive definition of L1 attrition as encompassing "all L1 phenomena that stem from the co-activation of languages, crosslinguistic transfer, or disuse, at any stage of L2 development and use" (pp. 637–638). In line with this view, we use *attrition* and *restructuring* interchangeably to describe phonological changes to the L1 after L2 onset. Rather than implying loss, we interpret these changes as outcomes of representational restructuring, consistent with models that treat acquisition and attrition as parallel processes governed by shared cognitive mechanisms (e.g., Schmid and Köpke, 2017; Kubota, 2019; Opitz, 2011).

Consistent with this view, we examine restructuring as it manifests in both perception and production of L2 English and L1 BP word-medial coda stops. In perception, restructuring is reflected in increased mapping of auditory inputs (e.g., [p] in *optar* "to opt") to surface forms containing a coda stop (o/p./tar), rather than to a stop syllabified as an onset followed by an epenthetic vowel (o.pi.tar). In production, it is reflected in reduced reliance on canonical BP repair strategies (e.g., epenthesis) or, in the absence of epenthesis, shifts in repair type (e.g., coronal palatalization, deletion). We interpret such productions not as signs of loss, but as indications of a grammar in transition that is shaped by bilingual experience and evolving phonotactic representations.

To contextualize these modality-specific outcomes, we turn to recent models of bilingual speech that formalize the perception–production relationship. While the revised Speech Learning Model (SLM-r; Flege and Bohn, 2021) assumes a strong but imperfect bidirectional connection, Nagle and Baese-Berk (2022) highlight that this link varies across developmental stages. de Leeuw and Chang (2023) Attrition and Drift in Access, Perception, and Production Theory (ADAPPT) formalizes this relationship in L1 attrition, arguing that perception and production are separate due to differential engagement of domain-general cognitive mechanisms and distinct L1/L2 developmental trajectories. Principle 4 of ADAPPT states that accurate perception is neither necessary nor sufficient for accurate production, a claim supported by studies showing bilinguals who perceive L2 contrasts or structures accurately but fail to produce them with the same accuracy (and vice versa) (e.g., de Leeuw et al., 2021; Gorba and

Cebrian, 2021; Kim and Han, 2022)¹. A further relevant account is the Second Language Linguistic Perception (L2LP) model (Escudero and Boersma, 2004; Escudero, 2005; van Leussen and Escudero, 2015; see Escudero and Yazawa, 2024, for a recent overview of the L2LP and its research program). L2LP posits distinct perceptual grammars for each language, with perception as the driver of development: production accuracy is assumed to emerge indirectly from successful perceptual mappings established through cue-constraints. In addition, the model incorporates the role of activation (or perception modes) in shaping which grammar(s) is/are engaged at a given moment (e.g., Yazawa et al., 2020). While the model also extends to lexical encoding, our nonce-word design engages it at the pre-lexical level, where convergence and divergence across modalities can arise because perceptual restructuring does not always translate directly into production outcomes. This emphasis on perception-driven development and separate perceptual systems highlights how perceptual restructuring may advance more rapidly than production accuracy and why convergence is not guaranteed.

Because our data capture a synchronic snapshot of bilingual grammars, we formalize these modality-specific outcomes using Optimality Theoretic Parallel Bidirectional Phonology and Phonetics (BiPhon; Boersma, 2011; see Section 6 for technical implementation). BiPhon models perception and production within a single grammar, where the structure of the constraint hierarchy determines, at a given evaluation point, whether the two modalities align (e.g., Alcorn, 2018) or diverge (Gorba and Cebrian, 2021). Although we do not directly capture real-time perception–production dynamics, we treat outcomes on both tasks—ABX discrimination data reported in Cabrelli et al. (2019) and concatenation production data reported in Section 9—as reflections of a shared constraint grammar. We return to this assumption in Sections 6 and 10.

3 Our test case: word-medial coda stops

Heterosyllabic consonant sequences ($C_1.C_2$) are rare in BP, comprising only ~2% of written words (Silveira, 2007) and ~22% of spoken data (Monareto, 2017). Among C_1 stop- C_2 stop sequences in BP and English alike, /p.t/ and /k.t/ are the most frequent while /b.g/ and /g.b/ are the least frequent (Brants and Franz, 2006; Estivalet, 2014–) and /p.k/ and coronal-initial sequences are unattested in BP (Estivalet, 2014–; see Supplementary materials for bigram frequencies in BP and English).

Because word-medial stop codas are dispreferred in BP, monolingual speakers often perceive and produce epenthetic vowels that resyllabify the stop as an onset. Word-medial epenthesis in monolingual perception reaches ~80% (Parlato-Oliveira et al., 2010), while production rates vary from ~35% (Quintanilha-Azevedo, 2016) to 88% (Cristófar-Silva and Almeida, 2008), with minimal regional differences reported (e.g., de Lucena and Alves, 2010).

¹ In evaluating these models, we operationalize “accurate” outcomes as perception or production outputs that reflect a stop in coda position.

Phonetic realizations of word-medial coda stops in monolingual BP vary based on proximity to the syllabic target (i.e., maintaining a coda consonant) and the segmental target (i.e., producing a coda stop), as illustrated via the example *ritmo* “rhythm.”

- **(Un)released stop** ([rit⁺.mo]/[rit.mo])
 - Approximates **both** the segmental and syllabic targets.
 - Unreleased stops are rare (~1%, Quintanilha-Azevedo, 2016, pp. 129–140).
- **Palatalized stop** ([ritʃ.mo])
 - Approximates **only** the syllabic target.
 - Limited to C1 coronal (/t, d/).
- **Vowel epenthesis** ([ri.ti.mo]/[ri.ti.mo])
 - Approximates **neither** target.
 - The epenthetic vowel may be **voiced or voiceless**, depending on voicing of the surrounding consonants. In one study, epenthesis occurred in 56% of sequences where both C1 and C₂ were voiceless stops, and 83% of those epenthetic vowels were voiceless (Quintanilha-Azevedo, 2016, p. 128).
- **Palatalized stop + vowel epenthesis** ([ri.tʃi.mo]/[ri.tʃi.mo])
 - Approximates **neither** target.
 - Represents the most phonotactically distant realization from the L2 target structure.

Multiple linguistic variables influence the rate of epenthesis in word-medial heterosyllabic C_1 stop- C_2 stop sequences, including manner of articulation (MOA), place of articulation (POA), voicing, and stress.

- **MOA:** Epenthesis is less frequent in stop-stop sequences than in stop-fricative and stop-nasal sequences (Quintanilha-Azevedo, 2016; Collischonn, 2003). In Monareto (2017) analysis of the VARSUL corpus, epenthesis occurred in 31% of stop-stop sequences, compared with 62% in stop-fricative and 61% in stop-nasal sequences.
- **POA:** Reported epenthesis rates vary across studies, but coronals typically show higher epenthesis rates than labials and dorsals in corpus data (Monareto, 2017), though read-aloud studies yield more variable patterns (Alcorn, 2018; Quintanilha-Azevedo, 2016).
- **Voicing:** Voiced stops are more likely to trigger epenthesis than voiceless stops (e.g., Alcorn, 2018), likely due to their marked status and the facilitation of vocalic production from vocal fold vibration.
- **Stress:** Epenthesis is more frequent in pretonic position (69% pretonic vs. 24% post-tonic in Collischonn, 2004; 74% vs. 33% in Alcorn, 2018). This is because post-tonic epenthesis results in antepenultimate stress, requiring a non-canonical foot structure (Hermans and Wetzel, 2012).

4 Acquisition of L2 English coda stops

Unlike BP, English permits all stops in coda position, and L1 English speakers do not perceive or produce epenthetic vowels after a coda stop (Alcorn, 2018; Cabrelli et al., 2019). The acquisition task for L1 BP learners of L2 English is to categorically license coda stops, requiring them to restructure their phonotactic representations. If such restructuring occurs in L2, we ask whether it extends to the L1 as well.

Epenthesis is the most common repair strategy among L1 BP speakers acquiring L2 English due to initial transfer of the L1 grammar (Alcorn, 2018; Cardoso, 2007; John and Cardoso, 2017), though other strategies such as deletion (Alves et al., 2008; Nascimento, 2019), aspiration (Cardoso, 2011), and palatalization (Alves et al., 2008; Bettioni-Techio, 2005) have been observed. These strategies vary by proficiency level, with epenthesis more common in early acquisition and less frequent as L2 exposure increases (Alcorn, 2018).

The likelihood of epenthesis may be modulated by whether a given sequence is shared between BP and English. Schneider (2009) posits that L2 epenthesis is more frequent in sequences found in both languages than in sequences found only in English, as bilinguals must revise entrenched L1 representations rather than develop entirely new ones. The revision of these representations, assumed to be error-driven and stochastic, may be slower for L2 sequences that have existing L1 epenthetic forms than for novel L2 sequences that are mapped directly onto new structures. In cases where learners lack stored epenthetic representations, markedness effects may emerge, leading to alternative repair strategies that are neither L1- nor L2-like (Eckman, 2008). These include deletion (Alves et al., 2008; John and Cardoso, 2017), spirantization (Flege and Davidian, 1984), devoicing (Broselow, 2018), aspiration (Cardoso, 2011), and palatalization of coronal stops (Alves et al., 2008; Bettioni-Techio, 2005). The presence of these patterns highlights the role of both universal linguistic principles and language-specific phonotactic constraints in shaping L2 repair strategies.

Alcorn's (2018) study of L1 BP/L2 English bilinguals in the U.S. and Brazil is the only to report word-medial stop coda data from advanced L2 English learners. Alcorn found that U.S.-based bilinguals produced significantly fewer epenthetic vowels (26.5%) than those in Brazil (37.8%) and a monolingual BP comparison group (74%). Epenthesis was predicted by self-rated L2 proficiency, place of articulation (/d/ and /g/ favored epenthesis, /b/ and /t/ disfavored it), stress (pre-tonic favored over post-tonic), and C2 voicing. A comparison of perception data and production data showed a negative correlation between perceptual acuity (higher d' sensitivity) and epenthesis rate, consistent with other findings that perception improves faster than production (Cardoso, 2011). However, studies on other L1–L2 pairs (Shin and Iverson, 2014) have found no correlation, highlighting the potential role of individual variability in bilingual perception-production relationships.

5 Restructuring of L1 BP coda stops

Once learners shift toward the English target grammar—licensing L2 coda stops—we want to know whether this shift extends to their L1 grammar. To inform our predictions and study design, we draw on research on L1 BP/L2 English bilingual production and the perception-production relationship within/across languages.

Two studies have examined attrition of L1 BP epenthesis: Alcorn (2018) and Cabrelli et al. (2019). Alcorn (2018) tested bilinguals in the U.S. and Brazil on illusory vowel perception via a forced-choice identification task with nonce words in BP and found both groups' d' sensitivity scores fell between BP and English monolinguals, indicating partial restructuring. In a sentence reading task with nonce words, bilinguals were more likely than BP monolinguals to produce coda stops, and epenthesis was predicted by C2 voicing and pre-tonic position. Place of articulation effects showed that C1 /b/ and /g/ and C2 /d/ favored epenthesis, while C1 /p/ and /k/ disfavored it. However, the statistical models that included linguistic variables as predictors did not include group as a predictor, leaving open whether L2 exposure modulated these patterns.

As with the perception data, between-group comparisons showed no significant differences between the bilingual groups, yet bilinguals overall were more likely than BP monolinguals to produce coda stops. A significant negative correlation between L1 BP epenthesis rates and L2 proficiency was observed, aligning with trends found in L2 English production. Additionally, perception and production were linked within each language, and L2 and L1 perception/production were positively correlated, reinforcing a crosslinguistic relationship between bilinguals' phonotactic grammars.

The perception data analyzed in the present study are the same data published in Cabrelli et al. (2019), collected from the participants whose production data we analyze here. Cabrelli et al. examined bilinguals' metalinguistic knowledge, phonetic processing, and phonological processing. Results demonstrated that L1 BP and L2 English coda stop perception exceeded the BP monolingual threshold observed in Dupoux et al. (2011), with perception remaining plastic into adulthood. Of particular interest to our study is an ABX phonological processing task, which most closely matches our production task in its reliance on surface phonological representations. Of particular interest to our study is an ABX phonological processing task reported in Cabrelli et al., which most closely matches our production task in its reliance on surface phonological representations. It is these ABX task data that we use in our analyses of perception–production relationships herein. These phonological processing data revealed a positive correlation between L2 and L1 d' sensitivity [$r(13) = 0.51, p = 0.055$], indicating that greater L2 discrimination of VC.CV and V.Ci.CV nonce pairs predicted stronger L1 discrimination.

Taken together, Alcorn (2018) and Cabrelli et al. (2019) suggest that L2-induced L1 phonotactic restructuring is evident in perception. Alcorn's data indicate that this restructuring extends to production, at least in a lexical reading task and that there is a perception-production link despite substantial task differences.

The present study examines whether similar restructuring emerges in pre-lexical perception and production of nonce words and to provide a theoretical account of the observed patterns, which we present in the following section.

6 Modeling bilinguals' L1 and L2 perception and production grammars

To illustrate how bilingual grammars can yield the observed mappings across perception and production, we use BiPhon (Boersma, 2011), a framework that models both modalities within a unified constraint system. This section provides a concise overview of the BiPhon architecture and its stochastic implementation, focusing on how constraint reranking and cue interaction conceptually account for the variable and sometimes asymmetrical patterns that emerge in the empirical data.

Though the modeling does not generate independent empirical findings, it supports interpretation of observed interlanguage patterns by illustrating how a unified probabilistic grammar can yield asymmetrical and variable mappings across modalities.

6.1 Optimality theoretic bidirectional phonetics and phonology (BiPhon)

BiPhon is a model of grammar and bidirectional processing that assumes a unified constraint ranking for both comprehension and production, allowing us to model bilingual phonological restructuring holistically.

The *comprehension* module of BiPhon (bottom-up, Figure 1) consists of *phonological perception* (mapping auditory input to abstract surface forms) and *recognition* (mapping surface forms to stored underlying forms). Since our study examines nonce words without lexical representations, we focus solely on phonological perception. Perception involves phonetic forms comprising auditory cues (e.g., formants, bursts, durations), evaluated against cue constraints (CUE) and structural constraints (STRUCTURE).

Cue constraints were first introduced in Escudero and Boersma, 2003, 2004 and elaborated in Escudero (2005, 2009) as the driver of perceptual development in L2 learning. In practice, they take the form “[x] (phonetic form) cannot map onto /y/ (surface form),” and their ranking determines the surface form to which acoustic input is mapped. While cue constraints operate over continuous auditory features rather than discrete symbolic segments, for clarity and brevity, we use IPA symbols in square brackets (e.g., [t], [i]) as readable shorthand for the characteristic acoustic profiles associated with those segments. For example, [t] indicates a transient burst without formant structure, and [i] indicates high F2/F3 formant values. We also use the symbol C(stop) as a cover term to represent the class of oral stops ([p, t, k, b, d, g]); on the phonological side, /C(stop)/ refers to any stop consonant, and on the phonetic side, [C(stop)] stands for the range of acoustic cues typically associated with stop bursts. To illustrate, the interaction of the cue constraints *[C(stop)]/C(stop)./ (a burst-like cue cannot map onto a coda stop; here, the full stop indicates coda position) and *[]/i/ (absence of acoustic energy cannot be parsed as /i/) determines whether

epenthesis occurs in phonological perception.² Readers interested in more fine-grained auditory cue representations, such as the decomposition of [tʃ] into [burst + noise], may refer to Quintanilha-Azevedo (2016, pp. 185–190) for a detailed implementation of that approach.

Because the mapping from continuous acoustic input to phonological form is not categorical, it requires a grammar that permits variable outcomes. Stochastic Optimality Theory (OT) provides one such framework, allowing variation to emerge from interactions among ranked constraints.

6.1.1 Variable grammars

In stochastic OT, constraints occupy a continuous ranking scale, and constraint ranking values are perturbed with random noise during evaluation, resulting in probabilistic output selection. When ranking values of relevant constraints are sufficiently close, this can yield output variation; when they are widely separated, outputs are categorical. To model variable monolingual BP perception, we used the Gradual Learning Algorithm (GLA; Boersma and Hayes, 2001) in Praat (v6.4.13, Boersma and Weenink, 2021) to derive a grammar that approximates the input-output distribution reported in Dupoux et al. (2011, p. 206), where 42% of mappings of the input [C(stop)] exhibit epenthesis (/C(stop)i./) and 58% do not (/C(stop)./). In this case, the GLA is not simulating an individual speaker's error-driven learning but rather converging on a constraint ranking that produces a grammar consistent with observed perceptual variability.

As illustrated with the nonce test item *latpa* in Figure 2, overlapping constraint distributions and fluctuating disharmonies in Evaluate A vs. Evaluation B explain the variable outputs. In contrast, monolingual English grammars, in which [C(stop)] is consistently mapped to a coda stop, exhibit categorical behavior (see Supplementary materials).

6.2 L2 development and L1 change

6.2.1 Phonological perception in BiPhon

In line with the Full Transfer/Full Access hypothesis (Schwartz and Sprouse, 1996), we assume that the initial state of an L2 grammar is a copy of the L1 grammar. Specific to the perceptual grammar, we follow e.g., Escudero (2005), who posits “full copying” (Escudero and Boersma, 2004) of the L1 perceptual grammar as the starting point of the L2 perceptual grammar. That is, L2 English learners begin with the BP constraint ranking and gradually rerank constraints based on L2 input. For an L1 BP speaker acquiring L2 English, this process involves promoting *[]/i/ while demoting *[C(stop)]/C(stop)./. Learners with ceiling L2 perceptual acuity (Cabrelli et al., 2019) exhibit constraint rankings mirroring English, whereas those with partial restructuring continue to exhibit variable mappings. Training the GLA with the bilingual sample's mean rate of coda stop perception (L2 English: 88%, L1 BP: 84%) yields a grammar where *[]/i/ dominates but whose proximity to the other constraints allows for epenthetic mappings 12% and 16% of the time, respectively (Supplementary materials).

² The symbol * marks a prohibited input-output mapping.

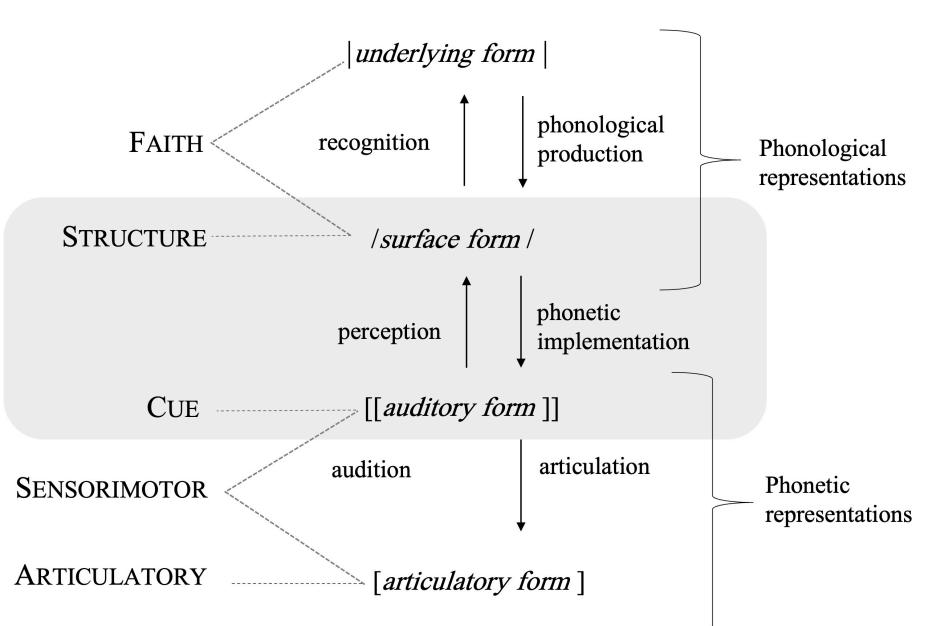


FIGURE 1
BiPhon (Boersma, 2011).

Evaluation A 100.30/ 99.70/ 98.14		
	[latpe]	*[]/i/ *[C(stop)]/C(stop)./
58%	☞ /lat.pa/	*
42%	/la.ti.pa]	*!
Evaluation B 99.70/ 100.47 99.74		
	[latpe]	*[C(stop)]/C(stop)./ *[]/i/
58%	/lat.pa/	*!
42%	☞ /la.ti.pa]	*

FIGURE 2
Variable outputs in perception due to overlapping constraint distributions (ranking values and disharmonies indicated by x/y; output distributions noted in %). Notation: * in evaluation cells marks a constraint violation; *! marks a fatal violation.

6.2.2 Production (phonetic implementation) in BiPhon

Optimal surface forms in perception serve as inputs to *phonetic implementation* (Boersma and Hamann, 2008), where cue constraints dictate input-output mapping of the/surface form/to [phonetic form]. Since monolingual BP hearers are not predicted to categorically perceive an epenthetic vowel, the same auditory input may be variably mapped to two distinct surface phonological forms: /C(stop)/ or /C(stop)i/. Each of these then serves as the input to phonetic implementation, resulting in corresponding variation in

production. Based on Quintanilha-Azevedo (2016), Table 1 details attested outputs (pp. 130–140) and the relevant cue constraint that each violates. While all coda stops in the study's auditory stimuli were released, we included a constraint penalizing unreleased coda stops to reflect the small number of unreleased realizations observed in Quintanilha-Azevedo (2016)³.

To illustrate phonetic implementation, consider a monolingual BP speaker who variably maps the same auditory input [t] to /t./ 42% of the time and to /t./ 58% of the time, reflecting variable perceptual mapping onto distinct surface forms. Suppose their production follows Quintanilha-Azevedo's BP sample of /tn/ and /tm/ clusters⁴.

$$\begin{aligned} [tʃi] \text{ (input /t./)} &= 21\% \\ [tʃi] \text{ (input /t./)} &= 2\% \\ [tʃ] \text{ (input /t./)} &= 77\% \end{aligned}$$

This distribution reflects the assumption that the same auditory input may be variably parsed in perception, yielding different surface forms—/t./ or /t./—which then serve as inputs to phonetic implementation. Using these distributions and the constraints in Table 1, we applied the GLA in Praat to model phonetic implementation for the nonce item *latpa*. The resulting evaluations (Figure 3) yield the categorical output [latpe] (input /lat.pa/) due to *[t]/t./> *[tʃ]/t./ and the highly probable [latʃpe] (input /la.ti.pa/) via domination of *[tʃ]/t./ over *[tʃi]/t./, which varies

³ Although cues to unreleased stops were not present in the experimental auditory input and are infrequent in the data, we acknowledge that the demotion of *[C(stop)]/C(stop)./ would be part of the broader learning trajectory for L1 BP learners acquiring English-like unreleased variants.

⁴ Quintanilha-Azevedo analyzed only two clusters with a voiceless coronal stop in C1 position: /tn/ and /tm/.

TABLE 1 Phonetic implementation output candidates for inputs /lat.pa/ and /la.ti.pa/ and corresponding cue constraint violation(s).

Realization	Output candidate	Violated cue constraint		Constraint definition
		Input /lat.pa/	Input /la.ti.pa/	
Released stop	[lat.p̚]	*[C(stop)]/C(stop)./		A phonetic form with burst-like cues characteristic of oral stops [C(stop)] cannot map to a coda stop /C(stop)./
Unreleased stop	[lat'.p̚]	*[C(stop)]/C(stop)./		A phonetic form with unreleased stop cues [C(stop)?] cannot map to a coda stop /C(stop)./
Deleted stop	[lap̚]	*[]/C(stop)./		A phonetic form with absence of acoustic energy [] cannot map to a coda stop /C(stop)./
Palatalized coronal	[latʃp̚]	*[C(aff)]/C(stop)./		A phonetic form with burst-plus-frication cues (i.e., an affricated burst) [C(aff)] cannot map to a coda stop /C(stop)./
Epenthetic [i]	[latip̚]		*[C(stop)i]/.C(stop)i./	A phonetic form with stop burst followed by high front formant structure (voiced) cannot map to a stop+i/ sequence /.C(stop)i./
Epenthetic voiceless [i]	[latip̚]		*[C(stop)v]/.C(stop)i./	A phonetic form with stop burst followed by high front formant structure (voiceless) cannot map to a stop+i/ sequence /.C(stop)i./
Palatalized coronal + epenthetic [i]	[la.t̚.p̚]		*[C(aff)i]/.C(stop)i./	A phonetic form with an affricated burst followed by high front formant structure (voiced) [C(aff)i] cannot map to a stop+i/ sequence /.C(stop)i./
Palatalized coronal + voiceless epenthetic [i]	[la.t̚ʃi.p̚]		*[C(aff)v]/.C(stop)i./	A phonetic form with an affricated burst followed by high front formant structure (voiceless) [C(aff)i] cannot map to a stop+i/ sequence /.C(stop)i./

As Quintanilha-Azevedo (2016) notes, an input/output pair in which a less marked input form such as /la.ti.pa/ is phonetically implemented as the more marked [latp̚] is illogical, as is its inverse input/output pair /lat.pa/-[latip̚].

with [latʃi p̚] due to the proximity of *[C(aff)i]/.C(stop)i./ and *[C(aff)v]/.C(stop)i./.

Since the comprehension and production modules in BiPhon share a single constraint ranking, it is logical to predict that phonological perception and phonetic implementation will align, as observed in Alcorn (2018). That is, phonetic implementation in the grammar of a participant with ceiling-level L2 English perceptual accuracy should categorically yield a coda stop (i.e., /latp̚/-[latp̚]). Yet, we will see in our results that this is not the case. BiPhon accounts for this asymmetry via the violation of cue constraints that overlap with the one violated by [latp̚] (*[t]/t./), as illustrated in Figure 3.

Predictions (P1) (based on Alcorn, 2018):

(P1a) Learners will produce L2 coda stops at ~74% accuracy, as observed in the US bilinguals in Alcorn (2018).

(P1b) Coronal stops will be produced with lower frequency than labials/dorsals due to palatalization as a repair strategy.

(P1c) L2 perception accuracy will predict L2 production accuracy, consistent with theoretical models that posit perception-driven or shared representational development (e.g., SLM-r, ADAPPT, L2LP) and supported by Alcorn's (2018) empirical finding.

Research Question 2 (L1 production):

(RQ2a) To what extent do L1 BP speakers accurately produce L1 coda stops?

(RQ2b) Which linguistic variables modulate L1 production?

(RQ2c) Does L1 perception accuracy predict L1 production accuracy?

Predictions (P2) (following Alcorn, 2018):

(P2a) Production rates will align with Alcorn's US bilinguals (~34% coda stop production).

(P2b) Coronal stops will be produced less frequently than labial/dorsal stops since affricates preserve stress.

7 Research questions and predictions

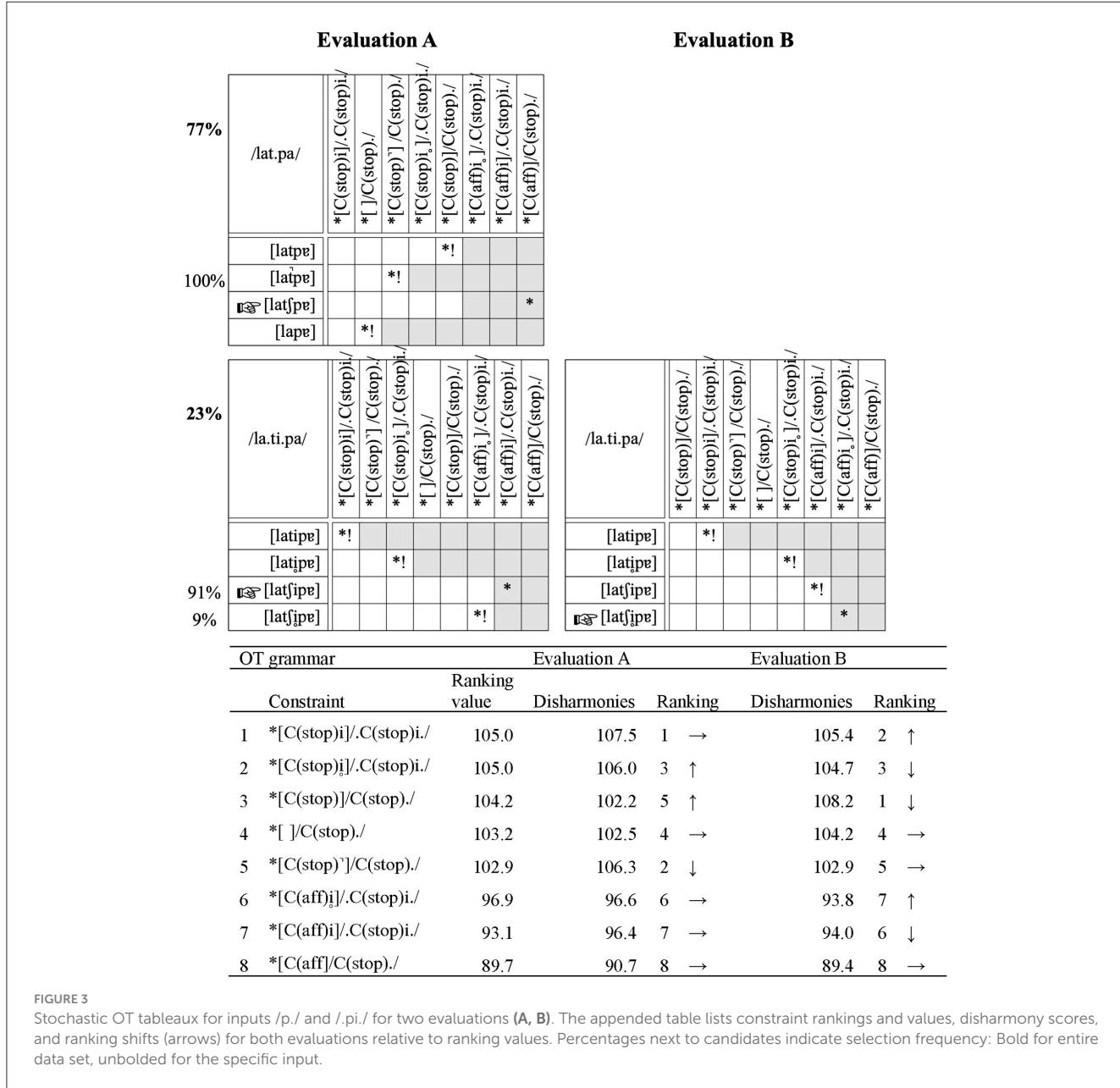
To investigate the acquisition and modification of phonotactic structure and the interplay between perception and production, we address the following research questions:

Research Question 1 (L2 acquisition):

(RQ1a) To what extent do L1 BP speakers accurately produce L2 coda stops?

(RQ1b) Which linguistic variables modulate L2 production?

(RQ1c) Does L2 perception accuracy predict L2 production accuracy?



(P2c) Post-tonic position will reduce epenthesis likelihood.
(P2d) L1 perception accuracy will predict L1 production accuracy.

Research Question 3 (L2–L1 relationship):

(RQ3) Does L2 production predict L1 production?
(P3) Given the L2–L1 perception relationship in Cabrelli et al. (2019) and Alcorn (2018), L2 production will predict L1 outcomes.

Research Question 4 (Individual variation):

(RQ4) How do participants' repair strategies for illicit coda stops vary between and within individuals, both in production and perception?

(P4) Participants will primarily rely on epenthesis, consistent with prior work (e.g., Alcorn, 2018; John and Cardoso, 2017). However, both between- and within-subject variability is expected, including perception-production mismatches. To illustrate how such patterns might reflect distinct underlying grammars or asymmetrical learning, we use BiPhon modeling to simulate selected participants' mappings using their own input-output distributions.

8 Methods and materials

8.1 Participants

The study included the same participants from Cabrelli et al. (2019): Fifteen L1 BP/L2 English university students (10 female) who had lived in a large urban center in the Midwest US for

TABLE 2 Participant characteristics.

Variable	<i>M</i>	<i>SD</i>
Age (years)	23.13	2.72
Age of first exposure to L2	13.93	2.60
Length of residence in US (years)	0.87	0.40
English proficiency score (0–50)	39.33	2.29
Language dominance [-218 (Portuguese-dominant) – 218 (English dominant)]	-79.19	22.16

~10 months at the time of testing. They primarily used English in university settings and BP elsewhere. Dominance was assessed via the Bilingual Language Profile (BLP; [Birdsong et al., 2012](#)), and English proficiency via a 50-point written measure adapted from the Oxford Placement Test. All participants had met the minimum TOEFL requirements for US university admission (≥ 550 or 525–549 with an intensive English course).

As shown in [Table 2](#), participants were BP-dominant, with minimal variance in age, L2 onset, length of residence, English proficiency, and BLP scores. Six were from Brazil's Southeast region, six from the Northeast region, and one each from the North, Central-West, and South regions.

8.2 Production task and stimuli

To assess coda stop production in the L1 and L2 after intensive English exposure, participants completed a concatenation task ([Wayland et al., 2006](#)) in BP and English on separate days. In each trial, they saw a carrier phrase (e.g., *Eu digo _____ para você*/I'm saying _____ to you) and heard two unstressed monosyllables (e.g., [lag], [ba]) separated by 500 ms. Based on the two isolated auditory inputs, participants were then prompted to produce them as a concatenated disyllabic nonce word embedded in the carrier phrase (e.g., I'm saying *lagba* to you).

Each block comprised 90 pseudorandomized trials plus three practice items. The English block featured penultimate stress (e.g., [lag.bə]). BP included two blocks, one with penultimate ([lag.bə]) and one with final stress ([lag.'bə]), with order counterbalanced across participants (see "Stimuli motivation" for rationale). Participants were explicitly instructed at the beginning of each BP block which stress pattern to use, and practice trials ensured compliance with the target stress pattern. All blocks comprised the same monosyllable pairs.

8.3 Task motivation

Comparing perception and production is inherently challenging ([Nagle and Baese-Berk, 2022](#)). We selected the concatenation task because, like the ABX perception task in [Cabrelli et al. \(2019\)](#), it (a) avoids orthographic input, (b) uses comparable disyllabic nonce words, and (c) minimizes reliance on auditory memory.

A concatenation task was preferred over alternatives (e.g., delayed repetition) as it allowed us to examine another phenomenon modulated by lexical stress (vowel reduction) without priming effects. While an ideal production task would exclude a perceptual component, auditory input was necessary to minimize orthographic influence, which enhances L2 consonant cluster accuracy (e.g., [Davidson, 2010](#); [Zjakic, 2017](#)). Tasks without perceptual or orthographic elements would require real words, which we avoided in favor of nonce words (see next section for details). To address potential task-related effects introduced by the perceptual component, the design was constructed to minimize such influences. Given the asymmetry in crosslinguistic epenthesis patterns reported in Section 9 despite stimuli with consistent acoustic cues across languages, we consider it unlikely that the perceptual component introduced such an effect. See Section 10.5 for further discussion.

8.4 Stimuli motivation

We selected nonce words to align with the perception data in [Cabrelli et al. \(2019\)](#) and to examine participants' productive syllable structure without interference from stored lexical representations. Word-medial stop codas were chosen for comparability with perception data and because L2 English speakers produce them more accurately than word-final codas ([John and Cardoso, 2017](#)). Since we aimed to capture L2 influence on L1 production, we focused on the word position where L2 stop codas are most likely to be acquired (see [Cabrelli, 2023](#)).

Coronal stops were excluded from [Cabrelli et al. \(2019\)](#) perception tasks due to their susceptibility to palatalization-induced epenthesis ([Quintanilha-Azevedo et al., 2017](#)). In production, however, we prioritized coronals to assess palatalization as an interlanguage repair strategy that conforms with target syllabic structure but not with segmental quality. Given the availability of palatalization, we predicted lower accuracy for coronal stops compared to dorsal and labial stops.

Learners produced disyllabic words in BP and English with penultimate stress to maintain comparability with perception data. This decision also stemmed from [John and Cardoso \(2017\)](#) findings that coda stops are more accurately produced in stressed penultimate syllables, thus increasing the likelihood of L2 acquisition and subsequent effects on the L1. Given the lower rate of epenthesis in penultimate-stressed items and the potential for learners to reach ceiling in both L2 and L1, we included a pre-tonic (final stress) block in BP to capture possible relics of L1 epenthesis in a more favorable context. This condition was limited to BP, as English prohibits final stress in words with heavy penultimate syllables, making such items unnatural.

See [Supplementary materials](#) for summary of the production stimuli in relation to the perception data from [Cabrelli et al. \(2019\)](#).

8.5 Stimuli description

Each block contained 20 critical items with a CV [C1_{stop}.C2_{stop}]V structure. C1 and C2 were hetero-organic

stops that matched in voicing and included all possible hetero-organic voiceless ($n = 10$) and voiced ($n = 10$) bigrams (see [Supplementary materials](#) for stimuli table and bigram frequencies). As noted in “Stimuli motivation,” 10 of the 20 items contained a coronal stop coda to assess participants’ use of palatalization.

The C1_{stop}.C2_{stop} clusters differ from the [Cabrelli et al. \(2019\)](#) perception stimuli, where C2 was a fricative (e.g., *abza*, *akfa*). Sonority-based predictions suggest that stop-stop clusters, forming a sonority plateau, should be less marked than rising-sonority (stop-fricative) clusters ([Quintanilha-Azevedo, 2016](#)), leading to lower epenthesis rates. However, Quintanilha-Azevedo’s data do not support the posited markedness effect and prior evidence from [Collischonn \(2002\)](#), which found greater epenthesis with fricative-initial clusters, is confounded by the exclusive analysis of pretonic items. We thus assume no substantive qualitative differences between the perception and production items due to sonority.

Monolingual BP and English monosyllables were recorded by the same phonetically trained speakers as in [Cabrelli et al. \(2019\)](#): BP stimuli were produced by a male L1 BP speaker from São Paulo, a near-native L2 English speaker who does not produce epenthetic vowels in the L2. English stimuli were recorded by a female L1 Midwestern American English speaker. All coda stops were released, with syllable offsets marked 10 ms after release. Although English exhibits variability in coda stop release (see [Davidson, 2011](#)), we opted for a consistent, released realization across stimuli to control the input signal.

Phonetic voicing status followed [Davidson \(2016\)](#): voiceless codas had <10% voiced frames during closure, voiced codas had >90%, and partially voiced codas (10%–90%) were recategorized based on voiced-frame distribution. Segments recategorized as phonetically voiced exhibited either (a) prevoicing, with more voiced frames in the final third of closure than the first third, or (b) a trough pattern, with higher voicing in the initial and final thirds than in the middle third. Of the 10 phonologically voiced codas in each language, two BP codas and one English coda were phonetically voiced. Phonological voicing was primarily cued by significantly longer duration of the preceding vowel in both English [$t_{(16)} = 9.06, p < 0.001$] and BP [$t_{(16)} = 6.67, p < 0.001$] (see [Supplementary materials](#) for means and SDs).

8.6 Procedure

Participants provided informed consent under the University of Illinois Chicago IRB (Protocol #2015-0040). Before the first of two in-person sessions, they completed the English proficiency measure and BLP online. The BP and English concatenation tasks were conducted in a sound-attenuated booth on separate days, with BP always tested first. Each session began with a 10–15-min interview to establish language mode. The task was run in E-Prime 2.0.10.356 (Psychology Software Tools, Inc.), with auditory stimuli delivered via a MOTU UltraLite-mk3 audio interface and AKG K240 MKII headphones. Speech was recorded using a Shure 10A head-mounted microphone and a Marantz PMD 661 steady-state recorder (44.1 kHz sampling rate).

8.7 Analysis

Of 900 items (600 BP, 300 English), 24 (2.67%) were excluded due to recording errors ($n = 2$), consonant or syllable epenthesis ($n = 3$; e.g., /map.ta/-[ma.pi.tʃi.te], /map.ka/-['mapt.ke]), syllable metathesis ($n = 9$; e.g., /map.ta/-['ta.map]), unnatural pauses ($n = 5$), or uncategorizable productions ($n = 5$; e.g., /mat.ka/-[ma.pe]).

To maximize data retention, we included items with incorrect stress ($n = 15$), recoding them accordingly, as well as items where C1 had a non-target POA ($n = 73$), C2 had a non-target POA ($n = 23$), or C2 was realized as a fricative ($n = 13$). The final dataset comprised 876 items: 586 in BP (296 with initial stress, 288 with final stress) and 292 in English.

8.7.1 Acoustic analysis and data coding scheme

Each item was analyzed in Praat ([Boersma and Weenink, 2021](#)) using separate TextGrid tiers to mark C1 onset and offset, C1 burst (if present), and any subsequent vocalic material. C1 onset was marked at the offset of clear formant structure and the last periodic pulse in the waveform, while C1 offset was placed at C2 closure onset.

C1 was coded as one of the following:

- **Released stop:** Closure with burst present ([Figure 4a](#)).
- **Unreleased stop:** Closure without burst.
- **Palatalization** (coronals only): Closure followed by aperiodic frication and/or concentrated frequency around 4,000 Hz ([Figure 4b](#)).
- **Lenition:** Frication, or periodicity with intensity consistent with an approximant.
- **Lengthened segment (gemination):** C1 or C2 replacing C1.C2, with homorganic POA confirmed via formant transitions.
- **Semivowel (diphthongization):** Clear formant structure, periodicity, and formant transitions consistent with [ai] (BP) or [aɪ] (English). This coding was used in cases where C1 weakened to a glide that resyllabified into the nucleus, forming a diphthong.
- **Deletion:** C1 or C2 absent.

Epenthetic vowels were marked from C1 release to C2 onset. Voiced vowels were identified by C1 release, clear formant structure, and waveform periodicity ([Figure 4c](#)), while voiceless vowels displayed aperiodic frication often extending from C1 or a lengthened offset with vocalic coarticulation ([Quintanilha-Azevedo, 2016](#)) ([Figure 4d](#)). All epenthetic vowels were confirmed auditorily.

The first coding phase was completed independently by four phonetically trained researchers (including the second and third authors), with each item double-coded. In the second phase, coding was confirmed by unanimous agreement among the first author and two additional phonetically trained researchers.

8.7.2 Statistical analysis

All analyses and visualizations were conducted in R (v4.2.2, [R Core Team, 2022](#)) with reproducible code available at <https://>

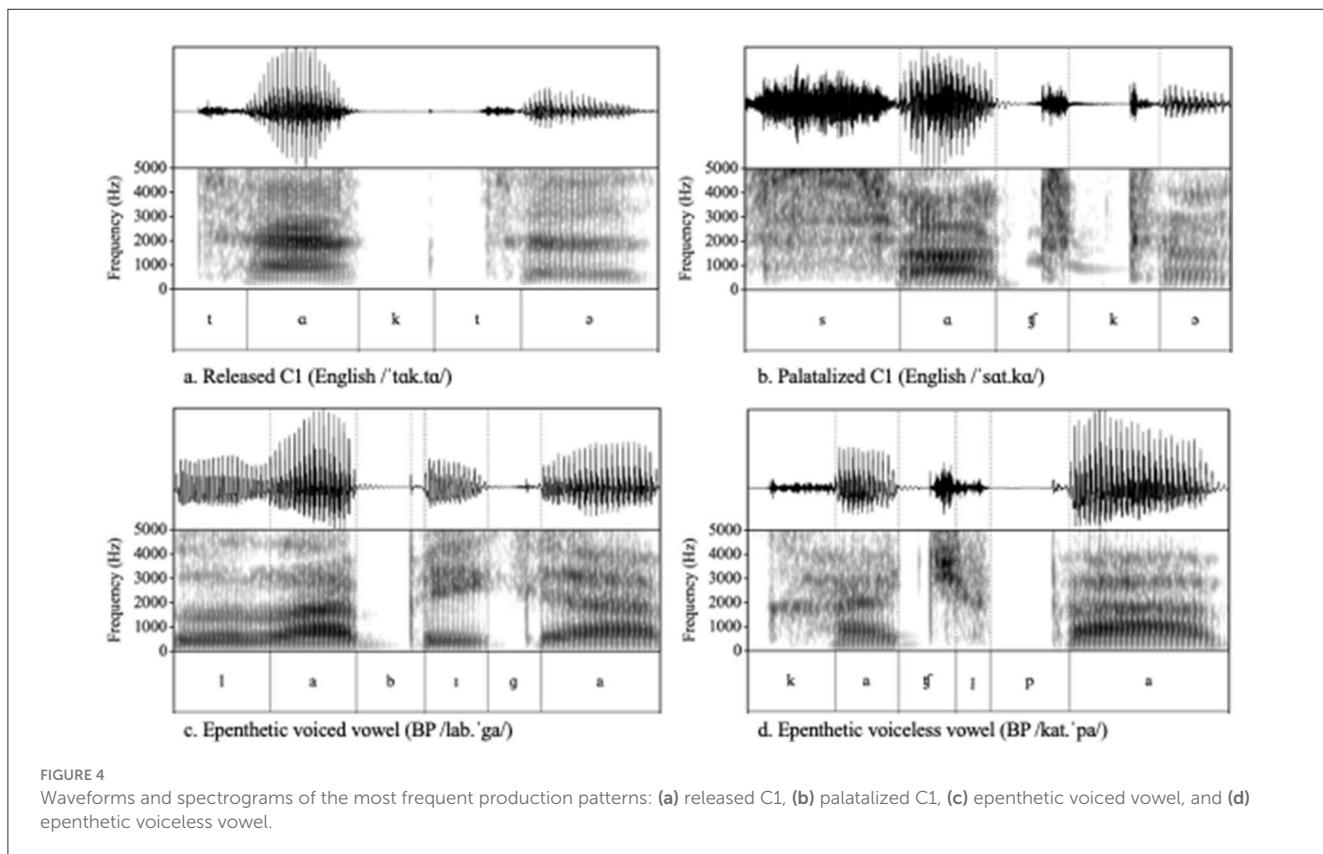


FIGURE 4

Waveforms and spectrograms of the most frequent production patterns: (a) released C1, (b) palatalized C1, (c) epenthetic voiced vowel, and (d) epenthetic voiceless vowel.

[/osf.io/zugb8/?view_only=b6ab9dc50b304154b9e62d4f4095f54f](https://osf.io/zugb8/?view_only=b6ab9dc50b304154b9e62d4f4095f54f).

Descriptive analysis included proportions and SDs of production types by Language and Language*POA, providing group-level patterns and repair strategy rates.

Inferential analyses addressed the research questions via mixed-effects logistic regression models using `glmer` (`lme4`, v1.1.31; [Bates et al., 2015](#)) and `buildmer` (v2.11; [Voeten, 2023](#)). Two binary dependent variables were examined:

- `syllabicTarget` (1 = C1 maintained in coda position or weakened without altering syllabic structure; 0 = C1 deleted or realized in onset via epenthesis), and
- `C1target` (1 = C1 produced as a released or unreleased stop in coda position; 0 = all other outcomes).

The `syllabicTarget` variable reflects preservation of the target syllabic structure and includes lenited, lengthened, and semivowel realizations of C1, as these maintain the number of segments and syllables. In contexts of L2 acquisition or L1 restructuring, such productions may represent intermediate stages in the reorganization of syllable structure: structurally faithful in terms of syllable count and segmental presence but not (yet) reflecting categorical stop articulation in coda. The `C1target` variable applies a stricter criterion, requiring both structural preservation and articulatory realization of a stop in coda position.

Models were optimized using backward stepwise elimination based on log-likelihood changes, retaining by-participant random intercepts via `buildmerControl`. Alpha was set to 0.05 for all significance testing. Estimated marginal means and pairwise

contrasts, adjusted for multiple comparisons via Minimum Variance Quadratic Unbiased Estimation, were computed using `emmmeans` (v1.8.3, [Lenth, 2020](#)).

For RQs 1a–1c (L2 English) and 2a–2c (L1 BP), the initial model included fixed effects of Language, POA, [+/-voice] (voice), and d' sensitivity scores (perceptionEN, perceptionBP), with random intercepts and slopes for Item and ID:

`syllabicTarget [or C1target] ~ Language * POA * voice * perceptionEN * perceptionBP + (1 + POA * voice * perceptionEN * perceptionBP | ID) + (1 + Language * POA * voice * perceptionEN * perceptionBP | Item)`

The final models were (M1) `syllabicTarget ~ 1 + Language + voice + POA + POA:voice + (1 | ID)` and (M2) `C1target ~ 1 + POA + voice + POA:voice + (1 | ID)`.

Perceptual accuracy in L2 English and L1 BP were not part of the maximal models. To determine whether the exclusion of perceptual accuracy reflected asymmetries in the perception and production stimuli, we fit the same full models for each language with C1 target as the dependent variable to data sets that excluded the coronal and final stress conditions that were exclusive to the production stimuli. When fitting the L2 English data, a maximal intercept-only model led us to re-estimate the model with a generalized linear model (GLM) using the `glm` function. The model including the fixed effect most critical to the RQ, `perceptionEN`, did not provide a significantly better fit than the intercept-only model ($p = 0.750$)

(M3). The final L1 BP model was (M4) $C1target \sim 1 + \text{voice} + \text{perceptionEN} + \text{voice:perceptionEN} + (1 | ID)$.

Stress effects (RQ 2d) were modeled in L1 BP data, with English CI target accuracy (ENacc) added as a fixed effect to the C1 target model to assess L2 effects on L1 production (RQ 3)⁵. The maximal models were (M5) $\text{syllabicTarget} \sim 1 + \text{voice} + \text{Stress} + \text{POA} + \text{POA:voice} + (1 | ID)$ and (M6) $C1target \sim 1 + \text{POA} + \text{ENacc} + \text{POA:ENacc} + \text{Stress} + (1 | ID)$. For RQ 4, individual variation was explored descriptively with visualizations.

The descriptive analysis (proportions and SDs) included production type by Language and by Language*POA to provide a general overview of the group-level patterns while also illustrating the rate of application of the various repair strategies in each language by POA.

9 Results

In this section, we start with an overview of the general production patterns in each language, followed by relevant model results for RQs 1, 2, and 3, and end with individual patterns of repair. All full model outputs can be found in [Supplementary materials](#).

9.1 Overview of production patterns

[Figure 5](#) presents the distribution of production outcomes across languages and repair strategies.

Panels A and B focus on syllabification patterns, showing the proportion of productions that preserved the target syllable structure (C.C) in labial/velar (Panel A) and coronal (Panel B) contexts, regardless of whether the coda stop was faithfully produced. In L2 English, speakers overwhelmingly preserved syllabic structure in both POA types, with syllabic targets produced at 96% accuracy (SD = 18%). In L1 BP, syllabification-preserving responses were less frequent and more variable, with overall syllabic target production at 68% (SD = 47%). This variability was especially pronounced in coronal contexts, which showed increased use of structure-disrupting strategies such as epenthesis and deletion.

Panels C and D isolate the subset of productions in which the target segment (a coda stop) was not produced, illustrating the full range of repair strategies used in each language and POA type. Structure-preserving strategies include glide formation (diphthongization), gemination, and lenition, while structure-disrupting strategies include epenthesis and deletion. In L1 BP, epenthesis was the most common repair (50%, 55% of which were voiceless), followed by deletion (25%). Coronal items in L1 BP also showed frequent palatalization and palatalization + epenthesis, contributing to greater variability in these items compared to labials and dorsals. In L2 English, non-target productions were rare overall, with stops produced at a rate of 81% (94% released; SD = 39%). When repairs occurred, they typically preserved syllabic

structure, and structure-disrupting strategies made up only 4% of all productions.

In L1 BP, stops were produced at a lower rate of 56% (95% released; SD = 50%), with repair strategies varying considerably across POA. Production of coda stops was lower for coronals ($M = 0.68$, SD = 0.46 in L2 English; $M = 0.48$, SD = 0.50 in L1 BP) than for labials and dorsals ($M = 0.94$, SD = 0.24 in L2 English; $M = 0.66$, SD = 0.48 in L1 BP). These asymmetries in repair patterns and accuracy rates are examined in detail in the inferential models presented in Section 9.2.

9.2 Predictors of production accuracy in L2 English (RQ 1) and L1 BP (RQ 2a–c)

Our first two RQs center on production accuracy of coda stops in the L2 and L1 and the variables that modulate production patterns in each language (perceptual acuity, POA, and voice).

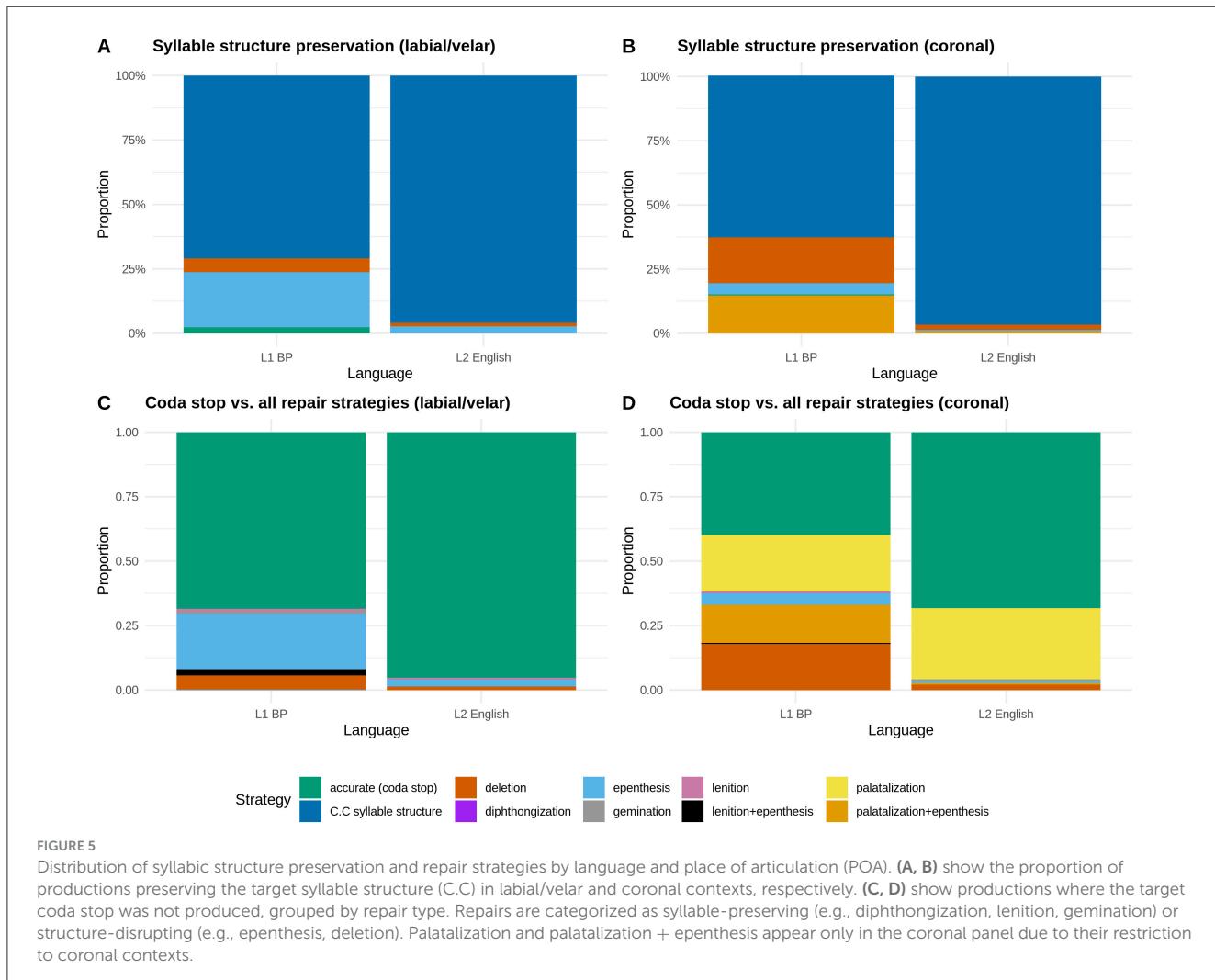
Starting with perceptual acuity as measured via d' sensitivity, the L2 English intercept-only model (M3) confirmed a lack of a significant relationship between L2 English d' sensitivity and $C1target$ production. In L1 BP, the maximal model (M4: $C1target \sim 1 + \text{voice} + \text{perceptionEN} + \text{voice:perceptionEN} + (1 | ID)$), revealed significant main effects and a significant interaction ($p < 0.05$). The fixed effects accounted for 77% of the variance, with substantial individual variation, (subject-level clustering accounted for 82% of the total variance, $ICC = 0.82$). The significant interaction ($OR = 0.00$, CI: 0.00–0.15, $p = 0.012$) indicates that, as L2 English perceptual acuity increased, the likelihood of producing a voiceless coda stop in L1 BP increased, while its relationship to production of voiced coda stops remained neutral. However, visual inspection of the data in [Figure 6](#) reveals that this relationship is not gradual or linear. Instead, for voiceless stops, L1 coda production remains low at lower levels of L2 perceptual acuity, but then shows a sharp, threshold-like increase as acuity improves, after which production quickly reaches ceiling. For voiced stops, the relationship remains flat across the range of perceptual acuity.

Turning to the remaining predictors, language, voice, and POA each were included in the maximal models M1 and M2. [Figure 7](#) illustrates the L2 English and L1 BP data in terms of syllabicTarget (M1) and $C1target$ (M2) and the significant interactions yielded by the data, respectively.

For both dependent variables, the probability of target-like production was greater in L2 English compared to L1 BP. However, the strength of effects differed between the two measures in ways that illuminate their hierarchical relationship. While both measures showed a significant POA:voice interaction, this effect was substantially stronger for C1 target ($z = 7.29$, $p = 0.001$) than for syllabic target ($z = 2.98$, $p = 0.019$).

Visualization of the syllabic target data indicates L2 English production at ceiling regardless of POA or voicing, while L1 BP showed greater variation. The POA:voice interaction revealed greater probability of C1.C2 productions in voiceless dorsals vs. coronals ($z = 2.98$, $p = 0.019$) and in voiced vs. voiceless dorsals ($z = -3.09$, $p = 0.013$).

⁵ This analysis was limited to C1 target due to ceiling performance in the syllabic target data.



For the C1 target, in addition to a stronger POA:voice interaction, the pattern of effects also differed, with greater probability of target productions in dorsals vs. coronals, particularly in voiceless segments ($z = 7.29, p < 0.001$) compared with voiced segments ($z = 3.29, p = 0.007$).

9.3 Effects of stress (RQ 2d) and rate of L2 English target production on L1 BP (RQ 3)

Stress was a significant predictor across both syllabic [M5, OR = 1.75, CI (1.14, 2.67), $p = 0.010$] and C1 target [M6, OR = 1.50, CI (1.01, 2.21), $p = 0.044$] analyses in BP production. In line with [Alcorn \(2018\)](#) and other previous research, initial (tonic) stress was associated with higher odds of syllabic target and C1 target production.

The relationship between L1 BP and L2 English production was examined through the inclusion of L2 English rate of accuracy (ENacc) in the C1 target analysis. A significant POA:ENacc interaction ([Figure 8](#)) demonstrated that an increased rate of L2 English coda stop production had a particularly strong effect on the

probability of BP coda stop production in non-coronal consonants. Specifically, a one-unit increase in ENacc increased the odds of C1 coda stop production by 390.03 times in dorsal items and 900.78 times in labial items, compared with 4.16 times in coronal items.

9.4 Individual variation in strategy implementation and perception-production alignment (RQ 4)

This section examines individual differences in bilinguals' production strategies and their alignment with perceptual accuracy. We first report participant-level variability in target structure production and repair strategies across L1 and L2, then turn to the relationship between individual perception and production outcomes.

9.4.1 Strategy implementation

[Figure 9](#) presents participant-level distributions of syllabic structure preservation and repair strategies by language and POA, using the same visualization scheme as [Figure 5](#). This format

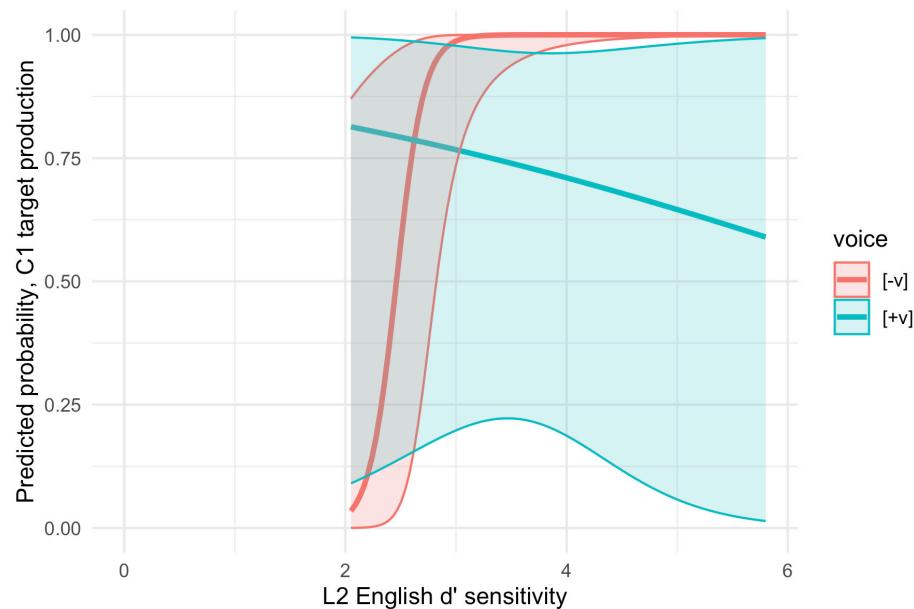


FIGURE 6
Interaction effect of L2 English perceptual sensitivity and [+/- voice] on probability of L1 BP coda stop production.

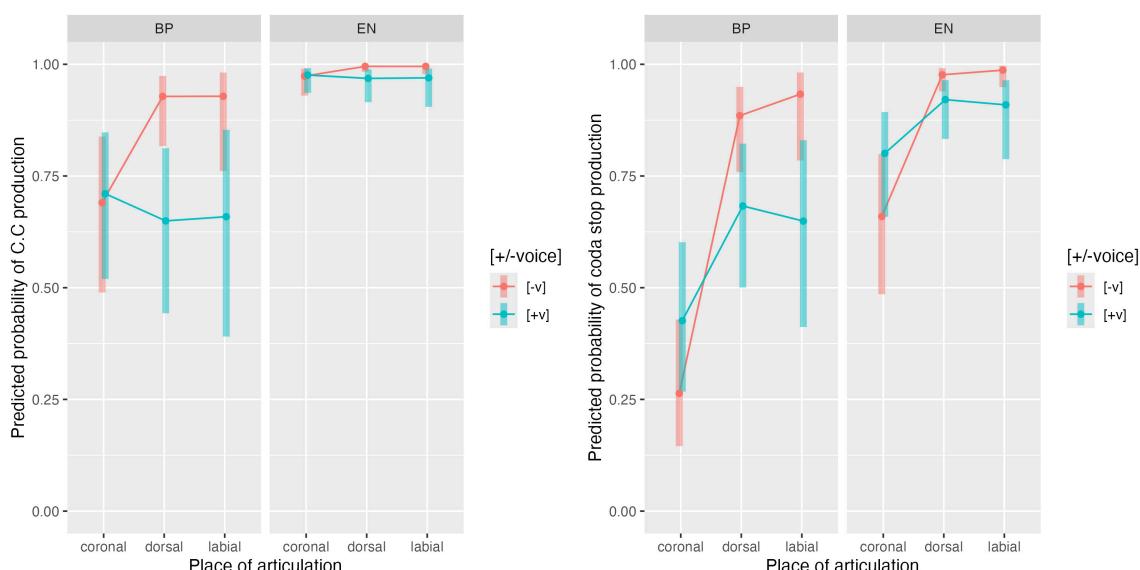


FIGURE 7
RQs 1 and 2, syllabic target and C1 target analysis: voice: POA interaction by language. Error bars represent 95% CI.

allows for direct comparison of target accuracy and repair type across individuals while preserving the subset-superset relationship between segmental and syllabic targets.

Across participants, L2 English showed little individual variation in syllabic structure preservation: eleven participants produced C.C structures at ceiling (100%), and the lowest accuracy observed was 84%. C1 stop targets in L2 English showed greater

variability, particularly in coronals ($M = 0.68, SD = 0.46$) compared to labials and dorsals ($M = 0.94, SD = 0.24$).

In L1 BP, syllabic target accuracy was more variable ($M = 0.68, SD = 0.47$). Seven participants produced C.C structures with $>80\%$ accuracy, while the remaining eight showed lower and more variable performance (range = 25%–73%). C1 stop targets were more challenging overall: only two of the seven high-performers

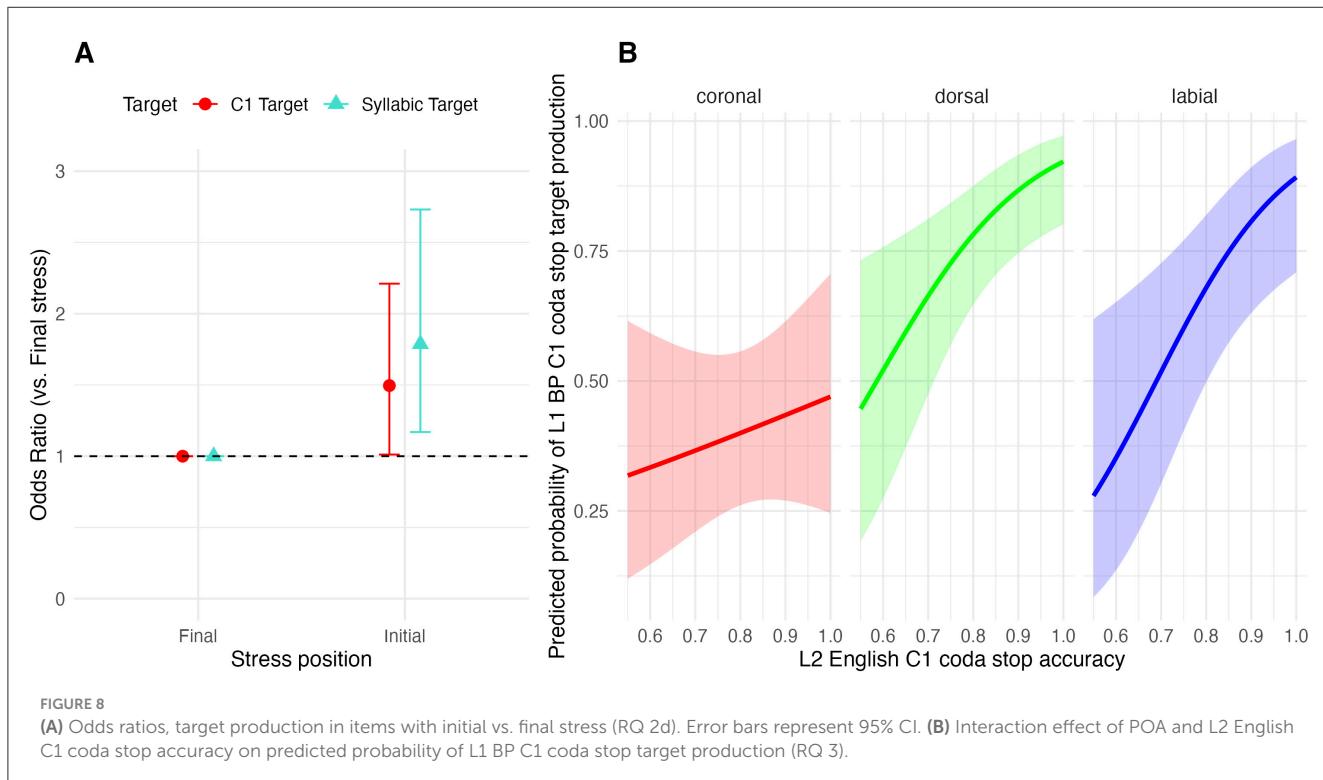


FIGURE 8

(A) Odds ratios, target production in items with initial vs. final stress (RQ 2d). Error bars represent 95% CI. (B) Interaction effect of POA and L2 English C1 coda stop accuracy on predicted probability of L1 BP C1 coda stop target production (RQ 3).

on the syllabic target also produced >80% C1 stop accuracy in L1 BP. Stop target accuracy in L1 BP averaged 56% ($SD = 0.50$), with coronals again showing the lowest rates.

Among participants producing BP coda stops in over half their items, nine of eleven exhibited broadly parallel production patterns in the two languages, despite differences in the specific strategies used. Several participants who showed ceiling-level production of stops in L2 English relied on different repair strategies in L1 BP. While some (e.g., P3) frequently used palatalization, others relied more heavily on deletion, epenthesis, or lenition. All participants used at least two distinct strategies in L1 BP, with one participant (P14) using five. These patterns deviate from the prediction of a dominant reliance on epenthesis (P4) and instead point to individualized repertoires of repair strategies shaped by POA, segmental context, and crosslinguistic experience.

9.4.2 Perception-production alignment

While these production patterns reveal robust individual variation, the relationship between perception and production accuracy was inconsistent across participants. As reported in Section 9.2 and shown in Figure 6, perceptual sensitivity was not a significant predictor of production accuracy at the group level in either language, and participant-level correlations between perception and coda stop production were weak and non-significant. Although some individuals with higher perceptual sensitivity demonstrated relatively higher production accuracy, this pattern was not systematic. A participant-level scatterplot illustrating the perception-production relationship is available in the [Supplementary materials](#).

10 Discussion

10.1 Summary

Participants produced syllabic and coda stop targets at greater rates in L2 English than in L1 BP and production was predicted by an interaction of POA and voicing that was independent of language. Perception did not predict production outcomes in either language, while initial stress and L2 English stop production increased BP target probabilities in non-coronals. L2 English C1 was produced as a coda consonant 95% of the time (81% stops), with coronal palatalization as the most prominent repair strategy. L1 BP C1 production was more variable: C1 coda consonants comprised only 68% of the data (56% stops) due to greater rates of epenthesis, deletion, and palatalization. At the individual level, L1 BP production varied widely, with participants using diverse repair strategies, some of which have not been reported in previous research. We address each of the first three research questions in turn, addressing individual patterns (RQ 4) as they pertain to each question. To conceptually account for these individual and cross-modal differences, we incorporate BiPhon modeling as an illustrative tool, showing how variable output patterns may arise from a unified phonological grammar shaped by stochastic learning.

10.2 Research question 1: coda stops in L2 English

Recall that Research question 1 asked (a) to what degree L1 BP learners of L2 English accurately produce L2 coda stops, (a) which



FIGURE 9

Individual distributions of syllabic structure preservation and repair strategies by language and place of articulation (POA). Each panel presents a single participant's productions in L1 Brazilian Portuguese and L2 English, separated by POA. The top row for each participant shows the proportion of productions preserving the target syllable structure (C.C.), while the bottom row shows the distribution of repair strategies when the coda stop target was not produced. Repair types are grouped by whether they preserved syllabification (e.g., lenition, gemination) or disrupted it (e.g., epenthesis, deletion). Palatalization strategies appear only in coronal contexts.

linguistic variables modulate production patterns, and (c) whether L2 perception accuracy predicts L2 production accuracy. The difference in syllabic target vs. C1 target suggests that convergence on the target syllable structure is an intermediate developmental stage, and that segmental accuracy remains the greater challenge due to the precise phonetic realization required.

Our results generally supported Alcorn's (2018) predictions regarding coda stop production, with production rates mirroring those of Alcorn's US bilinguals. However, notable POA-based differences emerged and L2 perception did not predict L2 production outcomes, contra Alcorn (2018) and Cardoso (2011). We note that the observed dimensional asymmetry could be attributed to individual variation (Figure 9), as proposed by Shin and Iverson, 2014.

The implications of these results for the theoretical analysis are threefold and relate specifically to the role of POA and voicing, the modeling of the repair strategies used, and the modeling of asymmetrical perception/production.

10.2.1 POA and voicing

Stronger POA: voice interaction effects in the C1 target model suggest greater influence on segmental accuracy than syllabic structure, largely due to the implementation of coronal palatalization, particularly in voiceless coronals.

The finding that coronals are less likely to surface as stops across languages squares with Collischonn (2002) claim that coronal stop codas are dispreferred despite their crosslinguistic unmarked status and their status as the only licit coda POA in BP. Of the 11 speakers who produced non-target forms, six did so exclusively with coronals. The low epenthesis rate across POA suggests that that the input to phonetic implementation reliably contained a coda stop (e.g., /lat.pə/) rather than a stop syllabified in onset (e.g., /la.ti.pə/). Why would a coronal coda pose a unique challenge? One hypothesis is articulatory: tongue-tip precision in the alveolar region is more difficult (Brownman and Goldstein, 1992), leading to gestural undershoot and coronal lenition or deletion (e.g., Kirchner, 2001). While BP data exhibit lenition and deletion, L2 coronal data predominantly show palatalization (without epenthesis) in nine of 15 speakers. Given the increased articulatory complexity of an affricate over a stop, palatalization is incompatible with a lenition account. Instead, we propose that L2 input, which includes coda stop coronals absent from the L1, prompted learners to preserve the coda but deploy a familiar L1 coronal strategy: palatalization.

Palatalization typically follows a coronal stop before a high front vowel in BP but also occurs in coda. Quintanilha-Azevedo (2016) found that word-medial /t/ was palatalized and licensed in coda at a rate of 77%. The production of [tʃ]/[dʒ], confirmed in all participants' L1 grammars via interview data, reflects a redeployment strategy (e.g., Archibald, 2005) that allows coda retention without segmental weakening and represents an intermediate acquisition stage. Individual patterns of coronal production show that exposure to optimal English phonetic forms has variably triggered L1 restructuring: While some speakers produced coda stops categorically, others alternated between stops, affricates, and (rarely) deletion.

L2 and L1 palatalization patterns diverged: in L1 BP, 40% of palatalized codas were syllabified in the onset, while L2 English palatalized /d/ was syllabified as an onset only once. The near absence of L2 coronal epenthesis aligns with Schneider (2009) prediction that novel L2 C1.C2 sequences yield lower epenthesis rates. However, their inverse prediction—that shared BP-English sequences (/p.t/-/k.t/) would show higher epenthesis—was not borne out, with production at/near ceiling. Two explanations are possible: (1) High-proficiency participants may have revised initial L2 representations, or (2) the sequences belong to an emergent BP syllabic pattern (ESP) in which voiceless non-coronal stops are increasingly produced without epenthesis (Cristófaro-Silva, 2024). This pattern, driven by gradient epenthetic vowel reduction and eventual deletion (e.g., Souza et al., 2020), explains why the BP /p.t/ and /k.t/ show higher coda realization than other bigrams. However, /b.d/, previously linked to ESP (Nascimento, 2016), showed 33% epenthesis, contrasting with /p.t/ (4%) and /k.t/ (13%) (see Supplementary materials for syllabic target and C1 target accuracy by bigram). Resolving this “chicken-or-egg” question requires longitudinal tracking of L1 BP representations prior to L2 acquisition and their redeployment over time.

Regarding cue constraints in our analysis, our results suggest that (a) separate constraints are necessary for each POA and (b) coronal cue constraints against affricate *and* vocalic material must rank higher than constraints militating against cues associated with affricates alone. Moreover, the interaction of POA and voicing suggests that (c) relevant cue constraints encode voicing information. Thus, rather than a single constraint set with /C(stop)./ as the phonological surface form, six distinct constraint sets are required, one for each stop /p t k b d g/. Since perception and weighting of cues to POA and voicing⁶ vary across L1, L2, and individuals, we adopt IPA notation to represent relevant cues. Although our revised constraint set comprises six sets, for reader clarity, we illustrate our analysis with /t./ and the *latpa* dataset to examine palatalization's role. In the absence of coronal-specific perception data, we extrapolate from /p/, using *mapta* to highlight perception-production distinctions and strategies available with vs. without palatalization. Given high L2 English accuracy, this contrast will be more apparent in the L1 BP data.

Together, these refinements help specify the constraint architecture underlying the BiPhon simulations that follow, which illustrate how a single phonological grammar shaped by linguistic factors and stochastic learning can produce the within-class, between- and within-speaker, and cross-modal variation observed in our data.

10.2.2 Repair strategies and the emergence of the unmarked

To explore how these and other repair strategies might emerge probabilistically, our original L2 English constraint set, based

⁶ We assume acoustic cues to POA in coda stops include spectral and temporal aspects of burst release when present, closure duration, and VC formant transitions, which vary by voicing and vowel quality (see Kent and Read, 2015). Cues to voicing include the consonant's F0, duration of voicing during closure, preceding vowel's duration and F1, and release force (see Alves, 2015; Kent and Read, 2015).

on prior research, included only epenthesis and palatalization (with/without epenthesis) as repair strategies. However, we observed a broader range, including deletion, lenition, and gemination, reflecting the influence of linguistic universals, which is more evident in the L1 BP grammar (see Section 10.3).

When a hearer maps an auditory input $[C(\text{stop})]$ to the surface form $/C(\text{stop})./$, mapping from the surface form to the same phonetic form does not always occur due to the marked status of coda stops in L1 BP, which we assume to form the initial state of the L2 English grammar (Schwartz and Sprouse, 1996). Instead, outputs can favor less-marked structures, aligning with The Emergence of the Unmarked (TETU; McCarthy and Prince, 2004), where universal principles shape outputs from novel L2 inputs.

For expository clarity, we introduce these less-marked repair strategies using cover constraints that represent the broader constraint families introduced in Section 6.2.2 (e.g., $*[C(\text{stop})]/C(\text{stop})./$), together with examples of the segment-specific constraints we implement in the remaining analyses to reflect the distinct phonetic behavior of each stop. These grammar-specific⁷ cue constraints penalize inputs mapped in phonetic implementation to:

- null outputs (deletion): $*[]/C(\text{stop})./$ (e.g., $*[]/t./$)
- approximants (lenition): $*[C(\text{approx})]/C(\text{stop})./$ (e.g., $*[\delta^*]/d./$)
- lengthened (gemination) stops: $*[C1(\text{stop})]/C1(\text{stop}).C2(\text{stop})/$ (e.g., $*[t:]/t.C(\text{stop})/$)

Because these constraints were violated in only seven instances, their rankings are higher than those of more frequently violated constraints (e.g., $*[t]/t./$, $*[t]/t./$). Still, their proximity in ranking permits probabilistic surfacing aligned with the observed output distribution. Feeding the constraint set and input-output distributions to the GLA yielded the following (Figure 10):

The two most frequent outputs for $/t./$ account for 94% of the data, with their corresponding cue constraints ranking lowest. In this evaluation, disharmonies have caused the lowest-ranked constraint to outrank the one favoring $[t]$, explaining why palatalization is more likely than assimilation, lenition, or deletion. The three instances in which the input contained an epenthetic vowel ($/t.i./$) each yielded a different optimal output, with each output uniquely violating a different cue constraint. As a result, the ranking values for those constraints were nearly identical.

Regarding L2 target grammar alignment, two phonetic implementation elements are key: palatalization and released stops. For palatalization, coronal data show 69% C1 target accuracy, with only 3% of non-target responses involving palatalization. In contrast, labial and velar C1s reach 98% accuracy, suggesting that palatalization persists where possible and convergence thus requires promoting the $*[t]/t./$ constraint. For released stops, 94% of C1 stops had a release burst, consistent with BP monolinguals (98%; Quintanilha-Azevedo, 2016) but contrasting with American English medial stop clusters (11%; Davidson, 2011). This reflects L1 BP influence, as BP exhibits less C1–C2 gestural overlap than

English. BiPhon attributes this gestural mismatch to non-optimal L2 sensorimotor knowledge, maintaining an L1-like association between auditory cues and gestures.

10.2.3 A note on phonetic implementation and articulatory coordination

Thus far, we have followed a simplified BiPhon model in which phonetic implementation directly links auditory and articulatory forms. However, a full account must consider the additional articulatory evaluation process, or *merely-phonetic-articulation*, where sensorimotor and articulatory constraints influence gestural coordination (see Figure 1). This is particularly relevant for understanding how L1 BP speakers acquire English-like $C1_{\text{stop}}.C2_{\text{stop}}$ transitions, as well as distinguishing between phonological epenthesis and intrusive vowels. Notably, the near-absence of unreleased coda stops in our bilingual participants—despite their high proficiency in English, where unreleased stops are common—suggests that articulatory factors may override perceptual or grammatical knowledge in production. Similar asymmetries between perception and production in L2 learners have recently been explored by Zhou and Hamann (2024). Due to space limitations, a detailed exploration of this articulatory coordination, including constraint rankings, gestural overlap patterns, and their implications for L2 acquisition, is available in the Supplementary materials.

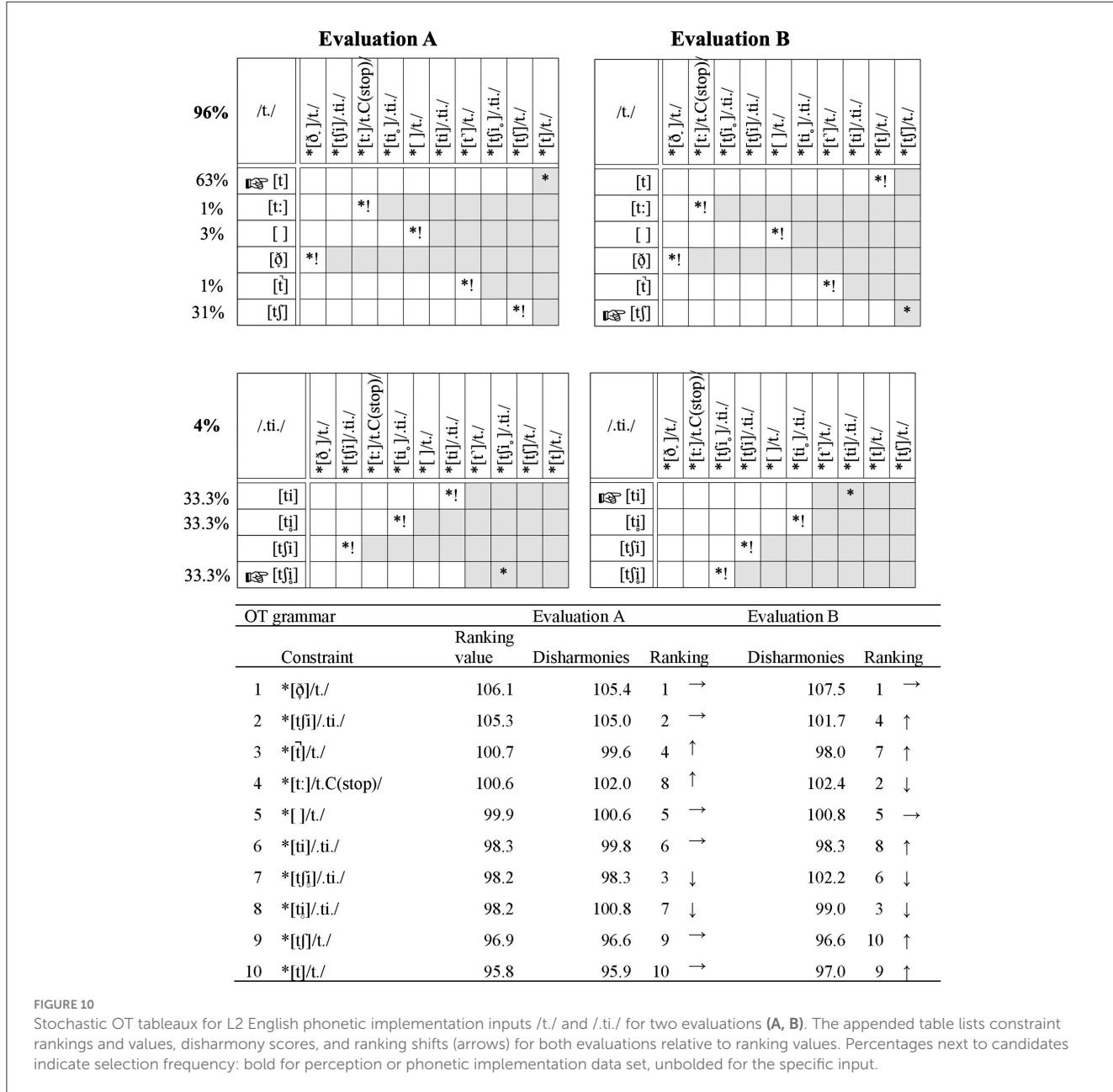
10.2.4 Perception and production

While the lack of a relationship between perception and production reported in Section 9 may reflect individual variation or, for example, learners' stage of acquisition, our findings align with theoretical accounts (e.g., de Leeuw et al., 2021) that emphasize the potential dissociation between perception and production mechanisms and with models such as L2LP, which posit that production accuracy emerges indirectly from perceptual restructuring and may therefore lag behind it. The data suggest that variation in production outcomes cannot be straightforwardly attributed to perceptual accuracy alone and underscore the need for a single theoretical analysis that can accommodate potential dimensional asymmetry. To this end, we model the group-level data here to illustrate (Figure 11) how perception and production outcomes may diverge even within a unified grammar. The L1 BP grammar is more illustrative of this asymmetry given that the labial and velar data are not as close to ceiling as they are in L2 English.

To integrate Cabrelli et al. (2019) perception data for $/p/$ into the L2 grammar, we introduce an input $[p]$ with output candidates $/p./$ and $/p./$ and add the cue constraint $*[p]/.p./$, which was inactive in phonetic implementation. Since GLA modeling requires distributions, we used mean perceptual accuracy (85%) whereby cue constraints $*[p]/p./$ and $*[p]/.p./$ were violated at respective rates of 85% and 15%.

Interestingly, production accuracy exceeded perception accuracy and there were no epenthetic productions in the L2 English labial data. Given that, when $[p]$ maps to $/p./$, learners should produce a predicted output of $/p./$, between-task design differences may explain this disparity. One possibility is that the production task's stimuli presentation enhanced perceptual acuity, whereby the stop in the monosyllable [map] is more salient and

⁷ Deletion, lenition, and complete assimilation are all phenomena that have been observed in spontaneous American English and/or Brazilian Portuguese speech (see e.g., Bagno, 1997; Davidson, 2011), so it is logical to assume that these constraints are part of the L2 grammar either as a relic of L1 transfer or as an acquired L2 constraint.



easier to map to /p./ than in word-medial position in an ABX trial, which requires multiple comparisons for fit and binning (see [Nagle and Baese-Berk, 2022](#)). Another is that lower perceptual accuracy stems from the markedness difference between C1_{stop}.C2_{fricative} sequences in the perception task and C1_{stop}.C2_{stop} sequences in production (cf. 5.2). The current dataset does not allow us to distinguish between these explanations, reinforcing [Nagle and Baese-Berk's \(2022\)](#) argument for greater cross-modal task similarity (see Section 10.5).

10.3 Research question 2: coda stops in L1 BP

RQ 2 examined (a) the degree to which participants produce coda stops in L1 BP, (b) linguistic variables modulating production

patterns, and (c) whether L1 perception accuracy predicts L1 production accuracy. The observed pattern, in which syllable structure is more accurate than C1 target, suggests that L2 influence first facilitates restructuring at the phonotactic level, with segmental realization following a less robust trajectory. This mirrors the L2 acquisition pattern in 10.2, where convergence on syllable structure serves as an intermediate developmental step while segmental accuracy remains a more persistent challenge. Thus, while L2 exposure may promote a more permissible L1 syllable structure, the extent to which it reshapes segmental articulation appears to be comparatively limited.

Recalling that our predictions for rates of stop production and the roles of predictive variables were based on [Alcorn \(2018\)](#), unlike RQ 1, most predictions did not hold. Alcorn's US bilinguals produced stops 34% of the time, compared to 56% in our sample. Repair strategies differed, such that epenthesis was the exclusive

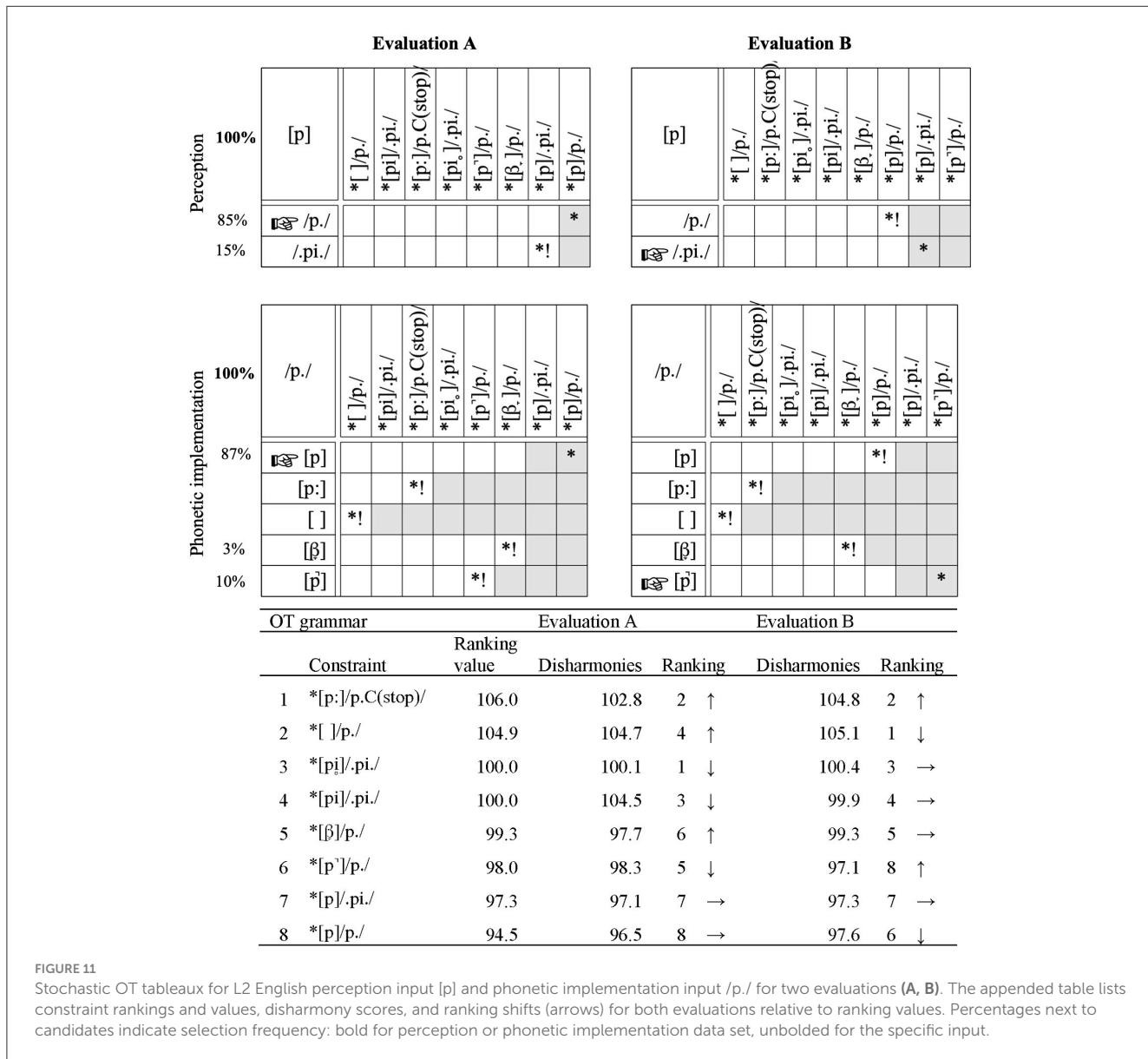


FIGURE 11

Stochastic OT tableaux for L2 English perception input [p] and phonetic implementation input /p./ for two evaluations (A, B). The appended table lists constraint rankings and values, disharmony scores, and ranking shifts (arrows) for both evaluations relative to ranking values. Percentages next to candidates indicate selection frequency: bold for perception or phonetic implementation data set, unbolded for the specific input.

strategy documented in Alcorn's sample but accounted for only half of all repairs in ours, with highly variable implementation.

10.3.1 POA, voice, and stress

Coronals were the least likely to surface as stops, regardless of L2 coronal accuracy, and coronal and /b/ stop production (<50%) contrasted sharply with /p/ and /k/ stop production (~80%). While voiceless non-coronals and /b/ patterned similarly between Alcorn and the present study, the studies' coronal patterns do not align. The coronals' behavior does not appear to be driven by L2 frequency: Such an account would predict that the least accurate L1 sequences (/t.k/ and /d.g/) would also be the least frequent L2 sequences, yet, they ranked among the most frequent after /p.t/ and /k.t/. Instead, we posit that this contrast can be explained via a markedness account, which we address in 10.3.2.

The role of stress also diverged from predictions: While the post-tonic context favored coda stops, the magnitude of the difference was smaller in our sample (58% pre-tonic vs. 53% post-tonic, OR = 1.22) than in Alcorn's (33% pre-tonic vs. 67% post-tonic, OR = 4.12). This raises the question of whether stress plays a weaker role in these bilinguals' epenthesis patterns. However, given task and stimulus differences, and the fact that Alcorn's stress data collapse across monolingual and bilingual groups (with monolinguals epenthizing more frequently), we refrain from drawing any conclusions. Further research, ideally including monolingual BP and longitudinal bilingual data, is needed to determine whether this reflects a true difference in stress-related restrictions in this sample.

Despite differences in outcomes in these two studies (likely due in part to methodological differences) both data sets indicate coda stop production that diverges from a monolingual BP grammar. We discuss this divergence as it relates to L2 acquisition in 10.4, after first addressing repair

strategies and the perception-production asymmetry observed in L1 BP.

10.3.2 Strategies and the emergence of the unmarked

The preference for epenthesis in [Alcorn's \(2018\)](#) participants contrasts with the variable repair strategies in our study. We attribute this difference to the use of real words vs. nonce words and the role of markedness. In BiPhon, lexical access determines whether a surface form maps to an entrenched underlying representation or is processed as a novel input ([Figure 1](#)). In L1 BP acquisition, certain lexemes will be stored with epenthetic /i/ (e.g., *obter* “to obtain” → /o.bi.'te∅/ → |obite∅|). This entrenched representation remains largely unchanged despite L2 English exposure, meaning an underlying form such as |obite∅| consistently maps to /o.bi.'te∅/ in *phonological production*, preserving the epenthetic vowel ([Quintanilha-Azevedo, 2016](#)) rather than favoring alternative repair strategies.

In contrast, novel words lack an established underlying form and these bilinguals more often map them to a coda stop surface form (/C(stop)./) rather than a form with epenthesis (/C(stop)i./) during *phonological perception* ([Cabrelli et al., 2019](#)), which in turn serves as the input in *phonetic implementation*. The greater availability of alternative repairs for novel words suggests that depth of lexical access constrains repair selection: words with entrenched representations favor epenthesis, while novel items are more susceptible to other repairs based on our sample's current cue constraint rankings. This explains why our participants, who rely on surface forms for nonce words, showed greater variability in repair strategies than Alcorn's participants, who accessed deeper underlying forms. Supporting this, [Nascimento \(2019\)](#) found that epenthesis rates were lower in nonce words than in lexical items as L2 proficiency increased and reported a greater range of strategies than in other studies that used real words.

Ranking values for *latpa* confirm this pattern ([Figure 12](#)): the highest probability outputs for input /t./ are released stops (Evaluation A), null outputs (Evaluation B), and palatalized segments (Evaluation C). Constraints prohibiting less marked outputs are ranked high enough to prevent selection, but the constraint penalizing null outputs is close in rank to those restricting palatalized and released stops, reflecting their similar distribution. For analytic completeness, we also include two additional constraints: one against approximants followed by epenthesis—*[C(approx)i]/C(stop)./(e.g., *[ði]/d./)—and another against semivowel realizations—*[i]/C(stop)./(e.g., *[i]/g./). Both were attested only in L1 BP and in very limited instances (nine and one token(s), respectively). Their rankings are therefore high and do not influence the modeled output distributions, but are included in the constraint set to reflect the full range of observed variation. As in the L2, coronal items exhibit a preference for palatalization when maintaining syllable structure, suggesting that L2 /t./-/d./ input has not overridden the ranking favoring this strategy. However, the relatively low ranking of the null output constraint suggests that a CV.CV repair remains similarly viable for these bilinguals, unlike in [Alcorn \(2018\)](#). Thus, lexical access depth provides a unifying explanation for differences in repair strategies across studies.

10.3.3 Perception and production

We initially expected to compare presence/absence of epenthesis in production with perceptual discrimination of /i/ in the acoustic signal. However, this approach oversimplifies the data.

A binary analysis disregards cases where epenthesis was absent and the syllabic target was produced, yet the segmental target was unmet. These outputs, constituting half of all productions without a coda stop, represent intermediate realizations that preserve syllable structure but fall short of full stop production in coda position. As such, they cannot be treated as simple “absence” cases, invalidating a direct comparison with [Cabrelli et al. \(2019\)](#). Instead, what we *can* conclude based on epenthesis rates, is that learners predominantly map a [C(stop)] acoustic signal to a coda stop surface form at rates exceeding those of monolingual BP speakers. However, when phonetic implementation receives a coda stop surface form input, alternative strategies emerge, favoring less marked outputs.

These asymmetrical mappings align with Principle 4 of ADAPPT ([de Leeuw and Chang, 2023](#)), which states that accurate perception does not guarantee accurate production. They are also consistent with the pre-lexical predictions of the L2LP model (e.g., [van Leussen and Escudero, 2015](#)), which holds that accurate production depends on accurate perception but predicts temporary mismatches as part of a developmental trajectory, since perceptual restructuring precedes stable changes in phonetic implementation. BiPhon captures this same asymmetry in our analytic framework by synchronically modeling the state of the grammar at the time of data collection. In this snapshot, bilinguals show stable perceptual mappings, yet their phonetic outputs remain variable due to the probabilistic evaluation of competing constraints.

In practice, to implement this bidirectional grammar, we introduced the cue constraint *[p]/.pi./ and added a [p] input to phonological perception, with an output distribution violating *[p]/p./ 77.5% of the time and [p]/.pi./ 22.5% of the time. The evaluation depicted in [Figure 13](#) shows how, while learners successfully map phonetic cues to a coda stop representation in perception and [p] remains the dominant phonetic output, phonetic implementation retains variability in output selection.

10.4 Research question 3: L2 effects on the L1 and the mechanisms through which an L1 grammar approximates an L2 grammar

Given the L2–L1 perception relationship found in [Cabrelli et al. \(2019\)](#), we predicted a similar relationship in production. This was partially confirmed: Greater L2 coda stop production increased L1 BP coda stop production for non-coronals, regardless of voicing, and L2 English perception predicted L1 BP production of voiceless coda stops. These results suggest the possibility that L2 acquisition has influenced the L1 grammar, though the likelihood of coda stop production in the L1 remains lower than in the L2, indicating an intermediate state between a monolingual L1 BP grammar and their L2 English grammar.

Several theoretical mechanisms could account for this pattern. [Repiso Puigdelliura \(2021\)](#) suggests that L2 rankings influence L1 grammar early in acquisition but are suppressed as bilingual language control strengthens. This view predicts that L1 BP coda

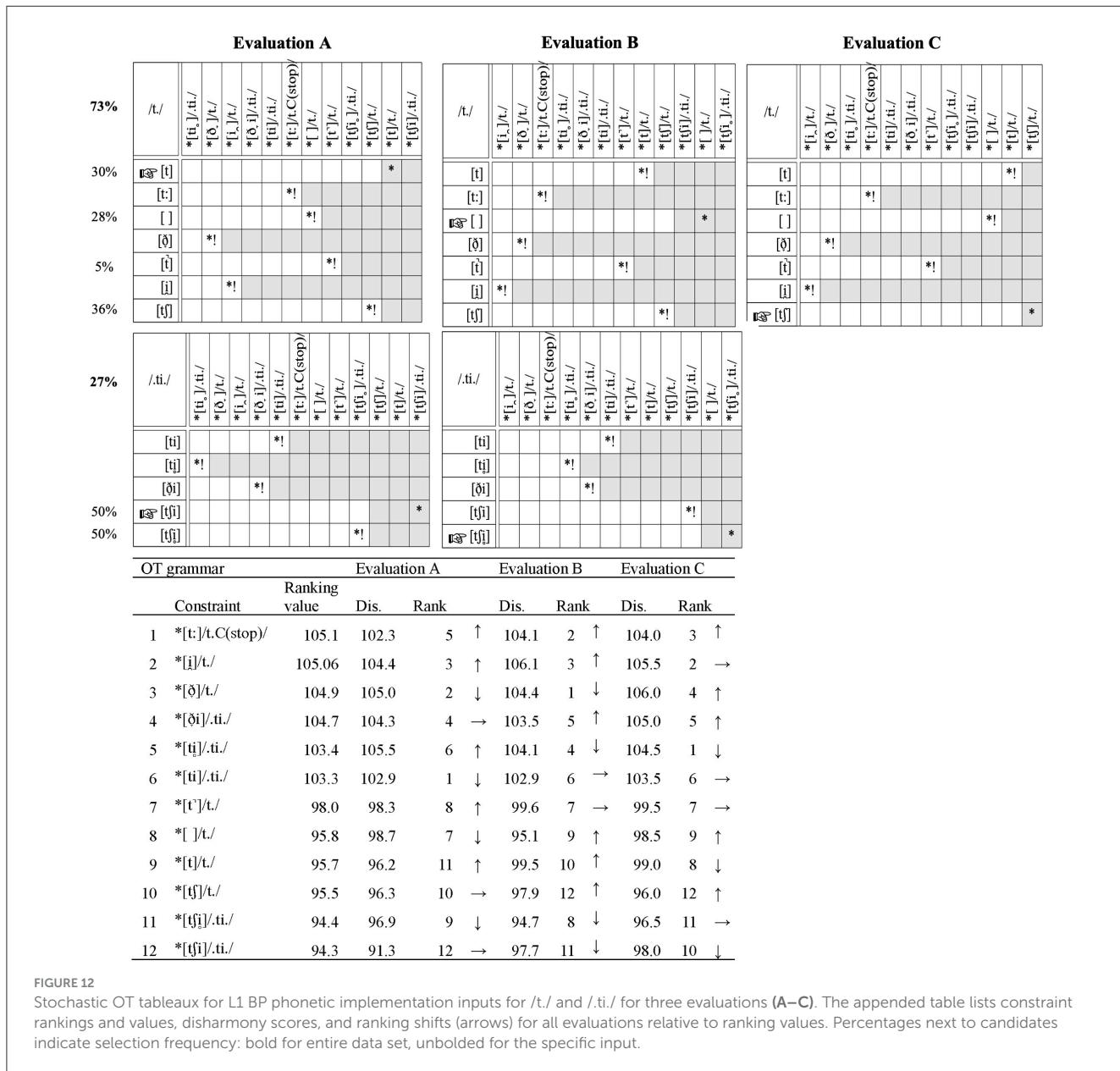


FIGURE 12

Stochastic OT tableaux for L1 BP phonetic implementation inputs for /t./ and /.ti./ for three evaluations (A–C). The appended table lists constraint rankings and values, disharmony scores, and ranking shifts (arrows) for all evaluations relative to ranking values. Percentages next to candidates indicate selection frequency: bold for entire data set, unbolded for the specific input.

production may initially approximate L2 English patterns and later stabilizes. Tetzloff's (2022) tandem-updating model instead allows both grammars to adjust dynamically: Contradictory L1–L2 outputs trigger a mechanism that penalizes the non-intended language's winner but penalizes the intended language's losing candidates even more, leading to a scenario where probability is split between both grammars' winners. A related possibility is that differences in plasticity between L1 and L2 grammars (Boersma and Hayes, 2001) allow L1 updates to occur, albeit at a slower rate than in the L2. A third possibility, consistent with Boersma and Escudero (2008) and in line with L2LP's account of language activation/perception modes (Escudero, 2005; Yazawa et al., 2020), is that our results do not reflect L1 restructuring at all, but rather the temporary activation of the L2 grammar while evaluating input in an L1 context. In such a scenario, variability comes from selective routing between parallel grammars during evaluation, depending on relative activation and bilingual

control. This account builds on earlier developments in the L2LP framework (Escudero, 2005), which explicitly incorporates language activation as a mechanism shaping bilingual speech patterns, later extended to mode-based perception in Boersma and Escudero (2008) and subsequent work (e.g., Yazawa et al., 2020). L1 perception and production that diverges from an L1 baseline may therefore be a surface reflection of mixed evaluations across grammars, rather than L1 grammar restructuring.

Determining whether L1–L2 approximation results from weaker bilingual control early in acquisition (Repiso Puigdelliura, 2021), differential penalty structures in error-driven learning (Tetzloff, 2022), disparities in grammatical plasticity (Boersma and Hayes, 2001), or the parser's routing of input to parallel phonological systems (Boersma and Escudero, 2008) requires longitudinal data and computational modeling. Individual variation may help adjudicate between these models: Does stronger bilingual control

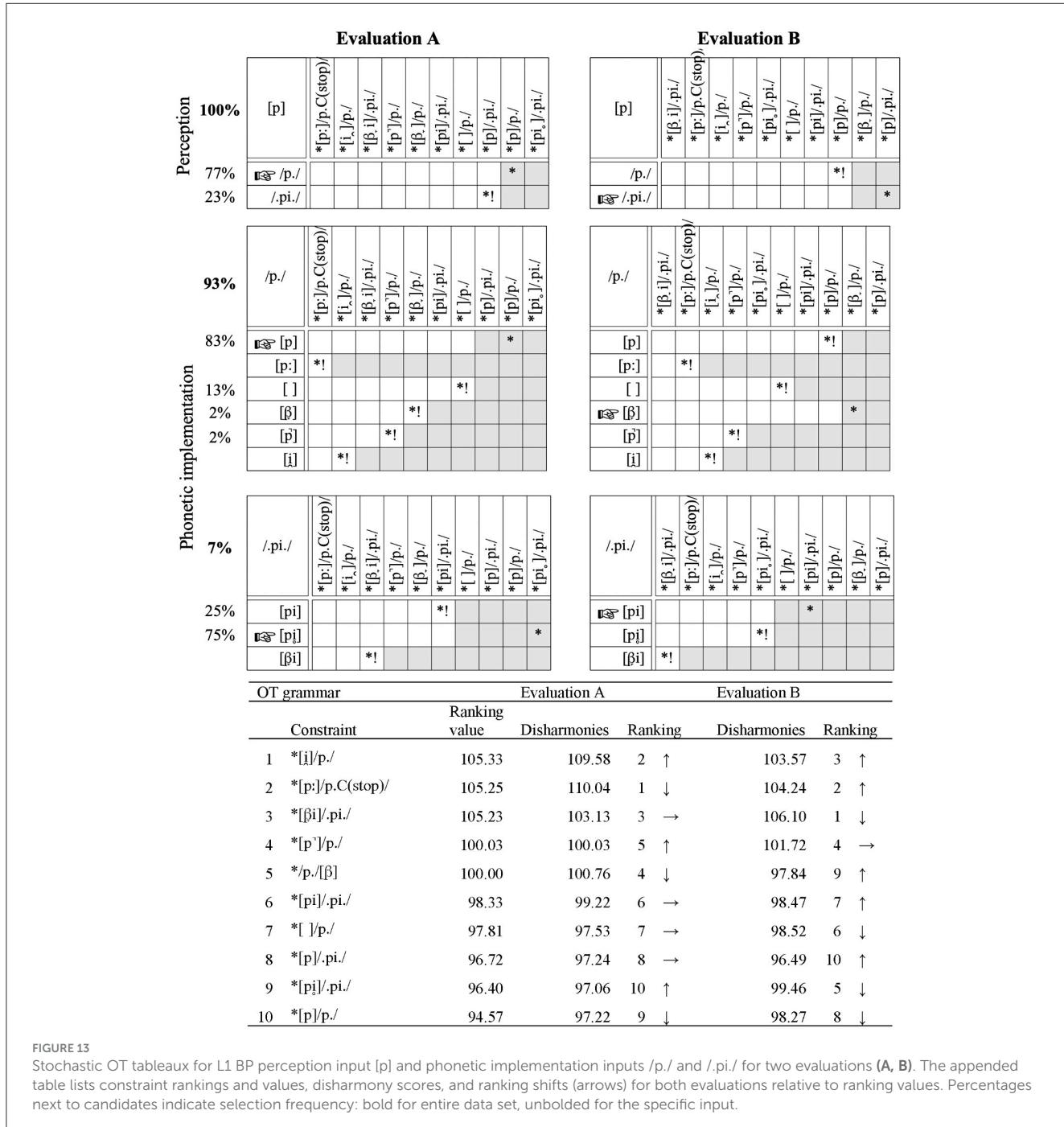


FIGURE 13

Stochastic OT tableaux for L1 BP perception input [p] and phonetic implementation inputs /p./ and /.pi./ for two evaluations (A, B). The appended table lists constraint rankings and values, disharmony scores, and ranking shifts (arrows) for both evaluations relative to ranking values. Percentages next to candidates indicate selection frequency: bold for entire data set, unbolded for the specific input.

limit L1 restructuring, as Repiso Puigdelliura predicts? Does greater L2 dominance accelerate L1 adaptation, aligning with tandem updating (Tetzloff)? Or do apparent L1 intrusions instead reflect greater relative L2 grammar activation, as in mode-based accounts of L2LP (Boersma and Escudero, 2008; Yazawa et al., 2020)?

These competing mechanisms can be distinguished through both modeling and empirical approaches. Model comparisons can evaluate whether the data are best explained by a single restructured grammar, a mixture of parallel grammars, or a tandem-updating model, while empirical studies—for example, contrasting automatic

vs. decisional measures or examining L1 recovery after short-term L2 input suppression—can inform whether observed L2 influence in the L1 reflects structural change or transient activation. We return to these questions in Section 10.5, where we outline how developmental simulations and longitudinal learner data can be combined to test these mechanisms directly.

As a point of departure, we can examine the asymmetries we observed across POA and stress, two domains where the bilingual system appears less stable. The selective effects for non-coronals and the elevated coronal deletion rates, together with differences

between pretonic and post-tonic contexts relative to [Alcorn's \(2018\)](#) L1 BP data, indicate specific areas where L1 outcomes diverge from monolingual patterns. These domains provide targeted test cases for determining whether such divergence reflects structural change or activation-based variability.

10.5 Limitations and future directions

This study presents several limitations that provide a foundation for future research.

First, data from a single time point cannot capture the bidirectional trajectories of L1–L2 interaction over time, which are critical to our understanding of the timing and degree of constraint grammar interactions and the variables that modulate them.

Within the L2LP framework ([Escudero and Boersma, 2003, 2004](#); [Boersma and Escudero, 2008](#)), longitudinal simulations have shown how reranking and category restructuring unfold. These studies have demonstrated that, even when the copied L1 grammar serving as the initial state of the L2 grammar changes, the original L1 system can remain stable, supporting the need for distinct perceptual grammars across languages. Extensive computational work in this framework has modeled how variation in input quantity and quality, proficiency, and bilingual mode influence speech learning over time (e.g., [Escudero and Boersma, 2003, 2004](#); [Boersma and Escudero, 2008](#); [Yazawa et al., 2020](#); see [Escudero and Yazawa, 2024](#) for an overview). Together, these studies illustrate how perceptual grammars develop through exposure and how activation patterns and language mode moderate perceptual outcomes.

Future research should link the synchronic modeling of bilingual grammars in the present study to developmental simulations to test how variation in input and bilingual experience predicts perception–production (a)symmetries and the relative stability of the L1 and L2 grammars. Computational simulations can predict how changes in input, activation, or learning rate drive development over time, offering insights that would be logically unfeasible with behavioral data alone. Longitudinal data from real learners, in turn, can be used to evaluate these modeled trajectories and to determine how individual differences modulate them. In such studies, measures of language control will help clarify whether L1–L2 parallelism reflects shared developmental plasticity or incomplete suppression of L2 rankings in L1 contexts, while additional cognitive measures (e.g., phonological short-term memory, auditory processing) can isolate the task-related demands shown to influence outcomes. Expanding samples to include a wider range of dominance and usage profiles will also make it possible to test how experience-based variability predicts patterns of bilingual phonological adaptation.

Finally, combining simulation-based predictions with longitudinal learner data can inform whether the effects observed here are reversible when L2 input declines ([Cabrelli, 2023](#); [Chamorro et al., 2016](#)) and whether the relationship between perception and production changes as a function of bilingual experience ([Nagle and Baese-Berk, 2022](#)). Together, these approaches can reveal how input, control, and activation interact over time to shape (in)stability in bilingual phonological systems.

Second, differences between perception and production tasks pose inherent challenges. The ABX task required complex comparisons of perceptual fit ([Nagle and Baese-Berk, 2022](#)) not shared by the concatenation task. Additionally, [Saito and Plonsky \(2019\)](#) posit that controlled production is connected to declarative pronunciation knowledge while spontaneous production taps into procedural knowledge. While spontaneous production may reveal patterns not captured in controlled data and thus a different relationship to phonological perception, we used controlled production for its greater alignment with the perception data ([Nagle, 2021](#)) and future analyses will evaluate our participants' guided interview data.

Third, the concatenation task itself introduces a potential confound, involving both perceptual and production components. While this could obscure true production ability ([Davidson, 2010](#); [de Jong et al., 2009](#); [Kato and Baese-Berk, 2020](#)), several aspects of the task design mitigate concerns. The task included a 500 ms silent interval between syllables and explicitly instructed participants to concatenate two independent inputs. This temporal separation reduces the likelihood of perceptual continuity effects that typically lead to epenthesis in real-time processing. In addition, the structure of the task discourages reinterpretation of the first monosyllable as part of a disyllabic unit, as there were no coarticulatory or prosodic cues linking the two. Together with the fact that the ultimate target was a disyllabic nonce word (e.g., *lagba*), these factors reduce the likelihood that perceptual illusions bleed into production. Participants' high perceptual accuracy in [Cabrelli et al. \(2019\)](#) further mitigates concerns. Crucially, we observe substantially more epenthesis in BP responses but not in English, despite comparable acoustic input across languages, further indicating that a task effect is unlikely to explain our results.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/zugb8/?view_only=c3bb065ddd474fcebd596bf799570caa.

Ethics statement

The studies involving humans were approved by University of Chicago Illinois Institutional Review Board (Protocol 2015-0040). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JCa: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JCr: Data curation, Writing – original draft, Writing – review & editing. JE: Data curation, Writing – original draft, Writing – review & editing. IF: Data curation, Project administration, Writing

– original draft, Writing – review & editing. AL: Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was supported by internal funds provided by the College of Liberal Arts and Sciences at the University of Illinois Chicago. It was also supported by the César Nombela Talent Attraction Grant (2023-T1/PH-HUM-2909), awarded to AL and funded by the General Directorate of Research and Technological Innovation of the Regional Government of Madrid (Spain), which partially covered the open-access publication fees.

Acknowledgments

The authors thank Carrie Pichan, Jess Ward, Ricardo Brum, and Vuong Nguyen for their assistance with data curation. Additional thanks to Carrie Pichan for her contributions to stimuli analysis. We are especially grateful to all the participants for generously sharing their time.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. In the preparation of this manuscript,

we utilized generative AI tools to assist with specific technical aspects of the research. The AI was employed to help generate and debug code for statistical analyses, create visualization scripts for data representation, and refine text to meet word count requirements while maintaining clarity. All AI-generated content was carefully reviewed, edited, and verified by authors to ensure accuracy and alignment with the research objectives. The AI served as a supplementary tool that enhanced efficiency in data processing and presentation, but all interpretations, conclusions, and theoretical frameworks remain the product of our own scholarly judgment and expertise. The data analysis procedures, including the R code for statistical modeling and visualization, were independently verified to ensure reproducibility and validity of the results.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

All supplementary materials can be found in osf.io: https://osf.io/o zugb8/?view_only=c3bb065ddd474fcebd596bf799570caa.

References

Alcorn, S. M. (2018). *The role of L2 experience in L1 phonotactic restructuring in sequential bilinguals* [Doctoral dissertation]. University of Texas, Austin, TX, United States.

Alves, M. A. (2015). *Estudo dos parâmetros acústicos relacionados à produção das plosivas do português brasileiro na fala adulta: análise acústico-quantitativa* [MA thesis]. Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

Alves, M. A., Seara, I. C., Pacheco, F. S., Klein, S., and Seara, R. (2008). “On the voiceless aspirated stops in Brazilian Portuguese,” in *Computational Processing of the Portuguese Language*, eds. A. Teixeira, V. L. S. de Lima, L. C. de Oliveira, and P. Quaresma (New York: Springer), 248–251.

Archibald, J. (2005). Second language phonology as redeployment of L1 phonological knowledge. *Can. J. Ling./Rev. Can. Ling.* 50, 285–314. doi: 10.1017/S0008413100003741

Bagno, M. (1997). *A Língua de Eulália: Novela Sociolinguística*, 17th Edn. São Paulo: Contexto.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bettoni-Techio, M. (2005). *Production of final alveolar stops in Brazilian Portuguese/English interphonology* [MA thesis]. Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

Birdsong, D., Gertken, L. M., and Amengual, M. (2012). *Bilingual language profile: An Easy-to-use instrument to assess bilingualism*. Available online at: <https://sites.la.utexas.edu/bilingual/>.

Boersma, P. (2011). “A programme for bidirectional phonology and phonetics and their acquisition and evolution,” in *Bidirectional Optimality Theory*, eds. A. Benz and J. Mattausch (Amsterdam: Benjamins), 33–72.

Boersma, P., and Escudero, P. (2008). “Learning to perceive a smaller L2 vowel inventory: an optimality theory account,” in *Contrast in Phonology: Theory, Perception, Acquisition*, eds. P. Avery, B. E. Dresher, and K. Rice (Berlin: Mouton de Gruyter), 271–301.

Boersma, P., and Hamann, S. (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25, 217–270. doi: 10.1017/S0952675708001474

Boersma, P., and Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguist. Inq.* 32, 45–86. doi: 10.1162/002438901554586

Boersma, P., and Weenink, D. (2021). Praat: doing phonetics by computer, versión 6.1.40. Available online at: <https://www.Fon.Hum.Uva.Nl/praat/> (Accessed August 19, 2022).

Brants, T., and Franz, A. (2006). *Google Web Trillion Word Corpus* [Dataset]. Google Web Trillion Word Corpus. Available online at: <https://catalog.ldc.upenn.edu/LDC2006T13>

Broselow, E. (2018). "Laryngeal contrasts in second language phonology," in *Phonological Typology*, eds. L. Hyman and F. Plank (Berlin: Mouton de Gruyter), 312–340.

Brown, C. P., and Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica* 49, 155–180. doi: 10.1159/000261913

Cabrelli, J. (2023). "Language attrition and L3/Ln," in *The Cambridge Handbook of Third Language Acquisition*, eds. A. Chaouch-Orozco, E. Puig-Mayenco, J. Rothman, J. Cabrelli, J. González Alonso, and S. M. Pereira Soares (Cambridge: Cambridge University Press), 317–353.

Cabrelli, J., Luque, A., and Finestrat-Martínez, I. (2019). Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception. *J. Phon.* 73, 55–69. doi: 10.1016/j.wocn.2018.10.006

Cardoso, W. (2007). Word-final stops in Brazilian Portuguese English: acquisition and pronunciation instruction. *Ilha Do Desterro* 55, 153–172. doi: 10.5007/2175-8026.2008n55p153

Cardoso, W. (2011). The development of coda perception in second language phonology: a variationist perspective. *Second Lang. Res.* 27, 433–465. doi: 10.1177/0267658311413540

Cellata, C. (2019). "Phonological attrition," in *The Oxford Handbook of Language Attrition*, eds. M. S. Schmid and B. Köpke (Oxford: Oxford University Press), 218–227.

Chamorro, G., Sorace, A., and Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Biling. Lang. Cogn.* 19, 520–532. doi: 10.1017/S1366728915000152

Collischonn, G. (2002). "A epéntese vocalica no português do sul do Brasil," in *Fonologia e variação: Recortes do português brasileiro*, eds. L. Bisol and C. Brescancini (Alegre: EDIPUCRS), 205–230.

Collischonn, G. (2003). Epéntese vocalica no português do sul do Brasil: Variáveis extralingüísticas. *Rev. Let.* 61, 285–297. doi: 10.5380/rel.v61i0.2892

Collischonn, G. (2004). Epéntese vocalica e restrições de acento no português do sul do Brasil. *Signum Estud. Ling.* 7, 61–78. doi: 10.5433/2237-4876.2004v7n1p61

Cristófaro-Silva, T. (2024). "Current issues in Portuguese syllable structure," in *The Routledge Handbook of Portuguese Phonology*, eds. A. Zampaulo (London: Routledge), 95–114.

Cristófaro-Silva, T., and Almeida, L. (2008). "On the nature of epenthetic vowels," in *Contemporary Phonology in Brazil*, eds. L. Bisol and C. Brescancini (Cambridge: Cambridge University Press), 193–212.

Davidson, L. (2010). Phonetic bases of similarities in cross-language production: evidence from English and Catalan. *J. Phon.* 38, 272–288. doi: 10.1016/j.wocn.2010.01.001

Davidson, L. (2011). Characteristics of stop releases in American English spontaneous speech. *Speech Commun.* 53, 1042–1058. doi: 10.1016/j.scom.2011.05.010

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *J. Phon.* 54, 35–50. doi: 10.1016/j.wocn.2015.09.003

de Jong, P. F., Bitter, D. J. L., van Setten, M., and Marinus, E. (2009). Does phonological recoding occur during silent reading, and is it necessary for orthographic learning? *J. Exp. Child Psychol.* 104, 267–282. doi: 10.1016/j.jecp.2009.06.002

de Leeuw, E., and Chang, C. B. (2023). "Phonetic and phonological L1 attrition and drift in bilingual speech," in *Cambridge Handbook of Bilingual Phonetics and Phonology*, ed. M. Amengual (Cambridge: Cambridge University Press), 32–52.

de Leeuw, E., Stockall, L., Lazaridou-Chatzgoga, D., and Gorba Masip, C. (2021). Illusory vowels in Spanish–English sequential bilinguals: evidence that accurate L2 perception is neither necessary nor sufficient for accurate L2 production. *Second Lang. Res.* 37, 587–618. doi: 10.1177/0267658319886623

de Lucena, R. M., and Alves, U. K. (2010). Implicações dialetais (dialeto gaúcho vs. paraibano) na aquisição de obstruintes em coda por aprendizes de inglês (L2): uma análise variaconista. *Let. Hoje* 45.

Dupoux, E., Parlato, E., Frotta, S., Hirose, Y., and Peperkamp, S. (2011). Where do illusory vowels come from? *J. Mem. Lang.* 64, 199–210. doi: 10.1016/j.jml.2010.12.004

Eckman, F. R. (2008). "Typological markedness and second language phonology," in *Phonology and Second Language Acquisition*, eds. J. G. Hansen Edwards and M. L. Zampini (Cambridge: Cambridge University Press), 95–115.

Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. Utrecht: Utrecht University and LOT.

Escudero, P. (2009). "The linguistic perception of similar L2 sounds," in *Phonology in Perception*, eds. P. Boersma and S. Hamann (Berlin: Mouton de Gruyter), 151–190.

Escudero, P., and Boersma, P. (2003). "Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm," in *Proceedings of the 25th Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics* (Philadelphia, PA: Penn Graduate Linguistics Society, University of Pennsylvania), Vol. 8, 71–85.

Escudero, P., and Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Stud. Second Lang. Acquis.* 26, 551–585. doi: 10.1017/S0272263104040021

Escudero, P., and Yazawa, K. (2024). "The second language linguistic perception model," in *The Cambridge Handbook of Bilingual Phonetics and Phonology*, ed. M. Amengual (Cambridge: Cambridge University Press), 173–195.

Estivalet, G. L. (2014). *Léxico do Português Brasileiro* [Dataset]. Léxico do Português Brasileiro. Retrieved from <http://lexicodoportugues.com> (Accessed November 10, 2023).

Flege, J. E., and Bohn, O.-S. (2021). "The revised speech learning model (SLM-r)," in *Second Language Speech Learning: Theoretical and Empirical Progress*, ed. R. Wayland (Cambridge: Cambridge University Press), 3–83.

Flege, J. E., and Davidian, R. D. (1984). Transfer and developmental processes in adult foreign language speech production. *Appl. Psycholinguist.* 5, 323–347. doi: 10.1017/S014271640000521X

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

Hermans, B. J. H., and Wetzels, W. L. (2012). Productive and unproductive stress patterns in Brazilian Portuguese. *Rev. Let.* 28, 77–115.

John, P., and Cardoso, W. (2017). On syllable structure and phonological variation: the case of i-epenthesis by Brazilian Portuguese learners of English. *Ilha Do Desterro* 70, 169–184. doi: 10.5007/2175-8026.2017v70n3p169

Kato, M., and Baese-Berk, M. M. (2020). The effect of input prompts on the relationship between perception and production of non-native sounds. *J. Phon.* 79:100964. doi: 10.1016/j.wocn.2020.100964

Kent, R., and Read, C. (2015). *Analise acústica da fala*. São Paulo: Editora Cortez.

Kim, S. Y., and Han, J.-I. (2022). The relationship between perception and production of illusory vowels in a second language. *Second Lang. Res.* 40:2676583221135185. doi: 10.1177/02676583221135185

Kirchner, R. (2001). *An Effort-Based Approach to Consonant Lenition*. New York: Routledge.

Kubota, M. (2019). Language change in bilingual returnee children: Mutual effects of bilingual experience and cognition (Doctoral dissertation). University of Edinburgh, Edinburgh.

Lenth, R. (2020). *Emmeans: Estimated estimated marginal means, aka least-squares means*. R package version 1.4. 7. Available online at: <https://CRAN.R-project.org/package=emmeans>

McCarthy, J., and Prince, A. (2004). "The emergence of the unmarked: optimality in prosodic morphology," in *Optimality Theory: A Reader*, ed. J. McCarthy (Hoboken, NJ: Blackwell), 483–494.

Monareto, V. N. de O. (2017). Frequência lexical de sequências mediais de obstruintes no português brasileiro. *ReVEL* 15, 115–133.

Nagle, C. L. (2021). Revisiting perception–production relationships: exploring a new approach to investigate perception as a time-varying predictor. *Lang. Learn.* 71, 243–279. doi: 10.1111/lang.12431

Nagle, C. L., and Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: revisiting theoretical assumptions and methodological practices. *Stud. Second Lang. Acquis.* 44, 580–605. doi: 10.1017/S0272263121000371

Nascimento, G. C. A. (2019). *Estratégias de reparo na pronúncia de oclusivas em posição de coda por falantes brasileiros de inglês como língua estrangeira*. [Doctoral dissertation]. Universidade Estadual Paulista, São Paulo, Brazil.

Nascimento, K. (2016). *Emergência de padrões silábicos no português brasileiro e seus reflexos no inglês língua estrangeira*. [Doctoral dissertation]. Universidade Estadual Ceará, Fortaleza, Brazil.

Opitz, C. (2011). First language attrition and second language acquisition in a second language environment (Doctoral dissertation). Trinity College Dublin, Dublin.

Parlato-Oliveira, E., Christophe, A., Hirose, Y., and Dupoux, E. (2010). Plasticity of illusory vowel perception in Brazilian–Japanese bilinguals. *J. Acoust. Soc. Am.* 127, 3738–3748. doi: 10.1121/1.3327792

Quintanilha-Azevedo, R. (2016). *Formalização fonético-fonológica da interação de restrições na produção e na percepção da epéntese no Português Brasileiro e no Português Europeu*. [Doctoral dissertation]. Federal University of Pelotas, Pelotas, RS, Brazil.

Quintanilha-Azevedo, R., Matzenauer, C. L. B., and Alves, U. K. (2017). Formalização fonético-fonológica da interação de restrições na percepção da epéntese vocalica no português brasileiro. *ReVEL* 15, 289–313.

R Core Team (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>

Reipo Puigdelliura, G. (2021). *The development of cross-linguistic transfer: the case of word-external repairs of empty onsets in Spanish heritage speakers*. [Doctoral dissertation]. UCLA, Los Angeles, CA, United States.

Saito, K., and Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: a proposed measurement framework and meta-analysis. *Lang. Learn.* 69, 652–708. doi: 10.1111/lang.12345

Schmid, M. S., and Köpke, B. (2017). The relevance of first language attrition to theories of bilingual development. *Linguist. Approaches Biling.* 7, 637–667. doi: 10.1075/lab.17058.sch

Schmid, M. S., and Köpke, B. (2019). *The Oxford Handbook of Language Attrition*. Oxford: Oxford University Press.

Schneider, A. (2009). *A epéntese medial em PB e na aquisição de inglês como LE: Uma análise morfológica*. [MA thesis]. Federal University of Rio Grande do Sul, Farroupilha, Brazil.

Schwartz, B. D., and Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Lang. Res.* 12, 40–72. doi: 10.1177/026765839601200103

Shin, D.-J., and Iverson, P. (2014). An experimental study of vowel epenthesis among Korean learners of English. *Phon. Speech Sci.* 6, 163–174. doi: 10.13064/KSSS.2014.6.2.163

Silveira, F. (2007). *Vocal epentética no português brasileiro: Um estudo acústico em encontros consonantais*. [Doctoral dissertation]. Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

Souza, A., Barboza, C., and Barra, A. R. (2020). Uma visão multirrepresentacional dos padrões silábicos emergentes do português brasileiro. *Entrepalavras* 10:121. doi: 10.22168/2237-6321-11735

Tetzloff, K. A. (2022). *Examining variability in Spanish monolingual and bilingual phonotactics: a look at sC-clusters*. [Doctoral dissertation]. UMass, Amherst, MA, United States.

van Leussen, J.-W., and Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Front. Psychol.* 6:1000. doi: 10.3389/fpsyg.2015.01000

Voeten, C. C. (2023). *buildmer: Stepwise elimination and term reordering for mixed-effects regression* (Version 2.11). Available online at: <https://cran.r-project.org/web/packages/buildmer/index.html>

Wayland, R., Landfair, D., Li, B., and Guion, S. G. (2006). Native Thai speakers' acquisition of English word stress patterns. *J. Psycholinguist. Res.* 35, 285–304. doi: 10.1007/s10936-006-9016-9

Yazawa, K., Whang, J., Kondo, M., and Escudero, P. (2020). Language-dependent cue weighting: an investigation of perception modes in L2 learning. *Second Lang. Res.* 36, 557–581. doi: 10.1177/0267658319832645

Zhou, C., and Hamann, S. (2024). Modelling the acquisition of the Portuguese tap by L1-Mandarin learners: a BiPhon-HG account for individual differences, syllable-position effects and orthographic influences in L2 speech. *Glossa* 9, 1–39. doi: 10.16995/glossa.9692

Zjakic, H. (2017). *Effects of orthography on monolingual and bilingual perception of non-native consonant clusters*. [MA thesis]. Western Sydney University, Penrith, Australia.