



The acquisition of L2 English complex onsets by L1 Farsi speakers

Noah Khaloo*, Department of Linguistics, University of California, San Diego, United States, nkhaloo@ucsd.edu

Connor Mayer, Department of Language Science, University of California, Irvine, United States

*Corresponding author.

Much previous work has shown that sibilant-initial complex onsets (SC onsets) differ in their typological, phonological, articulatory, and acquisitional properties from other onsets. The exact mechanism(s) underlying these differences are poorly understood. In this study, we investigate the acquisition and production of L2 English complex onsets by L1 Farsi speakers, focusing on differences between SC onsets and other onset types. Results from an experimental study corroborate past reports that Farsi speakers repair most SC onsets using epenthesis before the cluster and other onsets using epenthesis into the cluster. The results also support the claim that SC onsets are more difficult to produce and are acquired more slowly than other onsets. A phonological modeling study suggests that the epenthesis asymmetries observed in the experimental study are best accounted for by a pressure to maximize perceptual similarity to the unepenthesized forms. We close with speculation on how the variegated behavior of SC onsets can be given a holistic explanation under a perceptual account.



1. Introduction

A longstanding observation in language research is that complex onsets that begin with sibilants such as /s/ (*SC onsets*) are different from other types of complex onsets. These differences have so vexed researchers that, at the CASTL 2003 conference, phonologist Tobias Scheer was moved to comment that “if your theory can do s + C clusters, it must be wrong” (quoted in Parker, 2017). SC onsets are exceptional in a variety of ways. First, they are typological outliers. In general, languages prefer to make use of complex onsets where the difference in sonority between the first and second segment is large (e.g., Frisch, 2015). Within the class of SC onsets, however, the opposite preference is observed: SC onsets with lower sonority differences, such as /sp/ or /sn/, are used more commonly than those with greater sonority differences like /sl/ or /sɹ/ (Goad, 2012). *ST* onsets, a subset of the SC onsets where the second element is a voiceless stop, constitute a particularly striking example, as they are the most common across languages (Morelli, 2003), despite being ill-formed according to the Sonority Sequencing Principle (Selkirk, 1984; Clements, 1990).

Second, SC onsets frequently exhibit differences in their phonological behavior from other onset types, such as in deletion patterns (e.g., Kristoffersen & Simonsen, 2006), epenthesis patterns (e.g., Gouskova, 2001, 2004; Fleischhacker, 2001, 2005), and resyllabification in the presence of a preceding vowel (e.g., Chierchia, 1986; Kaye et al., 1990; Goad, 2012).

Finally, there appear to be differences in the articulation and acquisition of SC onsets. Articulatory studies have found that SC onsets tend to be produced with greater gestural overlap between the component segments (Poupplier et al., 2022) and more cohesive timing relative to the following vowel than other types of clusters (e.g., Marin, 2013). Some acquisition studies have shown that SC clusters are more difficult to acquire in L1 (e.g., Yavas et al., 2008; Jarosz, 2017) and L2 learning (Carlisle, 1991, 2001), though other studies have not found such effects.

Although much has been learned about SC clusters and how they differ from other types of clusters, there is still uncertainty about the mechanism(s) that drives their idiosyncratic phonological properties, their articulatory differences, and their acquisition trajectory. The goal of the current paper is to address some of these questions by investigating the production of L2 English by L1 Farsi speakers using an experimental study. There are two motivations for studying Farsi speakers. First, Farsi does not permit any complex onsets, which allows us to compare the acquisition of SC and other onset clusters in L2 directly, since neither is present in L1. Additionally, Farsi speakers repair the majority of SC onsets (excluding /sw/ and /sɹ/; Karimi, 1987; Fleischhacker, 2001; Shademan, 2002) by epenthesis a vowel before the cluster. This is problematic for several accounts of epenthesis asymmetry (see Section 2.2) and thus provides a useful data set for comparing the predictions of these accounts.

This paper has two specific aims: First, we hope to provide additional clarity on whether SC clusters are indeed more difficult to acquire than other clusters by relating L2 proficiency to epenthesis rates in both types of clusters. Although past work on this question has shown mixed results, our results are consistent with the claim that SC clusters are acquired more slowly. Second, we hope to provide clarity on the mechanism underlying the differences in epenthesis repair patterns between SC and other clusters by comparing two phonological accounts: one which suggests these differences arise from a pressure to minimize sonority rises across syllables (Gouskova, 2001, 2004) and one where they arise from a pressure to maximize perceptual similarity to the unpenthesized form (i.e., the underlying form without epenthesis) (Fleischhacker, 2001, 2005). Although previous work has assessed these analyses (e.g., Krämer, 2021), to our knowledge this is the first attempt to compare them directly in a quantitative phonological framework against experimental data. The results will show that the perceptual analysis best predicts our experimental data.

The paper¹ is structured as follows: In Section 2 we provide a background on the differences between SC and other types of onsets, with a focus on asymmetries in their repair across languages. In Section 3, we present the results of an experiment that tests the acquisition rates of these different onsets. Section 4 presents a phonological modeling study where we compare the two analyses of epenthesis asymmetry and provide support for the perceptual similarity account. Finally, in Section 5, we provide a discussion of the implications of our results and speculate on how the phonological, articulatory, and acquisitional properties of SC onsets can be related under a perceptual account.

2. Background

2.1 L2 learning of complex onsets

It has been previously observed that an individual's L1 can have profound effects on the acquisition of an L2. Generally speaking, if a structure/contrast is present in L2 but lacks a correspondent in L1, it can be difficult to learn. This is referred to as transfer effects (e.g., Lado, 1957; Eckman et al., 2003; Zampini & Edwards, 2008). One example of transfer effects is the influence of phonotactic restrictions in L1 on L2 acquisition. If L1 disallows a certain phonotactic structure, second language learners will often transfer this restriction to L2, even if the structure is permitted under L2. Thus, in order to become a competent speaker of L2, speakers must learn to “undo” these L1 constraints in their L2.

A specific instance of this kind of phonotactic transfer effect comes from the acquisition of complex onsets in L2. Many languages differ with respect to the types of onsets they permit.

¹ All of the code and data used in this paper can be found at https://github.com/connormayer/persian_epenthesis.

English, Spanish, and Farsi, for example, all have the sounds /p s l/. They differ, however, in how these sounds can be sequenced into onsets. In English, the legal complex onsets that can be generated from this set are /sp/ (as in ‘speech’), /sl/ (as in ‘sleek’), and /pl/ (as in ‘please’). Spanish is more restrictive, allowing /pl/ (as in ‘playa’), but not */sp/ or */sl/. Farsi is even more restrictive, allowing no complex onsets at all. When a native Spanish or Farsi speaker learns a more phonotactically permissive language like English, they will often struggle specifically with those onsets that violate syllable structure constraints present in their L1.

In both L1 and L2 phonology, a common strategy for dealing with illicit onsets is to repair them via insertion, deletion, or other mechanisms, such that a licit sequence is formed. The focus of this paper will be on the use of vowel epenthesis as a repair strategy, since this is the most observed repair strategy for our population of interest, L1 Farsi/L2 English speakers. Vowel epenthesis is defined as the insertion of a vowel that is not present underlyingly (e.g., Hall, 2011). The outcome of epenthesis is such that an illicit complex onset will be split over two syllables into simple onsets or codas. For example, the onset in a word like /flap/ ‘flop’ might be repaired to [fe.lap],² restructuring the complex onset into two simple onsets in separate syllables. Epenthesis has been claimed to be a significantly more common repair strategy than deletion of segments in the onset during L2 acquisition (Brasington, 1981).

2.2 Asymmetries in epenthesis location in complex onset repair

The placement of the epenthetic vowel within a word often varies systematically based on the segments in the onset, resulting in an asymmetric patterning. A majority of languages that display this asymmetry make a distinction between ST onsets and all other onsets. ST onsets tend to be repaired via pre-epenthesis (also known as prothesis; e.g., /spuk/ ‘spook’ → [es.puk]), while other onsets are repaired with medial epenthesis (also known as anaptyxis; /brɪŋ/ ‘bring’ → [be.rɪŋ]; e.g., Broselow, 1983; Fleischhacker, 2001; Shademan, 2002; Hall, 2011; Goad, 2012).

To account for this asymmetric patterning of epenthesis, Broselow (1992) claims that ST onsets are structurally represented as a single complex segment, which prevents them from being split in repair. One source of evidence used in this argument is that these clusters are notable exceptions to the Sonority Sequencing Principle (SSP; Selkirk, 1984), which dictates that languages prefer syllable onsets that rise in sonority as they move away from the syllable-initial position. Sonority is defined, roughly, as the resonance or loudness of a sound (see e.g., Clements, 2009), and the concept has been widely used in analyses of syllable structure and certain phonological processes. Sonority is often thought of as a scalar value, as shown in **Figure 1**.

² We do not consider the phonetic quality of epenthesized vowels here. Previous work on Farsi has suggested this vowel is typically [e], though copy epenthesis occurs in some contexts (e.g., Shademan, 2002; Ackbari, 2013). We will use [e] as the epenthetic vowel throughout the paper.

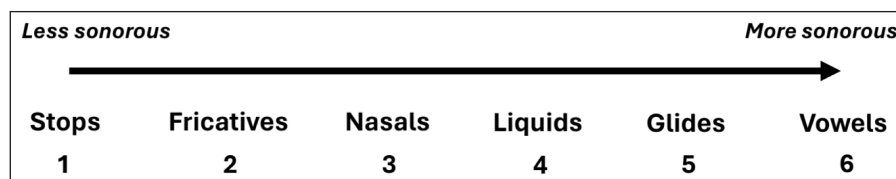


Figure 1: A scalar sonority hierarchy, showing that stops are considered less sonorous than fricatives, which are less sonorous than nasals, and so on.

Onsets consisting of a non-sibilant obstruent followed by a sonorant (henceforth, OR onsets) like /bl/, /fɹ/, and /tw/, satisfy the SSP because they rise in sonority: Their second sound is more sonorous than their first. ST onsets fall in sonority, violating the SSP, and should therefore be dispreferred. The fact that they are cross-linguistically common motivates their treatment as representationally distinct from other complex onsets. Other special representations for SC onsets have been proposed as well (for an overview, see Goad, 2012), but these all share the basic claim that SC onsets are somehow structurally different from other onset types.

Gouskova (2001, 2004) argues that, although this depiction of ST onsets does offer an explanation for their lack of “splittability” in interlanguage phonology, it does not sufficiently explain why an epenthetic vowel would be needed in the first place. In other words, if ST onsets comprise a single segment, then they should not violate L1 phonotactic constraints against complex onsets and should therefore not trigger epenthesis at all. Fleischhacker (2001, 2005) also argues that the single-segment account is unmotivated for several reasons (including some not repeated here). First, it does not explain the fact that some languages repair other types of sibilant-initial clusters (not just ST onsets), using pre-epenthesis as well (see also Krämer, 2021). Second, there is no obvious difference in markedness between ST clusters and OR clusters, despite the former’s purported higher degree of structural complexity. Both cluster types are commonly attested, and there is none of the implicational relationships that we often see between less marked and more marked structures.

We will not consider the single-segment account further in this paper, but we note that it is generally incompatible with our results, which show that Farsi speakers repair all SC onsets except for /sw/ with pre-epenthesis, not just ST onsets as Broselow’s theory might predict.

Other theories have been proposed to account for the epenthesis asymmetry phenomenon across languages based on the sonority of the individual sounds (e.g., Singh, 1985). Gouskova (2001, 2004) proposes that the asymmetric patterning is due to an innate dispreference against rises in sonority across syllable boundaries. This is referred to as the Syllable Contact Law (SCL; Hooper, 1976; Murray & Vennemann, 1983; Vennemann, 1988; Seo, 2011). In these analyses, researchers similarly make use of a scalar sonority hierarchy (e.g., Figure 1) to represent the relative sonority levels of different types of segments (e.g., Hooper, 1976; Murray & Vennemann, 1983; Vennemann, 1988; Clements, 1990; and for an overview, see Parker, 2002).

Gouskova (2001, 2004) suggests that the SCL, along with a preference towards pre-epenthesis, account for the epenthesis asymmetry between SC and OR onsets. In an OR onset like in /pliz/, repair using pre-epenthesis to [ep.liz] results in a violation of the SCL, since [p] at the end of the first syllable has a lower sonority than [l] at the beginning of the second. Medial epenthesis to [pe.liz] results in no violation of the SCP and so is preferred. In an SC onset like in /stap/, neither pre-epenthesis to [es.tap] nor medial epenthesis to *[se.tap] results in an SCL violation. Gouskova motivates the preference for [es.tap] as reflecting a bias against medial epenthesis. In other words, pre-epenthesis is preferred unless doing so would violate the SCL, in which case medial epenthesis is used.

Farsi poses a problem for this analysis because it repairs almost all SC onsets using pre-epenthesis (see Fleischhacker, 2001, 2005; Krämer, 2021). For example, the SCL-based analysis predicts that Farsi speakers should repair /slæb/ ‘slab’ to [se.læb], since [es.læb] produces an SCL violation. Despite this, [es.lab] is the preferred repair strategy among Farsi speakers, suggesting that avoidance of SCL violations is not the mechanism driving epenthesis asymmetries in Farsi. Krämer confirms this empirically by asking Iranian Farsi speakers to produce attested loanwords and wug words with complex onsets of varying sonority profiles, including many not attested in English. Although he observes a high amount of variability between participants, Krämer finds several robustly attested epenthesis patterns that are not predicted by the SCL, casting further doubt on its ability to account for these epenthesis asymmetries.³

Farsi is not exceptional in how it repairs SC onsets. Fleischhacker (2001, 2005) points out that variability in the treatment of SL (sibilant + liquid) and SN (sibilant + nasal) clusters has also been displayed in Hindi (Bharati, 1994), as well as in Russian loanwords produced in Kazakh (Sulejmanova, 1965), among other languages. With this in mind, Fleischhacker (2001, 2005) proposes that epenthesis repair asymmetries emerge from pressures to minimize perceptual distance from the unepenthesized form. In a series of studies, she observes that English speakers find medial epenthesis to be more perceptually disruptive than pre-epenthesis into ST clusters, while in other clusters, including SR (sibilant + sonorant) and OR clusters, medial epenthesis was judged to be less disruptive. Based on these results, she proposes that complex onsets across languages are repaired in such a way as to minimize perceptual distance from the unepenthesized form. The perceptual perspective is supported by independent evidence showing that L1 speakers

³ Krämer (2021) does not consider these results in light of the perception-based model described in the next paragraph but suggests that the SCL analysis could be salvaged by positing a constraint such as *sV (no sequences of [s] followed by a vowel) or a constraint that penalizes stops in coda position, which would be violated by medial epenthesis into SC clusters and pre-epenthesis into OR clusters respectively. We leave this open as a possibility but note a challenge for this analysis. These constraints cannot stem from L1 factors, since both /sV/ sequences (e.g. /salom/ ‘Hello’, /suxt/ ‘burned’, /soal/ ‘question’, /se/ ‘three’) and coda stops (see Krämer) are quite common in Farsi. Because of this, Krämer suggests this could be a TETU effect (McCarthy & Prince, 2004). However, it is unclear that coda [s] is less marked than coda stops (e.g., VanDam, 2004; Krämer & Zec, 2020), or that an sV sequence is more marked than a stop-vowel sequence.

of Spanish, which does not allow ST onsets and repairs them using pre-epenthesis, commonly perceive pre-epenthesized vowels in such clusters, even when they are not present (e.g., Cuetos et al., 2011; Gibson, 2012; Carlson, 2019). Thus, not only is pre-epenthesis into SC clusters less perceptually disruptive than medial epenthesis, it also aligns with a common misperception of these clusters by speakers whose languages do not allow them.

The phonological modeling study in Section 4 will directly compare the SCL analysis (Gouskova, 2001, 2004) and the perception-based analysis (Fleischhacker, 2001, 2005). We will discuss the formalization of these analyses in more detail in that section.

2.3 The acquisition and articulation of complex onsets

In addition to the differences in their preferred epenthesis locations, SC onsets appear to behave differently from OR onsets in terms of their acquisition and their articulation.

Given the marked status that the SSP ascribes to ST clusters, one might expect them to be acquired more slowly than other clusters in both L1 and L2. The evidence for this is equivocal. In L1, there is some evidence that SC onsets are indeed acquired later than other types of clusters, in the sense of being repaired more frequently by children (e.g., Barlow, 2001; Yavaş, et al. 2008), though other studies have failed to find this difference or found that SC onsets were acquired earlier (e.g., Smit, 1993; Gierut, 1999; Yavaş & Core, 2006). It has also been shown that generalization (i.e., the learning of one complex onset facilitating the learning of another) occurs within the SC and OR classes but not between them (e.g., Gierut, 1999, 2001). Finally, in both SC and OR onsets, the greater the difference in sonority between the first and second elements in the cluster (the better it abides by the SSP), the less likely learners are to repair that cluster (e.g., Jarosz, 2017).

Similarly, conflicting results have been found for L2 speakers. Epenthesis rates in general have been shown to decrease as L2 proficiency increases (Boudaoud & Cardoso, 2009; Yazawa et al., 2015). Some studies have found that L2 learners repair SC onsets more frequently than OR onsets (e.g., Carlisle, 2001; Yildiz, 2005), but other papers have failed to find such an effect (Sherwin, 1999; Tessier et al., 2013). As well, within the classes of SC and OR onsets, learners are generally less likely to repair onsets with greater sonority differences between the first and second segments (Carlisle, 2001; Cardoso, 2007; Boudaoud & Cardoso, 2009).

Studies investigating articulatory timing have also found differences between the articulation of SC clusters and OR clusters. Pouplier et al. (2022) show that across seven languages, sibilant-initial clusters exhibit greater articulatory overlap between the two segments and decreased variability in timing when compared to stop-initial onsets. With regards to temporal organization, SC onsets appear to display a more robust “c-center” effect (Browman & Goldstein, 2000) in comparison to singleton onsets (Pastätter & Pouplier, 2014), as well as coda /s/ clusters (Marin, 2013). This does

not seem to be the case with stop-initial clusters. Although this heightened articulatory coupling between the segments in SC onsets is generally consistent with the notion that they constitute a single complex segment, the relationship between these results and the sonority and perceptual accounts of SC cluster behavior is less clear. We will return to this question in the discussion at the end of the paper.

2.4 The current study

In the remainder of the paper, we analyze epenthesis asymmetry patterns in native Farsi speakers who have learned English as their L2. As noted above, we choose this population because Farsi has no complex onsets, allowing us to compare SC and OR onsets directly and because Farsi speakers have been shown to repair a majority of SC onsets using pre-epenthesis, which poses a challenge for several of the theoretical analyses described above. We also seek to provide additional clarity on whether the cross-linguistic asymmetry in repair that has been displayed between SC and OR onsets corresponds to an asymmetry in acquisition as well.

Section 3 presents the results of an experimental study looking at the production of English complex onsets by L1 Farsi/L2 English speakers of varying abilities. Section 4 presents a phonological analysis and modeling study of the experimental data. Section 5 discusses the implications of these results and speculates on how a perceptual account of SC clusters can unify some of their heterogeneous properties.

3. Experimental Study

The following describes an experimental study done to better understand the effects of L2 proficiency, which we assess quantitatively, and onset type on epenthesis patterns. Based on the discussion in the previous section, we expect to see: (a) that (most) SC onsets will be repaired with pre-epenthesis and OR onsets with medial epenthesis; (b) that SC onsets will undergo more epenthesis than OR onsets; (c) that higher English proficiency will lead to less epenthesis; and (d) that participants with lower English proficiency will show relatively higher rates of medial epenthesis (predicted for OR onsets) relative to pre-epenthesis (predicted for SC onsets), suggesting delayed acquisition of SC onsets. In other words, we expect that rates of epenthesis for OR onsets will decrease more rapidly than rates of epenthesis for SC onsets as English proficiency increases.

3.1 Methods

We recruited 20 native Farsi speakers living in Orange County, California who learned English as a second language (14 male; 36–80 years old⁴) and tasked them with producing 78 English target

⁴ The age range of the participants is on the higher side because a majority of the Farsi speaking population in the US immigrated from Iran during the 1970s–1990s. This means that much of the population of L1 Farsi/L2 English speakers in the US is above the age of 35.

words with complex onsets. Thirty-seven of these target words contained SC onsets, while the other 41 contained OR onsets. Note that all SC onsets were /s/-initial, given the marginal status of /ʃ/-initial complex onsets in English. Participants were recruited by word of mouth. After providing consent, they were asked to fill out a language background questionnaire (described in the following section). Participants were then asked to read three passages containing the target words (see Appendix A). Finally, participants were asked to read each target word in the context of a carrier phrase (*'_ is a good word'*). The orders of the three passages and the carrier sentences were varied for each participant.

Audio recordings were made for 13 of the participants using a Rode Lavalier Go microphone and Focusrite Scarlet 2i2 USB audio interface. Due to technical issues, the remaining seven participants were recorded using the integrated microphone on a 2017 Macbook Pro. Although this resulted in decreased audio quality for these participants, recordings were only used qualitatively to identify the presence of epenthesis, and the quality of these recordings was sufficient for this purpose. In both cases, the recordings were made with a sampling rate of 44.1 kHz and a bit depth of 16. Recordings were made in quiet rooms at participants' homes or places of work.

Due to variable English reading ability among our participants, four were only able to read the carrier phrases and not the passage. In addition, two participants declined to read the carrier phrases after reading the passages. One participant was entirely excluded due to poor audio quality, and two participants' carrier phrase productions were excluded for the same reason. This left a total of 19 participants, with four having read only the carrier phrases, four only the passages, and eleven both. Words that participants were unable to read or that were read incorrectly were omitted, leaving a total of 2,129 tokens. See Appendix B for a full list of onset types and their production counts.

For each target word, epenthesis type (pre-epenthesis, medial epenthesis, or no epenthesis) was recorded. The presence of an epenthetic vowel was confirmed by examination of the waveform and spectrogram for each token, with the criteria that epenthetic vowels must display clear formant structure in the expected location (i.e., either before the onset or within the onset) and differences in amplitude sufficient to distinguish them from surrounding context.

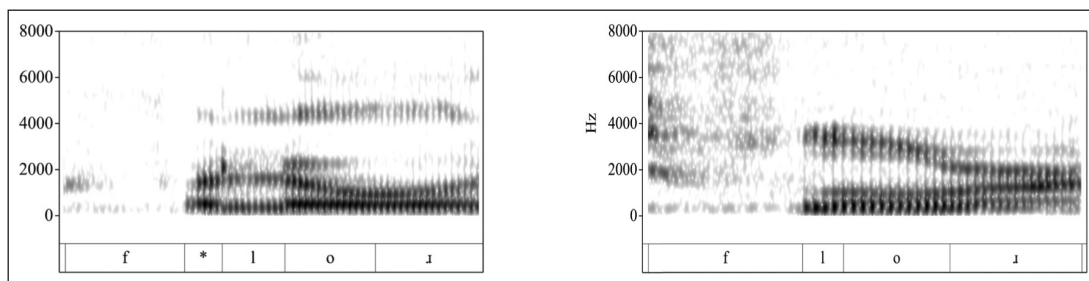


Figure 2: Tokens of the word /flor/ 'floor/ produced with medial epenthesis (left) and without medial epenthesis (right).

Figure 2 shows productions of the word /floʌ/ “floor” by two different participants. The image on the left shows a production with medial epenthesis, with the epenthetic vowel indicated by *. The presence of this vowel is reflected in the spectrogram by increased amplitude and distinct formant structure from both the initial /f/ and the following /l/. The image on the right shows a production without epenthesis.

Given the structure of the passages that participants were prompted to read, some of the onsets were preceded by a word-final vowel. For example, “the stove.” Though we expect such configurations to reduce rates of epenthesis, since the initial segment of the complex onset can be syllabified as the coda of the preceding syllable (e.g., Boudaoud & Cardoso, 2009), these configurations pose a challenge for detecting pre-epenthesis: cases where there is word-final vowel followed by pre-epenthesis (e.g., “the estove”) could be difficult to distinguish from cases without epenthesis. In cases of pre-epenthesis, we observed that there was often a glottal stop between the offset of the word-final vowel and the epenthetic vowel. In other cases, this juncture was less apparent. In these more ambiguous cases, we relied on F2 transitions to determine whether the preceding material constituted a single vowel (“the stove”) or a pair of adjacent vowels (“the estove”). **Figure 3** provides an example of the latter case, with the epenthetic vowel indicated using “*”. Here we can see differences in the formants between the underlying and epenthetic vowel. The glottal pulses in the epenthetic vowel are also slower and less periodic, indicating weak glottalization, which is also consistent with an epenthetic vowel. In our dataset of roughly 2,100 tokens, there were only 18 cases where a word-final vowel preceded a pre-epenthesized onset.

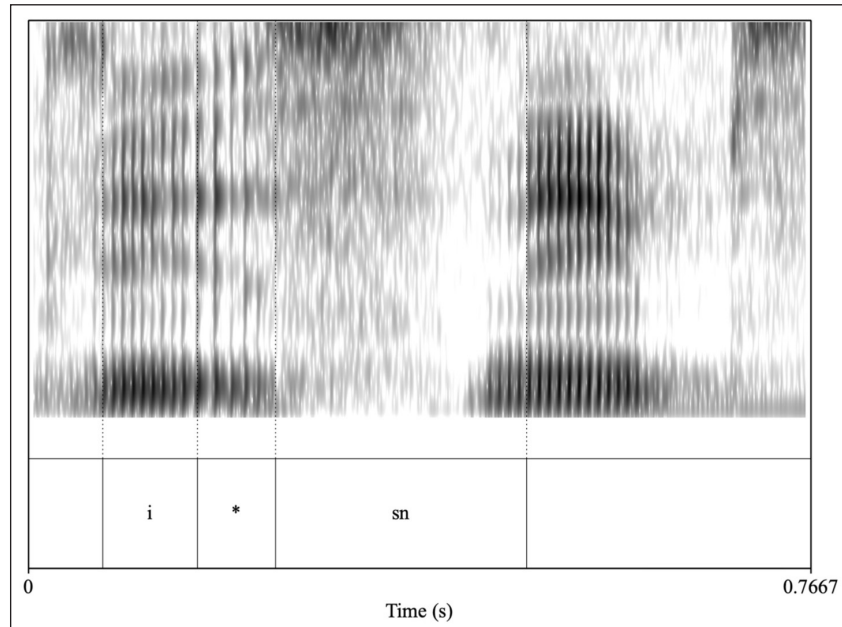


Figure 3: An example of pre-epenthesis in a post-vocalic context in the phrase “he sniffs.”

3.2 Scalar measurement of L2 proficiency

Before reading the passages and carrier phrases, participants completed the LEAP-Q (Marian et al., 2007), an experimentally-validated questionnaire measuring the language profiles of multilingual speakers. To create a scalar measure of language proficiency suitable for both the statistical model described here and the phonological analyses presented in the next section, we retained only the LEAP-Q questions with numeric responses and conducted a Principal Components Analysis to reduce these responses to a single dimension (for a similar approach, see Danielson, 2012). In each case described below, we used the first Principal Component (PC1) as the measure of interest. This component always displayed a strong positive correlation with Farsi ability and exposure. In order to ease interpretation of this value, we negated it so that higher positive values represent greater L2 English proficiency. We refer to this negated score as Relative English Dominance (RED).

Following Marian et al. (2007), who advise against running a PCA on the entirety of the LEAP-Q data, we classified the numeric questions in the questionnaire into one of three categories based on the primary question categories described in the paper and ran separate PCA's on each of these bins, as well as on the full set of numeric questions. This produced four different versions of RED:

- (1) **RED_{acquisition}**: based on questions related to age of L2 acquisition and L1/L2 exposure
- (2) **RED_{immersion}**: based on questions related to L1/L2 immersion
- (3) **RED_{self-report}**: based on questions related to self-reported L1/L2 proficiency
- (4) **RED_{full}**: based on all numeric questions combined

The specific questions from the LEAP-Q that were categorized into each of these bins are provided in Appendix C.

In order to determine which implementation of RED best predicts epenthesis patterns, we fit nine instantiations of our statistical model (specific details of this model are given in the following section). Eight of these had as fixed effects each possible combination of RED_{acquisition}, RED_{immersion}, and RED_{self-report}, while the final used only RED_{full}.

We compared these models using the Bayesian Information Criterion (BIC; Schwarz, 1978; Raferty, 1995), which rewards model fit and penalizes model complexity, with lower scores being better. The results of this analysis showed that the model containing only RED_{acquisition} performed the best. Accordingly, the results we report below come from the model that uses this score to characterize English proficiency, which we will simply abbreviate as RED going forward. The full details of the model comparison can be found in Appendix D, and Appendix E presents the loadings of the RED_{acquisition} principal component.

3.3 Results

Figure 4 shows the proportion of each response type (no epenthesis, medial epenthesis, or pre-epenthesis) broken down by onset type. The total number of tokens for each response type are given above each bar. Most tokens were produced without epenthesis. In cases of epenthesis, all SC onsets are repaired with pre-epenthesis, except for /sw/, which is repaired with medial epenthesis, and all OR onsets are repaired with medial epenthesis. Overall epenthesis rates are greater for SC onsets than OR onsets. **Figure 5** shows epenthesis rates broken down by individual onsets (e.g., /st/, /pl/ /br/ etc.).

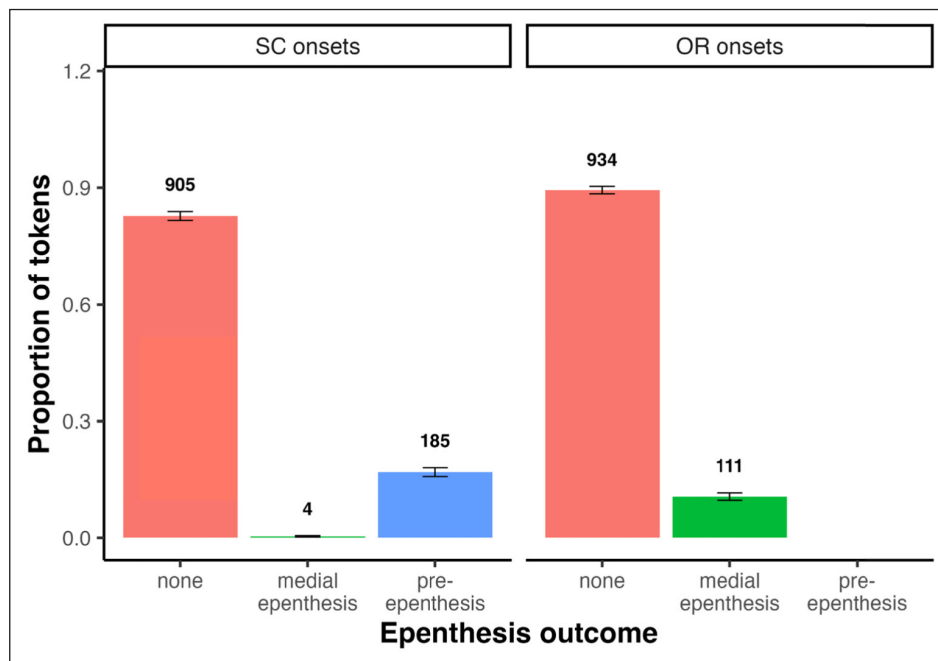


Figure 4: Proportion of each epenthesis outcome for SC onsets and OR onsets. Error bars indicate standard error.

Figure 6 shows how overall epenthesis rates and the proportion of pre- vs. medial epenthesis correlate with RED. A higher RED score is associated with lower overall rates of epenthesis and with a higher rate of pre-epenthesis repair relative to medial epenthesis repair.

Epenthesis rate is calculated as the proportion of complex onsets in our dataset that are produced with an epenthesized vowel, and pre-epenthesis rate is calculated as the proportion of cases of epenthesis where pre-epenthesis is used. There is an inverse relationship between English dominance and overall epenthesis rates (**Figure 6**, left). However, as proficiency rises, speakers' repaired productions show a greater proportion of pre-epenthesis relative to medial epenthesis (**Figure 6**, right).

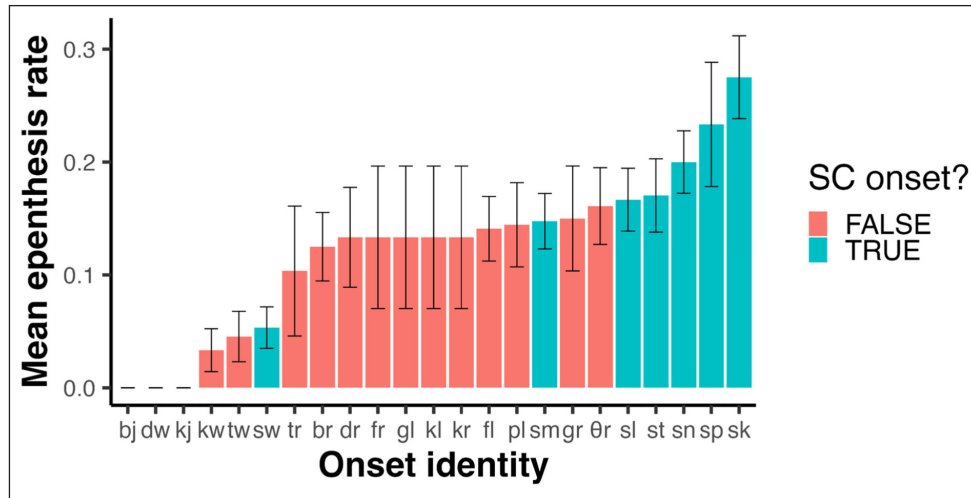


Figure 5: Mean epenthesis rates for each onset in the experimental study. Error bars indicate standard error.

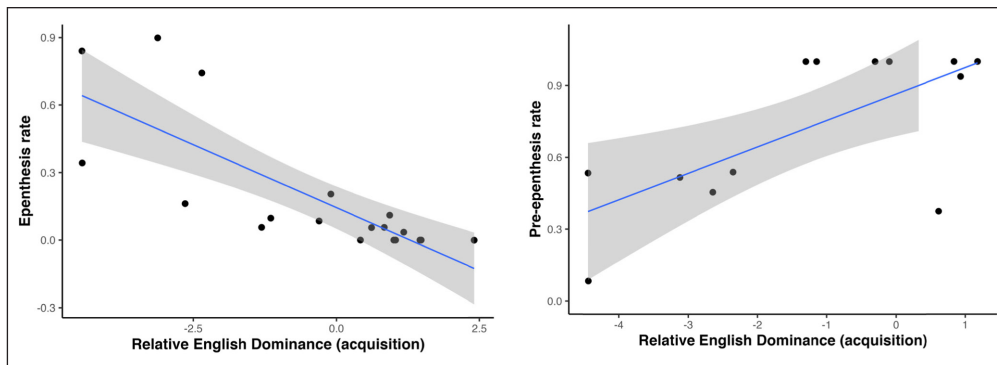


Figure 6: (Left) Epenthesis rate by RED_{acquisition}. (Right) Pre-epenthesis rates by RED.

We fit a mixed-effects logistic regression model using the lme4 package in R to provide quantitative support for these observations. Because epenthesis location was almost perfectly predicted by onset type (with /sw/ the exception), we binned tokens of pre-epenthesis and medial epenthesis together, such that the dependent variable was whether epenthesis occurred (1) or not (0). The independent variables were whether the onset was preceded by a vowel, onset type (SC or OR), context (passage or carrier phrase), RED (scaled and centered), as well as an interaction between RED and onset type. The model also included random intercepts for both speaker and word.

The details of the fitted model are shown in **Table 1**. The results indicate that the presence of a preceding vowel is not a significant predictor of epenthesis ($\beta = 0.297$, $p = 0.47$). We also see that SC onsets are more likely to be epenthesized than OR onsets ($\beta = 4.21$, $p < 0.001$), and

	Coefficient estimate	Std. Error	Z-value	p-value
Intercept	−8.3544	1.0146	−8.234	<0.001
Preceding Vowel (0 = no, 1 = yes)	0.2968	0.4108	0.722	0.47
Onset type (0 = TR, 1 = sC)	4.2147	1.0640	3.961	<0.001
Context (0 = carrier phrase, 1 = passage)	1.2867	0.2924	4.400	<0.001
RED	−4.4311	0.5916	−7.490	<0.001
RED * Onset type (sC)	2.1514	0.3064	7.020	<0.001

Table 1: The details of the logistic regression model fit to the experimental data.

a complex onset is more likely to be repaired in a passage than in a carrier sentence ($\beta = 1.287$, $p < 0.001$). The model also suggests that a higher RED is associated with lower overall rates of epenthesis ($\beta = -4.33$, $p < 0.001$). The interaction between RED and Onset Type ($\beta = 2.15$, $p < 0.001$) shows that the effect of RED is modulated by onset type: increases in RED cause the probability of epenthesis into OR clusters to decline more quickly than for SC clusters.

3.4 Intermediate summary

The results⁵ of the experiment presented above indicate that:

1. L1 Farsi/L2 English speakers repair OR onsets using medial epenthesis, and SC onsets using pre-epenthesis. The exception to this is /sw/, which was repaired using medial epenthesis.
2. SC onsets displayed overall higher rates of repair than OR onsets.
3. Overall epenthesis rates decreased as English proficiency increased, as measured using the RED score calculated from the LEAP-Q survey.
4. More proficient English speakers showed higher rates of pre-epenthesis relative to medial epenthesis.

These results corroborate the asymmetries in epenthesis location found in previous research on Farsi speakers and support the claim that SC onsets are more difficult and slower to acquire than OR onsets. In the following section, we will use these experimental results to test the predictions of two phonological accounts of epenthesis asymmetries.

⁵ We also conducted a pilot study looking at recordings in the Speech Accent Archive (<https://accent.gmu.edu/>), which contains recordings of L1 Farsi/L2 English speakers reading passages. We do not present these results here, but note that they were qualitatively similar to the experimental study.

4. Phonological modeling

In this section we will present the results of a phonological modeling study. The primary goal of this study is to evaluate the analyses of epenthesis asymmetry proposed by Gouskova (2001, 2004) and Fleischhacker (2001, 2005) to determine which best predicts the patterns in the experimental data we collected. We will do so using the maximum entropy constraint grammar framework (henceforth MaxEnt; Goldwater & Johnson, 2003; Hayes & Wilson, 2008), which is well-suited to modeling the kinds of variability observed in the production data. Variants of each analysis will be implemented in this framework, and their empirical predictions compared against the experimental data from the previous section.

In carrying out this modeling study, we will also demonstrate a simple technique that uses constraint weight scaling to account for the influence of speaker proficiency on epenthesis rates. This allows us to factor out participant-specific differences to more clearly observe global patterns. Additionally, we will show that employing separate markedness constraints for SC and TR onsets better predicts the data, as do analyses that encode separate learning rates for each.

We begin this section with a formal description of the two analyses under comparison.

4.1 The Syllable Contact analysis

As discussed above in Section 2, Gouskova (2001, 2004) accounts for epenthesis asymmetries by proposing that pre-epenthesis is the default strategy cross-linguistically but that medial epenthesis is used to avoid repairs that would violate the Syllable Contact Law (SCL), which penalizes sonority rises across syllable boundaries. This process is demonstrated in the pseudo-tableau in **Table 2**, in which the word /pliz/ ‘please’ is repaired with either medial epenthesis or pre-epenthesis. Δ_{sonority} refers to the difference in sonority between the first segment of the second syllable, and the final segment of the first syllable, using the scale in **Figure 1**:

/pliz/ ‘please’	Δ_{sonority}
☞ [pe.liz] (medial epenthesis)	son(l) – son(e) = 4–6 = –2
[ep.liz] (pre-epenthesis)	son(l) – son(p) = 4–1 = +3

Table 2: Calculation of Δ_{sonority} for ‘please’ repaired with medial epenthesis or pre-epenthesis.

If ‘please’ were to be repaired using pre-epenthesis (/pliz/ → [ep.liz]), there would be a rise in sonority ($\Delta_{\text{sonority}} = +3$) across the new syllable boundary, violating the Syllable Contact Law. On the other hand, if ‘please’ were to be repaired with medial epenthesis (/pliz/ → [pe.liz]), the output would have a fall in sonority across syllable boundaries ($\Delta_{\text{sonority}} = -2$), which does not violate this law. Thus, the latter strategy emerges as the preferred choice. Conversely, **Table 3** shows a pseudo-tableau for the word ‘steep’ repaired by either medial or pre-epenthesis.

/stip/	Δ_{sonority}
☞ [es.tip]	$\text{son}(\text{t}) - \text{son}(\text{s}) = 1 - 2 = -1$
[se.tip]	$\text{son}(\text{t}) - \text{son}(\text{e}) = 1 - 6 = -5$

Table 3: Calculation of Δ_{sonority} for ‘steep’ repaired with medial epenthesis or pre-epenthesis.

Here, if ‘steep’ were repaired using pre-epenthesis (/stip/ → [es.tip]), there would be a fall in sonority across the resulting syllable boundary ($\Delta_{\text{sonority}} = -1$). Medial-epenthesis (/stip/ → [se.tip]) would also yield a fall in sonority ($\Delta_{\text{sonority}} = -5$). From the perspective of the SCL, therefore, both forms are acceptable. Because [es.tip] emerges as the winner, Gouskova takes this as evidence for pre-epenthesis as the default repair strategy, with medial epenthesis only used when pre-epenthesis would result in an SCL violation, as in **Table 2**.

Gouskova implements this analysis using the following constraints:

SyllableContact: Sonority must not rise across a syllable boundary.

***Complex:** No tautosyllabic consonant sequences.

Dep: Don’t epenthesize.

Contiguity: Elements adjacent in the input must be adjacent in the output.

Tables 4 and 5 showcase two tableaux demonstrating the ranking proposed in Gouskova (2001, 2004). **Table 4** shows that the analysis predicts medial epenthesis for OR onsets.

/frut/	*Complex	Dep	SyllableContact	Contiguity
a. ☞ fe.rut		*		*
b. ef.rut		*	*! (f.r)	
c. frut	*!			

Table 4: A tableau showing medial epenthesis under the Syllable Contact analysis.

The high-ranked *Complex constraint rules out the unepenthesized candidate 4c. Although Candidate 4b satisfies the Contiguity constraint, it does so at the cost of producing a violation of SyllableContact because of the sonority rise from [f] to [r] across the syllable boundary. Candidate 4a is accordingly chosen as the winner, despite its violation of Contiguity.

Table 5 below shows that the analysis predicts pre-epenthesis for ST clusters. Candidate 5c is again ruled out because of its violation of *Complex. In contrast to the pre-epenthesized candidate in the previous tableau, Candidate 5b in this tableau does not violate the SCL, because there is a sonority drop from [s] to [p] across the syllable boundary. Although Candidate 5b

/spitʃ/	*Complex	Dep	SyllableContact	Contiguity
a. sepitʃ		*		*!
b. ɛspitʃ		*		
c. spitʃ	*!			

Table 5: A tableau showing pre-epenthesis under the Syllable Contact analysis.

also satisfies the SCL, its violation of Contiguity casts the deciding vote against it, and the pre-epenthesized candidate 5b comes out as the winner.

We will refer to the analysis above as the **Simple Syllable Contact** analysis. This constraint hierarchy is structured such that, in the absence of a SyllableContact violation (i.e., neither pre- nor medial epenthesis results in a rise in sonority across syllable boundaries), Contiguity will ensure that the pre-epenthesized form is chosen. In other words, this analysis predicts that all onsets with flat or falling sonority should be repaired by pre-epenthesis (ST onsets) and all others with medial epenthesis (OR and SR onsets).

Gouskova (2001, 2004) also presents an extension of this analysis that divides the SyllableContact constraint into separate constraints for each degree of sonority rise/fall across syllables. She argues that these constraints are intrinsically ranked, such that higher sonority rises are always penalized more than lower sonority rises.

$$\text{SyllableContact}_{\Delta_4} > > \text{SyllableContact}_{\Delta_3} > > \dots > > \text{SyllableContact}_{\Delta_4} > > \text{SyllableContact}_{\Delta_5}$$

This hierarchy can be interspersed with other markedness or faithfulness constraints, such that the degree of sonority rise tolerated and the repair strategy used for different sonority rises can differ. We will refer to this analysis as the Complex Syllable Contact analysis.

As mentioned previously, the major issue with the Syllable Contact analysis is that it fails to account for the behavior of subtypes of SC onsets, such as SR (sibilant + sonorant) onsets in certain languages (Krämer, 2021; Fleischhacker, 2001, 2005). This phenomenon is evident in Farsi, which repairs SN (/s/ + nasal) and SL (/s/ + liquid) onsets using pre-epenthesis, despite this resulting in a violation of the SCL. **Table 6** shows an example of how the Syllable Contact analysis predicts SR clusters should behave.

Here, the Syllable Contact analysis predicts that SN onsets will be repaired with medial epenthesis, since pre-epenthesis produces a violation of the SCL. The same outcome is also predicted for SL onsets. Though some languages are known to repair SR onsets using medial epenthesis, the Syllable Contact Law does not predict Farsi speakers' repair of these onsets with pre-epenthesis (Fleischhacker, 2001, 2005; Krämer, 2021). These issues persist even in the Complex Syllable Contact analysis (see Section 4.7 for discussion as to why).

/snɪf/	*Complex	Dep	SyllableContact	Contiguity
a.  se.nɪf		*		*
b.  es.nɪf		*	*! (s.n)	
c. snɪf	*!			

Table 6: A tableau showing that medial epenthesis of SN clusters is predicted under the Syllable Contact analysis. The sad face indicates the candidate who should have won but did not, while the bomb indicates the candidate that should not have won, but did.

4.2 The Perceptual Cost analysis

An alternative account of epenthesis asymmetries is presented by Fleischhacker (2001, 2005), who argues that these patterns are the results of pressure to maximize perceptual similarity between the epenthesized and unepenthesized forms.

Fleischhacker provides experimental evidence to show that the perceptual cost of splitting an ST onset with medial epenthesis is greater than that of splitting an SR (sibilant + sonorant) or OR onset. In her experiments, Fleischhacker presented English-speaking participants with words that began with initial ST, SN (sibilant + nasal), SL (sibilant + liquid), SW (sibilant + glide), and OR onsets. These words were modified with either pre-epenthesis or medial epenthesis. Participants were then asked which modified form of the word sounded more like the unepenthesized word. The results demonstrate that speakers found pre-epenthesis less disruptive than medial epenthesis for SC onsets, while medial epenthesis was generally preferred for other onset types. Subsequent studies have corroborated aspects of this proposal (e.g., Davidson & Shaw, 2012). Based on these experimental results, as well as typological considerations, Fleischhacker proposes the following ranking of different onset types based on how perceptually disruptive medial epenthesis is:

$$\begin{array}{ccccccc}
 \text{ST} & > > & \text{SN} & > > & \text{SL} & > > & \text{SW} & > > & \text{OR} \\
 \{ \text{medial epenthesis less likely} \} & & \rightarrow & & \{ \text{medial epenthesis more likely} \}
 \end{array}$$

In other words, the perceptual difference between inserting and not inserting a vowel in the context S_T (e.g., medial epenthesis in a word like ‘stop’ [se.tɒp]) is greater than the perceptual difference between inserting and not inserting a vowel in the context S_N (e.g., medial epenthesis in a word like ‘snap’ [se.næp]), and so on. The pressure to maximize perceptual similarity between epenthesized and non-epenthesized forms means that onsets that are higher in this ranking should be more likely to be repaired via pre-epenthesis, since it results in a relatively smaller degree of perceptual disruption than medial epenthesis.

The core constraints of this analysis are implemented using P-Map (perceptual map) constraints (Steriade, 2000, 2001; Zuraw, 2007, 2013), which penalize candidates based on

the perceptual distance between underlying and surface forms.⁶ These constraints typically penalize greater perceptual deviations more strongly than smaller ones (though this effect may be modulated by language exposure).

We follow Fleischhacker in using contextual Dep constraints of the form Dep-V/X_Y. These constraints are violated when a vowel is inserted into the context X_Y, where X and Y refer to segmental classes. We adopt the following specific constraints:

Dep-V/S_T: Do not insert a vowel between a sibilant and a stop.

Dep-V/S_N: Do not insert a vowel between a sibilant and a nasal.

Dep-V/S_L: Do not insert a vowel between a sibilant and a liquid.

Dep-V/S_W: Do not insert a vowel between a sibilant and a glide.

Dep-V/O_R: Do not insert a vowel between a non-sibilant obstruent and a sonorant.⁷

Fleischhacker suggests that these constraints are ranked in a hierarchy in accordance with the scale of perceptual disruptiveness presented above.

Dep-V/S_T >> Dep-V/S_N >> Dep-V/S_L >> Dep-V/S_W >> Dep-V/O_R

This hierarchy requires that the cost of insertion into an ST cluster is higher than into an SN cluster, which is higher than into an SL cluster, and so on. In other words, epenthesis into a cluster is penalized proportionally to how perceptually distant the resulting epenthesized form is from the unepenthesized or underlying form.

Fleischhacker also uses the following additional constraints, some of which are shared with Gouskova's (2001, 2004) analysis:

***Complex:** No tautosyllabic consonant sequences.

C/V: No consonant sequences at all.

Contiguity: Elements adjacent in the input must be adjacent in the output.

L-Anchor: Elements on the left edge of the input must be on the left edge of the output.

⁶ McCarthy (2009) raises a problem for P-Map constraints: Underlying forms are considered to be underspecified for many perceptually-relevant features, such as whether a stop is released, making it unclear how perceptual similarity to an abstract underlying form should be determined. It is beyond the scope of this paper to address this issue, but one could assume in this case that P-Map constraints enforce a perceptually-based paradigm uniformity (Steriade, 2000) with the unrepaired candidate (the native pronunciation target) rather than between underlying and surface representations.

⁷ Although Fleischhacker (2001, 2005) finds that /f/-initial onsets pattern with TR onsets in her experimental studies, she does not deal with /f/- and /θ/-initial onsets directly in her formal analysis. In our analysis, clusters beginning in non-sibilant fricatives and clusters beginning in stops both violate Dep-V/O_R.

Example tableaux based on Zuraw (2013) are shown below. In the first tableau (**Table 7**), the faithful realization of /frut/ (candidate 7c) is ruled out because it violates *Complex. Candidate 7b satisfies *Complex, but violates L-Anchor because the /f/ at the left edge of the input is no longer at the left edge in the output. The medial epenthesis form (candidate 7a) is chosen because it satisfies *Complex while also avoiding a violation of L-Anchor.

/frut/	*Complex	Dep-V/S_T	L-Anchor	Dep-V/O_R
a.  ferut				*
b. efrut			*!	
c. frut	*!			

Table 7: A tableau showing medial epenthesis under the Perceptual Cost analysis.

In the second tableau (**Table 8**), the pre-epenthesis form (candidate 8b) is chosen because it avoids a violation of Dep-V/S_T, which strongly penalizes epenthesis into ST clusters because of its perceptual disruptiveness.

/spitʃ/	*Complex	Dep-V/S_T	L-Anchor	Dep-V/O_R
a. sepiʃ		*!		
b.  espiʃ			*	
c. spitʃ	*!			

Table 8: A tableau showing pre-epenthesis under the Perceptual Cost analysis.

In general, any structures whose corresponding contextual Dep constraint is ranked below L-Anchor will be repaired by medial epenthesis, while those whose Dep constraint is ranked above L-Anchor will be repaired by pre-epenthesis. In our data, both Dep-V/S_W and Dep-V/O_R must be ranked lower than L-Anchor, since /sw/ and OR onsets are always repaired using medial epenthesis.

In the following section we will describe the MaxEnt framework, which we will use to implement variants of these two analyses.

4.3 Maximum entropy constraint grammars

MaxEnt is a variant of Harmonic Grammar (Legendre et al., 1990; Prince & Smolensky, 1993/2004; Pater 2009), which is a generalization of Optimality Theory (OT; Prince & Smolensky, 1993/2004) where constraints are assigned numeric weights instead of being ranked. Higher weights indicate

stronger constraints. Given some input $/x/$, each output candidate $[y]$ is assigned a *harmony score* H based on its constraint violations:

$$H(x, y) = \sum_{i=1}^K w_i C_i(x, y),$$

where K is the number of constraints, w_i is the weight of the i^{th} constraint, and $C_i(x, y)$ is the number of times the mapping $/x/ \rightarrow [y]$ violates the i^{th} constraint. A candidate that violates no constraints has a harmony score of 0, and a more positive score indicates a less preferred candidate.

For an input $/x/$, MaxEnt calculates a probability distribution over all possible output candidates $y \in Gen(x)$ as follows:

$$P(y|x;w) = \frac{e^{-H(x,y)}}{\sum_{z \in Gen(x)} e^{-H(x,z)}}.$$

In other words, the probability of a candidate is proportional to its (exponentiated) negative harmony. This means that candidates that violate more constraints, or constraints with higher weights, are assigned lower probabilities.

Because MaxEnt assigns a probability to each observed form, we can use MaxEnt models to calculate a conditional log likelihood for a data set by summing the log probabilities of each of the N tokens in the data set:

$$LL_w(D) = \sum_{i=1}^N \ln P(y_i | x_i; w)$$

A model that perfectly predicts a data set will assign it a conditional log likelihood of 0. As the predictive accuracy of the model decreases, the conditional log likelihood will become increasingly negative.

Given a set of constraints and a data set consisting of input-output pairs with violation profiles, the weights that optimally predict the data (i.e., that maximize the conditional log likelihood) can be learned using gradient descent (see Hayes & Wilson, 2008). This means the model weights can be fit directly to the experimental data rather than determined manually.

4.4 Enforcing constraint hierarchies

Both the Complex Syllable Contact and the Perceptual Cost analyses contain constraint hierarchies, which put certain constraints into a fixed relative ranking. The hierarchy in the Complex Syllable Contact analysis requires that constraints penalizing greater sonority rises across syllable boundaries should be ranked higher than those targeting smaller sonority rises. Similarly, the hierarchy in the Perceptual Cost analysis requires that the contextual Dep constraints be ranked in accordance with the relative perceptual disruptiveness of epenthesis into each context.

In order to enforce these hierarchies in MaxEnt models, we implement them as stringency relationships (e.g., de Lacy, 2004, 2006). Under such a relationship, if a constraint in the hierarchy is violated, then every constraint below it in the hierarchy is also violated. This is shown in the tableau below.

	Dep-V/S_T	Dep-V/S_N	Dep-V/S_L	Dep-V/S_W	Dep-V/O_R
a. SVT	*	*	*	*	*
b. SVN		*	*	*	*
c. SVL			*	*	*
d. SVW				*	*
e. OVR					*

Table 9: A tableau illustrating a stringency relationship in classical OT.

Because of the strict constraint ranking used in classical OT, the stringency relationship shown in **Table 9** does not seem to do much for us: So long as we keep the constraints in the ranking shown above, violations of lower ranked constraints by candidates that also violate higher ranked ones do not affect the outcome. In the context of the numeric constraint weights used in MaxEnt, however, the usefulness of a stringency relationship becomes apparent: It allows constraint weights to be fit freely (so long as they are positive), while still imposing the restriction that forms that violate higher ranked constraints in the hierarchy cannot be penalized less than forms that violate only lower ranked constraints. An example MaxEnt tableau with toy weights is shown below in Table 10 to illustrate this. Constraint weights are given under each constraint name. The numbers in the cells below are the number of times the corresponding candidate violates each constraint (once in all cases here).

	H	Dep-V/S_T w = 1	Dep-V/S_N w = 1	Dep-V/S_L w = 1	Dep-V/S_W w = 1	Dep-V/O_R w = 1
a. SVT	5	1	1	1	1	1
b. SVN	4		1	1	1	1
c. SVL	3			1	1	1
d. SVW	2				1	1
e. OVR	1					1

Table 10: A tableau illustrating a stringency relationship in harmonic grammar.

Under a stringency relationship with weighted constraints, Candidate 10d (for example), with medial epenthesis between a sibilant and a glide, must receive a harmony score that is either equal to the harmony of Candidate 10e (if the weight of Dep-V/S_W is 0) or greater than

the harmony of 10e (if the weight of Dep-V/S_W > 0). In other words, the stringency relationship requires that epenthesis into SW clusters will *always* be penalized to an equal or greater extent than epenthesis into OR clusters, and so on up the hierarchy, regardless of how the constraints are weighted. Were this hierarchy not in place, arbitrary constraint weights could be used to generate typologically bizarre predictions. The weight of the lowest weighted constraint reflects the penalty assigned to violating that constraint (in this case, epenthesis into an OR onset), while the weights of every higher ranked constraint reflect the *difference* in penalty between violating that constraint and violating the next-lowest ranked constraint in the hierarchy. Note, crucially, that the stringency relationship is not meant to claim that epenthesis into an SW cluster as in 10d somehow also constitutes epenthesis into an OR cluster, despite the violation of Dep-V/O_R. Rather, the stringency relationship between these two Dep constraints mandates that forms that violate Dep-V/S_W must be penalized *at least* as much as forms that violate Dep-V/O_R.

4.5 Analysis variants

In addition to comparing the two Syllable Contact analyses and the Perceptual Cost analysis described above, we also compare several variants of each that differ in (a) whether *Complex constraint weights are scaled by English dominance; (b) whether *Complex can have different weights for SC vs. OR onsets; and (c) whether the relationship between English dominance and the weight of *Complex differs for SC vs. OR onsets. These variants are described in the following sections.

4.5.1 Constraint scaling and language dominance

Because our participants vary in their English proficiency, which in turn affects their epenthesis rates, we would like to encode this information in the MaxEnt models we form. Classic work encodes L2 acquisition as constraint re-ranking from L1 rankings to L2 rankings (e.g., Broselow et al., 1998; Hancin-Bhatt, 1997). We adopt a similar approach here using constraint scaling, which has been used to model lexical and other hierarchical effects in weighted grammars (e.g., Coetzee & Kawahara, 2013; Gouskova & Linzen, 2015; Zymet, 2018a,b; Hugto, et al., 2019; Shih, 2020). We do not attempt to model the learning process itself, but rather its outcome: More proficient English speakers will have a lower-weighted *Complex constraint, meaning complex onsets become easier to produce faithfully.

In order to implement this, we calculate a speaker-specific weight of *Complex for speaker j by scaling its weight by the speaker's Relative English Dominance based on acquisition (RED), as follows:

$$w_{*Complex_j} = w_{*Complex} - \rho \cdot RED_j$$

where $w_{*Complex}$ is the global weight of *Complex, RED_j is the j th speaker's RED score, and ρ is a blending factor that relates RED to the scaled weight.

Note that since MaxEnt is fundamentally a (multinomial) logistic regression model, this participant-specific scaling of constraint weights serves the same purpose as a random slope in a hierarchical model, though we estimate these effects using the LEAP-Q questionnaire rather than during the process of model fitting. That is, rather than trying to model a population-level grammar, we attempt to incorporate (at least some) participant-specific differences into the model and test the prediction that Relative English Dominance has a determinative effect on epenthesis patterns. For more on the role of hierarchical modeling in the MaxEnt framework see, e.g., Zymet (2018a,b), and Garcia (2019).

4.5.2 Differentiating SC and OR onsets

The experimental results from the previous section suggest that SC onsets are repaired more frequently and acquired more slowly than OR onsets. To account for their overall differences in repair rates, we follow Fleischhacker (2005) and split *Complex into two separate constraints penalizing each cluster type separately:⁸

- *Complex_{SC}: No tautosyllabic SC sequences.
- *Complex_{OR}: No tautosyllabic OR sequences.

To account for different learning rates, we can also split ρ from the scaling equation in the previous section into two parameters specific to each onset type, ρ_{SC} and ρ_{OR} .

This leads to a total of 15 MaxEnt models based on four choice points:

- Which of the three base MaxEnt models do we start with (Simple/Complex Syllable Contact, Perceptual Cost)?
- Do we scale *Complex by speakers' RED scores?
- Do we split *Complex into SC- and OR-specific variants
- If the answers to the latter two questions are “yes,” do we split ρ into SC- and OR-specific variants?

4.6 Model comparison

All MaxEnt models were fit to the experimental data set using the *maxent.ot* R package (Mayer et al., 2024). Constraint weights were fit to the data using gradient descent with a weak Gaussian prior of $\mu = 0$ and $\sigma = 1000$. This implements a soft preference for lower constraint weights (see

⁸ A reviewer wonders whether this could be implemented using a single constraint with violation counts or weight scaled according to cluster type (e.g., the treatment of SSP violations in Linzen et al., 2013). A similar proposal could also be made for the constraint hierarchies described above. This is certainly possible, but we suspect it would not result in a meaningful difference in the analysis. We use split constraints here for consistency with the original analyses as well as ease of interpretability.

Mayer et al., 2024, for more detail). Optimal ρ or ρ_{SC}/ρ_{OR} was then calculated using grid search over the range [0.1, 1] in increments of 0.1. Model comparison was done using BIC, which, again, rewards model fit to the data while penalizing model complexity. When evaluating models based on BIC, it is the relative difference in BIC scores that is important, not their particular values. Lower BIC scores are preferred. A rule of thumb, proposed in Raftery (1995), is that a difference in BIC scores of 0 to 2 is weak positive evidence for the model with the lower score; a difference of 2 to 6 is positive evidence; a difference 6 to 10 is strong positive evidence; and a difference of 10 or more is very strong evidence.

Table 11 below shows the conditional log likelihood, number of parameters (reflecting model complexity), and BIC score for each model we tested. The BIC scores of each model are also presented graphically in **Figure 7**.

<i>Base Model</i>	<i>Scaled by RED?</i>	<i>Split *Complex?</i>	<i>Split ρ?</i>	<i>Log Likelihood</i>	<i>Num. Params</i>	<i>BIC</i>
Perceptual Cost	Y	Y	Y	-658	12	1408
Perceptual Cost	Y	Y	N	-671	11	1427
Syllable Contact – Complex	Y	Y	Y	-729	15	1573
Syllable Contact – Complex	Y	Y	N	-743	14	1593
Perceptual Cost	Y	N	N	-777	10	1631
Syllable Contact – Complex	Y	N	N	-822	13	1745
Syllable Contact – Simple	Y	Y	Y	-915	7	1885
Syllable Contact – Simple	Y	N	N	-926	5	1891
Syllable Contact – Simple	Y	Y	N	-926	6	1898
Perceptual Cost	N	Y	n/a	-989	10	2055
Syllable Contact – Complex	N	Y	n/a	-1061	13	2222
Syllable Contact – Complex	N	N	n/a	-1089	12	2270
Perceptual Cost	N	N	n/a	-1109	9	2288
Syllable Contact – Simple	N	N	n/a	-1172	4	2375
Syllable Contact – Simple	N	Y	n/a	-1171	5	2381

Table 11: Results of each model. The best performing model from each of the three base models (Simple/Complex Syllable Contact and Perceptual Cost) is in boldface.

These results demonstrate several things. First, scaling *Complex by speaker RED scores substantially improves model fit across the board. Second, splitting *Complex into SC- and OR-specific variants also improves model fit across the board, with the exception of the Simple

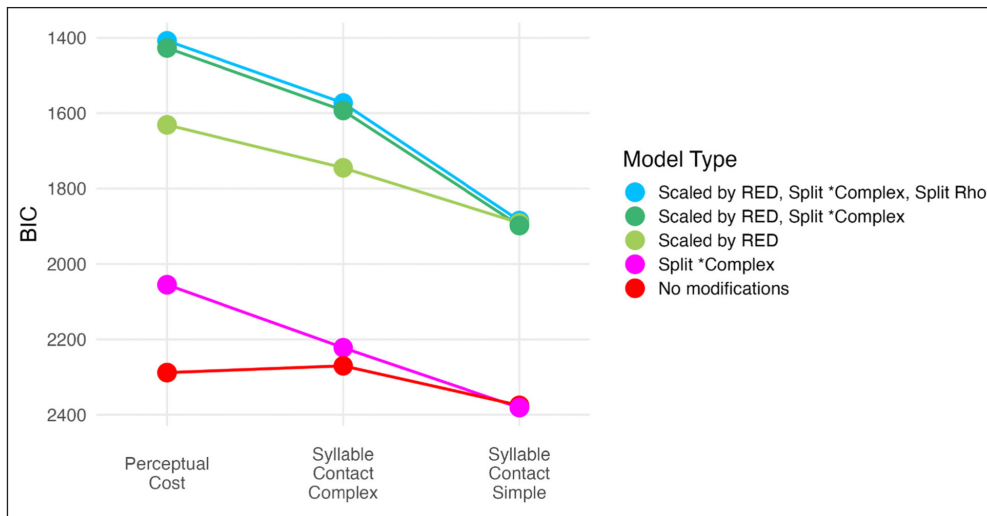


Figure 7: BIC scores for each model type. The x-axis represents the base model type, and the colored points and lines correspond to the variations on each model described in the previous section.

Syllable Contact model with no scaling: Though Fleischhacker (2005) justified this split primarily on typological grounds, this provides additional evidence for its necessity. Finally, introducing separate scaling rates for SC and OR onsets results in a modest increase in model fit across the board. These results provide further support for the claims from the previous sections about the production and acquisition of complex onsets and demonstrate that the innovations to these models presented here (scaling and splitting *Complex) are useful mechanisms for modeling these factors.

Additionally, and most importantly, these results provide broad support for the Perceptual Cost model over the Syllable Contact model in accounting for patterns of repair in the experimental data. In every variant of the model where either *Complex is split by onset type *or* the weight of *Complex is scaled by speakers' English dominance, the Perceptual Cost model achieves the best performance, followed by the Complex Syllable Contact model and the Simple Syllable Contact model. Only in the model that does not account for interspeaker and inter-onset differences does the Complex Syllable Contact model outperform the Perceptual Cost model. This underscores the importance of controlling for these factors in our model comparison.

4.7 Why does the Perceptual Cost model succeed?

We can gain some insight into the success of the Perceptual Cost model by plotting the mean errors made by the three best performing models in each class. These errors are shown in **Figure 8**. The bar corresponding to errors on tokens without epenthesis is negative for all models, indicating that all models underpredict epenthesis rates. The difference between the models becomes

apparent when looking at their predictions for pre- vs. medial epenthesis. The Perceptual Cost model successfully predicts that SC onsets (with the exception of /sw/ onsets) should be repaired by pre-epenthesis and OR onsets should be repaired by medial epenthesis. On the other hand, both Syllable Contact models predict that SC onsets should sometimes be repaired by medial epenthesis (which is unattested in our data except for a small number of /sw/ tokens) and that OR onsets should sometimes be repaired by pre-epenthesis (which is completely unattested in our data). Thus, both Syllable Contact models are unable to capture the clear separation of repair strategies by onset type.

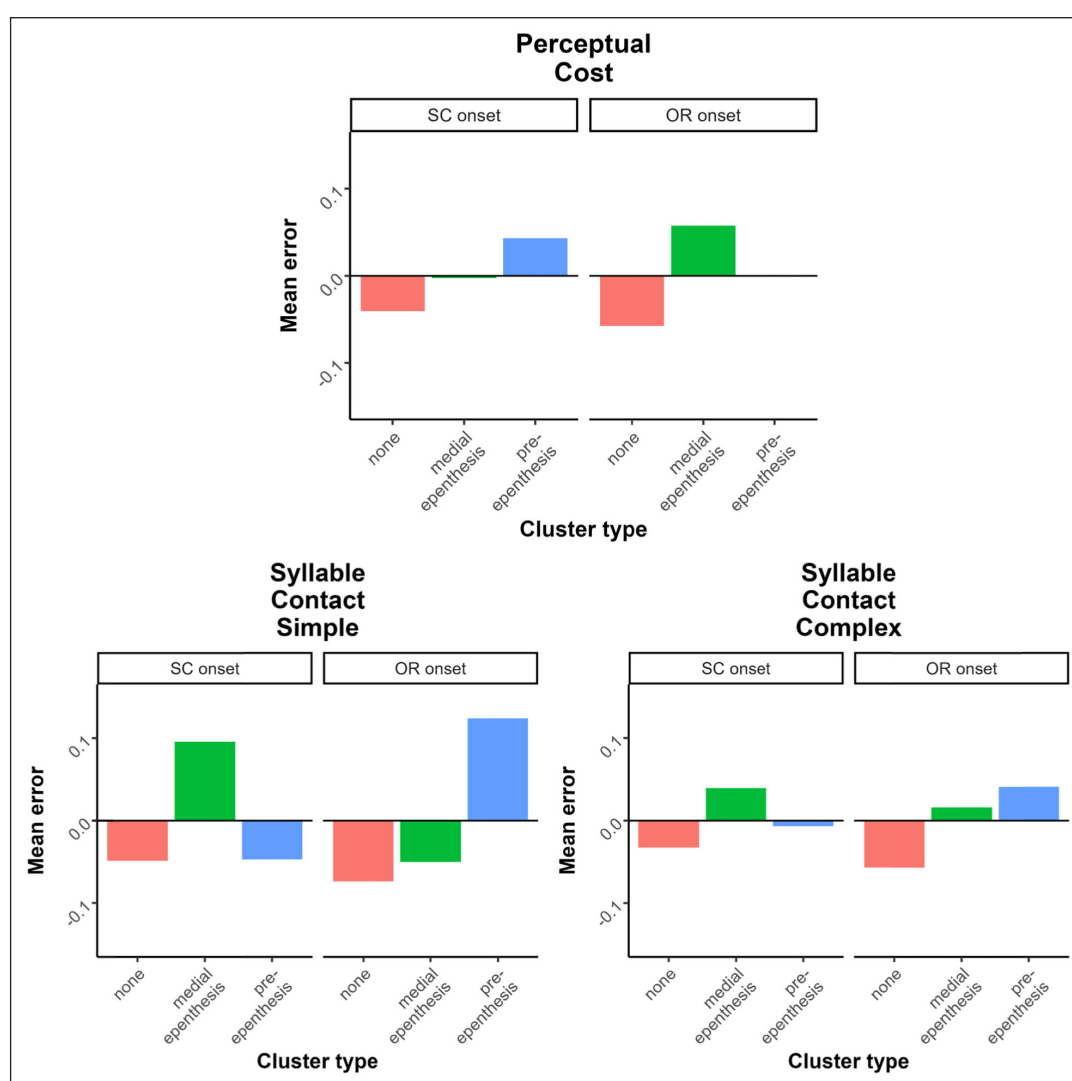


Figure 8: Error rates from each model, broken down by onset type (SC vs. OR) and epenthesis type (none vs. pre-epenthesis vs. medial epenthesis).

Why do the Syllable Contact models predict unattested repair strategies? This issue arises because some OR and SR onsets have identical sonority deltas, such as /fl/ and /sl/ where $son(/f/) = son(/s/) = 1$ and $son(/l/) = 4$. Because the Syllable Contact models cannot distinguish between onsets with the same sonority delta, the model necessarily predicts that these onsets should display equivalent behavior. An example of this is shown for the words ‘fly’ and ‘sleep’ in the following two tableaux, which uses the unscaled constraint weights from the best performing Simple Syllable Contact model.

/flaɪ/	Predicted Frequency	Harmony	Dep w = 1.13	Contiguity w = 1.34	*Complex _{OR} w = 0.20	Syllable-Contact w = 1.68
a. felai	0.09	2.47	1	1		
b. eflai	0.06	2.81	1			1
c. flai	0.85	0.2			1	

Table 12: A tableau for /flaɪ/ ‘fly’ under the Simple Syllable Contact model.

In Table 12, Candidate 12c violates *Complex_{OR} due to its OR onset, but it is still the most probable candidate due to the low weight of this constraint. Although both the repaired forms 12a and 12b violate two constraints each (Dep plus Contiguity or SyllableContact respectively), Candidate 12a is predicted to be more probable because the violation of Contiguity is less costly than the violation of SyllableContact.

/slɪp/	Predicted Frequency	Harmony	Dep w = 1.13	Contiguity w = 1.34	*Complex _{SC} w = 0.04	Syllable-Contact w = 1.68
a. selip	0.08	2.47	1	1		
b. eslip	0.05	2.81	1			1
c. slip	0.87	0.04			1	

Table 13: A tableau for /slɪp/ ‘sleep’ under the Simple Syllable Contact model.

The situation is essentially identical in **Table 13**, except that Candidate 13c is expected to be relatively more probable than Candidate 12c (and, accordingly, Candidates 13a and 13b relatively less probable than Candidates 12a and 12b) due to the lower weight of *Complex_{SC}. Crucially, however, the Harmony of the two repaired candidates a and b is the same in both cases due to their common sonority profile. This means that their relative frequency of occurrence remains the same under any weighting of these constraints. Decreasing the weight of Contiguity or increasing the weight of SyllableContact to increase the rate of medial epenthesis for /fl/ onsets will have

the same effect on /sl/ onsets, and vice versa. The same issue occurs for the Complex Syllable Contact model and for other pairs of onsets with identical sonority deltas such as /sw/ and /pl/.

In the Perceptual Cost model, on the other hand, different repair strategies are predicted due to differences in the violations of the contextual Dep-V constraints. This is illustrated in **Tables 14** and **15**, again using the unscaled weights of the *Complex constraints.

/flaɪ/	Pred. Freq.	Harm.	C/V w = 5.44	L-Anchor w = 23.30	Contiguity w = 5.02	*Complex _{OR} w = 2.87	Dep-V/S_L w = 11.62	Dep-V/S_W w = 19.21	Dep-V/O_R w = 5.02
a. feɪaɪ	0.15	10.05			1				1
b. eɪaɪ	0	28.74	1	1					
c. flaɪ	0.85	8.31	1			1			

Table 14: A tableau for /flaɪ/ ‘fly’ under the Perceptual Cost model.

In **Table 14**, Candidate 14b receives a probability of zero due to its violation of the highly-weighted L-Anchor constraint. The relatively low weight of *Complex_{OR} leads to 14c being the most probable outcome, while 14a, the form repaired by medial epenthesis, receives the remainder of the probability mass.

/slɪp/	Pred. Freq.	Harm.	C/V w = 5.44	L-Anchor w = 23.30	Contiguity w = 5.02	*Complex _{SC} w = 21.73	Dep-V/S_L w = 11.62	Dep-V/S_W w = 19.21	Dep-V/O_R w = 5.02
a. seɪp	0	40.87			1		1	1	1
b. esɪp	0.17	28.74	1	1					
c. slɪp	0.83	27.17	1			1			

Table 15: A tableau for /slɪp/ ‘sleep’ under the Perceptual Cost model.

In **Table 15**, the unepenthesized form 15c again receives the majority of the probability, but now the only predicted repair strategy is pre-epenthesis, in 15b. This shift is due to the additional violations 15a incurs for the constraints Dep-V/S_L and Dep-V/S_W. The violations of these two constraints are collectively costlier than Candidate 15b’s violation of L-Anchor, and so only pre-epenthesis repairs are predicted in this case.

An additional difference between the Syllable Contact and Perceptual models is in their weightings and scalings of the *Complex constraints. These values from the most successful model of each type are shown in **Table 16**.

Model	*Complex _{OR} global weight	*Complex _{SC} global weight	ρ_{OR}	ρ_{SC}
Perceptual Cost	2.87	21.73	0.9	0.6
Syllable Contact Complex	1.81	0.59	0.9	0.6
Syllable Contact Simple	0.20	0.04	0.8	0.5

Table 16: Weightings and scalings of the *Complex constraints.

Note that although all models have a higher learning rate for OR onsets than for SC onsets, the two Syllable Contact models assign lower weights to *Complex_{SC} than to *Complex_{OR}. Weights in a MaxEnt model represent the relative strength of each constraint, such that violations of constraints with higher weights are more strongly penalized. Concretely, because the probability of a candidate is inversely proportional to its harmony score, and because the harmony score is the sum of all the weights of the constraints the candidate violates, violating a constraint with a higher weight will lower the probability of a candidate more than a violation of a lower weighted constraint. Both Syllable Contact models predict that (all else being equal) forms with SC onsets should be *more* likely to surface without repair than forms with OR onsets, while the Perceptual Cost model predicts the opposite. Thus, the optimal weights for the Syllable Contact models fail to directly encode the observation that SC onsets are more difficult than OR onsets, while the Perceptual Cost model succeeds.

5. Discussion

This paper has presented an investigation of the acquisition of complex onsets in L2 English by L1 Farsi speakers, as well as the pressures that drive epenthesis asymmetries by these speakers in the repair of OR vs. SC onsets.

The results of the experimental and phonological modeling studies provide evidence for several properties of the acquisition process. First, the data are consistent with past accounts of epenthesis by L1 Farsi/L2 English speakers, which suggest that these speakers repair all SC onsets (except /sw/) with pre-epenthesis and all other onset types with medial epenthesis (e.g., Karimi, 1987; Fleischhacker, 2001). Second, SC onsets appear to be repaired more frequently than OR onsets. This is also consistent with some past work on L2 acquisition (e.g., Carlisle, 2001; Yildiz, 2005). Third, although epenthesis rates in general decrease as L2 proficiency increases (e.g., Boudaoud & Cardoso, 2009; Yazawa et al., 2015) the relationship between English ability, measured by English onset age in the corpus study and the Relative English Dominance score calculated from the LEAP-Q survey in the experimental study, suggests that SC onsets are acquired more slowly than OR onsets.

The modeling study in Section 4 also demonstrates that a phonological model of epenthesis asymmetries that is based on minimizing the perceptual distance between epenthesized and unepenthesized forms better predicts our experimental data than a model that is based

on minimizing sonority rises across syllables. Additionally, this study shows how language proficiency can be integrated into a harmonic grammar analysis using constraint scaling to factor out individual variation. In the remainder of this section, we will discuss some of the implications of these results, as well as avenues for future research.

5.1 Relating acquisition, articulation, and perception

A question that emerges from the current study as well as past research is how to reconcile the idiosyncratic phonological behavior of SC onsets with their articulatory and acquisitional properties. In other words, SC onsets tend to be repaired using pre-epenthesis while other onsets use medial epenthesis. Does this relate at all to the fact that SC onsets display greater articulatory coupling between the component sounds (see Section 2.3), are more difficult to produce and are slower to acquire?

Although it is not as apparent, we will speculate here on a possible connection between the Perceptual Cost account of epenthesis asymmetries and the other properties of SC onsets. Hall (2006) demonstrates that an “intrusive” vowel may be perceived between the segments of a complex onset as the result of a mistiming of their articulatory gestures. This mistiming increases the amount of time between the offset of the initial consonant and the onset of the second consonant, resulting in a (possibly voiceless) vowel-like transition in between them. This observation, combined with Fleischhacker’s claim (2001, 2005) that vowel insertion into SC clusters is generally more perceptually disruptive than into other cluster types, suggests that the perceptual consequences of mistiming an SC cluster are greater than those of mistiming an OR cluster.

The articulatory differences between SC and OR clusters follow from this hypothesis: SC clusters are articulated more precisely because the perceptual consequences of mistiming are greater. This idea can be clarified by thinking of it from the perspective of “good variance” vs. “bad variance,” a dichotomy proposed in the motor control literature (see, e.g., Latash, 2012). The production of any task-oriented movement, speech included, is subject to variability in its execution. Some of this variability is “bad” because it interferes with the goals of the task (e.g., I articulate a word in a way that differs from my intended perceptual target), while some of it is “good” because it does not (e.g., I may articulate a word in different ways, but all variants meet the perceptual target). Researchers have found that skilled performers exhibit high degrees of good variability and low degrees of bad variability; for an example from speech, see Kang et al., (2019). The maximization of good variability, which is irrelevant to task outcomes, imparts flexibility and allows the same movement goals to be achieved under different conditions; the minimization of bad variability ensures that the task goals are consistently achieved.

Skilled production of an SC onset requires minimization of variability in the relative timing of the two component segments because excrescent vowels arising from mistiming are particularly salient. That is, for an SC onset, variability in relative timing is (mostly) bad variability. OR

onsets, on the other hand, can tolerate more slop: The production of an excrescent vowel is less likely (relative to an SC onset) to be perceptually disruptive. In this sense, at least some of the variability in the relative timing of OR onset can be seen as good variability, in that it allows flexibility in articulation while still achieving perceptual goals. The slower acquisition of SC onsets could then be a consequence of their more demanding timing requirements. Although we do not provide direct evidence for this connection in the current paper, we believe it is a useful perspective to explore to provide a more holistic account of SC vs. OR onsets. This account is also broadly consistent with the idea that ST onsets (and perhaps other SC onsets to varying extents) constitute single, complex segments (e.g., Broselow, 1992; Goad, 2012), in that they exhibit a higher degree of articulatory coupling between their component segments.

The connection between the Syllable Contact account of epenthesis asymmetries and the other properties of SC clusters is less clear. Because the driving force is the avoidance of sonority rises across syllables, the composition of an onset is unimportant except for the sonority values of its component segments. Under such an account, it is difficult to see how the articulatory and acquisitional differences between SC clusters and other clusters can be explained as anything more than a coincidence.

5.2 Prospects for scalar analyses of sonority

One of the major reasons that the Syllable Contact account fails to capture the epenthesis patterns found in this paper is that it bins together onsets that have the same sonority profile but different epenthesis patterns (e.g., /sl/ vs. /fl/). This is a consistent problem for analyses that treat sonority as a scalar value because onsets that share sonority profiles are often alike in some ways and different in others. For example, the SSP, described in Section 2 above, predicts that ST onsets should be marked because of their falling sonority profile. Although other onsets with a falling profile, such as /ft/, are indeed uncommon, ST onsets are not. Other definitions of sonority, such as the NAP model (Albert & Nicenboim, 2022), attempt to circumvent this issue by modifying the definition of sonority profiles to assign ST onsets similar sonority profiles to other unmarked onsets. Although binning the sonority profiles of ST onsets together with certain OR onsets solves the markedness problem, it also means that we can no longer use differences in sonority profile to account for the distinct behavior of ST onsets: If we assign /st/ the same sonority profile as /bɹ/ (which NAP does), we can no longer use their sonority profiles to distinguish their behavior in other respects, such as acquisition trajectories or epenthesis patterns. It seems unlikely that we will converge upon a single scalar definition of sonority/sonority profiles that is sufficient to capture the variety of ways in which onsets clusters are similar or different.

Although analyses that use scalar definitions of sonority and sonority profiles have provided much insight into the behavior of complex onsets, as well as many other domains, they struggle with providing a consistent and holistic account of the rich variety of phenomena that are related to

onset type. We expect that additional progress in this area will come from adopting a more nuanced view of sonority and its phonological reflexes that is grounded in perception and articulation.

5.3 Future research

In this paper, we showed that the perceptual analysis of epenthesis asymmetry outperforms a sonority-based analysis. However, it is generally not clear whether Fleischhacker's perceptual analysis really reflects language-general perceptual constraints or something that is specific to English and/or Farsi. A more robust series of experimental studies would be useful to confirm that her perceptual hierarchy is truly language-general.

In addition, the presence or absence of an epenthetic vowel is a coarse measurement of a speaker's skill in producing different onset types. Simply producing the onset without epenthesis does not necessarily mean that the onset is produced in a native-like fashion. We will get greater insight into the acquisition trajectory of complex onsets by investigating acoustic or articulatory data to get finer-grained measurements of timing and investigate how these change with language ability.

Lastly, it is unclear whether the patterns found in this paper also apply to languages that use medial epenthesis more broadly within the class of SC onsets. Farsi is a language that treats almost all SC onsets differently from OR onsets. As described earlier, however, this pattern is relatively rare cross-linguistically. It is more common for only ST onsets to be repaired with pre-epenthesis. In the current paper we do not differentiate between onsets within the class of SC clusters, but find that, on average, SC clusters are more difficult to acquire than OR onsets. This raises the question of whether all SC onsets are equally difficult to acquire, and whether the phonological repair of these onsets in a language (pre-epenthesis vs. medial epenthesis) tracks with their acquisition difficulty. The epenthesis rates across individual onsets shown in **Figure 5** in Section 3 suggest some within-class differences, where SC onsets with greater sonority rises are repaired less frequently than those with smaller rises. This is consistent with some past work on L2 learning (see Carlisle, 2001). We leave these as exciting questions for future research.

6. Conclusion

This paper has presented an experimental study and a phonological modeling study that investigate the acquisition and production of complex onsets by L1 Farsi/L2 English speakers. This work makes two primary contributions: The first is to provide additional evidence that SC onsets are repaired more frequently and acquired more slowly than OR onsets in L2 speech. The second is to provide support for an account of epenthesis asymmetries between SC vs. OR onsets that is based on the perceptual consequences of epenthesis rather than constraints on sonority rises across syllable boundaries. We expect that future research in this area will provide important insights into sonority and its phonological reflexes.

Appendix A: Passages read by participants

Smith wants to go to the **store** and **snag** a nice **small** energy **drink** so that he doesn't need **sleep**. He looks at his watch and realizes it is time to go. He then **scoops** up his bag and leaves. On his way over, he hears a **plethora** of loud **quacks** and sees a **flock** of ducks **flying** back to their **creek** to go and **swim**. **Three** ducks **fly** into his face and he **freaks** out. His **skin** **twitches** and he **swats** the duck away. He finally gets to the **store** and **grabs** his **drink**. He **slugs** it down and begins his **trot** back home. On his way home he sees a man walking hand in hand with a **dwarf** holding a **bugle**. He then sees the **queen** of England carrying a **cute** little **sloth** in her arms. He wonders what was in that energy drink. Then, suddenly, a tornado comes and **sweeps** him off his feet. Then, right as he is about to get **smacked** onto the **ground**, he wakes up and **smiles** as he realizes he was **dreaming** all along.

Please ask **Scarlett** if she can **snatch** a big, **green** mop so that she can **clean** up the dirty **floor**. She can wear her **small**, purple, **gloves** while she tries **sweeping** the **snow** off the **driveway**. She can ask her **twin brother** to come help her, even though he might **slow** her down. If she feels **skeptical** about **stopping** what she is doing to help, then she can **skip** over to me, and I can help **swoon** her **quickly**.

Dwayne is working on **tweaking** his car so that it will **speed** faster. He also wants to replace his tires because they are losing their **thread** and new ones will make his car **beautiful**. While he is **slouched** over working, he accidentally **snips** at the wrong wire. He then sees a big **spark** and his lights **start flickering**. He also **sniffs** out some oil leaking, as it begins **smelling** like **smog**. The oil is **spreading** all over, and he realizes that he damaged his **throttle** cable, but knows that fixing it will be a **breeze**. His little **brother** **brings** him the tools that he needs, and gets to work. He then **snaps** all the new parts into **place**. However, when he goes and **starts** his engine, he hears a loud **thrashing** noise. He then **steps** out of the car and **snoops** what is going on. He **quickly** figures out the issue, **slips** the key into the ignition, and **starts** his car.

Appendix B: Onset types and counts

Onset	Count
bj	44
br	120
dr	60
dw	47
fl	149
fr	30
gl	30
gr	60
kj	30
kl	30
kr	30
kw	90
pl	90
sk	149
sl	180
sm	210
sn	210
sp	60
st	135
sw	150
tr	29
tw	88
θr	118

Table B.1: Production counts of each onset type in the experimental study.

Appendix C: LEAP-Q questions and bins

Age when you <u>began acquiring</u> English	
Age you <u>became fluent in (speaking)</u> English	
Please list what percentage of the time you are <u>currently and on average exposed</u> to English. (Your percentages should add up to 100%):	Acquisition/Exposure bin
Please list what percentage of the time you are <u>currently and on average exposed</u> to Farsi. (Your percentages should add up to 100%):	
Please list the <u>number of years and months</u> you spent in a country in which Farsi is spoken	
Please list the <u>number of years and months</u> you spent in a country in which English is spoken	Immersion bin
On a scale from zero to ten, please select your level of <u>proficiency</u> in: 1. speaking 2. understanding 3. reading <i>Farsi</i>	Self-Reported Proficiency bin
On a scale from zero to ten, please select your level of <u>proficiency</u> in: 4. speaking 5. understanding 6. reading <i>English</i>	

Appendix D: Model comparison of different instantiations of RED

In the following table, each model we compared is referenced by the version of RED (or combinations of RED) used. We also report the number of parameters for each model, as well as the BIC and log-likelihood. Each of these models includes an interaction term between each of the RED variables and the onset type variable (sC vs. TR).

Model	Number of Parameters	BIC	Log-likelihood
$RED_{\text{acquisition}}$	9	864.4883	-397.7377
$RED_{\text{acquisition}} + RED_{\text{self-report}}$	11	871.9208	-393.7859
RED_{full}	9	872.8642	-401.9257
$RED_{\text{acquisition}} + RED_{\text{immersion}}$	11	878.4888	-397.0699
$RED_{\text{acquisition}} + RED_{\text{immersion}} + RED_{\text{self-report}}$	13	886.9061	-393.6104
$RED_{\text{immersion}}$	9	894.4273	-412.7072
$RED_{\text{immersion}} + RED_{\text{self-report}}$	11	894.687	-405.169
$RED_{\text{self-report}}$	9	898.8027	-414.8949

Table D.1: Model comparison between different RED scores.

Appendix E: Loadings from PCA on acquisition/exposure bin

The table below shows the loadings for the PCA based on the acquisition/exposure questions from the LEAP-Q survey. PC1, which we negate and use as the scalar value representing English dominance, accounts for about 79% of the variance. Note that in PC1, the positive loadings correlate with greater Farsi exposure and later English acquisition/fluency, while the negative loadings correspond to greater English exposure.

	PC1	PC2	PC3	PC4
Current Farsi Exposure	0.507	0.512	-0.211	-0.660
Age of English Acquisition	0.455	-0.661	-0.596	0.027
Age of English Fluency	0.510	-0.353	0.774	-0.130
Current English Exposure	-0.526	-0.420	0.031	-0.739

Table E.1: Loadings for the PCA used to calculate $RED_{\text{acquisition}}$.

Acknowledgements

We thank Tim Hunter, two anonymous reviewers, and the attendees of the 2024 LSA Annual Meeting for their valuable feedback.

Ethics and consent

This work was carried out under UCI IRB #3014 “Investigating the acquisition of English complex onsets by Farsi speakers.”

Competing interests

The authors have no competing interests to declare.

Author contributions

NK and CM jointly conceived of the research, analyzed and interpreted the data, and wrote the manuscript. NK designed the experimental study and carried out the data collection. CM implemented the MaxEnt modeling study. The authors are listed in alphabetical order.

References

- Albert, A., & Nicenboim, B. (2022). Modeling sonority in terms of pitch intelligibility with the Nucleus Attraction Principle. *Cognitive Science*, 46(7). <https://doi.org/10.1111/cogs.13161>
- Barlow, J. A. (2001). The structure of /s/-sequences: Evidence from a disordered system. *Journal of Child Language*, 28(2), 291–324. <https://doi.org/10.1017/S0305000901004652>
- Bharati, S. (1994). *Aspects of the Phonology of Hindi and English*. Arnold.
- Boudaoud, M., & Cardoso, W. (2009). Vocalic [e] epenthesis and variation in Farsi-English interlanguage speech. *Concordia Working Papers in Applied Linguistics*, 2, 1–34.
- Brasington, R. W. (1981). Epenthesis and deletion in loan phonology. *Work in Progress* 3. University of Reading Phonetics Laboratory, 97–103.
- Broselow, E. (1992). Transfer and universals in second language epenthesis. In S. M. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning*. John Benjamins, 71–86. <https://doi.org/10.1075/lald.5.07bro>
- Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20(2), 261–280. <https://doi.org/10.1017/S0272263198002071>
- Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP. Bulletin de la communication parlée*, 5, 25–34.
- Cardoso, W. (2007). The development of sC onset clusters in interlanguage: Markedness vs. frequency effects. In R. Slabakova, J. Rothman, P. Kempchinsky, & E. Gavrusseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*. Cascadia Press, 15–29.
- Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics*, 12(1), 76–95. <https://doi.org/10.1093/applin/12.1.76>
- Carlisle, R. S. (2001). Syllable structure universals and second language acquisition. *International Journal of English Studies*, 1(1), 1–19.
- Carlson, M. T. (2019). Now you hear it, now you don't: Malleable illusory vowel effects in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 22(5), 1101–1122. <https://doi.org/10.1017/S136672891800086X>

- Chierchia, G. (1986). Length, syllabification and the phonological cycle in Italian. *Journal of Italian Linguistics*, 8, 5–34.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I*. Cambridge University Press, 283–333. <https://doi.org/10.1017/CBO9780511627736.017>
- Clements, G. N. (2009). Does sonority have a phonetic basis? Comments on the chapter by Bert Vaux. In E. Raimy & C. Cairns (Eds.), *Contemporary views on architecture and representations in phonology*. MIT Press. <https://doi.org/10.7551/mitpress/9780262182706.003.0007>
- Coetzee, A. W., & Kawahara, S. (2013). Frequency biases in phonological variation. *Natural Language & Linguistic Theory*, 31(1), 47–89. <https://doi.org/10.1007/s11049-012-9179-z>
- Cuetos, F., Hallé, P. A., Domínguez, A., & Segui, J. (2011). Perception of Prothetic /e/ in #sC Utterances: Gating Data. In *Proceedings of the 17th International Congress on Phonetic Sciences*, 540–543.
- Davidson, L., & Shaw, J. A. (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics*, 40(2), 234–248. <https://doi.org/10.1016/j.wocn.2011.11.005>
- De Lacy, P. (2004). Markedness conflation in Optimality Theory. *Phonology*, 21(2), 145–199. <https://doi.org/10.1017/S0952675704000193>
- De Lacy, P. (2006). *Markedness: Reduction and preservation in phonology* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486388>
- Eckman, F. R., Elreyes, A., & Iverson, G. K. (2003). Some principles of second language phonology. *Second Language Research*, 19(3), 169–208. <https://doi.org/10.1191/0267658303sr2190a>
- Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics*, 7, 71–116.
- Fleischhacker, H. A. (2005). *Similarity in phonology: Evidence from reduplication and loan adaptation* [Doctoral dissertation, University of California, Los Angeles].
- Frisch, S. A. (2015). A preliminary investigation of quantitative patterns in sonority sequencing. *Italian Journal of Linguistics*, 27(1), 9–27.
- Garcia, G. D. (2019). When lexical statistics and the grammar conflict: Learning and repairing weight effects on stress. *Language*, 95(4), 612–641. <https://doi.org/10.1353/lan.2019.0068>
- Gibson, M. (2012). A gestural-based analysis of /e/ prosthesis in word-initial /sC/ loanwords in Spanish. *Ianua. Revista Philologica Romanica*, 12, 35–56.
- Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research*, 42(3), 708–726. <https://doi.org/10.1044/jslhr.4203.708>
- Gierut J. A. (2001). Complexity in phonological treatment: Clinical factors. *Language, Speech, and Hearing Services in Schools*, 32(4), 229–241. [https://doi.org/10.1044/0161-1461\(2001/021\)](https://doi.org/10.1044/0161-1461(2001/021))
- Goad, H. (2012). sC Clusters are (almost always) coda-initial. *The Linguistic Review*, 29(3). <https://doi.org/10.1515/tlr-2012-0013>

- Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In J. Spenader, A. Eriksson & O. Dahl (Eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm University Department of Linguistics, 111–120.
- Gouskova, M. (2001). Falling sonority onsets, loanwords, and syllable contact. *Chicago Linguistic Society*, 37(1), 175–185.
- Gouskova, M. (2004). Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology*, 21(2), 201–250. <https://doi.org/10.1017/S095267570400020X>
- Gouskova, M., & Linzen, T. (2015). Morphological conditioning of phonological regularization. *Linguistic Review*, 32(3). <https://doi.org/10.1515/tlr-2014-0027>
- Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology*, 23(3), 387–429. <https://doi.org/10.1017/S0952675706000996>
- Hall, N. (2011). Vowel epenthesis. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*. Wiley-Blackwell, 1576–1596. <https://doi.org/10.1002/9781444335262.wbctp0067>
- Hancin-Bhatt, B., & Bhatt, R. M. (1997). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, 19, 331–378. <https://doi.org/10.1017/S0272263197003033>
- Hayes, B., & Wilson, C. (2008). A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry*, 39(3), 379–440. <https://doi.org/10.1162/ling.2008.39.3.379>
- Hooper, J. B. (1976). *An introduction to natural generative phonology*. Academic Press.
- Hughto, C., Lamont, A., Prickett, B., & Jarosz, G. (2019). Learning exceptionality and variation with lexically scaled maxent. *Proceedings of the Society for Computation in Linguistics*, 2(1), 91–101.
- Jarosz, G. (2017). Defying the stimulus: Acquisition of complex onsets in Polish. *Phonology*, 34(2), 269–298. <https://doi.org/10.1017/S0952675717000148>
- Kang, J., Nam, H., Chen, W., & Whalen, D. (2019). Benign vs. harmful variability in second language vowel production. In S. Calhoun, P., Escudero, M., Tabain, and P. Warren (Eds.), *Proceedings of the 19th International Congress on Phonetic Sciences*, 1749–1753.
- Karimi, S. 1987. Farsi speakers and the initial consonant cluster in English. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system*. Newbury House, 305–318.
- Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, 7(1), 193–231. <https://doi.org/10.1017/S0952675700001184>
- Krämer, M. (2021). Complex onsets and coda markedness in Persian. *Nordlyd*, 45(1), 95–118. <https://doi.org/10.7557/12.6239>
- Krämer, M., & Zec, D. (2020). Nasal consonants, sonority and syllable phonotactics: The dual nasal hypothesis. *Phonology*, 37(1), 27–63. <https://doi.org/10.1017/S0952675720000032>
- Kristoffersen, K. E., & Simonsen, H. G. (2006). The acquisition of #sC clusters in Norwegian. *Journal of Multilingual Communication Disorders*, 4(3), 231–241. <https://doi.org/10.1080/14769670601110556>

- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
- Latash, M. L. (2012). The bliss (not the problem) of motor abundance (not redundancy). *Experimental Brain Research*, 217, 1–5. <https://doi.org/10.1007/s00221-012-3000-4>
- Legendre, G., Miyata, Y., & Smolensky, P. (1990). Harmonic Grammar – A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 388–395.
- Linzen, T., Kasyanenko, S., & Gouskova, M. (2013). Lexical and phonological variation in Russian prepositions. *Phonology*, 30(3), 453–515. <https://doi.org/10.1017/S0952675713000225>
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics*, 41(3–4), 211–227. <https://doi.org/10.1016/j.wocn.2013.02.001>
- Mayer, C., Tan, A., & Zuraw, K. (2024). Introducing maxent.ot: An R package for maximum entropy constraint grammars. *Phonological Data and Analysis*, 6(3), 1–44. <https://doi.org/10.3765/pda.v6art4.88>
- McCarthy, J. J. (2009). *The p-map in Harmonic Serialism* (Publication no. ROA-1052) [Masters Thesis, University of Massachusetts, Amherst]. Rutgers Optimality Archive.
- McCarthy, J. J., & Prince, A. (2004). The emergence of the unmarked. In J. J. (Ed.), *Optimality theory in phonology: A reader*. Blackwell, 483–494. <https://doi.org/10.1002/9780470756171.ch26>
- Morelli, F. (2003). The relative harmony of /s + stop/ onsets: Obstruent clusters and the sonority sequencing principle. In C. Fery & R. van de Vijver (Eds.), *The syllable in optimality theory*. CUP, New York, 356–371. <https://doi.org/10.1017/CBO9780511497926.015>
- Murray, R. W., & Vennemann, T. (1983). Sound change and syllable structure in Germanic phonology. *Language*, 59(3), 514–528. <https://doi.org/10.2307/413901>
- Parker, S. (2002). *Quantifying the sonority hierarchy* [Doctoral dissertation, University of Massachusetts, Amherst].
- Parker, S. G. (2017). Sounding out sonority. *Language and Linguistics Compass*, 11(9), e12248. <https://doi.org/10.1111/lnc3.12248>
- Pastätter, M., & Pouplier, M. (2014). The temporal coordination of Polish onset and coda clusters containing sibilants. In *Proceedings of the 10th International Seminar on Speech Production*, 312–315.
- Pater, J. (2009). Weighted constraints in generative linguistics. *Cognitive Science*, 33(6), 999–1035. <https://doi.org/10.1111/j.1551-6709.2009.01047.x>
- Pouplier, M., Pastätter, M., Hoole, P., Marin, S., Chitoran, I., Lentz, T. O., & Kochetov, A. (2022). Language and cluster-specific effects in the timing of onset consonant sequences in seven languages. *Journal of Phonetics*, 93, 101153. <https://doi.org/10.1016/j.wocn.2022.101153>

- Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Blackwell. <https://doi.org/10.1002/9780470759400>
- Raftery, A. E. (1995). Bayesian Model selection in social research. *Sociological Methodology*, 25, 111–163. <https://doi.org/10.2307/271063>
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language and sound structure: Studies in phonology presented to Morris Halle by his teacher and students*. MIT Press.
- Seo, M. (2011). The syllable contact law. In M. van Oostendorp, C. J. Ewen, E. V. Hume & K. Rice (Eds.), *Blackwell companion to phonology* (Vol. 2). Wiley-Blackwell, 1245–62. <https://doi.org/10.1002/9781444335262.wbctp0053>
- Shademan, S. (2002). *Epenthetic vowel harmony in Farsi* [Master's thesis, University of California, Los Angeles].
- Sherwin, S. (1999). The sonority sequencing principle in interlanguage phonology. *George Mason University Working Papers in Linguistics*, 6, 55–74.
- Shih, S. (2020). Gradient categories in lexically-conditioned phonology: An example from sound symbolism. In H. Baek, C. Takahashi, & A. Yeung (Eds.), *Proceedings of the 2019 Annual Meeting on Phonology*. <https://doi.org/10.3765/amp.v8i0.4689>
- Singh, R. (1985). Prosodic adaptation in interphonology. *Lingua*, 67(4), 269–282. [https://doi.org/10.1016/0024-3841\(85\)90001-4](https://doi.org/10.1016/0024-3841(85)90001-4)
- Smit, A. B. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Consonant singletons. *Journal of Speech, Language, and Hearing Research*, 36(3), 533–547. <https://doi.org/10.1044/jshr.3603.533>
- Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. *Papers in Laboratory Phonology*, 5, 313–334.
- Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume & K. Johnson (Eds.), *The role of speech perception phenomena in phonology*. Academic Press. https://doi.org/10.1163/9789004454095_013
- Sulejmenova, B. A. (1965). O foneticheskom osvoenii leksiki, zaimstvovannoj iz russkogo jazyka. Progressivnoe Vlijanie Russkogo Jazyka na Kazakhskij. [On the phonetic adaptation of vocabulary borrowed from Russian. The progressive influence of the Russian language on Kazakh.] In *Akademija Nauk Kazakhskoj SSR* (Ed.), Institut Jazykoznanija, 60–95.
- Tessier, A. M., Duncan, T. S., & Paradis, J. (2013). Developmental trends and L1 effects in early L2 learners' onset cluster production. *Bilingualism: Language and Cognition*, 16(3), 663–681. <https://doi.org/10.1017/S136672891200048X>
- VanDam, M. (2004). Word final coda typology. *Journal of Universal Language*, 5, 119–148. <https://doi.org/10.22425/jul.2004.5.1.119>

- Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change*. Mouton de Gruyter. <https://doi.org/10.1515/9783110849608>
- Yavaş, M., Ben-David, A., Gerrits, E., Kristoffersen, K. E., & Simonsen, H. G. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics*, 22(6), 421–441. <https://doi.org/10.1080/02699200701875864>
- Yazawa, K., Konishi, T., Hanzawa, K., Short, G., & Kondo, M. (2015). Vowel epenthesis in Japanese speakers' L2 English. In *Proceedings of the 18th International Congress on Phonetic Sciences*.
- Yildiz, Y. (2005). The structure of initial/s/-clusters: evidence from L1 and L2 acquisition. Developmental paths in phonological acquisition. *Special issue of Leiden Papers in Linguistics*, 2, 163–187.
- Zampini, M. L., & Hansen Edwards, J. G. (Eds.) (2008). *Phonology and second language acquisition*. John Benjamins. <https://doi.org/10.1075/sibil.36>
- Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language*, 83(2), 277–316. <https://doi.org/10.1353/lan.2007.0105>
- Zuraw, K. (2013). *Map constraints [Unpublished manuscript]. Department of Linguistics, University of California, Los Angeles.
- Zymet, J. (2018a). Learning a frequency-matching grammar together with lexical idiosyncrasy: Maxent versus hierarchical regression. In K. Hout, A. Mai, A. McCollum, S. Rose, & M. Zaslansky (Eds.), *Proceedings of the 2018 Annual Meeting on Phonology*. <https://doi.org/10.3765/amp.v7i0.4495>
- Zymet, J. (2018b). *Lexical propensities in phonology: Corpus and experimental evidence, grammar, and learning* [Doctoral dissertation, University of California, Los Angeles].

