# Evidence for a Learning Bias Against Saltatory Phonological Alternations

James White

University of California, Los Angeles

Currently at: University of Ottawa

Address for correspondence:

James White
Department of Linguistics
University of Ottawa
Arts Hall, Room 401
70 Laurier Ave. East
Ottawa, ON Canada
K1N 6N5

Email:

jwhit6@uottawa.ca

Phone:

(613) 562 – 5800 ext. 1758

## Abstract

This study provides new experimental evidence that people learn phonological alternations in a biased way. Adult participants were exposed to alternations between phonetically dissimilar sounds (i.e., those differing in both voicing and manner, such as [p] and [v]). After learning these alternations, participants assumed, without evidence in the input, that more similar sounds (e.g., [b] and [v]) also alternated (Exp. 1). Even when provided with explicit evidence that dissimilar sounds (e.g., [p] and [v]) alternated but similar sounds ([b] and [v]) did not, participants tended to make errors in assuming that the similar sounds also alternated (Exp. 2). By comparison, a control group of participants found it easier to learn the opposite pattern, where similar sounds alternated but dissimilar sounds did not. The results are taken as evidence that learners have a soft bias, considering alternations between perceptually similar sounds to be more likely.

**Keywords**: phonology; artificial language; substantive bias; phonological alternations

## 1. Introduction

A phonological alternation occurs when a form is pronounced differently depending on its phonological context. In American English, for instance, the verb root *pat* is pronounced with a final [t] in the word *pats* [pæ**t**s] but with a tap sound [ɾ] in the word *patting* [pæ**ɾ**ɪŋ]. Native speakers of English know that the words *pats* and *patting* are related to the same verb root *pat* even though the root itself is pronounced differently in the two words. More generally, adult speakers tacitly know the distribution of phonological variants in their language and they are able to map multiple surface variants to the same representation at an abstract level (Lahiri & Marslen-Wilson, 1991). Learning the alternations of one's language must therefore be part of the language acquisition process, but there have been few studies directly looking at how they are learned.

Part of the process of acquiring such phonological mappings likely involves tracking statistical properties of the linguistic input. Distributional learning is undeniably a powerful tool available to the language learner. Research has indicated that it plays a role in several aspects of early phonological acquisition, including discrimination of speech sounds (Maye, Werker, & Gerken, 2002; Anderson, Morgan, & White, 2003), phonotactic learning (Chambers, Onishi, & Fisher, 2003), and word segmentation (Saffran, Aslin, & Newport, 1996). Indeed, 12-month-old infants can learn novel alternations in an artificial language based solely on distributional information (White et al., 2008).

A plausible starting point for learning alternations is by looking for complementary distributions among speech sounds, that is, by looking for cases where two speech sounds never occur in the same phonological environment (e.g., Peperkamp et al., 2006a). For instance, infants exposed to English may notice that [t] and [ɾ] never occur in the same environment, leading them to analyze the two sounds as alternating variants of the same phoneme. However, this process is unlikely to be based on distributional information alone. In English, for instance, the sounds [h] and [ŋ] happen to have completely non-overlapping distributions because [h] only occurs at the beginning of syllables (as in [hæt] *hat*) and [ŋ] only occurs at the ends of syllables (as in [sɪŋ] *sing*). No phonological analysis, however, would claim that [h] and [ŋ] are context-dependent variants of the same underlying sound because, other than being consonants, the two sounds are phonetically distinct in almost every possible way (Trubetzkoy, 1939/1969, pp. 49–50; see Peperkamp et al., 2006a for a similar case in French).

If learning which sounds alternate involves more than just tracking their distributions, which biases play a role in constraining this learning? The current study focuses on one way in which distributional learning might be constrained according to the similarity of the sounds involved. In the remainder of this section, I will first discuss the potential role of similarity in learning alternations and introduce the principle of minimal modification (section 1.1). I will then define saltatory alternations and explain why they represent a counter-example that is problematic for phonological theory (section 1.2). Finally, this section will conclude with a brief introduction to two experiments designed to test for a learning bias against such alternations (section 1.3).

### *1.1. The principle of minimal modification as the basis for a learning bias*

Languages differ in which alternations they allow, yet we see that some alternations occur commonly across languages whereas other possibilities are uncommon or even unattested. It has long been noted that alternations between dissimilar sounds are uncommon relative to alternations between more similar sounds (e.g., see Trubetzkoy, 1939/1969). Steriade

(2001/2008) expressed this principle in terms of *minimal modification,* arguing that alternations typically result in variant forms that are minimally different from each other, perceptually speaking. For example, many languages, such as German, do not allow voiced obstruents at the ends of words (where obstruents are sounds with a high degree of constriction in the oral tract, such as stops and fricatives). Such restrictions have a functional aerodynamic explanation because it is difficult to sustain the subglottal pressure necessary to maintain voicing in word-final obstruents (see Westbury & Keating, 1986; Kirchner, 1998, pp. 56-57). Steriade shows that in languages with such restrictions, word-final voiced obstruents are overwhelmingly “repaired” by being devoiced rather than by being deleted, nasalized, moved, etc. She provides evidence that devoicing represents the minimal perceptual change for voiced obstruents in that position.

Because of this cross-linguistic tendency for minimal modification, Steriade (2001/2008) has proposed that humans approach the language learning process with an *a priori* bias, which causes learners to assume that alternations between highly similar sounds are more likely than alternations between dissimilar sounds. This bias acts in a scalar way: the more similar the sounds, the better the alternation. Underlying this bias, Steriade proposes, is an implicit awareness on the part of learners, possibly based on their prior experience, of the relative perceptual similarity between pairs of sounds in any given phonological context. They can draw upon this mental representation of relative similarity, which she calls the perceptibility map (P-map), to facilitate their learning of phonological patterns.

Two studies provide experimental evidence that language learners are indeed sensitive to the similarity of sounds when learning novel alternations. Skoruppa, Lambrechts, and Peperkamp (2011) trained adults on arbitrary alternations between sounds differing in one, two, or three phonological features. They found that alternations between sounds differing in only one feature (e.g., [p ~ t]) were more quickly learned and more readily extended to new words of the same type relative to the other alternations. There was no difference in learning, however, for alternations between sounds differing in two features (e.g., [p ~ s]) and those differing in three features (e.g., [p ~ z]). These results suggest that learners find alternations between highly similar sounds easier to learn, but the effect of similarity beyond a single feature remains unclear.

Wilson (2006) used a different artificial language paradigm. Participants were trained on novel alternations involving palatalization, with one group learning that it occurred before high vowels (e.g., [ki] → [tʃi]) and a different group learning that it occurred before mid vowels (e.g., [ke] → [tʃe]). Velar stops are more acoustically and perceptually similar to palato-alveolar affricates before high vowels than before mid vowels (e.g., [ki] and [tʃi] are more similar than [ke] and [tʃe]). When each group was later tested on the cases that were not presented during training, the results suggested an asymmetry in generalization. The high vowel group ([ki] → [tʃi]) only rarely generalized the alternation to the mid vowel context ([ke] → [tʃe]) where the alternating sounds were less similar to each other. In comparison, the mid vowel group ([ke] → [tʃe]) was more willing, relatively speaking, to generalize to the high vowel context ([ki] → [tʃi]) where the alternating sounds were more similar to each other. With the support of a computational model, Wilson suggested that phonetic similarity was responsible for the observed asymmetry in degree of generalization (but see Moreton & Pater, 2012, for potential problems with the interpretation of these results).

The current study takes a novel approach to investigating the role of similarity in phonological learning. It focuses on a specific type of alternation, called a saltatory alternation, which represents a striking counterexample to the principle of minimal modification proposed by

Steriade and others. The experimental results presented below demonstrate a clear case where learners respond in a way that is not based solely on their input, but is consistent with a learning bias based on the principle of minimal modification.

### *1.2. Saltatory phonological alternations: A case of excessive modification*

A *saltatory alternation*[1] refers to a phonological alternation in which an intermediate, non-alternating sound must be "leaped over" (White, 2013). An illustrative example of a saltatory alternation comes from Campidanian Sardinian (Bolognesi, 1998). In this language, voiceless stops [p, t, k] become voiced fricatives [β, ð, ɣ] after vowels (as in 1a), but voiced stops [b, d, g] remain unchanged in that context (as in 1b). Crucial sounds are denoted with bold font:

| (1) | Isolation Form | | Post-vowel Form | Meaning |
|---|---|---|---|---|
| a. | [**p**ãi] | → | [sːu **β**ãi] | 'the bread' |
| | [**t**rintaduzu] | → | [sːu **ð**rintaduzu] | 'the thirty-two' |
| | [**k**uatːru] | → | [dɛ **ɣ**uatːru] | 'of four' |
| b. | [**b**ĩu] | → | [sːu **b**ĩu], | 'the wine' |
| | [**d**ɔmu] | → | [dɛ **d**ɔmu] | 'of house' |
| | [**g**ɔma] | → | [dɛ **g**ɔma] | 'of rubber' |

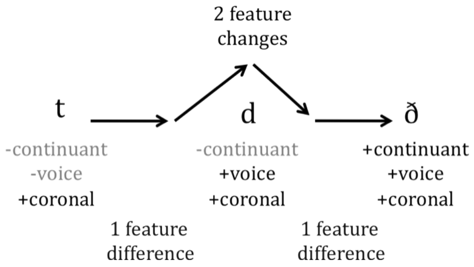
A sound is considered *intermediate* if it is more similar to each of the two alternating sounds than the two alternating sounds are to each other (for a similar use of the term, see Peperkamp et al., 2006a). Determining whether a sound is intermediate or not by this definition requires some measure of similarity. One way to determine similarity is by using phonological features (e.g., Skoruppa et al., 2011). Speech sounds are commonly characterized based on

---

[1] The term "saltatory" (from Latin *saltus* 'leap') is borrowed from Lass (1997, ch. 5), who used it to refer to similar types of "jumping" diachronic sound changes. Similarly, Minkova (1991) has used the term "leapfrogging", as well as the term "saltatory" (p. 222), to refer to such sound changes.

binary features corresponding to aspects of their articulation or aspects of their acoustics (Chomsky & Halle, 1968; Hayes, 2009, ch. 4). Features relevant to the consonants discussed in this paper are their place of articulation (e.g., lips [+labial], blade of tongue [+coronal], etc.), whether the airflow is completely stopped [–continuant] or continues to flow through the mouth [+continuant], and whether the vocal folds are vibrating [+voice] or not [–voice]. Considering these features, Figure 1 illustrates why the [t ~ ð] alternation in Campidanian Sardinian can be considered saltatory with an unchanging intermediate [d]. For each feature, the value of that feature for [d] is the same as the value of that feature for either [t] or [ð]: [d] is a stop like [t], it is voiced like [ð], and it has a coronal place of articulation like both [t] and [ð]. Because [d] is only different from either [t] or [ð] by a subset of the features that differentiate [t] and [ð] from each other, [d] can be considered intermediate between them. The other cases in (1) are analogous at other places of articulation.

Figure 1. Example of a saltatory alternation as seen in Campidanian Sardinian.



Another way of defining similarity is in terms of perception. Similarity is likely evaluated by listeners on the basis of multiple sources of information, and forming a complete understanding of how people judge the similarity of two speech sounds is not trivial (see Steriade, 2001/2008; Wilson, 2006; Mielke, 2012; Cristia et al., in press). For simplicity, I take the relatively straightforward notion of *mutual confusability* as an approximation of perceptual similarity.

The mutual confusability of two sounds can be computed by taking the rate that each sound is confused for the other in a perceptual identification task (e.g., Miller & Nicely, 1955; Singh & Black, 1966; Wang & Bilger, 1973). In such tasks, participants hear several speech sounds in a particular context, and they must decide which sound they heard on any given trial. The results are commonly organized into a matrix of values expressing the number of times that each sound label was chosen for each stimulus item. The mutual confusability (MC) of two sounds, x and y, can then be computed according to the following formula:

$$\text{MC }(x,y) = \frac{\text{proportion } x \text{ confused for } y \ + \ \text{proportion } y \text{ confused for } x}{2}$$

that is, the average of the proportion of times that the sounds are confused for each other. Greater MC values thus imply that two sounds are more confusable and therefore more similar to each other. Using the values reported by Wang and Bilger, for instance, [d] is more similar to [t] (MC = .029) and more similar to [ð] (MC = .083) than [t] and [ð] are to each other (MC = .011).[2] Thus by perceptual similarity, as well as by features, [d] can be considered intermediate between [t] and [ð].

---

[2] These values were calculated using the average of Wang and Bilger's CV and VC contexts across all signal-to-noise ratios (Tables II and III). The confusability of two speech sounds likely varies depending on one's language experience. The participants in Wang and Bilger's task were English speakers, like the participants in the current study.

It is also worth noting that Singh and Black (1966) obtained values for consonants in VCV contexts with no noise, which more closely reflects the conditions of the experiments in the current study. However, the error rates in the Singh and Black study were quite low, making it difficult to draw conclusions about the relative differences in confusability for the sounds relevant here. If the values from Miller and Nicely (1955) or the conditions without noise from Wang and Bilger are used, we find comparable similarity patterns, indicating that the main point does not depend on using any particular confusion matrix. At any rate, confusion rates in noise may be the more realistic basis for the measure. It is reasonable to assume that people use their prior experience perceiving speech sounds (in which much of the input is very noisy), rather than idealized lab speech, as the basis for judging the similarity of speech sounds.

Saltatory alternations, like the ones in Campidanian Sardinian, represent striking counterexamples to the principle of minimal modification. The fact that [t] is changed when it occurs after vowels is presumably driven by a ban on post-vowel voiceless stops in Campidanian Sardinian. The change from [t] to [ð] already represents an alternation between sounds that are not minimally different, but this alternation would not be particularly troubling if the language also banned post-vowel voiced stops (e.g., [d]). However, because intermediate [d] appears unchanged after vowels in Campidanian Sardinian (i.e., it does not alternate), the alternation between [t] and [ð] represents excessive modification. It is unclear why [t] changes into [ð] when instead changing to [d], which is legal in that phonological context, would require a less extreme modification. Looking at it from a different angle, if speakers of Campidanian Sardinian tolerate an alternation between sounds as dissimilar as [t] and [ð], why would they not tolerate an alternation between similar sounds, such as [d] and [ð]? Intuitively, saltatory alternations represent a contradiction to the principle of minimal modification because long journeys are allowed but short journeys are not. Indeed, due to this atypical characteristic, some theories of phonology (e.g., classical Optimality Theory; Prince & Smolensky, 1993/2004) predict that saltatory alternations should not exist (Lubowicz, 2002; Ito & Mester, 2003).

As mentioned, alternations with non-minimal modification are reported to be less common cross-linguistically (Trubetzkoy, 1939/1969; Steriade, 2001/2008). As special cases of such non-minimal alternations, saltatory alternations appear to be particularly uncommon. In addition to the case in Campidanian Sardinian, White (2013, ch. 2) provides a brief summary of other cases of saltation that have been reported in the literature, including in Colloquial Northern German (Ito & Mester, 2003) and Polish (Lubowicz, 2002). But relative to non-saltatory alternations, which abound in languages, saltatory alternations appear to be uncommon cross-linguistically.

### *1.3. Current study: Testing for a learning bias*

In the current study, artificial language experiments were used to see if learners exhibit a bias against saltatory alternations. The patterns tested were intended to be artificial analogues of those found naturally in Campidanian Sardinian (see section 1.2). The experiments each consisted of three phases: exposure, verification, and generalization. In the exposure phase, participants learned alternations by listening to pairs of nonwords representing singular and plural nouns in an artificial language, paired with pictures of singular and plural items, respectively. Plural words were always formed by adding [i] to the end of singular words. The singular words ended in target sounds, some of which changed when the [i] suffix was added in the plural word (e.g., singular [kamap], plural [kamavi]), providing the basis for the phonological alternations that participants were learning. Participants were also exposed to words ending in filler sounds (sonorants or sibilant fricatives) that did not change in the plural form (e.g., singular [luman], plural [lumani]). In the verification phase, participants were tested on <u>a subset of words from the exposure phase</u> using a two-alternative forced-choice task. Participants heard a singular word and then chose between two possible plural forms, one with a changed final target sound and one with an unchanged target sound. For each trial, one of the plural options (either the changing or non-changing option) followed the pattern learned during exposure and the other option was a foil that did not follow the learned pattern. The generalization phase was similar to the verification phase, except participants were tested on <u>novel words</u> that were not presented during the exposure phase. Some of the novel words ended in the same target sounds from the exposure phase, but to test for an inductive bias, a subset of the novel words in the generalization phase contained new target sounds that were not presented during exposure.

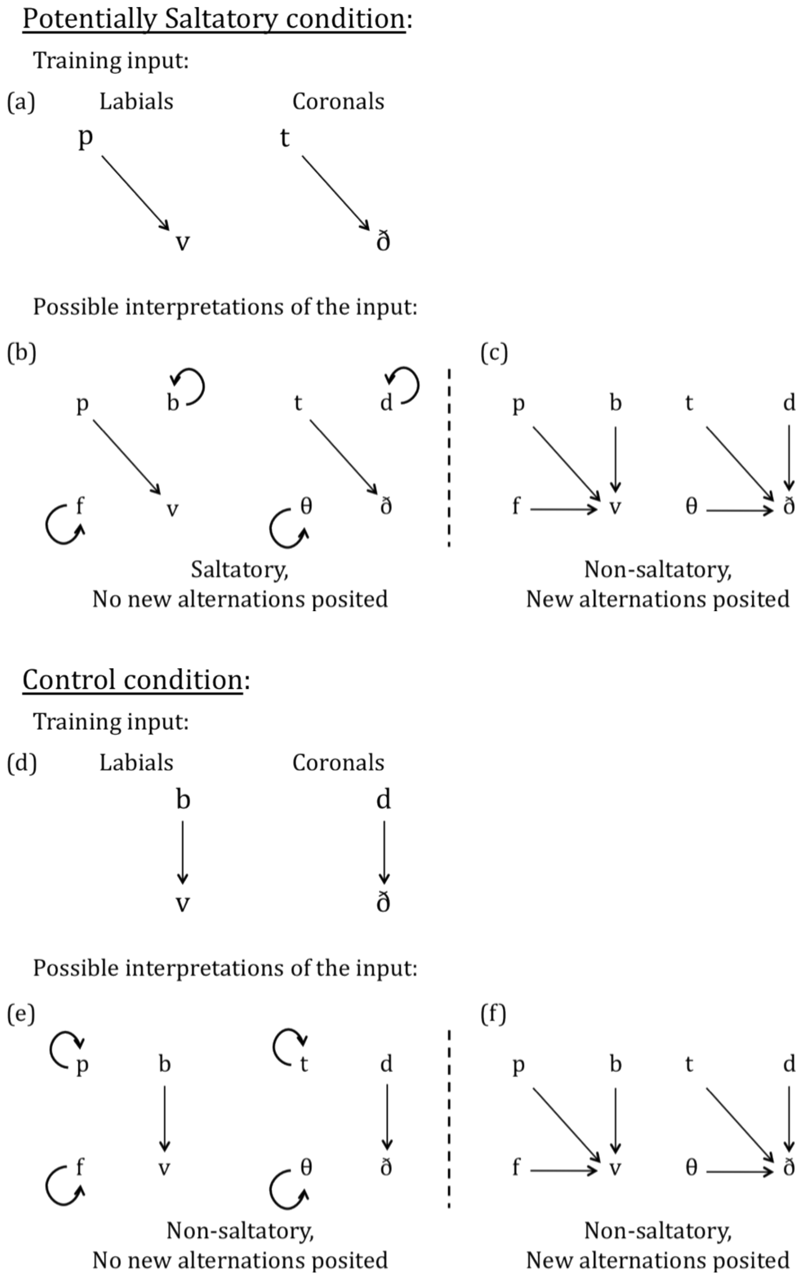
## 2. Experiment 1

Experiment 1 was designed to test an implicational question: given that a learner has acquired a *potentially* saltatory alternation, which assumptions does the learner make about untrained, intermediate sounds? Participants were randomly assigned to two conditions, Potentially Saltatory or Control. Figure 2 summarizes the input provided during the exposure phase for each of these conditions. Participants in the Potentially Saltatory condition were trained on alternations between voiceless stops and voiced fricatives ([p ~ v] and [t ~ ð]) during the exposure phase, but crucial examples of the intermediate consonants [b, f, d, θ] were withheld. The alternations learned during exposure in the Potentially Saltatory condition were thus ambiguous: they would be saltatory if intermediate sounds remained unchanged, but non-saltatory if intermediate sounds also alternated with voiced fricatives. In the Control condition, participants were instead trained on the alternations [b ~ v] and [d ~ ð] with examples of [p, f, t, θ] withheld. The alternations in the Control condition were unambiguously non-saltatory because none of the withheld sounds were intermediate between the alternating sounds.

The Control condition acted as the baseline for comparison in this experiment. As Figure 2 demonstrates, participants in the Control condition could choose to treat the untrained sounds as alternating (2f) or non-alternating (2e), but in both cases, the system learned would be non-saltatory. Assuming that learners are generally reluctant to extend patterns to unseen sounds (e.g., Peperkamp & Dupoux, 2007), participants were predicted to change untrained sounds relatively infrequently in the Control condition because they had no evidence for doing so. In the

Potentially Saltatory condition, participants had the same choice between changing untrained sounds without evidence (2c) or leaving them unchanged (2b). However, changing untrained sounds in the Potentially Saltatory condition (unlike in the Control condition) would avoid a saltatory alternation. Thus, if learners disfavor saltatory alternations, participants should change untrained sounds more often in the Potentially Saltatory condition than in the Control condition. In other words, saltation avoidance should lead participants in the Potentially Saltatory condition to counteract any default inclination (relevant to both conditions) towards being conservative.

Figure 2. Summary of the input during the exposure phase and possible interpretations of the input for the Potentially Saltatory and Control conditions in Experiment 1.



### *2.1. Method*

#### *2.1.1. Participants*

Forty undergraduate students in introductory psychology or linguistics classes at UCLA completed the experiment for partial course credit. Seven additional participants (2 in the

Potentially Saltatory condition, 5 in the Control condition) began the experiment but did not complete it because they failed to reach the criterion in the verification phase within the allotted time (see section 2.1.3. below). These participants received credit, but their data were not used in the analysis.

### *2.1.2. Materials and apparatus*

Exposure phase. For the exposure phase, 72 nonwords of the form CVCVC (e.g., [kamap]) were created as singular stimuli for the Potentially Saltatory condition. Half of the nonwords ended in the target sounds {p, t}, 18 of each, and half of the nonwords ended in one of the filler sounds {m, n, l, ɹ, s, ʃ}, 6 of each. The initial consonant sounds were drawn from the set {p, b, t, d, k, g, f, θ, s, ʃ, m, n, l, ɹ}. Because the crucial context for the alternations was between vowels, the medial consonants were chosen from the more limited set of filler sounds {m, n, l, ɹ, s, ʃ} so that the middle consonants would not provide unintended distributional information. Vowels were drawn from the set {i, a, u}. Nonwords were created by combining the possible consonants and vowels for each slot in a pseudorandom manner. Each consonant and vowel was used an approximately equal number of times in any given word position with the exception of the word-final position, which followed the proportions described above. Resulting nonwords were thrown out and replaced if they closely resembled real English words as judged by a native speaker (the author) or if they contained the same consonant in all positions (e.g., [ʃuʃuʃ]).

For each of the 72 singular nonwords, a plural form was also created. For nonwords ending in fillers sounds, plural forms were created by adding the vowel [i] to the end with no change in the final consonant (e.g., singular [luman], plural [lumani]). For nonwords ending in {p, t}, a final [i] was added and the final consonant was changed to the corresponding voiced fricative, either [v] or [ð] (e.g., singular [kamap], plural [kamavi]).[3]

Stress was placed on the second syllable of all words, that is, on the final syllable of CVCVC singular words and on the middle syllable of CVCVC-i plural words. This pattern, consistent with a system in which stem-final vowels receive stress (e.g., as attested in Albanian; Bevington, 1974, p. 24), was chosen so that stress would be on the same syllable of the stem in the singular and plural forms.

For the singular nonwords ending in {p, t}, corresponding nonwords for the Control condition were created by changing each final [p] to [b] and each final [t] to [d]. The same list of 36 singular nonwords ending in filler sounds from the Potentially Saltatory condition was used in the Control condition without alteration. Plural forms for the Control condition were created in the same manner described above. Thus, the list of stimuli for the Potentially Saltatory condition and the Control condition differed only in the final target sound of the non-filler items. For example, singular [kamap] and plural [kamavi] in the Potentially Saltatory condition corresponded to singular [kamab] and plural [kamavi] in the Control condition.

Each set of nonwords was randomly paired with one of 72 pictures showing singular objects (e.g., a strawberry) and 72 corresponding pictures showing multiple objects (e.g., two strawberries). The pictures were made up of clipart-style images or small photographs of everyday nouns taken from the Internet. The number of objects in the plural pictures was always greater than one, but otherwise varied. Corresponding nonwords in the Potentially Saltatory

---

[3] For the coronal sounds, [ð] was chosen as the voiced fricative rather than logically possible [z] for two reasons. First, [z] changes the extra phonological feature [strident], which is not changed by [ð]. Second, [ð] was chosen to remain as close as possible to the attested case in Campidanian Sardinian (described in section 1.2.) while still using sounds that are English phonemes, which should be more easily distinguished by the English-speaking participants. The main predictions related to saltation avoidance still hold if [s] and [z] had been used instead of [θ] and [ð].

condition and the Control condition were paired with the same pictures. The word-picture pairings were the same for all participants in all conditions.

Verification phase. For the Potentially Saltatory condition, 32 of the singular nonwords (8 *p*-final, 8 *t*-final, and 16 fillers), along with their associated pictures, were randomly chosen from the set of nonwords in the exposure phase for use in the verification phase. For each nonword in this phase, it was necessary to have both a changing plural option and a non-changing plural option. To make a changing plural option for the singular nonwords ending in filler sounds, the following correspondences were used: [m, ʃ, ɹ] changed to [v], and [n, s, l] changed to [ð]. As examples, singular [kamap] had plural options [kamapi] and [kamavi], and singular [luman] had plural options [lumani] and [lumaði]. In the Control condition, the corresponding set of nonwords was used, and the plural options were formed in the same manner.

Generalization phase. For the generalization phase, 72 new singular nonwords were created in the same manner described above. For the Potentially Saltatory condition, one-third ended in {p, t} (12 of each), one-third ended in the filler sounds {m, n, l, ɹ, s, ʃ} (4 of each), and one-third ended in the intermediate sounds {b, d, f, θ} (6 of each). For the Control condition, the same set of words were used except word-final [p] was changed to [b], word-final [t] was changed to [d], and vice versa. Thus, one-third of the Control nonwords ended in {b, d}, one-third ended in the filler sounds {m, n, l, ɹ, s, ʃ}, and one-third ended in the sounds {p, t, f, θ}. Changing and non-changing plural forms were created in the same manner described above. The nonwords for the generalization phase were randomly assigned to 72 new pairs of pictures showing singular and plural objects. For reference, a sample of the nonwords used in all phases is provided in Table 1.

Table 1. Representative sample stimuli from Experiment 1, transcribed in IPA.[4]

*Exposure phase*

| *Potentially Saltatory condition* | | | *Control condition* | | |
|---|---|---|---|---|---|
| *Type* | *Examples* | | *Type* | *Examples* | |
| | *Singular* | *Plural* | | *Singular* | *Plural* |
| *18 p → v* | *kamap* | *kamavi* | *18 b → v* | *kamab* | *kamavi* |
| | *nisup* | *nisuvi* | | *nisub* | *nisuvi* |
| *18 t → ð* | *ɹamit* | *ɹamiði* | *18 d → ð* | *ɹamid* | *ɹamiði* |
| | *kunit* | *kuniði* | | *kunid* | *kuniði* |
| *36 fillers (same in both conditions)* | *luman* | *lumani* | | | |
| | *gunam* | *gunami* | | | |
| | *misil* | *misili* | | | |
| | *kaluɹ* | *kaluɹi* | | | |
| | *kuɹas* | *kuɹasi* | | | |
| | *ɹamiʃ* | *ɹamiʃi* | | | |

*Verification phase (subset from exposure phase, correct answer in bold font)*

| Potentially Saltatory condition | | | | Control condition | | | |
|---|---|---|---|---|---|---|---|
| Type | Examples | | | Type | Examples | | |
| | Singular | Non-changing plural option | Changing plural option | | Singular | Non-changing plural option | Changing plural option |
| 8 p-final | kamap | kamapi | **kamavi** | 8 b-final | kamab | kamabi | **kamavi** |
| | nisup | nisupi | **nisuvi** | | nisub | nisubi | **nisuvi** |
| 8 t-final | ɹamit | ɹamiti | **ɹamiði** | 8 d-final | ɹamid | ɹamidi | **ɹamiði** |
| | kunit | kuniti | **kuniði** | | kunid | kunidi | **kuniði** |
| 16 fillers (same in both conditions) | luman | **lumani** | lumaði | | | | |
| | gunam | **gunami** | gunavi | | | | |
| | gimal | **gimali** | gimaði | | | | |
| | kaluɹ | **kaluɹi** | kaluvi | | | | |
| | kuɹas | **kuɹasi** | kuɹaði | | | | |
| | ɹamiʃ | **ɹamiʃi** | ɹamivi | | | | |

---

[4] A full list of stimuli is available at the author's website: http://artsites.uottawa.ca/james-white/papers/.

*Generalization phase*

| Potentially Saltatory condition | | | | Control condition | | | |
|---|---|---|---|---|---|---|---|
| Type | Examples | | | Type | Examples | | |
| | Singular | Non-changing plural option | Changing plural option | | Singular | Non-changing plural option | Changing plural option |
| 12 p-final | sulap<br>kiʃap | sulapi<br>kiʃapi | sulavi<br>kiʃavi | 12 b-final | sulab<br>kiʃab | sulabi<br>kiʃabi | sulavi<br>kiʃavi |
| 12 t-final | gumut<br>ʃaɹut | gumuti<br>ʃaɹuti | gumuði<br>ʃaɹuði | 12 d-final | gumud<br>ʃaɹud | gumudi<br>ʃaɹudi | gumuði<br>ʃaɹuði |
| 6 b-final | talab | talabi | talavi | 6 p-final | talap | talapi | talavi |
| 6 d-final | masid | masidi | masiði | 6 t-final | masip | masiti | masiði |
| 6 f-final | tunuf | tunufi | tunuvi | 6 f-final | tunuf | tunufi | tunuvi |
| 6 θ-final | paɹuθ | paɹuθi | paɹuði | 6 θ-final | paɹuθ | paɹuθi | paɹuði |
| 24 fillers (same in both conditions) | niʃin<br>tasam<br>baɹul<br>funiɹ<br>ɹanus<br>θanaʃ | niʃini<br>tasami<br>baɹuli<br>funiɹi<br>ɹanusi<br>θanaʃi | niʃiði<br>tasavi<br>baɹuði<br>funivi<br>ɹanuði<br>θanavi | | | | |

<u>Stimuli recording</u>. A male native speaker of English with phonetic training, who was unaware of the purpose of the experiment, recorded the nonwords in a soundproof booth using a Shure SM10A head-mounted microphone, whose signal ran through an XAudioBox pre-amplifier and A-D device. The recordings were done using PcQuirerX at a sampling rate of 22,050 Hz. In order to make the relevant contrasts as perceptible as possible, a relatively careful speech style was used. The speaker was asked to release all word-final consonants and to fully voice all voiced segments. The spectrogram for each token was inspected using Praat (Boersma & Weenink, 2010) to confirm that voicing and frication was present in voiced sounds and fricatives, respectively.

Stimuli for the Potentially Saltatory condition and the Control condition were recorded on separate occasions, but every effort was made to ensure that the stimuli were produced in a consistent way. To ensure that the target sounds were perceived equally well in the two conditions, five native speakers of English listened to the singular nonwords ending in the target sounds [p, t, b, d, f, θ] from the generalization phase and the corresponding changing and non-changing plural options (288 total nonwords). They were asked to identify the target sound (i.e., the last consonant) for each nonword by choosing one of eight options, [p, t, b, v, d, f, θ, ð]. The sounds [θ] and [ð] were described as "<th> in *thick*" and "<th> in *the",* respectively. Accuracy was very high overall, and crucially, it was comparable between the Potentially Saltatory condition (94.3%) and the Control condition (94.7%), indicating that the target sounds were perceived equally well across the two conditions.

### *2.1.3. Procedure*

The experiment consisted of three phases: exposure, verification, and generalization. In the exposure phase, participants were instructed that they would be learning words in a foreign language. They were told that they should try their best to remember the words because they would be tested on them later. Participants were encouraged to repeat each word out loud after hearing it because doing so would help them remember. Participants heard 72 unique, self-paced trials in this phase. Each trial began with a picture showing a singular object appearing in the center-left part of the computer screen. After the picture had been displayed for one second, the singular nonword for that item was played over headphones. The singular picture disappeared 2.5 seconds after the sound file began playing, and the corresponding plural picture immediately appeared in the center-right part of the screen. The plural nonword for that picture was played over headphones one second after the plural picture appeared. After hearing both nonwords, participants pressed the spacebar, which initiated the next trial. The plural picture remained on the screen until the participant pressed the spacebar. Nonwords were only presented in auditory form, never in orthography. Participants were given no further instructions. The order of trials in this phase, as well as in the following two phases, was randomized anew for each participant by E-prime. The exposure phase lasted approximately 10-25 minutes depending on how quickly a given participant pressed the spacebar.

In the verification phase that followed, participants were tested on 32 words that they had heard during the exposure phase. The purpose of the verification phase was to ensure that participants had successfully learned the pattern presented during exposure. A singular picture appeared on the left side of the screen and the singular nonword for that picture was played over headphones. Once the singular picture disappeared, the plural picture was displayed on the right side of the screen along with a row of question marks located just under the picture. Up to this point, the trial was identical to trials in the exposure phase (except for the question marks), including the timing of the stimuli. After the plural picture was on the screen for 1.5 seconds, participants heard two plural options—the changing plural option and the non-changing plural option—with a one second pause in between them. Order of the two plural options was counterbalanced such that the changing option and the non-changing option occurred first an equal number of times for each type of singular word. Participants were asked to choose the correct word for the plural picture by pressing the appropriate keys on the keyboard: a key marked “1” for the first option and a key marked “2” for the second option (the “f” and “j” keys, respectively, were used for this purpose). The next trial started immediately after a response key was pressed. The verification phase lasted approximately five minutes.

At the end of the verification phase, a screen appeared showing the participant's accuracy. If participants did not achieve an accuracy of at least 80%, they were told that they needed to reach 80% to continue in the experiment. They then repeated the exposure and verification phases until they reached an accuracy of at least 80%. Participants heard the same trials on subsequent exposure and verification phases, but in a different random order. Participants continued cycling through the exposure phase and verification phase until they either reached 80% or 50 minutes had elapsed (typically after two cycles). Those who did not reach 80% after 50 minutes did not complete the generalization phase.

Participants who achieved 80% accuracy on the verification phase moved into the generalization phase, where they were tested on 72 novel words, including words ending in untrained target sounds. Otherwise, trials were identical to those in the verification phase. Participants were instructed that they would be hearing new words in the same language and that they should make their best guess based on their experience so far with the language. The generalization phase lasted approximately 10 minutes.

The experiment was conducted in a quiet room on a Dell computer equipped with a 20-inch monitor and Sony MDR-V200 headphones. The experimental software E-prime (version 2.0) was used to present the stimuli and record the responses.

## *2.2. Results*

Only responses in the generalization phase (i.e., responses to words not encountered during exposure) were included in the analysis. The data were analyzed using mixed-effects logistic regression models (see Jaeger, 2008), implemented in R (R Core Development Team, 2008) using the *lme4* package (Bates, Maechler, & Dai, 2008). To compare models, likelihood ratio tests were conducted using the anova() function. The likelihood ratio test compares the log likelihoods of two models with different numbers of factors (in a subset relationship) and determines if the added factors are justified based on a chi-squared test (see Baayen, 2008, ch. 7). The random effect structure of the final model was determined by backwards stepwise comparison. As a starting point, random intercepts and random slopes were included for participants and items. Each random effect factor was then taken out one a time, the simpler model was compared to the more complex model using likelihood ratio tests, and the random factor was removed if it did not significantly improve model fit (see Baayen, Davidson, & Bates, 2008).

### *2.2.1. Trained alternations and filler sounds*

I first consider how well participants extended the patterns learned during the exposure phase to new words of the same type in the generalization phase. There was no reason to expect a difference between the Potentially Saltatory condition and the Control condition on trained sounds because participants in both conditions had to reach the 80% criterion in the verification phase to move on to the generalization phase. Indeed, the results show that accuracy on trained stops was similar in the Potentially Saltatory condition (96.3% of [p, t] correctly changed to [v, ð]) and the Control condition (88.8% of [b, d] correctly changed to [v, ð]). A mixed logit model predicting log odds of an accurate response, with random intercepts for subjects and a fixed effect of Condition (Potentially Saltatory condition vs. Control condition), found that Condition was not a significant predictor (Wald $z = 1.12$, $p = .27$), and including it in the model did not significantly improve model fit, $\chi^2(1) = 1.29$, $p = .26$.

The percent of filler sounds that were changed incorrectly was likewise comparable in the Potentially Saltatory condition (10.2% changed incorrectly) and the Control condition (4.4% changed incorrectly). A mixed logit model predicting log odds of an accurate (unchanging) response, with random intercepts for subjects and words and a fixed effect of Condition (Potentially Saltatory condition vs. Control condition), found that Condition was not a significant predictor (Wald $z = .03$, $p = .97$), and including it in the model did not significantly improve model fit, $\chi^2(1) = .001$, $p = .97$. Overall, accuracy on the trained alternations and filler sounds was high, indicating that participants in both conditions readily extended the patterns that they learned during exposure to new words of the same type.

### *2.2.2. Untrained sounds*

The primary purpose of Experiment 1 was to see how participants would treat untrained, intermediate sounds given that they had learned a potentially saltatory alternation. Recall that if learners have a bias that disfavors saltatory alternations, participants were predicted to change untrained sounds more often in the Potentially Saltatory condition than in the Control condition. Figure 3 shows the results for untrained sounds according to Condition and Sound Type (untrained stops vs. untrained fricatives). Because participants received no information about words ending in these sounds during the exposure phase, there was no “correct” answer, so accuracy cannot be calculated. Instead, the mean percent of trials in which participants chose the changing plural option is reported (e.g., for the singular word [talab], how often did they choose [talavi] rather than [talabi]). Overall, we see that participants tended to change the untrained intermediate sounds more often in the Potentially Saltatory condition than in the Control condition, for both untrained stops (70.0% vs. 20.8%) and untrained fricatives (45.0% vs.

15.8%). Within the Potentially Saltatory condition, participants also showed a tendency to change untrained stops more often than untrained fricatives (70.0% vs. 45.0%).

To evaluate these differences, a mixed logit model was fitted, predicting log odds of having a changing response for words ending in untrained target sounds. The final model included fixed effects for Condition (Potentially Saltatory vs. Control), Sound Type (stops vs. fricatives), and a Condition x Sound Type interaction. Random intercepts for subjects and by-subject random slopes for Sound Type were also included. By-subject random slopes were included because they significantly improved model fit according to a likelihood ratio test, $\chi^2(3) = 75.62$, $p < .001$. Random intercepts for individual words were not included in the final model because they did not significantly improve model fit, $\chi^2(1) = .12$, $p = .72$.

The fixed effects for the final model are provided in Table 2. The significant negative intercept indicates that untrained fricatives in the Control condition (coded as the baseline in this model) were changed infrequently. The non-significant main effect of Sound Type follows from the fact that untrained stops were also changed infrequently in the Control condition. Condition was a significant predictor in the model, indicating that participants chose the changing option for words in the Potentially Saltatory condition (i.e., those with final intermediate sounds) significantly more often than for words in the Control condition. These results are consistent with the main prediction: participants changed untrained sounds more often when they were intermediate between a potentially saltatory alternation. There was also a significant interaction, indicating that untrained stops were changed more frequently than untrained fricatives, but only in the Potentially Saltatory condition.

Figure 3. Results for untrained sounds in Experiment 1 by Condition and Sound Type. Individual results (diamonds) and overall means (bars) are provided.

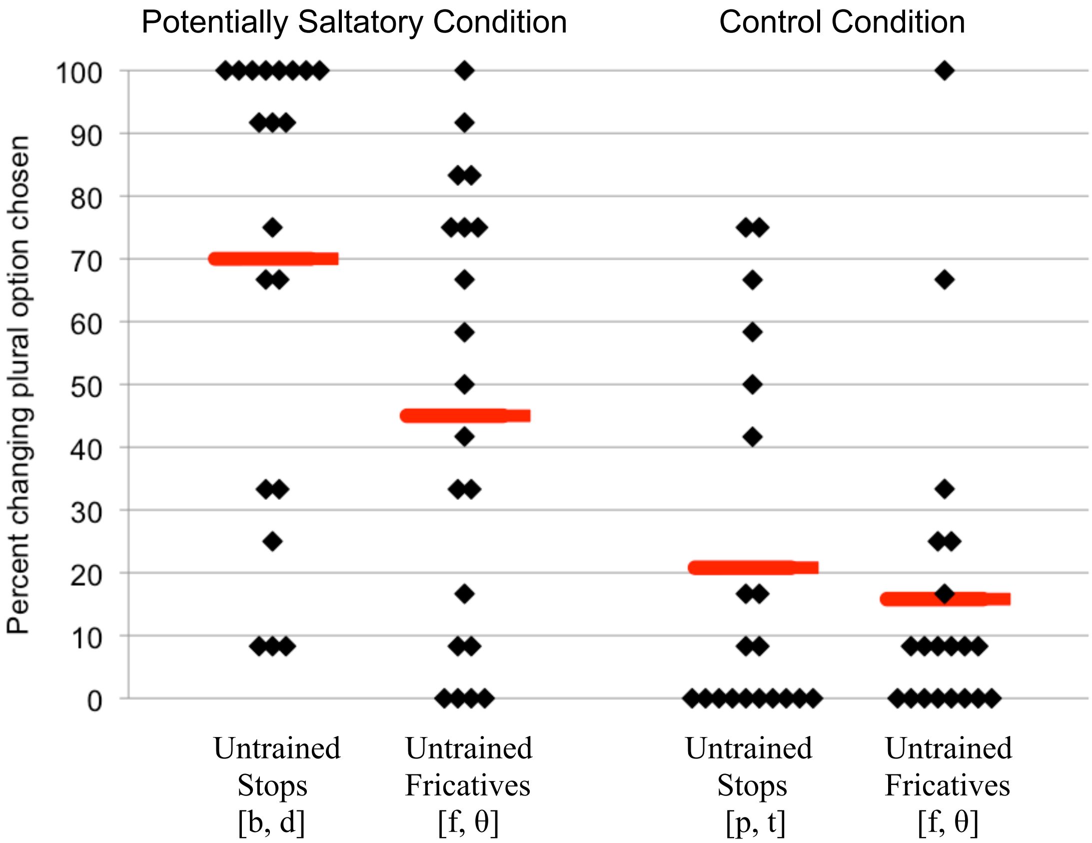


Table 2. Summary of the fixed effects for untrained sounds in Experiment 1.

| Predictor | Estimate | Standard error | Wald *z* | *p*-value |
|---|---|---|---|---|
| Intercept | -2.80 | .57 | -4.87 | <.001 |
| Condition = *Potentially Saltatory* | 2.35 | .78 | 3.02 | .002 |
| Sound Type = *Untrained stops* | - .33 | .72 | - .46 | .65 |
| Interaction = *Potentially Saltatory & Untrained stops* | 2.80 | .97 | 2.89 | .004 |

### *2.2.3. Effect of amount of exposure*

Due to the experimental design, participants received variable amounts of training, either completing one or two cycles of the exposure phase.[5] This design choice was made because of the implicational nature of the hypothesis: given that a participant has learned a potentially saltatory alteration, how does the participant treat untrained, intermediate sounds? To answer this question, it was more critical to ensure that participants had actually learned the potentially saltatory alternations (or the comparable non-saltatory alternations in the Control condition) before being tested on new cases, as opposed to ensuring that all participants received the same amount of exposure. On average, participants in the Potentially Saltatory condition completed 1.40 cycles of the exposure phase whereas participants in the Control condition completed 1.75 cycles. Thus, it is possible that the amount of exposure, rather than the intermediate status of the untrained sounds, can explain the differences observed between the Potentially Saltatory condition and the Control condition.

To address this possibility, a new mixed logit model for untrained sounds was run with an added fixed effect for number of cycles through the exposure phase (either one or two). In the model, amount of exposure was not a significant predictor (Wald $z = -.73$, $p = .46$) and its inclusion did not significantly improve model fit according to a likelihood ratio test, $\chi^2(1) = .50$, $p = .48$. An additional model was run with interaction effects between amount of exposure and each of the other fixed effects. None of the effects related to amount of exposure reached significance in the model. Moreover, including the added fixed effects for amount of exposure and the associated interactions did not significantly improve model fit, $\chi^2(4) = 5.19$, $p = .27$. In both models, the effect of Condition (Potentially Saltatory vs. Control) remained significant

---

[5] In principle, it was possible to have three cycles, but no participant who completed Experiment 1 within the alotted hour had more than two cycles of the exposure phase.

(Wald $z > 2.5$, $p = .01$). Overall, these models indicate that the amount of exposure did not have a significant effect on how often participants chose the changing option for untrained sounds in Experiment 1.

### *2.3. Discussion*

The results from Experiment 1 were consistent with a learning bias that disfavors saltatory alternations. Participants who learned potentially saltatory alternations during the exposure phase (Potentially Saltatory condition) changed intermediate sounds at a high rate despite having no direct evidence for such changes in the input. By changing intermediate sounds, participants avoided the dispreferred saltatory alternations that would result if the intermediate sounds remained unchanged. Participants in the Control condition learned comparable alternations, but were not under the same pressure to avoid saltatory alternations. As predicted, they changed the untrained sounds much less frequently than participants in the Potentially Saltatory condition. Intuitively, the results can be summarized in the following way: participants learning alternations between dissimilar sounds extended the pattern to alternations between more similar sounds, consistent with the principle of minimal modification.

The Control condition served as an important baseline for comparison because it provided us with an idea of how often participants would change untrained sounds when saltation was not a factor. We see that participants were not making responses at random for the untrained, non-intermediate sounds in the Control condition, which would have resulted in chance performance (50%). Rather, participants appear to have sensibly taken the more conservative approach, that is, they were reluctant to posit new alternations without evidence, consistent with previous findings (e.g., Peperkamp & Dupoux, 2007). The relatively high rate that participants chose the changing option in the Potentially Saltatory condition is even more striking compared to the low rate that the changing option was chosen in the Control condition. Together, the results indicate that the preference to avoid saltatory alternations was strong enough that learners were willing to go against their default preference to avoid positing new alternations without evidence.

The Control condition also rules out two possible alternative explanations for why participants in the Potentially Saltatory condition might have changed untrained sounds. First, participants may have been giving product-oriented responses (Bybee & Slobin, 1982). In other words, participants may have responded based on the form of the *product* (i.e., the plural word) rather than based on the form of the singular word. During exposure, participants heard a large proportion of plural words with either [v] or [ð] as the final consonant. In fact, half of the plural forms had [v] or [ð] as the final consonant whereas only one-twelfth of the total plural forms in the exposure phase had any one of the other possible final consonants (see section 2.1.3.). Participants may have responded to any untrained cases by matching the frequency of the plural endings that they heard during training. This strategy would have resulted in a preference for the changing option for any novel sound. However, the same proportion of changing and non-changing plural forms was heard during the exposure phase in both the Potentially Saltatory condition and the Control condition. Thus, if participants were using a product-oriented strategy, they should have changed an equal percentage of untrained sounds in both conditions, contrary to the results.

A second possibility is that participants may have been biased to target a more general class of phonetically similar sounds (e.g., Saffran & Thiessen, 2003; Skoruppa & Peperkamp, 2011). In particular, they may have learned that all stops (or all obstruents) changed to voiced fricatives between vowels, rather than limiting the alternations only to voiceless stops. It has been argued that phonological generalizations are easier to learn when they can be described using fewer features (e.g., Skoruppa & Peperkamp, 2011; Moreton & Pater, 2012). Because examples ending in voiced stops and voiceless fricatives were withheld from training, targeting a more general class of sounds would (a) be equally consistent with the input, (b) require fewer features for grouping the sounds, and (c) explain the extension to untrained sounds in the Potentially Saltatory condition. However, if participants were merely targeting general classes of similar sounds, it is left unexplained why they did not generalize to untrained sounds (at least to voiceless stops) at comparable rates in the Control condition. Because participants in the Control condition did not change untrained sounds at the same rate as those in the Potentially Saltatory condition, we can conclude that targeting a more general class of sounds cannot fully explain the results in Experiment 1. I return to a more nuanced version of the featural complexity account in the General Discussion.

Crucially, neither of these alternative explanations—product-oriented responding or targeting general classes of sounds—can account for the difference observed between the Potentially Saltatory condition and the Control condition. However, it is worth noting that participants did sometimes choose the changing option for untrained sounds in the Control condition (roughly 15-20% of the time). If participants are truly averse to positing new alternations without a reason, then we might expect this value to be very close to 0%. The low, but non-zero, values in the Control condition may simply be due to random noise, but it could indicate that the alternate explanations described above had some degree of influence on the responses, at least for some participants (see individual data points in Figure 3). The changing responses in the Control condition may also be due, in part, to task demands. In the two-alternative forced-choice task, some participants may have been motivated to use both response options for untrained sounds, even in the Control condition.

The final noteworthy aspect of Experiment 1 is that participants in the Potentially Saltatory condition preferred to change untrained stops more frequently than untrained fricatives. In terms of phonological features, these sounds are equally different from each other: voiced stops and voiced fricatives differ in one feature ([continuant]), and voiceless fricatives and voiced fricatives differ in one feature ([voice]).[6] If participants prefer alternations between similar sounds compared to alternations between less similar sounds (i.e., if they follow the principle of minimal modification), these results imply that abstract features may not provide a sufficient measure of similarity.[7] I return to this issue in the General Discussion.

Overall, Experiment 1 demonstrated that learners disprefer saltatory alternations when presented with ambiguous input. They changed intermediate sounds, without evidence, so that the alternations learned during exposure were rendered non-saltatory. Experiment 2 was designed to look further at the strength of the anti-saltation bias observed in Experiment 1. Does the bias appear only when the input is ambiguous, or does it endure even when there is unambiguous evidence during training for saltatory alternations? In Experiment 2, participants

---

[6] That is, the changes are equal if the features for place of articuation are construed broadly as labial or coronal. If labio-dentality is considered part of the feature system to differentiate between the slightly different places of articulation of [b] and [v], then /b/ → [v] requires changing two features, whereas /f/ → [v] requires only one (voicing). Under the similarity bias hypothesis, this would predict a preference to change /f/ → [v] <u>more</u> often, the opposite of the actual results. The same argument holds for the subtle place of articulation distinction for the coronal stops [t, d], which are alveolar, and the coronal fricatives [θ, ð], which are dental (see Hayes, 2009).

[7] An alternative explanation for the difference in untrained stops and fricatives in the Potentially Saltatory condition is that two of the non-changing filler sounds were [s] and [ʃ], which are voiceless fricatives like untrained [f] and [θ]. Even though [s] and [ʃ] are sibilants (unlike [f] and [θ]), it is possible that some participants were reluctant to change [f] and [θ] because they form a natural class of voiceless fricatives with two of the filler sounds. To address this possibility, a version of the Potentially Saltatory condition was run with only sonorants [m], [n], [l], [ɹ] as filler sounds (otherwise identical). A difference was once again found between untrained stops (65% changed) and untrained fricatives (47% changed), indicating that the difference was <u>not</u> due to the presence of [s] and [ʃ] in the set of filler sounds. Nevertheless, [s] and [ʃ] were not included in the set of filler sounds in Experiment 2.

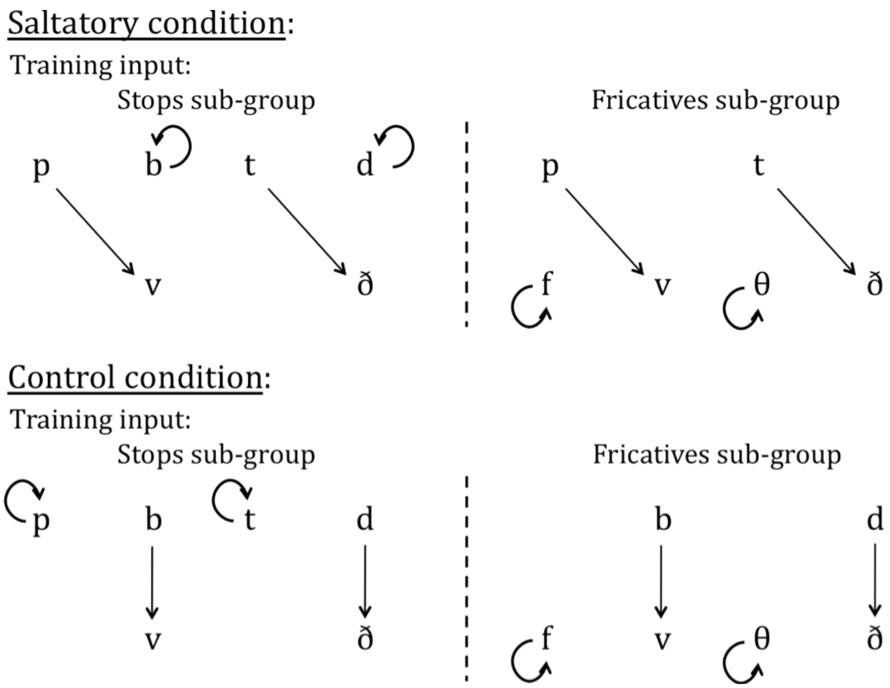
were exposed to the same alternations as in Experiment 1, but they received explicit evidence that intermediate sounds did not change.

## 3. Experiment 2

In Experiment 2, participants were randomly assigned to two conditions: a Saltatory condition and a Control condition. Participants in the Saltatory condition learned the same alternations as in the Potentially Saltatory condition in Experiment 1 (i.e., [p ~ v] and [t ~ ð]). This time, however, they also had cases of non-changing intermediate voiced stops [b, d] (Stops sub-group) or non-changing intermediate voiceless fricatives [f, θ] (Fricatives sub-group) during the exposure phase. As a result, the alternations were explicitly saltatory because there was evidence for an intermediate non-alternating sound. The Stops sub-group received no information about intermediate fricatives [f, θ] during exposure, and the Fricatives sub-group received no information about intermediate stops [b, d]. Participants in the Control condition learned the same alternations as in the Control condition of Experiment 1 (i.e., [b ~ v] and [d ~ ð]), but they too were trained that additional sounds did not alternate. The Stops sub-group had examples of non-changing [p, t], and the Fricatives sub-group had examples of non-changing [f, θ]. The input for each of the four resulting sub-groups (Saltatory/Stops, Saltatory/Fricatives, Control/Stops, and Control/Fricatives) is summarized in Figure 4.

The goal of Experiment 2 was to see if participants would find it difficult to learn (and remember) that intermediate sounds did not alternate, even with explicit evidence in exposure. Provided that the anti-saltation bias observed in Experiment 1 is sufficiently strong, participants in the Saltatory condition were predicted to make more errors on intermediate sounds (i.e., by incorrectly choosing the changing plural option) relative to the number of errors made by participants in the Control condition on comparable sounds that were not intermediate.

Figure 4. Summary of input in Experiment 2.



*3.1. Method*

*3.1.1. Participants*

Eighty undergraduate students in introductory psychology or linguistics classes at UCLA completed the experiment for partial course credit. None of the participants had participated in Experiment 1. Twenty-one additional participants (13 in the Saltation condition, 8 in the Control condition) began the experiment but did not complete it because they failed to reach the 80% criterion in the verification phase within the allotted time. These participants received credit, but their data were not used in the analysis. In addition, the data of four participants (all from the

Saltatory condition) who completed the experiment were not used in the analysis because the participants had clearly not learned (or retained) at least one of the trained alternations (all had accuracy of less than 10% for one of the trained alternations in the generalization phase).[8] These participants were replaced by four new participants.

### *3.1.2. Materials*

Exposure phase. The exposure phase consisted of 72 singular nonwords. In the Saltatory condition, half of the singular nonwords ended in the target sounds {p, t}, 18 of each type. Half of the remaining nonwords (18) ended in the intermediate stops [b, d] (Stops sub-group) or the intermediate fricatives [f, θ] (Fricatives sub-group), 9 of each type. The final quarter of the nonwords ended in filler sounds, consisting of {m, n, l, ɹ}, 3-4 of each type. The singular nonwords were generated as described in Experiment 1. Changing and non-changing plural forms were created for each of the singular nonwords in the same way described for Experiment 1. The Control condition was analogous, except all target [p, t] sounds were substituted for [b, d], and vice versa. The same pairs of pictures from Experiment 1 were used in Experiment 2.

Verification phase. For the verification phase, 32 of the nonwords (8 *p*-final, 8 *t*-final, 4 *b*-final (Stops sub-group) or *f*-final (Fricatives sub-group), 4 *d*-final (Stops sub-group) or *θ*-final (Fricatives sub-group), and 2 ending in each of the four filler sounds) were chosen at random from the set of forms used in the exposure phase for the Saltatory condition. The corresponding words were used in the Control condition.

---

[8] The focus of the experiment was to determine how learners would perform on sounds that were intermediate in a saltatory alternation. These four participants clearly had not internalized the alternations that would make the target sounds intermediate, so their data were ultimately not relevant to the question in this study. The main overall pattern of results remains, however, even if their data are included.

Generalization phase. For the generalization phase, 64 new singular forms were created (24 ending in [p, t], 12 ending in [b, d] (Stops sub-group) or [f, θ] (Fricatives sub-group), 12 ending in filler sounds, and 16 ending in untrained sounds [f, θ] (Stops sub-group) or [b, d] (Fricatives sub-group)). The Control condition was analogous, except all target [p, t] sounds were substituted for [b, d], and vice versa.

Except as noted, the nonwords for each phase were created and recorded following the same procedure described for Experiment 1. Five native English speakers identified the target sounds of singular nonwords ending in {p, b, f, t, d, θ} and the corresponding changing and non-changing plural options (the task and the five speakers were the same as in Experiment 1, see section 2.1.2). Accuracy was comparably high in the Saltatory condition (97.0% correct) and the Control condition (96.5% correct), indicating that the target sounds were perceived equally well in the two conditions.

### *3.1.3. Procedure*

The procedure was identical to the procedure in Experiment 1, except that the generalization phase contained only 64 trials in Experiment 2 as opposed to 72 trials in Experiment 1. Note that participants were still required to reach 80% accuracy in the verification phase, which included cases of non-alternating intermediate sounds. Including intermediate sounds in the verification phase likely made the task harder, a point I will return to in the Discussion section.

### *3.2. Results*

Only results from the generalization phase were analyzed. Analyses were conducted using mixed-effects logistic regression models as in Experiment 1.

#### *3.2.1. Trained alternating and filler sounds*

Performance on these sounds was not the focus of this study and, in fact, was partially used as an exclusion criterion. Recall that four participants were excluded from the Saltatory condition even though they reached the 80% criterion in the verification phase because their accuracy on the trained alternations in the generalization phase was below 10% (see section 3.1.1.). Among the remaining participants, accuracy on the trained alternations was similar in the Saltatory condition (95.7%) and in the Control condition (94.2%). Moreover, filler sounds were changed (in error) at a comparably low rate in the Saltatory condition (6.7% changed in error) and in the Control condition (2.7% changed in error).

#### *3.2.2. Trained non-alternating intermediate sounds*

Recall that the primary purpose of Experiment 2 was to determine whether participants would find it difficult to learn that <u>intermediate</u> sounds did not alternate when provided with explicit evidence during training. If so, participants were predicted to incorrectly choose the changing plural option more often for the intermediate sounds in the Saltatory condition than for the comparable sounds in the Control condition. Figure 5 shows how often participants chose the changing plural option (in this case, an incorrect response) for trained non-alternating intermediate sounds (Saltatory condition) and comparable control sounds (Control condition) in Experiment 2, sorted by Exposure Group (Stops sub-group and Fricatives sub-group). Overall, we see that participants in the Saltatory condition made more errors than participants in the Control condition as predicted, both for the Stops sub-group (20.8% vs. 6.7% errors) and the Fricatives sub-group (38.8% vs. 18.3% errors).

To evaluate these differences, a mixed logit model was fitted to predict the log odds of an error (i.e., a changing plural response). The final model included a fixed effect for Condition (Saltatory vs. Control) and a fixed effect for Exposure Group (Stops sub-group vs. Fricatives sub-group). The Condition x Exposure Group interaction was not included because it was not a significant predictor (Wald $z$ = .31, $p$ = .76), and a likelihood ratio test indicated that it did not significantly improve the fit of the model when compared to the simpler model without an interaction, $\chi^2(1) = .09$, $p = .76$. Random intercepts for subjects and for items were included in the model because they significantly improved model fit ($\chi^2(1) = 20.68$, $p < .001$ for subjects, $\chi^2(1) = 128.01$ $p < .001$ for items). Random slopes were not included in the model because the comparisons were fully between-subjects.

The fixed effects for the final model are provided in Table 3. The significant negative intercept reflects the fact that participants had a low rate of errors in the Control condition of the Fricatives sub-group (which is coded as the baseline in this model). Condition was a significant predictor in the model, indicating that participants made significantly more errors (by incorrectly changing intermediate sounds) in the Saltatory condition compared to the Control condition, as predicted. Exposure Group was also a significant predictor, reflecting the fact that overall, participants in the Stops sub-group made fewer errors than participants in the Fricatives sub-group. The lack of a significant Condition x Exposure Group interaction effect indicates that the difference in accuracy between the Saltatory condition and the Control condition holds for both the Stops sub-group and the Fricatives sub-group. Indeed, Condition remains a significant predictor in models run on subsets of the data containing only the Stops sub-group (Estimate = 1.75, Wald *z* = 2.20, *p* = .03) or only the Fricatives sub-group (Estimate = 1.35, Wald *z* = 2.56, *p* = .01).

Overall, the results show that even with explicit evidence, participants found it harder to learn that sounds did not alternate if the sounds were intermediate between alternating sounds (Saltatory condition) than if they were not intermediate (Control group). This main effect was true regardless of whether the intermediate sounds were voiced stops (Stops sub-group) or voiceless fricatives (Fricatives sub-group).

Figure 5. Percent of trial in which the changing plural option was chosen for trained intermediate sounds (Saltatory condition) and comparable non-intermediate sounds (Control condition) in Experiment 2 by sub-group. Individual results (diamonds) and overall means (bars) are provided.

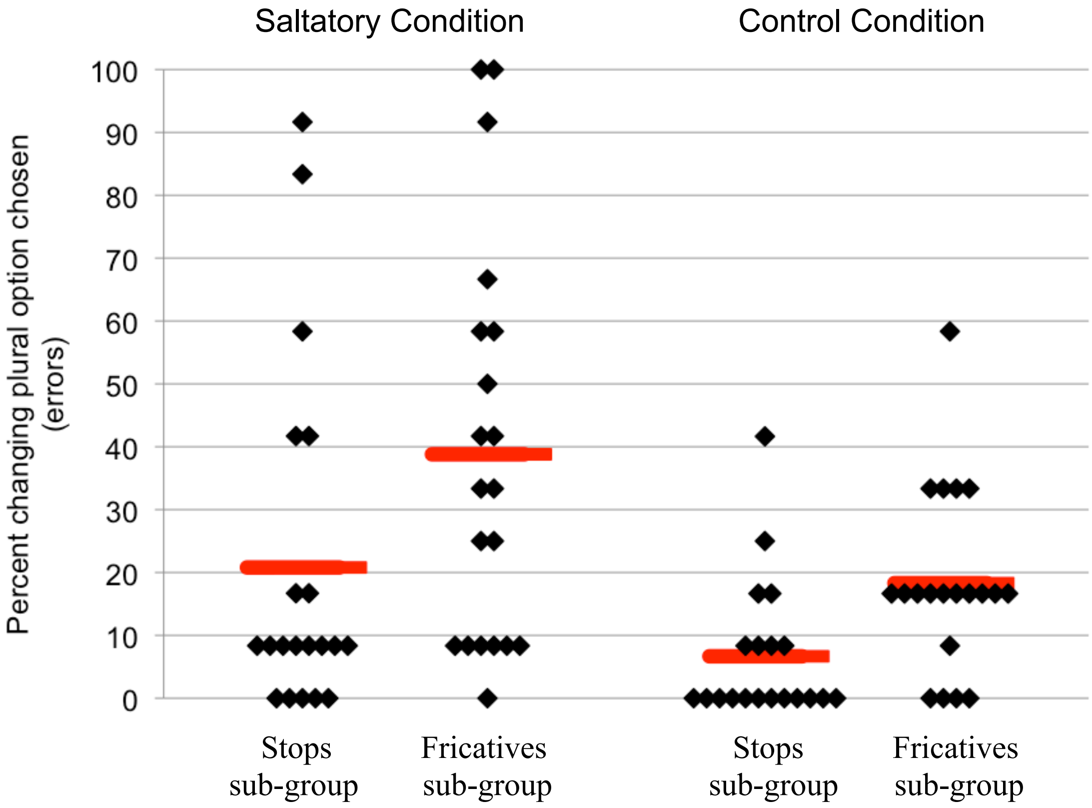


Table 3. Summary of the fixed effects in the final model for trained intermediate sounds (and control sounds) in Experiment 2.

| Predictor | Estimate | Standard error | Wald *z* | *p*-value |
|---|---|---|---|---|
| Intercept | -2.09 | .43 | -4.92 | <.001 |
| Condition = *Saltatory* | 1.49 | .45 | 3.31 | <.001 |
| Exposure group = *Stops sub-group* | -1.57 | .51 | -3.09 | .002 |

### *3.2.3. Effect of amount of exposure*

As in Experiment 1, participants in Experiment 2 received varying amounts of exposure due to the nature of the design. On average, participants in the Saltatory condition completed 1.35 cycles of the exposure phase whereas participants in the Control condition completed 1.65 cycles. To address the possibility that the amount of exposure affected participants' accuracy on intermediate sounds, the logit model in Table 3 was rerun with an added fixed effect for number of cycles through the exposure phase. Amount of exposure was not a significant predictor in the model (Wald $z = -1.28$, $p = .20$), and including the factor in the model did not significantly improve model fit, $\chi^2(1) = 1.58$, $p = .21$. Including amount of exposure along with its interactions with the other fixed effects also failed to significantly improve model fit, $\chi^2(3) = 4.60$, $p = .20$, and none of the new factors reached significance in the model. Thus, we may conclude that amount of exposure did not have a significant effect on how often participants changed intermediate sounds (and comparable sounds in the Control condition) in Experiment 2.

### *3.2.4. Untrained sounds*

A secondary goal for Experiment 2 was to replicate the results for untrained sounds found in Experiment 1. Like in Experiment 1, untrained sounds in Experiment 2 were intermediate between alternating sounds in the Saltatory condition, but were not intermediate in the Control condition, so the pattern of results for untrained sounds should be similar to the one found in Experiment 1. Recall that unlike in Experiment 1, however, each participant in Experiment 2 only had one type of untrained sound, either stops or fricatives, whereas participants in Experiment 1 had both types. Figure 6 presents the percent of trials that participants chose the changing plural option for words ending in untrained sounds in Experiment 2. Participants changed the untrained intermediate sounds more often in the Saltatory condition than in the Control condition both for untrained stops (65.0% vs. 26.9%) and for untrained fricatives (53.2% vs. 21.9%). Within the Saltatory condition, participants also showed a tendency to change untrained stops more often than untrained fricatives (65.0% vs. 53.2%). Overall, we see that the values in Experiment 2 are similar to those in Experiment 1 (compare Figure 6 to Figure 3).

A mixed logit model was run to check for potential differences in how participants treated untrained sounds in Experiment 1 and Experiment 2. The random effect structure included intercepts for subjects and by-subject slopes for Sound Type. The by-subject slopes were included because they significantly improved model fit, $\chi^2(2) = 71.83$ $p < .001$. Adding random intercepts for items did not significantly improve model fit, $\chi^2(1) = 3.74$ $p = .053$.

A first model was run with fixed effects for Condition (Potentially Saltatory/Saltatory vs. Control), Sound Type (untrained stops vs. untrained fricatives), and Experiment (Exp. 1 vs. Exp 2), as well as all interaction effects. Neither the factor for Experiment, nor its associated interaction effects were significant predictors in the model ($p > .16$). Moreover, none of the effects significantly improved model fit according to a likelihood ratio test when removed from the model one at a time using backwards stepwise comparison ($p > .13$), indicating that there were no significant differences between the two experiments.

The fixed effects of the final model (excluding Experiment and its interactions) are summarized in Table 4. The final model included a fixed effect for Condition (Potentially Saltatory/Saltatory vs. Control), a fixed effect for Sound Type (untrained stops vs. untrained fricatives), and a Condition x Sound Type interaction. The significant negative intercept indicates that untrained fricatives in the Control condition (which was coded as the baseline in this model) were changed at a low rate. Condition was a significant predictor in the model, indicating that participants in the Potentially Saltatory/Saltatory condition changed untrained sounds more frequently than those in the Control condition. The significant interaction effect indicates that untrained stops were changed more often than untrained fricatives, but only in the Potentially Saltatory/Saltatory conditions. This model looks very similar to the one conducted for Experiment 1 (Table 2). Overall, the results for untrained sounds in Experiment 2 replicated the basic findings from Experiment 1.

Figure 6. Percent of trials in which the changing plural option was chosen for untrained target sounds according to Condition (Saltatory or Control) and sub-group in Experiment 2. Individual results (diamonds) and overall means (bars) are provided.

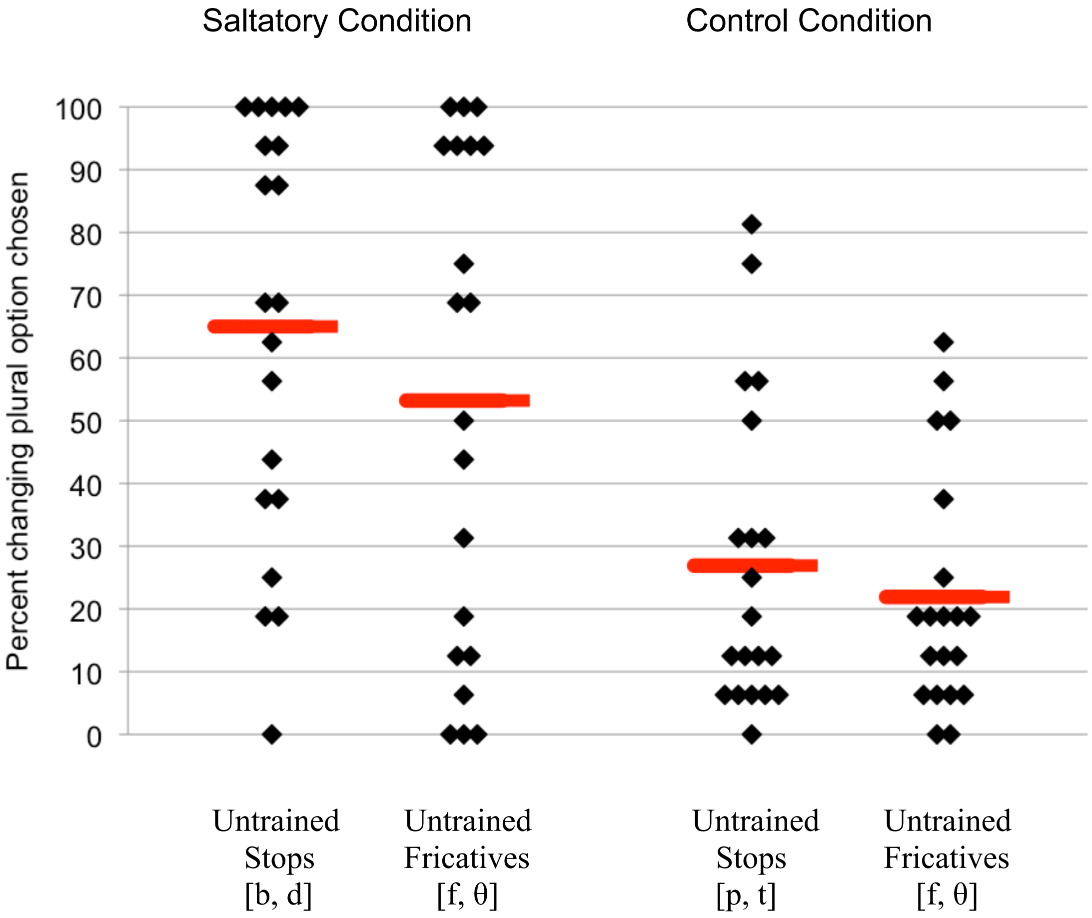


Table 4. Summary of the fixed effects in the final model for untrained sounds in Experiments 1 and 2 combined.

| Predictor | Estimate | Standard error | Wald *z* | *p*-value |
|---|---|---|---|---|
| Intercept | -2.11 | .36 | -5.87 | <.001 |
| Condition = *Potentially Saltatory/Saltatory* | 1.93 | .50 | 3.87 | <.001 |
| Sound Type = *Untrained stops* | .22 | .46 | .47 | .64 |
| Interaction = *Potentially Saltatory/Saltatory & Untrained stops* | 1.50 | .64 | 2.34 | .02 |

### *3.3. Discussion*

Experiment 2 accomplished two main objectives. First, it replicated the primary finding of Experiment 1: untrained sounds were changed more frequently when they were intermediate between alternating sounds (Saltatory condition) than when they were not intermediate (Control condition). Second, Experiment 2 expanded on the findings from Experiment 1 by showing that the desire to change intermediate sounds is strong enough to have an effect even when there is evidence in the input that intermediate sounds do not change. Thus, the anti-saltation bias not only affects how learners interpret ambiguous input, but it also affects how well learners acquire patterns provided explicitly in the input.

Two points related to the experimental design speak further to the robustness of the anti-saltation effect observed in Experiment 2. First, participants had to reach 80% accuracy in the verification phase even though it included the trained intermediate sounds. To advance to the generalization phase, participants had to respond correctly for at least some of the trials with intermediate sounds in the verification phase. Therefore, it is striking that participants still made a significant number of errors on intermediate sounds in the generalization phase.

Relatedly, participants who failed to reach the criterion in the verification phase within 50 minutes did not complete the experiment, and including the intermediate sounds in the verification phase increased the overall difficulty of the task. As a result, the attrition rate in Experiment 2 was fairly high. The attrition has the effect of potentially limiting the representativeness of the sample, but in fact, it does so in a way that is desirable. Those participants who reached the generalization phase (and thus were used in the analysis) consisted of the learners who were best able to learn the new patterns. Yet even among the best learners, there were significantly more errors on the intermediate sounds (Saltatory condition) than on the non-intermediate sounds (Control condition). If anything, excluding intermediate sounds from the verification phase or relaxing the 80% criterion would likely increase the difference in accuracy observed between the Saltatory and Control condition. It is quite possible that several participants failed to complete the experiment because they were unable to learn that intermediate sounds did not alternate, and if so, including them would only enhance the difference. The fact that a significant difference was found using the more restrictive design in the current study implies that the effect is quite robust.

One aspect of the results was unexpected—namely, there were more errors on intermediate fricatives than on intermediate stops. Recall that in Experiment 1, the opposite direction was found: participants changed intermediate stops more often than intermediate fricatives. In the case of Experiment 2, the higher amount of errors for intermediate fricatives is likely due to the orthographic ambiguity in English between [θ] and [ð], both written as <th>. During the post-experiment debriefing, the vast majority of participants reported what they had learned in terms of letters, that is, they would say things like "*p* becomes *v*" rather than making the [p] and [v] sounds. In the Fricatives sub-group of Experiment 2, this strategy would result in confusion at test. If participants exposed to [θ] becomes [θ] during training internalized the generalization as '*th* becomes *th*', they would have no way of knowing at test whether they had learned [θ] becomes [θ] or [θ] becomes [ð] (or potentially even [ð] becomes [θ] or [ð] becomes [ð]). This confusion may have led to more errors on these sounds in the Fricatives sub-group of Experiment 2. In Experiment 1, by contrast, the orthographic ambiguity of [θ] and [ð] was not an issue. In that experiment, [θ] was always an untrained sound, meaning that participants never needed to remember a generalization involving [θ]. At test, there was less of a need to consider the orthography of [θ] and [ð] because participants could compare the phonetic realizations of the two sounds directly.

Support for this explanation comes from the fact that the percentage of errors for [θ] far exceeded the percentage of errors for [f] in the Fricatives sub-group of Experiment 2, both in the Saltatory condition ([θ]: 49% changed vs. [f]: 28% changed) and in the Control condition ([θ]: 26% changed vs. [f]: 11% changed). This large asymmetry in errors between [θ] and [f] was not found in Experiment 1 (Potentially Saltatory condition, [θ]: 41% vs. [f]: 49%; Control condition, [θ] 18% vs. [f]: 14%). This suggests that the large number of errors for intermediate fricatives in Experiment 2 was due, in part, to greater confusion between the dental sounds [θ] and [ð]. Crucially, the saltation avoidance effect (i.e., more errors in the Saltatory condition than in the Control condition) still occurred in the Fricatives sub-group, independent of the greater number of errors that may have resulted from the orthographic ambiguity of [θ] and [ð].

Finally, there are two logically possible ways that participants could exhibit difficulty learning saltatory alternations: (1) they could wrongly change intermediate sounds, or (2) they could fail to properly change the alternating sounds. This experiment focused on the first case—how participants treated intermediate sounds. By ensuring that participants learned the alternations in question, the design did not permit a systematic investigation of how difficult it is to learn the alternation itself, but follow-up experiments could be designed to do so. There is indirect evidence from the current experiments supporting the possibility that the alternations themselves are harder to learn. First, the attrition rate (i.e., the number of participants never making criterion in the verification phase) in the Saltatory condition of Experiment 2 was far greater (n = 13) than the attrition rate in the Potentially Saltatory condition of Experiment 1 (n = 2), implying that the task was much harder with the addition of the non-alternating, intermediate sounds. Part of this difference is likely due to errors on the intermediate sounds themselves, as mentioned, but some of it may also be due to the increased difficulty of learning the alternating sounds.

Second, even among those participants who made it into the generalization phase in Experiment 2, four participants had to be excluded from the Saltatory condition because they apparently "forgot" the alternations that they learned during exposure (i.e., they had less than 10% accuracy on at least one of the alternating sounds in the generalization phase, see section 3.1.1). These observations suggest that the difficulty associated with learning saltatory alternations may not be limited to intermediate sounds (as demonstrated in this study), but may also be reflected in how quickly learners acquire the alternations themselves. A modified version of the current experiment designed to measure speed of acquisition (e.g., Skoruppa et al., 2011) would be useful to corroborate this prediction, but I leave a systematic investigation of this question to future research.

## 4. General discussion

The purpose of this study was to determine experimentally whether adults have a learning bias that disfavors saltatory phonological alternations. The most striking aspect of the results was that participants extended learned alternations to untrained intermediate sounds (but not to comparable non-intermediate sounds) without evidence in the input (Exp. 1), and in some cases *contrary* to evidence in the input (Exp. 2). To summarize, in Experiment 1 learners changed untrained sounds much more frequently when doing so would avoid a saltatory alternation (Potentially Saltatory condition) than when saltation was not at stake (Control condition). Experiment 2 showed that even with explicit training, participants had greater difficulty learning that sounds did not change if they were intermediate between two alternating sounds (Saltatory condition) than if they were not intermediate (Control condition). Taken together, the results from these experiments provide strong evidence that people learn novel alternations in a way that disfavors saltatory alternations. This study adds to a growing body of literature showing that phonological learning is constrained by biases (Saffran & Thiessen, 2003; Wilson, 2006; Zuraw, 2007; Finley & Badecker, 2008; Moreton, 2008; Hayes et al., 2009; Becker et al., 2011; Skoruppa, Lambrechts, & Peperkamp, 2011; Skoruppa & Peperkamp, 2011; Baer-Henney & van de Vijver, 2012; Becker et al., 2012; Finley & Badecker, 2012; Hayes & White, 2013).

Given what we know about saltatory alternations, our theory of phonological learning must be able to account for two facts: (1) saltatory alternations are attested in natural languages (e.g., Campidanian Sardinian) and therefore must be learnable, and (2) they are dispreferred during learning, as demonstrated in the current study. The results from these experiments raise several questions about how people learn phonological alternations. Which mechanism is responsible for the anti-saltation bias observed here? Can it be derived from some principle that is more general? Which type of learning model can predict that saltatory alternations are dispreferred, but learnable? I consider these issues below.

One way to deal with the problem of saltatory alternations is by implementing a hard bias, that is, by ruling them out altogether, as in Peperkamp et al. (2006a). They implemented a computational model that learned which sounds were context-dependent variants of other sounds by looking for complementary distributions. Two sounds were considered to be allophonic variants if they rarely occurred in the same phonological environment, that is, if their distributions in the input had little or no overlap. To prevent the model from learning spurious pairings (e.g., between [h] and [ŋ] in English, as discussed in the Introduction), Peperkamp et al. equipped the basic statistical model with two linguistic filters. One of these filters prohibited mappings between two sounds if an intermediate sound existed between them, where an intermediate sound was defined similarly to the way it is defined here. The filter, which effectively banned saltatory alternations, improved the model's performance because it was successful at excluding spurious mappings. However, it would also prevent the model from learning that [t] and [ð] alternate in Campidanian Sardinian due to the presence of intermediate [d]. That conclusion would, of course, be incorrect in the case of Campidanian Sardinian. Given the existence of languages with saltatory alternations, it must be possible for a child to learn them; thus, an absolute ban on saltatory alternations, like the one proposed by Peperkamp et al., is not tenable.

To account for the fact that saltatory alternations are dispreferred during learning but not unlearnable, we most likely need a soft bias. A soft bias is typically implemented in models using a prior, which makes certain outcomes, within the space of possible outcomes, have higher *a priori* likelihoods relative to others. The idea that such soft biases have a role in phonological learning has been growing in the literature (e.g., see Wilson, 2006; Zuraw, 2007; Finley & Badecker, 2008; Moreton, 2008; Hayes et al., 2009; Hayes & White, 2013). Implementing a fully functioning learning model to account for these results is beyond the scope of this paper. Instead, I will provide a brief sketch of some components that such a model might require.

One straightforward way to proceed would be to modify the filter banning saltations proposed in Peperkamp et al.'s statistical model so that it applies as a soft bias rather than an outright restriction. We could update the filter so that two sounds with complementary distributions are less likely to be considered allophonic variants if there is an intermediate sound between them. The model might then be able to learn saltatory alternations, but doing so would require more input than non-saltatory cases. However, this solution requires positing a mechanism devoted specifically to preventing saltatory alternations. It may be possible to account for the anti-saltation bias in a more parsimonious way by relying on principles that have been proposed for independent reasons, only resorting to new stipulations if necessary. Below, I consider two possibilities proposed in the literature that may also be capable of accounting for the behavior observed in the current study: a similarity bias and an anti-complexity bias.

### *4.1. Similarity bias*

As mentioned in the Introduction, it has long been noted that alternating sounds tend to be highly similar, and alternations between dissimilar sounds are less common than those between similar sounds (Trubetzkoy, 1939/1969). Recall that a saltatory alternation is a particularly striking counterexample to the principle of minimal modification, proposed by Steriade (2001/2008) as an explanation for why alternations between dissimilar sounds are uncommon. Skoruppa, Lambrechts, and Peperkamp (2011) provided experimental evidence that language learners are sensitive to the similarity between sounds when learning novel alternations in an artificial language. In their study, adults were able to learn novel alternations between sounds differing in a single feature (e.g., [p ~ t]) more easily than sounds differing in two or more features (e.g., [p ~ z]). These results are straightforwardly predicted by a similarity bias based on the principle of minimal modification expressed in the P-map (Steriade, 2001/2008). If learners are biased to prefer alternations between similar sounds, they should find [p ~ t] easier to learn than [p ~ z].

The current study did not focus on the similarity of the alternating sounds directly, but rather on how learners treated sounds that were intermediate between alternating sounds. Still, it is plausible that the same mechanism (i.e., a similarity bias) could explain why participants learning [p ~ v] were biased to assume [b ~ v]. Conceptually, the similarity account would work as follows. By learning that voiceless stops [p, t] changed to voiced fricatives [v, ð] between vowels, participants had evidence for a restriction on stops between vowels. When faced with voiced stops [b, d] at test, participants noticed that they also violated the restriction on stops between vowels; moreover, voiced stops are more similar to voiced fricatives than voiceless stops are, resulting in a tendency to change the voiced stops. Intuitively, if a phonological constraint (e.g, no stops between vowels) warrants alternations between dissimilar sounds, then it should also warrant alternations between more similar sounds.[9] In the Control condition, extending the alternation to untrained sounds is correctly predicted to occur less frequently because the same logic does not hold in the reverse direction, that is, evidence for an alternation between highly similar sounds does not necessarily warrant an alternation between less similar sounds.

How could we formally implement a similarity bias to test the hypothesis? Steriade's P-map was originally conceived within the phonological framework of Optimality Theory (Prince & Smolensky, 1993/2004), but its basic principles can be extrapolated to other models of phonological learning. Indeed, Wilson (2006) implemented a computational version of the P-map as a way to explain why learners extended novel alternations involving palatalization more often to contexts where it would be less perceptibly noticeable (before high vowels) compared to contexts where it would be more noticeable (before mid vowels). Using maximum entropy grammar models (see also Goldwater & Johnson, 2003; Hayes & Wilson, 2008), Wilson implemented the similarity bias as a prior. A similar approach could be taken to account for the results of the current study (as well as Skoruppa et al.'s (2011) results). In effect, alternations

---

[9] The same logic would apply to the voiceless fricatives [f, θ], except the relevant phonological constraint would be a ban on voiceless sounds between vowels.

between highly similar sounds would be assigned greater *a priori* likelihoods than alternations between less similar sounds. In Wilson's (2006) terms, this bias would be a type of *substantive bias* because it is based on prior (unconscious) knowledge of phonetic principles.

Wilson's approach has two further advantages in relation to the current study's results: the type of learning mechanism and the similarity metric. First, the learning model needs to account for the fact that saltatory alternations are initially dispreferred (or difficult to learn), but ultimately learnable. The general mechanism of weighting constraints according to the principle of maximum entropy, in a way that is constrained by a prior, is particularly well suited for deriving the type of learning behavior desired. The prior makes some outcomes more likely early in the learning process, but given enough input, the prior can ultimately be overcome.

Second, the results from the current study imply that perceptual similarity, rather than number of phonological features, is more appropriate as a measure of similarity for the purposes of this bias. Recall that in Experiment 1, participants in the Potentially Saltatory condition preferred to change untrained voiced stops [b, d] more often than untrained voiceless fricatives [f, θ] (70% changed vs. 45% changed), even though both types of sounds were intermediate. This asymmetry was also found in Experiment 2 (65% of untrained [b, d] changed vs. 53% of untrained [f, θ] changed).[10] To capture this preference, we cannot rely only on counting phonological features because both types of sounds differ from the voiced fricatives [v, ð] by the same number of features (i.e., one feature).

The preference for changing voiced stops more than voiceless fricatives is, however, consistent with the relative differences between the sounds in terms of perceptual similarity.

---

[10] The trained intermediate sounds in Experiment 2 seem to counter this argument because participants exhibited the opposite pattern—they made more errors on intermediate fricatives than intermediate stops (i.e., by changing them incorrectly). However, the high number of errors for fricatives in this case was likely due to the orthographic ambiguity of [θ] and [ð] (see section 3.3).

Table 5 provides the mutual confusability values for the untrained sounds from the Potentially Saltatory condition in Experiment 1 (calculated from Wang & Bilger 1973, as described in section 1.2) along with how frequently participants chose the changing plural option for each untrained sound. We see that the voiced fricatives [v, ð] are more confusable with the voiced stops [b, d] than with the voiceless fricatives [f, θ]. Thus, the preference to change voiced stops more than voiceless fricatives is predicted under the perceptual similarity account. It is also possible that phonological features would be sufficient as a similarity metric under a theory in which features can vary in their degree of perceptual salience (e.g., Stevens & Keyser, 1989; Clements, 2005). Further experimental and computational research would be necessary to evaluate the subtly different predictions of using raw perceptual similarity versus features with different degrees of perceptual salience.

In sum, the predictions of the perceptual similarity account appear to be on the right track. I leave it to future work to test the hypothesis in more detail using a fully implemented model, but preliminary computational work in this area looks promising (see White, 2012, 2013).

Table 5. Mutual confusability of untrained sounds and the corresponding voiced fricatives from the Potentially Saltatory condition in Experiment 1, with how frequently participants chose the changing plural option for each untrained sound. Higher mutual confusability implies higher similarity.

| Sound pair | Mutual confusability | % changing option chosen |
|---|---|---|
| [b] and [v] | .117 | 73.3 |
| [d] and [ð] | .083 | 66.7 |
| [f] and [v] | .033 | 49.3 |
| [θ] and [ð] | .038 | 41.1 |

### *4.2. Anti-complexity bias*

Another possible type of bias that could be relevant to the results of this study is an anti-complexity bias. There is extensive evidence that humans prefer simple solutions to more complex solutions within a number of cognitive domains, and indeed, the results of many artificial grammar studies showing differences in phonological learning can potentially be attributed to a preference for simpler patterns (for a review, see Moreton & Pater, 2012). Particularly strong examples of an anti-complexity bias come from studies showing that phonological patterns are learned more easily if they target classes of sounds that can be characterized using few features (often called "natural classes") than if the patterns refer to arbitrary classes of sounds that require several features or groups of features to describe. For example, Saffran and Thiessen (2003) showed that 9-month-old infants learned novel restrictions on which segments could appear in certain syllable positions when they involved natural classes (e.g., only voiceless stops [p, t, k] appear in syllable-initial position) but not when the targeted group of sounds was arbitrary (e.g., only [p, d, k], which includes two voiceless stops and one voiced stop). Likewise, Skoruppa and Peperkamp (2011) found that adult French speakers were able to adapt to a new "dialect" of French containing either vowel harmony (e.g., rounded vowels occur after rounded vowels) or vowel disharmony (e.g., unrounded vowels occur after rounded vowels), both of which target natural classes of sounds. However, they failed when the new "dialect" involved arbitrary vowel pairings that were not easily described in general terms (i.e., some rounded vowels resulted in harmony, some in disharmony).

Returning to the current study, neither of the alternations in the Potentially Saltatory condition or the Control condition in Experiment 1 is obviously more complex than the other at the segmental level (and likewise in Experiment 2). Each has one pair of segments alternating with one other pair of segments in a parallel way: [p, t] → [v, ð] vs. [b, d] → [v, ð]. However, if learners track the number of phonological features that change (as proposed, e.g., in Chomsky & Halle, 1968) rather than just the ones needed to target the class of sounds undergoing the change, the avoidance of saltatory alternations seen in Experiment 1 might be viewed as a sort of complexity avoidance, rather than being due to phonetic similarity per se.[11] This approach requires counting the number of features involved in each of the phonological rules; as a reference, I provide the rules and analysis in the Appendix.

However, even if the general anti-saltation bias could be explained by counting features, it is unclear how the anti-complexity account would explain the preference to change voiced stops more often than voiceless fricatives, observed for untrained sounds in Experiment 1 and Experiment 2. The feature-counting approach predicts no difference in these sounds (or perhaps the opposite difference, see fn. 8), implying that we would still need some type of bias based on perceptual similarity to fully account for the experimental data.

## 5. Conclusions

This study has provided experimental evidence that adults are biased against learning saltatory alternations when learning an artificial language, adding to the growing body of literature showing that language learners display biases against certain phonological patterns. Given these results, models of phonological learning must be able to account for the dispreferred status, yet ultimate learnability of saltatory alternations. Augmenting models of phonological learning with a substantive bias based on the principle of minimal modification, that is, one that assigns greater prior likelihoods to alternations between sounds with greater perceptual similarity, appears promising as a way to account for the data observed in this study.

Some issues remain for future work. First, these studies used an explicit two-alternative forced-choice task to test participants' learning. The advantage of this task is that the data analysis is straightforward, but it has two potential downsides: (a) participants may wish to

---

[11] Thanks to a reviewer for pointing out this possibility.

provide answers that are not among the response options, and (b) participants may not have considered one or both of the response alternatives if they had not heard them as response options. It would thus be interesting to expand on this study using a production task so that participants are free to provide any response.

Finally, the results from these experiments provide strong evidence that adults have a bias against saltatory alternations, but ultimately we are most interested in how children acquire language. Like any artificial language study with adult participants, this work faces the limitation that adults may use strategies in the experimental task that are not available to children during acquisition of a first language. In particular, adults may bring native language knowledge to bear on the task or they may use non-linguistic problem solving strategies. It is worth noting that English does not have the alternations tested here, with the exception of the marginally productive fricative voicing rules for plurals (e.g., *half* [hæf] ~ *halves* [hævz]) and noun-verb pairs (e.g., *bath* [bæθ] ~ *bathe* [beɪð]). Despite this marginal evidence for a fricative voicing alternation in English, however, participants in Experiment 1 actually preferred turning stops into fricatives more than voicing fricatives. Still, we cannot be sure that participants were not bringing native language knowledge to the task in some form. There is also no way to ensure that participants were employing linguistic mechanisms rather than non-linguistic problem solving strategies (although we would still be left explaining why these strategies resulted in an anti-saltation bias).

For these reasons, an important task for future research is to test for an anti-saltation bias in infants who are just beginning to learn alternations. White et al. (2008) showed that 12-month-old infants can learn novel alternations after brief exposure to an artificial language. Moreover, White and Morgan (2008) showed that 19-month-olds displayed a gradient response to 1-, 2-, and 3-feature mispronunciations in an object recognition task, indicating that infants develop a fine-tuned sensitivity to the similarity of speech sounds early in life. Thus, future work should investigate whether infants are subject to an anti-saltation bias like adult learners (for preliminary work testing infants, see White & Sundara, 2012).

## 6. Acknowledgements

I would like to sincerely thank Bruce Hayes, Megha Sundara, Robert Daland, Kie Zuraw, Sharon Peperkamp, Adam Albright, Jennifer Cole, Karen Campbell, Elliott Moreton, and two anonymous reviewers for comments and discussion that has improved this work in many ways. I would also like to thank audiences at the UCLA phonology seminar and the 2012 Meeting of the Linguistic Society of America for helpful discussion, Marc Garellek for recording the stimuli, and my undergraduate research assistants, Kelly Ryan, Kelly Stroh, Kelly Nakawatase, Ariel Quist, and Emily Ho for their assistance with various aspects of this project. Any remaining errors are my own. This work was funded in part by a UCLA Graduate Research Mentorship fellowship to the author and by NSF grant BCS-0951639 (to Megha Sundara).

## 7. References

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*(2-3), 155-182.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, MA: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Baer-Henney, D., & van de Vijver, R. (2012). On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology, 3*(2), 221–250.

Bates, D., Maechler, M., and Dai, B. (2008). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28. http://lme4.r-forge.r-project.org/.

Becker, M., Ketrez, N., & Nevins, A. (2011). The surfeit of the stimulus: Grammatical biases filter lexical statistics in Turkish voicing deneutralization. *Language, 87*(1), 84-125.

Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language, 88*(2), 231–268.

Bevington, G. L. (1974). *Albanian Phonology.* Otto Harrassowitz: Wiesbaden.

Boersma, P. & Weenink, D. (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.44, retrieved 20 October 2010 from http://www.praat.org/.

Bolognesi, R. (1998). *The phonology of Campidanian Sardinian: A unitary account of self-organizing structure.* The Hague: Holland Academics.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58*(2), 265-289.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition, 87*(2), 69-77.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English.* New York: Harper & Row.

Clements, G. N. (2005). The role of features in speech sound inventories. In E. Raimy & C. Cairns (Eds.), *Contemporary Views on Architecture and Representations in Phonological Theory.* Cambridge: MIT Press.

Cristia, A., Daland, R., Mielke J., & Peperkamp S. (in press). Constrained generalization of implicitly learned sound patterns. *Journal of Laboratory Phonology*.

Finley, S., & Badecker, W. (2008). Substantive biases for vowel harmony languages. In J. Bishop (Ed.), *Proceedings of the West Coast Conference on Formal Lingusitics 27,* 168-176.

Finley, S., & Badecker, W. (2012). Learning biases for vowel height harmony. *Journal of Cognitive Science, 13,* 287–327.

Goldwater, S. & Johnson, M. (2003). Learning OT constraint rankings using a Maximum Entropy model. *Proceedings of the Workshop on Variation within Optimality Theory.* Stockholm University.

Hayes, B. (2009). *Introductory Phonology.* West Sussex: Wiley-Blackwell.

Hayes, B., Zuraw, K., Siptár, P., & Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language, 85,* 822-863.

Hayes, B., & White, J. (2013). Phonological naturalness and phonotactic learning. *Linguistic Inquiry, 44,* 45–75.

Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry, 39*(3), 379-440.

Ito, J., & Mester, A. (2003). On the sources of opacity in OT: Coda processes in German. In C. Féry & R. van de Vijver (Eds.), *The Syllable in Optimality Theory* (pp. 271-303)*,* Cambridge: Cambrige Univ. Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434-446.

Kirchner, R. (1998). *An Effort-Based Approach to Consonant Lenition.* Ph.D. Dissertation, UCLA.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245-294.

Lass, R. (1997). *Historical Linguistics and Language Change.* Cambridge: Cambridge Univ. Press.

Lubowicz, A. (2002). Derived environment effects in Optimality Theory. *Lingua, 112,* 243-280.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cogntion, 82*(3), B101-B111.

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua, 122*(2), 145-163.

Miller, G., & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*(2), 338-352.

Minkova, D. (1991). On leapfrogging in historical phonology. In J. van Maarle (Ed.), *Historical Linguistics 1991: Papers from the 10th International Conference on Historical Linguistics.* Amsterdam: John Benjamins.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology, 25*(1), 83-127.

Moreton, E., & Pater, J. (2012). Structure and substance in artificial-phonology learning. *Language and Linguistics Compass, 6*(11), 686-718.

Peperkamp, S., & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology 9* (pp. 315-338). Berlin: Mouton de Gruyter.

Peperkamp, S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006a). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition, 101*(3), B31-B41.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006b). The role of phonetic naturalness in phonological rule acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 464-475). Somerville, MA: Cascadilla Press.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar.* Cambridge: Blackwell. (Published in 2004. First circulated in 1993.)

R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R- project.org.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental psychology, 39*(3), 484-494.

Singh, S. & Black, J. W. (1966). Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *The Journal of the Acoustical Society of America, 39*(2), 372-387.

Skoruppa, K., Lambrechts, A., & Peperkamp, S. (2011). The role of phonetic distance in the acquisition of phonological alternations. In S. Lima, K. Mullin, & B. Smith (Eds.), *Proceedings of the 39th Annual Meeting of the North Eastern Linguistic Society* (pp. 717-729). Somerville, MA: Cascadilla Press.

Skoruppa, K. & Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, 35, 348-366.

Steriade, D. (2001/2008). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. In S. Inkelas & K. Hanson (Eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky* (pp. 151-180). Cambridge: MIT Press. (Published in 2008. Originally circulated as ms. in 2001.)

Stevens, K. N., & Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language, 65*(1), 81–106.

Trubetzkoy, N. S. (1939/1969). *Principles of Phonology*. (C. A. M. Baltaxe, Trans.). Los Angeles: Univ. of California Press. (Original work published 1939.)

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *Journal of the Acoustical Society of America, 54*(5), 1248-1266.

Westbury, J., & Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*, 145-166.

White, J. (2012). On the learnability of saltatory phonological alternations. Poster presented at the 13th Meeting of the Association for Laboratory Phonology. Stuttgart, Germany.

White, J. (2013). *Bias in Phonological Learning: Evidence from Saltation*. Ph.D. dissertation, UCLA.

White, J. & Sundara, M. (2012). Infant biases in the learning of phonological alternations. Talk presented at the 37th Boston University Conference on Language Development. Boston, MA.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59*, 114–132.

White, K. S., Peperkamp, S., Kirk, C., & Morgan J. L. (2008). Rapid acquisition of phonological alternations by infants. *Cognition, 107*(1), 238-265.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science, 30,* 945-982.

Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language, 83*(2), 277-316.

## 8. Appendix: Anti-complexity analysis, using feature counting.

Under the feature-counting analysis, participants in Experiment 1 were considering (at least) two possible rules in each of the conditions, given below (in the general style of Chomsky & Halle, 1968). In the Potentially Saltatory condition, the more general rule (1) requires fewer total features whereas the more specific rule (2) requires two extra features. Thus, participants may have preferred the general rule, leading them to change intermediate sounds. In the Control condition, participants also had the option of a general rule (3) and a specific rule (4), but in this case the two rules involve an equal number of features. Thus participants may have been less inclined to learn the general rule over the specific rule. However, this analysis cannot account for the difference observed between voiced stops and voiceless fricatives.

<u>Potentially Saltatory condition</u>:

(1) General rule (fewer features):

$$[-\text{sonorant}] \longrightarrow \begin{bmatrix}+\text{continuant}\\ +\text{voice}\end{bmatrix} \Big/ \quad \text{V}\,\_\_\_\,\text{V}$$

Effect (for labials): "{p, f, b} become [v] between vowels."

(2) Specific rule (more features):

$$\begin{bmatrix}-\text{sonorant}\\ -\text{continuant}\\ -\text{voice}\end{bmatrix} \longrightarrow \begin{bmatrix}+\text{continuant}\\ +\text{voice}\end{bmatrix} \Big/ \quad \text{V}\,\_\_\_\,\text{V}$$

Effect (for labials): "{p} becomes [v] between vowels."

<u>Control condition</u>:

(3) General rule:

$$[-\text{sonorant}] \longrightarrow \begin{bmatrix}+\text{continuant}\\ +\text{voice}\end{bmatrix} \Big/ \quad \text{V}\,\_\_\_\,\text{V}$$

Effect (for labials): "{p, f, b} become [v] between vowels."

(4) Specific rule:

$$\begin{bmatrix}-\text{continuant}\\ +\text{voice}\end{bmatrix} \longrightarrow [+\text{continuant}] \Big/ \quad \text{V}\,\_\_\_\,\text{V}$$

Effect (for labials): "{b} becomes [v] between vowels."