Huang, Tingyu & Do, Youngah. 2023. Substantive bias and variation in the acquisition of vowel harmony. *Glossa: a journal of general linguistics* 8(1). pp. 1–28. DOI: https://doi.org/10.16995/glossa.9313

Open Library of Humanities

Substantive bias and variation in the acquisition of vowel harmony

Tingyu Huang, Sichuan University, CN, huangty@connect.hku.hk

Youngah Do, The University of Hong Kong, HK, youngah@hku.hk

This study investigates substantive bias, a phenomenon wherein learners exhibit a preference for phonetically motivated patterns in language acquisition. The study presents evidence that variable input, rather than categorical input, can activate substantive bias. Native speakers of Hong Kong Cantonese were randomly assigned to categorical or variable training conditions for vowel backness harmony or disharmony, or to a no-training control condition. Results indicate that participants in the categorical and control conditions did not exhibit a bias towards either pattern. However, participants in the variable conditions displayed a bias towards vowel harmony, suggesting that input variability can strengthen the effect of substantive bias. These findings contribute to our comprehension of the role of input variability in phonological learning and the mechanisms involved in the acquisition of phonetically motivated and unmotivated phonological patterns.

1 Introduction

The relationship between phonological typology and synchronic learning biases has been a topic of debate in phonological theory. According to the hypothesis on phonological learning biases posited by Moreton (2008), learners have a preference towards the acquisition of certain phonological patterns over others. Two extensively researched learning biases in phonology are structural bias, which is a bias against structurally complex patterns, and substantive bias, which favors phonetically natural or motivated patterns (Moreton & Pater 2012a; b). Artificial language learning (ALL) experiments have been a common approach to test learning biases, and studies reviewed in this article tested learning biases using ALL experiments, unless otherwise specified. The experimental evidence supporting the structural bias hypothesis is strong. A wide range of phonological learning experiments support this hypothesis, demonstrating, for instance, that a pattern involving a single feature is easier to learn than those involving multiple features (Saffran & Thiessen 2003; Peperkamp et al. 2006; Cristià & Seidl 2008; Kuo 2009; Skoruppa et al. 2009; Chambers et al. 2010). On the other hand, the experimental evidence for substantive bias, as summarized by Glewwe (2019), has yielded mixed results. Some studies have provided compelling evidence for substantive bias, demonstrating that participants are more inclined to learn and generalize natural patterns over unnatural ones, towards broader natural classes (Zhang & Lai 2010; Finley 2012; Do 2013; White 2014; van de Vijver & Baer-Henney 2014). However, other studies have produced mixed results, indicating that substantive bias only operates under specific learning conditions (Wilson 2006; see section 1.1.1 for the details). Yet other studies have discovered no difference in the learning of natural and unnatural patterns (Seidl & Buckley 2005; Lysvik 2018), or even better learning of unnatural patterns (Albright & Do 2017; Glewwe 2019). Additionally, in certain studies where evidence for substantive bias was observed, the results could be attributed to structural bias, since substantively-grounded patterns are often structurally simple as well (Peperkamp et al. 2006; Skoruppa et al. 2009). Moreton and Pater (2012b) attributed the mixed experimental evidence for substantive bias to the weak role it plays in phonological learning. However, there has yet to be a consensus on the reasons for the relative weakness of substantive bias.

The aim of this study is to explore why the effect of substantive bias has been weak or inconsistent in phonological learning experiments. Previous research has proposed several possible factors. Baer-Henney et al. (2015) investigated the relative roles of L1 and substantive bias and discovered that the effect of substantive bias was stronger under conditions of greater input variability and less training. Martin (2017) examined the learning of vowel (dis)harmony patterns and found that participants learned harmony more effectively than disharmony only when the input was inconsistent. Lysvik (2020) compared the learning of categorical final devoicing with final voicing and discovered no significant difference between experimental conditions. However, the author observed a high degree of individual variation in learning phonetically grounded versus ungrounded patterns, which may explain the mixed support for the substantive bias hypothesis. Do & Havenhill

(2021) demonstrated that the effect of substantive bias is sensitive to the methodological choices of artificial language learning paradigms. They found that learning a phonological pattern was substantively biased only when the pattern was presented as variable.

A common thread throughout these studies is the prevalence of *uncertainty* in input patterns: Martin (2017) and Do & Havenhill (2021) found the effect of substantive bias when the pattern was presented with variability, a learning condition that had a higher degree of uncertainty, but not when the pattern was presented without exceptions, a case where the learning condition was more certain. The significant degree of individual variability noted by Lysvik (2020) also suggests greater uncertainty in predicting global tendencies regarding substantively biased learning. Baer-Henney et al. (2015) suggest that when there is more uncertainty in the input, either in the form of variability or when participants were provided with less training, such uncertainty could increase the effect of substantive bias. Despite the observation that substantive bias appears sensitive to the (un)certainty of input data, this has never been a direct research question in the study of phonological learning biases.

In this study, we aim to investigate the effect of substantive bias under conditions of uncertainty by examining the learning of variable versus categorical phonological patterns, while controlling for other factors. Due to their probabilistic nature, variable phonological patterns are expected to provide greater uncertainty to learners compared to absolute and categorical phonological patterns. Based on the findings from previous studies reviewed above and further reviewed in section 1.2.2, we predict that the effect of substantive bias will be more pronounced in variable learning conditions. Conversely, we predict that the effect of substantive bias will be undetectable, or at least weaker, in categorical learning conditions. To test our hypothesis regarding the potential impact of uncertainty on substantive bias, we examine the learning of categorical and variable vowel (dis)harmony patterns by native Cantonese speakers from Hong Kong.

Vowel harmony is a form of phonological assimilation whereby vowels within a given phonological domain must conform to a specific phonetically based feature value, such as front and back or round and unround (Clements 1976). For instance, in Hungarian, the dative suffix has two allomorphs that depend on the backness of the stem vowel. When the stem vowel is front, the suffix appears as [nɛk], as in [ɛmbɛr-nɛk] 'person-DAT', whereas, when the stem vowel is back, the suffix takes the form [nɔk], as in [ɔblɔk-nɔk] 'window-DAT' (Hayes & Londe 2006; Martin & Peperkamp 2020). Vowel harmony is selected as the target pattern of the current study because it has a clear phonetic basis (Gordon 2016; Kaun 2004) and is widespread in natural languages (Gafos & Dye 2011). Furthermore, numerous experimental studies have shown that adult participants can acquire a vowel harmony pattern with minimal exposure to the pattern (Pycha et al. 2003; Moreton 2008; Finley & Badecker 2009; Skoruppa & Peperkamp 2011; White et al. 2018). Additionally, Hong Kong Cantonese lacks harmony patterns for both vowels and consonants, which may reduce the influence of L1 in the present study.

We conducted a study involving the design of six artificial languages with variations in phonetic grounding and pattern variability. The aim was to explore the influence of substantive bias on variable learning. Previous research on variable learning and substantive bias has primarily utilized production practice during either the exposure or test phase, as discussed below in Section 1.2.2. It has been suggested that production tasks may be more effective in revealing learning biases, likely due to their higher cognitive demand (Peperkamp et al. 2006). Consequently, it can be challenging to differentiate between in-experiment production and variable input as factors that may elicit substantive bias. To address this issue, our study employed a perception-only task to investigate whether input variability alone can activate substantive bias with a task that is less cognitively demanding. As anticipated, the results provide evidence that substantive bias emerges only when the alternation patterns are presented with variability. These findings validate that variable input, in the absence of a production task, can indeed trigger substantive bias.

Section 1.1 will provide an overview of the phonetic motivation of vowel harmony. Section 1.2 will review previous research on substantive bias in both categorical and variable training conditions, followed by our hypotheses and predictions. Section 2 will describe the experimental methodology used in this study. Section 3 will present the results and analyses, and section 4 will discuss the findings and provide a conclusion.

1.1 Phonetic grounding of vowel backness harmony

Vowel harmony is a prevalent linguistic phenomenon, as demonstrated by its high frequency across different language families, including the Uralic (e.g., Hungarian and Finnish) and Altaic (e.g., Turkish and Mongolian) families (Ohala 1994). According to the World Atlas of Language Structures (WALS) survey, 26 out of 100 languages exhibit vowel harmony, while only 2 out of 100 languages display vowel disharmony (Gordon 2016): Ainu represents an example of vowel backness disharmony (Ito 1984), and Maskelynes presents a vowel height disharmony pattern (Lynch 2003). Vowel harmony arises from the ease of articulating sounds that share certain features, and it results from vowel-to-vowel assimilation caused by coarticulation between adjacent sounds. Ohala (1994) proposed that misperception resulting from coarticulation leads to phonological reanalysis, and this idea has been supported by a modeling study (Mailhot 2013). Ohala's (1994) approach assumes that the misparsing of the speech signal caused by misperception drives vowel harmony, as listeners may attribute too little or too much of the acoustic signal of a target sound to coarticulatory context, resulting in the phonologization of vowel-to-vowel coarticulation. Conversely, vowel disharmony is relatively uncommon. Although disharmony increases the perceptual distinctiveness of sounds (Wilson 2003), it is less phonetically-based compared to vowel harmony.

Vowel harmony and disharmony patterns between two adjacent syllables can be compared in terms of their structural complexity (Moreton & Pater 2012b; Martin & White 2021). Some may argue that disharmony is more complex because it requires a negative operator to describe it.

Psychological research has indeed shown that the concept of *different* may be more challenging to learn than the concept of *same* (Hochmann et al. 2018). However, in constraint-based phonology, it is the harmony pattern, not the disharmony pattern, that is formalized with a negative operator (e.g., harmony: *[αF][-αF]). Therefore, general vowel harmony and disharmony rules have been considered equivalent in terms of complexity, making them an appropriate test case for the substantive bias hypothesis (Martin & Peperkamp 2020).

1.2 Substantive bias in categorical vs. variable training conditions

This section provides a review of artificial language learning studies that investigate the presence of substantive bias in the learning of categorical and variable patterns, respectively.

1.2.1 Substantive bias in learning categorical phonology

The majority of artificial language learning studies on substantive bias have focused on the acquisition of categorical patterns, where all critical items undergo the target phonological processes. Some of these studies have reported convincing evidence for substantive bias (Hayes et al. 2009; Finley 2012; Shapp 2012; White 2014). For instance, in Finley (2012), participants learned a rounding harmony pattern where stem vowels triggered rounding harmony on the suffix vowel (e.g., [kini-mi], [tudu-mu]). In one condition, the stem vowels were exclusively mid vowels, while in the other, they were high vowels. Mid vowels are a better trigger of rounding harmony because the round feature of high vowels is easier to perceive than mid vowels, and high vowels are produced with more rounding based on measurements of horizontal opening and lip opening (Linker 1982). Therefore, rounding harmony enhances the perceptibility of rounding features for mid vowels more than for high vowels. The findings indicated that participants in the mid vowel condition learned the rounding harmony better than those in the high vowel condition, supporting the perceptual motivation of the effect of substantive bias.

Martin & Peperkamp (2020) conducted a study on the learning of categorical vowel backness (dis)harmony and the effect of sleep on learning, providing clear evidence for substantive bias in vowel (dis)harmony learning. The study included 173 North American English speakers who were exposed to 96 CVCV stems containing French phonemes and a plural suffix, either [-tɛl] or [-tɔl], with half of the stems containing two front vowels and the other half containing two back vowels. During exposure, participants were required to click on the written form of TEL or TOL that corresponded to the plural form they heard. In the test, participants chose between the two possible plural forms (i.e., CVCV-[tɛl] or CVCV- [tɔl]). Vowel harmony was learned better than disharmony, and performance remained stable after 12 hours, with or without sleep. Martin & White (2021) also found a substantive bias effect in the generalization of categorical vowel (dis)harmony. They exposed 120 native English speakers to CVCV stems with a single suffix [-be/-bo] or [-fi/-fu] and tested them on new stems with double suffixes (e.g. [peti-be-fi]). Participants in the harmony condition generalized harmony to test items with double suffixes, while participants in the disharmony condition did not infer a general vowel disharmony pattern. They argue that substantive bias plays a role in vowel (dis)harmony learning, as harmony can be iterative while disharmony cannot.

Several studies have examined categorical pattern learning and have found evidence of partial effects of substantive bias (Wilson 2006; Baer-Henney 2015). For instance, Wilson (2006) reported an asymmetry in the generalization of velar palatalization (/k/→/tʃ/, /g/→/dʒ/) to new contexts. Participants who were exposed to palatalization before /e/ (a context in which [k] and [t͡ʃ] are less perceptually similar) generalized to the /i/ context (a perceptually more similar context). However, those who were exposed to palatalization before /i/ did not generalize to the /e/ context. This finding supports the substantive bias hypothesis. However, further results demonstrate inconsistencies. All participants were also taught that velar stops did not palatalize before /ɑ/. Participants who were exposed to palatalization before /i/ did not generalize to the /ɑ/ context. On the other hand, participants who were exposed to palatalization before /e/ not only extended to the /i/ context but also to the /ɑ/ context.[1] This tendency is not predicted by substantive bias because velar stops and palate-alveolar affricates are more dissimilar before /ɑ/ than before /e/.

Some studies investigating categorical pattern learning failed to find supporting evidence for substantive bias (Pycha et al. 2003; Peperkamp & Dupoux 2007; Finley 2008; Finley & Badecker 2009; Skoruppa & Peperkamp 2011; Do et al. 2016; Lysvik 2018; Glewwe et al. 2018). Pycha et al. (2003) compared the learning of categorical vowel harmony with vowel disharmony by English-speaking participants. They were exposed to CVC stems, where the vowel was either front or back, and the morphophonological alternation was triggered by a plural suffix that was either [-ɛk] or [-ʌk]. Three conditions were tested, varying featural complexity and phonetic grounding. In Condition 1, stem and suffix vowels agreed in backness. In Condition 2, stem and suffix vowels disagreed in backness. In Condition 3, stem and suffix vowels agreed or disagreed in backness, depending on what the stem vowel was. After [i, æ, ʊ], the suffix vowel should be front [ɛ]; after [ɪ, u, a], the suffix vowel should be back [ʌ]. Participants performed significantly better in Conditions 1 and 2 than in Condition 3, providing support for the effect of structural complexity. However, results did not support the role of phonetic substance, as vowel harmony was not learned better than vowel disharmony. Similar results were obtained by Skoruppa & Peperkamp (2011), who investigated the learning of vowel rounding (dis)harmony by 60 native speakers of French. The participants were exposed to 40-minute-long stories spoken in a novel accent of French that was categorically harmonic or disharmonic. Each participant was randomly assigned to one of the two accents. In the test, participants heard thirty pairs of nonwords and were asked to indicate which one was pronounced in the accent they had been exposed to. Participants performed above chance level for both experimental conditions, with no significant

[1] Wilson's particular implementation of a grammatical model incorporating substantive bias predicts this extension, so Wilson considers the extension to /ɑ/ context as a further proof of the model's accuracy.

differences between harmony and disharmony. The authors proposed that substantive bias in vowel harmony did not exist in a perception-only context and that a bias favoring harmony may be restricted to production.

In summary, the evidence for the presence of substantive bias in categorical learning is mixed, with some studies reporting partial effects and others failing to find supporting evidence. These inconsistent findings may be attributed to methodological choices. For instance, studies such as Pycha et al. (2003) and Skoruppa & Peperkamp (2011) have presented differing results, highlighting the potential influence of these methodological factors.

1.2.2 Substantive bias in learning variable phonology

Most studies on the acquisition of variable linguistic patterns in artificial language learning have focused on morphosyntactic pattern learning (Singleton & Newport 2004; Kam & Newport 2005; 2009; Culbertson & Newport 2015; Schuler et al. 2016). However, a few experimental studies have examined phonological variation (Baer-Henney et al. 2015; Martin 2017; Mooney & Do 2018; Do & Mooney 2021; Do & Havenhill 2021). Baer-Henney et al. (2015) investigated the interaction between L1 phonotactics and substantive bias in variable phonological pattern learning by simultaneously manipulating phonetic naturalness and L1 phonotactic relevance. They tested the ability of native German speakers to learn two types of alternation: vowel backness harmony, which is phonetically motivated but has no relevance in the phonotactics of native German speakers, and a tenseness-backness pattern, which is unnatural but reflects a phonotactic property of German. In German phonotactics, tenseness but not backness is a relevant feature: Lax vowels only appear in closed syllables, while tense vowels can appear in closed and open syllables. Lax vowels can also appear in syllables with a complex coda, while tense vowels do not. Moreover, tenseness provides grammatical information in German; for example, pseudowords with lax vowels tend to be judged as nouns but not verbs by German speakers (Ott 2011). Baer-Henney et al. (2015) manipulated the degree of variability (65% vs. 85%) and the length of training (two vs. three training blocks) and found that participants performed better on the vowel backness harmony than on the tenseness-backness pattern when the alternation was less regular and when the training was shorter, while they performed better on the tenseness-backness pattern when the alternation was more regular and when the training was longer. Thus, the acquisition of a less stable pattern was facilitated by substantive bias rather than L1, and the influence of L1 was stronger in the acquisition of a more stable pattern. These results suggest that the uncertainty of input affects the strength of substantive bias. However, it should be noted that the effect of substantive bias in this study may be attributed to structural bias as well, since the unnatural tenseness-backness pattern is more structurally complex than the vowel backness harmony. Specifically, the vowel backness harmony relates to two instances of the same feature [backness], while the tenseness-backness pattern relates to one instance of [tenseness] and one instance of [backness].

Martin (2017) examined the role of substantive bias in the learning of vowel (dis)harmony through a production task and by manipulating the proportion of harmonic items. The study included 78 adult native speakers of French who were randomly assigned to one of five exposure conditions. These conditions included training items that were all harmonic, mostly harmonic (i.e., 75% harmonic items), all disharmonic, mostly disharmonic (i.e., 75% disharmonic items), or half and half (i.e., control condition). The effect of substantive bias was found only in the inconsistent input conditions, specifically between the mostly harmonic and mostly disharmonic conditions. To investigate whether the results could be replicated, the same materials were used in a follow-up experiment with 32 adult native speakers of French who were randomly assigned to one of the two inconsistent input conditions. However, the results of the first experiment were not replicated, as participants in the mostly harmonic condition did not perform better than those in the mostly disharmonic condition. The author attributed the discrepancy to the sample size and noted that a marginal difference was found between the two conditions when the data from the two experiments were analyzed together.

Do & Havenhill (2021) examined the impact of production practice on the learning of postnasal voicing and devoicing, and whether this effect is dependent on the training context. They compared perception-only training with perception-with-production training in both categorical and variable training contexts. In the variable training context, the (de)voicing alternation occurred in 70% of the stems, while the remaining 30% maintained their underlying voicing. All participants listened to recordings of singular-plural pairs and were presented with their written forms. In the perception-with-production condition, participants were asked to produce the plural form after listening to a singular-plural pair. In the subsequent test phase, participants heard the singular form and were required to select the appropriate plural form from written options. However, they were not asked to produce the plural forms during testing. The results from their participants indicated that postnasal voicing was learned more effectively than postnasal devoicing when production practice was included in the training. However, this effect was observed only in variable training contexts and not in categorical training contexts.

Mooney and Do (2018) identified a significant bias in the learning of phonological variation among native English-speaking adults. Participants were exposed to languages with different proportions of rounding harmony and disharmony patterns. Results of a subsequent production task indicated that learners in both languages adjusted the probabilistic distribution of variables towards the phonetically-grounded harmony pattern, with a higher degree of modulation observed in the language with more ungrounded patterns. Mooney and Do (2018) proposed that patterns need to be sufficiently unnatural to cause learners to adjust their proportions. In a subsequent study, Do and Mooney (2021) replicated the experiment with Cantonese-native preschoolers and found similar results. Compared to children, adults in Mooney and Do's study did not significantly deviate from the input distribution in learning the language with more ungrounded patterns, with adults showing a preference for harmony patterns over disharmony patterns at a higher rate (49%) than its proportion in the input (33%). In contrast, children entirely reversed the distribution to favor harmony patterns, selecting them at a higher rate (54% for seen items and 72% for unseen items) than what was presented in the input (33%), suggesting a stronger substantive bias among children. The researchers attributed this heightened bias to the relatively smaller lexicons of children compared to adults in real-language contexts, which results in less certainty in their lexicons. Similarly, artificial language learners in variable training conditions tend to have less consolidated lexicons compared to those in categorical training conditions. This disparity in lexicon consolidation may contribute to a more pronounced substantive bias among learners in variable training conditions.

Overall, the studies on variable pattern learning have provided evidence of substantive bias, although the number of studies on this topic is limited compared to studies on categorical pattern learning in phonology, and the effect of substantive bias has not been consistent. These studies suggest that substantively biased learning occurs when the pattern is variable but not categorical (Martin 2017; Do & Havenhill 2021), and the effect of substantive bias is strong in a condition with more variable input (Baer-Henney et al. 2015). **Table 1** summarizes the previous phonological studies on variable pattern learning. However, previous studies have involved a

Studies on variable learning	**Types of task**		**Results**
Baer-Henney et al. 2015	Exposure	Perception	Substantive bias was observed to be more prominent in cases where the alternation was less regular and when the duration of training was shorter.
	Test	Production	
Martin 2017	Exposure	Perception	Substantive bias was detected exclusively in the conditions with inconsistent input; however, the outcome was not replicated in a subsequent experiment.
	Test	Production	
Do & Havenhill 2021	Exposure	Perception vs. Perception with production	Substantive bias was solely observed in the variable training conditions, specifically when production practice was included during the exposure phase.
	Test	Perception	
Mooney & Do 2018	Exposure	Perception	Substantive bias exhibited greater strength in cases where the variable input consisted of phonetically-ungrounded patterns.
	Test	Production	
Do & Mooney 2021	Exposure	Production	The findings from the follow-up experiment with children were largely consistent with the results of Mooney & Do (2018), revealing a more pronounced substantive bias among the younger participants.

Table 1: Previous studies on substantive bias in variable learning.

production task either in the exposure or the test phase, which may confound variable input and in-experiment production as factors. Therefore, the current study aims to investigate whether variable input alone can activate substantive bias by implementing a perception-only task.

1.2.3 Hypothesis and predictions

Based on the literature review, we hypothesize that the effect of substantive bias will be stronger in variable pattern learning conditions compared to categorical pattern learning conditions. In this study, we introduce variability through alternating and non-alternating suffixes, which is analogous to vowel harmony patterns found in Hungarian and Korean. For instance, although generally categorical, variability exists in Hungarian vowel harmony, with some affixes participating in vowel harmony and others not. For example, the temporal suffix /-kor/ in Hungarian is non-alternating, and the vowel always surfaces as [o], regardless of the context (Kenesei et al. 2002). This type of variability has been shown to be learnable in artificial language learning paradigms (Finley 2021). By implementing variability through alternating and non-alternating suffixes in this study, we aim to investigate whether variable input can activate substantive bias in phonological learning.

To examine the impact of pattern variability on substantive bias, we randomly assigned Hong Kong Cantonese native speakers in the Experimental conditions to one of six artificial languages (see **Table 2**). These languages included a categorically natural language (CNL), a categorically unnatural language (CUL), and four variably natural and unnatural languages (VNL1, VNL2, VUL1, VUL2) with differing non-alternating suffixal vowels.[2] Participants were taught singular, dual, and plural forms for each language's stems. In CNL and CUL, the input pattern was categorical, with CNL always featuring phonetically grounded backness harmony and CUL featuring 100% phonetically ungrounded alternations. In VNL1 and VUL1, the input patterns were variable, with the plural suffix always being non-alternating [-ge]. In VNL1, the dual suffix agreed with the backness of the stem, resulting in a 75% phonetically grounded dominant pattern. In VUL1, the dual suffix disagreed with the backness of the stem, resulting in a 75% phonetically ungrounded dominant pattern. VNL2 and VUL2 differed from VNL1 and VUL 1 only in the non-alternating suffix, which surfaced as [-go]. We created 24 CVCV stems, with half containing two front vowels and half containing two back vowels. Each stem was combined with a dual suffix ([-me] ~ [-mo]) and a plural suffix ([-ge] ~ [-go]). The Experimental conditions were compared to a no-training Control condition, in which participants completed the test trials in the Experimental conditions without learning an artificial language.

[2] Note that for a phonological process to be *unnatural*, it does not decrease articulatory effort and does not increase perceptual distinctiveness (Beguš 2018). Vowel disharmony increases the perceptual distinctiveness of sounds (Wilson 2003). While *phonetically-grounded* is the more appropriate term to describe vowel harmony, for simplicity, we use *natural/unnatural* in the condition names.

	Stem	Suffix	(Dis)harmony	Example
Categorically Natural Language (CNL)	12 front vowel	12 [-me] for duals 12 [-ge] for plurals	harmonic	tipe~tipeme~tipege
	12 back vowel	12 [-mo] for duals 12 [-go] for plurals		tonu~tonumo~tonugo
Categorically Unnatural Language (CUL)	12 front vowel	12 [-mo] for duals 12 [-go] for plurals	disharmonic	tipe~tipemo~tipego
	12 back vowel	12 [-me] for duals 12 [-ge] for plurals		tonu~tonume~tonuge
Variably Natural Language 1 (VNL1)	12 front vowel	12 [-me] for duals 12 [-ge] for plurals	harmonic	tipe~tipeme~tipege
	12 back vowel	12 [-mo] for duals 12 [-ge] for plurals	harmonic disharmonic	tonu~tonumo~tonuge
Variably Unnatural Language 1 (VUL1)	12 front vowel	12 [-mo] for duals 12 [-ge] for plurals	disharmonic harmonic	tipe~tipemo~tipege
	12 back vowel	12 [-me] for duals 12 [-ge] for plurals	disharmonic	tonu~tonume~tonuge
Variably Natural Language 2 (VNL2)	12 front vowel	12 [-me] for duals 12 [-go] for plurals	harmonic disharmonic	tipe~tipeme~tipego
	12 back vowel	12 [-mo] for duals 12 [-go] for plurals	harmonic	tonu~tonumo~tonugo
Variably Unnatural Language 2 (VUL2)	12 front vowel	12 [-mo] for duals 12 [-go] for plurals	disharmonic	tipe~tipemo~tipego
	12 back vowel	12 [-me] for duals 12 [-go] for plurals	disharmonic harmonic	tonu~tonume~tonugo

Table 2: The six artificial languages.

By comparing CNL and CUL, we aim to examine the role of substantive bias in categorical pattern learning, while comparing VNL1 vs. VUL1 and VNL2 vs. VUL2 allows us to explore the same question in a variable pattern learning context. Our hypothesis is that the variability of patterns will affect the impact of substantive bias on pattern learning, with the effect of substantive bias being more pronounced in variable pattern training conditions than in categorical pattern training conditions. We predict that learners in both types of conditions may be more likely to acquire harmonic patterns than disharmonic ones if they are substantively biased, but the preference for harmony is expected to be stronger in variable pattern training conditions.

2 Methodology

2.1 Stimuli

The participants in the Experimental conditions were exposed to 24 series of singular, dual, and plural forms exhibiting vowel backness harmony or disharmony (see Supplementary Materials A). The singular form consisted of a consonant-vowel-consonant-vowel ($C_1V_1C_2V_2$) structure, where the stem vowels included two front vowels (/i, e/) and two back vowels (/o, u/), and followed back harmony constraints. Specifically, if V1 was front, V2 was also front. Twelve stems contained front vowels, and twelve stems contained back vowels. To match transitional probability, the frequencies of the vowels at V1 and V2 were balanced, as were the frequencies of the stem consonants, which were chosen from /p, t, k, b, d, g, m, n/. The combination of $C_1V_1C_2V_2$ were adapted from Finley (2021). The dual form had two possible suffixes, [-me] or [-mo], which were determined by the backness of the stem, as explained in Section 1.2.3. In categorically and variably natural languages, the backness of the suffix vowel agreed with the backness of the stem vowels, resulting in forms such as [tipeme] and [podomo]. In categorically and variably unnatural languages, the backness of the suffix vowel and the stem vowels disagreed, resulting in forms such as [tipemo] and [podome]. The plural form had two possible suffixes, [-ge] or [-go]. In categorically natural languages, the backness of the suffix vowel harmonized with the backness of the stem vowels, resulting in forms such as [tipege] and [podogo]. In categorically unnatural languages, there was vowel disharmony, resulting in forms such as [tipego] and [podoge]. In variably natural or unnatural languages, the plural suffix was non-alternating, resulting in forms such as [tipege] and [podoge], or [tipego] and [podogo]. The visual stimuli consisted of cartoon monsters from Sporepedia (2009), with one, two, and four monsters used to create singular, dual, and plural forms, respectively (see Supplementary Materials B). The visual stimuli were used in both the exposure and test phases.

Learning was assessed using a two-alternative forced-choice test, which consisted of 32 test items, with 16 seen items and 16 unseen items (see Supplementary Materials C). The seen items contained stems that had been presented during the exposure phase, while the unseen items contained stems with phoneme combinations that were not heard in the exposure phase. For each test item, the participants heard the name of a monster and were asked to choose the correct forms for the dual or plural form, which contained the same stem and differed only in the suffix vowel ([e] or [o]). For both dual and plural form questions, two answer choices were given, e.g., [dinime] vs. [dinimo] or [dinige] vs. [dinigo]. The dual and plural test items were presented equally often for both seen and unseen items.

All stimuli were recorded by an adult male speaker of American English in a sound-proof booth at a sampling rate of 44100 Hz, at the second author's institute. The stimuli were initially stressed with English pronunciation, but the vowels were not reduced. Amplitude normalization was performed on all stimuli using Praat (Boersma & Weenink 2017).

2.2 Participants

A total of 143 self-reported native speakers of Hong Kong Cantonese (81 women & 62 men) participated in the experiment and completed the task. All our participants self-reported that they were fluent in English while all of them also reported their dominant language was Cantonese. The participants were 118 individuals between the ages of 18 and 25, 21 individuals between the ages of 26–35, and 4 individuals between the ages of 36–45. In the present study, we conducted an experiment involving native Cantonese speakers from Hong Kong who were exposed to non-native stimuli comprised of English phonemes, which represented their second language. It is important to note that these participants were not accustomed to a three-way contrast between singular, dual, and plural forms, as their native language does not exhibit such a distinction. The hypothesis put forward in this study was that this exposure to non-native stimuli would not hinder the learning process. This hypothesis was supported by the findings of Steele et al. (2015), who reported that participants in artificial language learning experiments showed comparable proficiency in acquiring both native and non-native stimuli. Building upon this, Martin & Peperkamp (2020) proposed that the processing of non-native stimuli is primarily influenced by phonetic characteristics and is less likely to be influenced by additional knowledge or factors that could affect speech perception and learning.

127 participants were assigned to the Experimental conditions, including 24 participants in the CNL condition, 20 participants in the CUL condition, 20 participants in the VNL1 condition, 21 participants in the VUL1 condition, 23 participants in the VNL2 condition, and 21 participants in the VUL2 condition. The Control condition included the remaining 16 participants. All participants were adult native speakers of Cantonese with no speech or hearing disorders and were eligible to participate. Each participant was paid 50 HKD upon completion of the task. All participants’ responses were included in the data analysis.

2.3 Procedures

A computer-based task was created using PsychoPy version 3.0 (Peirce et al. 2019) and was conducted online via Pavlovia (2021). Before the experiment, participants completed a questionnaire about their age range, gender, and native language using Qualtrics (2021) (see Supplementary Materials D). For participants in the Experimental conditions, the experiment consisted of a training phase and a testing phase. The stimuli were randomized within each phase for each participant. Neither orthography nor feedback was provided, making the experimental procedure a relatively implicit learning task. At the beginning of the training phase, participants saw instructions displayed in traditional Chinese characters explaining that they were going to learn a foreign language similar to English and would be asked to name dual and plural forms in this language. The training phase consisted of 24 items, with each trial presenting the corresponding visual stimuli and then the name of the monster in singular, dual, and plural form. The order of presentation was always the singular form, followed by the dual form, and finally the plural form. The 24 items were repeated three times, totaling 72 trials, with a different random order for each iteration of the training items. To ensure participants were focused on learning, 12 focus questions were presented orthographically in traditional Chinese characters, asking participants about the color of a monster they had just seen.

After completing the training phase, participants in the Experimental conditions moved on to the testing phase, which included 16 items that they had seen during the training phase and 16 new items that they had not seen before. These items were presented in a random order to each participant. During the testing phase, participants heard the name of a singular monster and then listened to two options for the corresponding dual form or plural form, which either showed vowel backness harmony or disharmony between the stem and the suffix. The two options for dual forms were stem-[me] and stem-[mo], while the two options for plural forms were stem-[ge] and stem-[go]. Half of the harmonic answers were given as the first option ("a" key) and the other half as the second option ("l" key), which were randomized across participants. Participants in the Experimental conditions were instructed to select the word that was more likely to belong to the language they had just learned, while participants in the Control condition were asked to select the item they preferred without learning a language. Participants were only allowed to make their answer choice after both options had been played. The Experimental conditions took approximately 25 minutes to complete, while the Control condition took on average 10 minutes.

3 Results

Firstly, we analyzed the data from the Control condition (no-training condition), followed by the analysis of data from the Categorical and Variable conditions in the Experimental conditions. To conduct our analyses, we employed logistic mixed-effects regression models, utilizing the glmer function of the lme4 package (Bates et al., 2015) in the R programming language (R Core Team, 2018). All models included random intercepts for Participants and Items.[3]

3.1 Control condition

To investigate the preference of participants in the Control condition for harmonic or disharmonic patterns, Harmony (0 = disharmonic answer; 1 = harmonic answer) was designated as the dependent variable in the model. The model results revealed that the rate of harmonic responses (M = 0.494, SD = 0.138) in the Control condition did not significantly

[3] Random slopes were not included in the models due to convergence issues.

differ from chance (β = –0.02463, SE = 0.14784, z = –0.167, p = 0.868). Additionally, **Figure 1** shows that there was a high level of agreement among participants. Therefore, it can be concluded that the Cantonese participants in our study did not show a pre-existing bias towards vowel harmony.

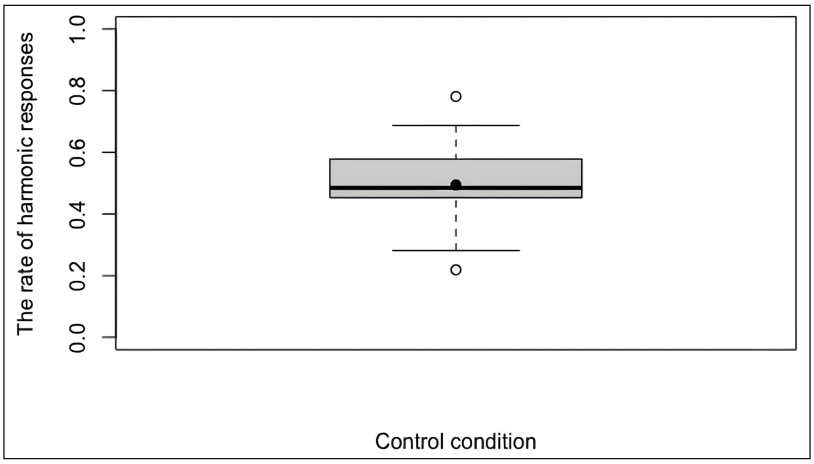

Figure 1: The rate of harmonic responses in the Control condition.

3.2 Categorical training conditions

The overall accuracy rates of participants in the categorical training conditions were 0.773 ± 0.176 for CNL (n = 22) and 0.730 ± 0.178 for CUL (n = 20), as illustrated in **Figure 2**. The data for seen and unseen items are presented in **Figure 3** and **Figure 4**, respectively. We fitted a model with Accuracy as the dependent variable and an intercept, but no other fixed effects for each condition (CNL, CUL) separately. Accuracy was converted to binary values, i.e., either as a correct or an incorrect response. The results showed that performance was above chance in both CNL (β = 1.5383, SE = 0.2704, z = 5.69, $p < .001$) and CUL (β = 1.2802, SE = 0.2866, z = 4.467, $p < .001$), indicating that learning had an effect in the categorical training conditions. To compare the two conditions, we further conducted a regression model with Accuracy as the dependent variable. We focused on two contrast-coded factors: CNL vs CUL (i.e., the contrast between CNL and CUL) and Seen vs Unseen (i.e., the contrast between seen and unseen items). The results, as presented in **Table 3**, indicated a lack of significant effect of CNL vs CUL (β = 0.2619, p = 0.493), Seen vs Unseen (β = 0.0108, p = 0.950), and their interaction (β = 0.1320, p = 0.627). These results suggest that participants did not exhibit a preference for learning harmony over disharmony, and that accuracy was not influenced by whether the items were seen or unseen. Therefore, we conclude that there was neither substantive bias, nor memory effect of individual items, observed in the categorical training.

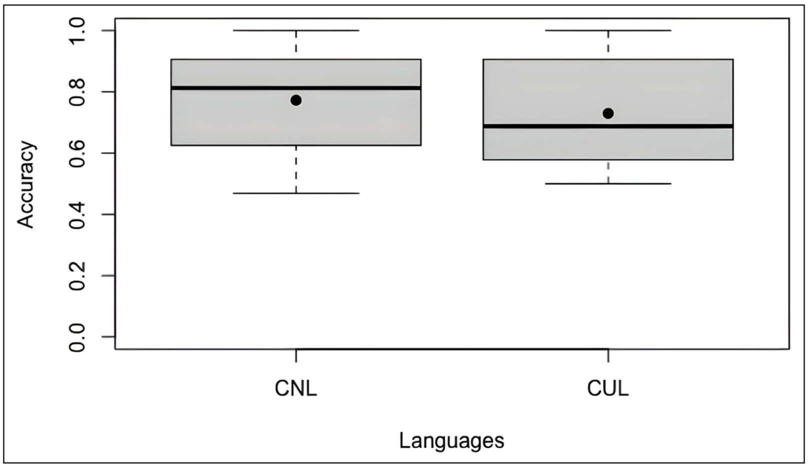

Figure 2: Accuracy rate in categorical training.

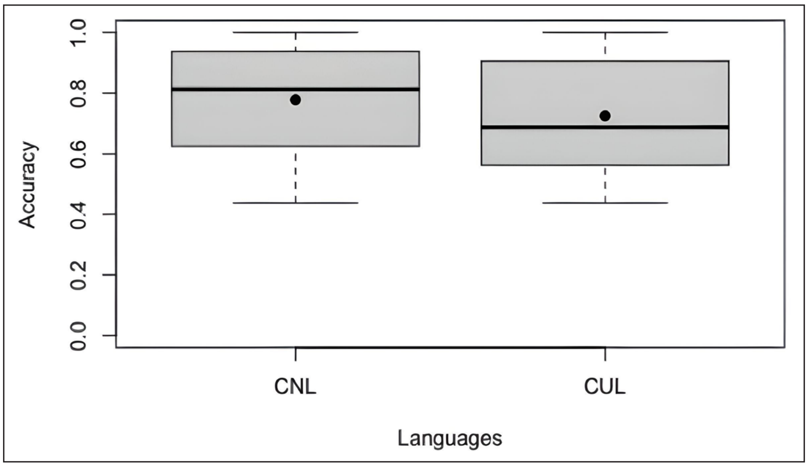

Figure 3: Seen items in categorical training

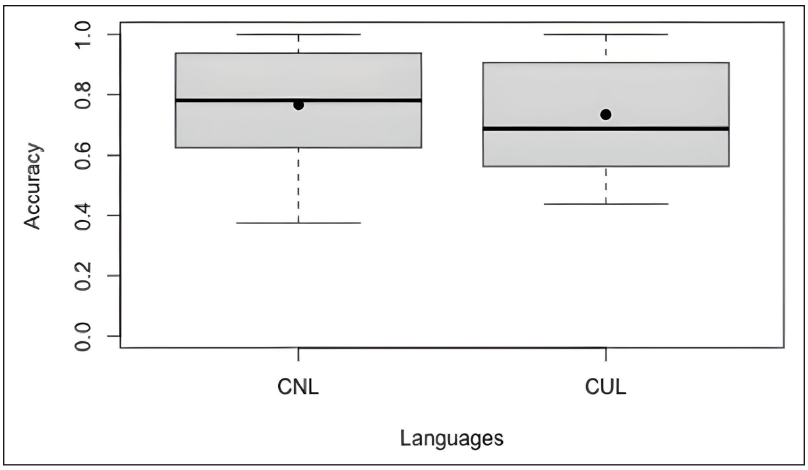

Figure 4: Unseen items in categorical training

	Estimate	SE	z value	p(z)
Intercept	1.4126	0.2014	7.014	2.32e-12 ***
CNLvsCUL	0.2619	0.3821	0.685	0.493
SeenvsUnseen	0.0108	0.1723	0.063	0.950
CNLvsCUL: SeenvsUnseen	0.1320	0.2713	0.486	0.627

Table 3: The results of a logistic regression model for the accuracy rates in categorical training.

3.3 Variable training conditions

Next, the data collected under variable training conditions were analyzed. In this context, 'accuracy' refers to a harmonic response in the dominantly natural languages (VNL1/VNL2) and a disharmonic response in the dominantly unnatural languages (VUL1/VUL2). In other words, 'accuracy' was measured as an application of (dis)harmony relative to the dominant input pattern. Participants in VNL2 (n = 23) with [-go] as the non-alternating suffix did not demonstrate a significant difference in accuracy rate (M = 0.686, SD = 0.142) compared to participants in VNL1 (n = 20) with [-ge] as the non-alternating suffix (M = 0.653, SD = 0.151) (β = 0.3091, SE = 0.2266, z = 1.364, p = 0.172602). Similarly, participants in VUL2 (n = 21) with [-go] as the non-alternating suffix did not demonstrate a significant difference in accuracy rate (M = 0.564, SD = 0.110) compared to participants in VUL1 (n = 21) with [-ge] as the non-alternating suffix (M = 0.546, SD = 0.088) (β = 0.06739, SE = 0.14887, z = 0.453, p = 0.65080). This suggests that the non-alternating suffix with a non-back vowel /e/ or a back vowel /o/ did not affect learning. Consequently, the data in VNL1 and VNL2 were merged into VNL, and VUL1 and VUL2 were merged into VUL.

The accuracy rate for participants who learned VNL and VUL was 0.671 ± 0.145 and 0.555 ± 0.099, respectively, as depicted in **Figure 5**. **Figures 6** and **7** show the data for seen and unseen items, respectively. Similar to the analysis conducted with the categorical training data, we began by developing a model with the dependent variable Accuracy, with an intercept for each condition (VNL, VUL) individually. Accuracy was converted into binary values. The results of the analysis revealed that performance was significantly above chance in both VNL (β = 0.7934, SE = 0.1209, z = 6.561, $p < 0.001$) and VUL (β = 0.2228, SE = 0.0612, z = 3.64, $p < 0.001$), indicating that learning had a positive effect in both variable training conditions. This result is consistent with the findings from the categorical training conditions.

Similar to the analysis conducted with the categorical training data, a logistic regression model was utilized to investigate the impact of substantive bias in variable training, where the dependent variable was Accuracy. Two contrast coded factors, VNL vs VUL (i.e., the contrast between VNL and VUL) and Seen vs Unseen (i.e., the contrast between seen and unseen items), were examined. The outcomes, displayed in **Table 4**, revealed a significant effect of VNL vs

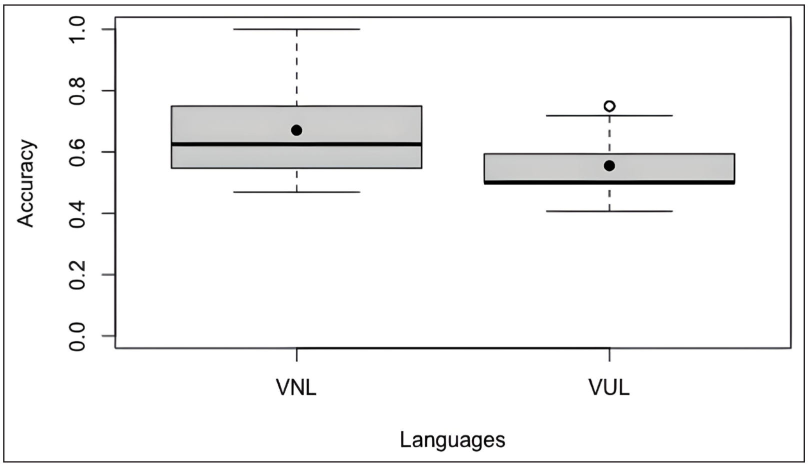

Figure 5: Overall accuracy rate in variable training.

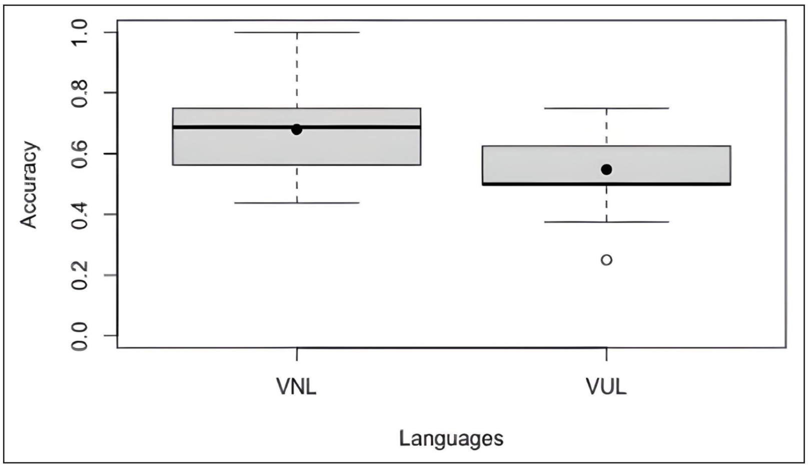

Figure 6: Seen items in variable training.

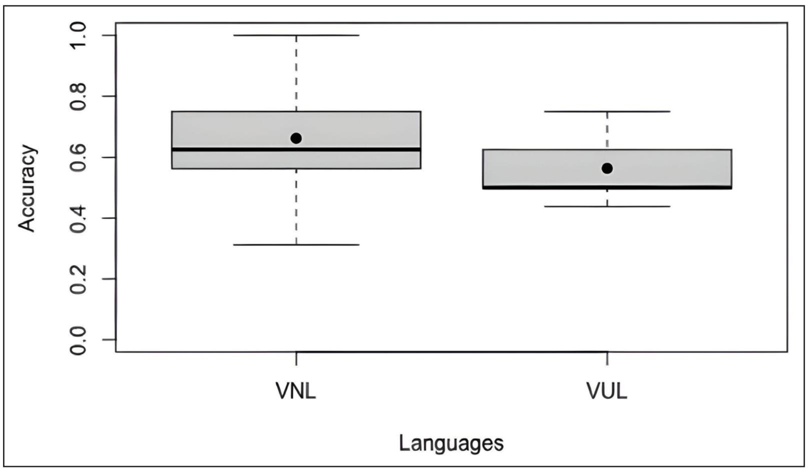

Figure 7: Unseen items in variable training.

VUL (β = 0.5249, p < .001), but no impact of Seen vs Unseen (β = 0.01368, p = 0.866) or their interaction (β = 0.15179, p = 0.348). The findings indicated that when provided with an equivalent proportion of harmonic patterns in VNL and disharmonic patterns in VUL (i.e., 75%), participants learned disharmony less successfully than harmony. This suggests the presence of a phonetic naturalness effect within the variable training context.

	Estimate	SE	z value	p(z)
Intercept	0.49207	0.06247	7.877	3.36e-15 ***
VNLvsVUL	0.52490	0.12465	4.211	2.54e-05 ***
SeenvsUnseen	0.01368	0.08091	0.169	0.866
VNLvsVUL: SeenvsUnseen	0.15179	0.16184	0.938	0.348

Table 4: The results of a logistic regression model for the accuracy rates in variable training.

It is important to note that the dual suffix ([-me] ~ [-mo]) in the variable training involved (dis)harmony, while the plural suffix ([-ge] ~ [-go]) was insensitive to (dis)harmony patterns. To examine whether participants in VNL and VUL performed differently on the two suffixes, a pairwise comparison analysis was conducted between the two suffixes (dual vs. plural) and the two variable languages (VNL vs. VUL) using the emmeans package (Lenth et al. 2020). The results in **Table 5** revealed a significant difference in the application of (dis)harmony to the dual suffix ([-me] ~ [-mo]) by participants in the VNL condition and those in the VUL condition (β = 0.694, SE = 0.157, z = 4.419, p < .001). Specifically, the analysis revealed that participants in the VNL condition had a higher accuracy rate (75.73%) in applying harmony to the dual suffix compared to those in the VUL condition, who applied disharmony to the dual suffix with an accuracy rate of 62.05%. This finding further indicates that the substantive bias effect was observed in the learning of the suffix alternating according to its vowel harmony with the stem vowels, i.e., ([-me] ~ [-mo]).

	Estimate	SE	z.ratio	p(z)
VNL dual suffix – VUL dual suffix	0.694	0.157	4.419	0.0001***
VNL plural suffix – VUL plural suffix	0.409	0.149	2.741	0.0311*

Table 5: Pairwise comparisons between the two suffixes and the two variable languages.

A significant difference was found between the VNL group and the VUL group in their ability to apply the plural suffix ([-ge] ~[-go]) (β = 0.409, SE = 0.149, z = 2.741, p < .05). Participants in the VNL condition were more accurate in applying harmony to the plural suffix

(58.43%) compared to those in the VUL condition who showed disharmony in applying the plural suffix (48.96%). Since the plural suffix was non-alternating, half of the training items containing the plural suffix exhibited either harmony or disharmony. Thus, perfect learning of the plural suffix would result in a 50% harmonic response in both the VNL and VUL conditions. The rate of harmonic responses for participants in the VUL condition was 51.04%, which could indicate either perfect learning or chance performance. For participants in the VNL condition, the rate of harmonic response was 58.43%, which was significantly higher than chance ($\beta = 0.3599$, $SE = 0.1049$, $z = 3.432$, $p < .001$). Therefore, participants in the VNL condition "modulated" the non-alternating plural suffix towards harmony, providing additional evidence for a substantive bias effect in variable training. Furthermore, as mentioned in section 2.2, native Cantonese speakers are not familiar with a three-way contrast among singular, dual, and plural. The current results suggest that participants were able to learn both the dual and plural suffixation. Participants had a general tendency to alternate the dual suffix, compared to the plural suffix, in line with the input.

In summary, participants performed better than chance in all Experimental conditions, indicating that they were able to learn the novel vowel harmony patterns. However, the effect of substantive bias was only observed in the variable training condition. Specifically, learners demonstrated a bias towards acquiring harmony patterns better than disharmony patterns when the pattern was variable during training. These findings highlight the important role that variability plays in learning phonetically motivated and unmotivated patterns and the impact it can have on the acquisition of phonological patterns.

4 Discussion and conclusion

The current study discovered no evidence of substantive bias in categorical learning. However, upon introducing variability in the input, the substantive bias effect emerged, as evidenced by the acquisition of vowel backness harmony by native Hong Kong Cantonese speakers. This finding implies that participants learned and generalized the alternation pattern in a substantively biased manner when exposed to *uncertain input* (or variability) during the learning process.

The absence of evidence for substantive bias in categorical learning, as observed in the current study, is consistent with the findings of two previous studies conducted by Pycha et al. (2003) and Skoruppa & Peperkamp (2011). These earlier studies also investigated the learning of vowel harmony, but with different participant populations and language contexts. Specifically, Pycha et al. (2003) examined English speakers, Skoruppa & Peperkamp (2011) investigated French speakers, and our study focused on Hong Kong Cantonese speakers. The fact that all three studies did not find a substantive bias effect in the categorical training conditions suggests that the absence or weak presence of substantive bias is not limited to certain language-specific contexts. However, contrary to the present findings, a previous study by Martin & Peperkamp (2020), which tested English native speakers and employed stems and a single suffix, revealed that categorical vowel backness harmony was learned more effectively than disharmony. These mixed results may suggest that the influence of substantive bias may be inconsistent in categorical learning in phonology.

The current findings on variable learning are in accordance with previous studies on substantive bias involving variable phonological patterns, conducted by Baer-Henney et al. (2015), Martin (2017), Do & Havenhill (2021), Mooney & Do (2018), and Do & Mooney (2021). Similar to previous studies, the current study indicates that substantive bias is activated when the alternation is variable and thus less certain. As the previous studies did not directly investigate the role of certainty in substantive bias, it is important to highlight that further research is needed to fully understand the extent to which variability triggers substantive bias. Future experiments could explore variable patterns with varying proportional distributions, as our study did not examine how the degree of variability influences the activation of substantive bias. In this study, our primary focus was to compare the presence of substantive bias between categorical pattern learning conditions and variable pattern learning conditions, in order to examine whether variability affects the manifestation of substantive bias.

Do & Havenhill (2021) uncovered that learning an articulatory-based pattern was substantively biased only when production practice was involved in learning. This production effect emerged only from variable training conditions, not from categorical training conditions. In line with Do & Havenhill (2021), we also found that the effect of substantive bias is sensitive to methodological choices of variable phonological pattern learning, i.e., categorical vs. variable input data. Hence, our findings align with the argument put forth by Moreton and Pater (2012a; b) that, unlike structural bias which consistently emerges in experiments, the impact of substantive bias is relatively weak. Substantive bias was observed in learning only under specific circumstances. In the present study, it was the presence of input variability, indicating uncertainty, that rendered phonetic substance relevant to the learning process. This suggests that the shape and distribution of patterns in the input can influence the manifestation of substantive bias.

The current result also aligns with the findings in a previous study on the learning of rounding harmony variation by Mooney & Do (2018), albeit with some discrepancies. In Mooney & Do (2018), learners in both experimental conditions (67% rounding harmony and 33% disharmony vs. 33% rounding harmony and 67% disharmony) modulated the probabilistic distribution of variables towards the phonetically-grounded harmony pattern, and the degree of modulation was higher from the artificial language with more ungrounded patterns (33% to 49%). In the current experiment, harmony in VNL and disharmony in VUL were both underlearned, but disharmony in VUL was learned worse. Participants in VUL modulated the proportion of harmony from 25% to 44.5%. Thus, the finding that the languages with dominantly ungrounded patterns are more likely to undergo proportional changes of variants is consistent across the two studies. However, languages with dominantly grounded patterns may or may not be subject to changes, as the diverging tendencies shown from the two studies. The discrepancy between the current result and the result of Mooney & Do (2018) may be influenced by the different types of variability between the two studies. The variability in Mooney & Do (2018) was implemented by free variation of rounding harmony, whereas the variability in our study was realized by the alternating and non-alternating suffixes. The non-alternating suffix in our study may have made participants less prone to alternation when the language was dominantly natural (i.e., VNL).

The results of our study on variable phonological learning differ from the tendency reported in studies on variable learning in morphosyntax. In those studies, adults reproduced approximately the proportion of variants (e.g., determiners) present in the input, while children regularized the language, either systematically using or omitting determiners (Kam & Newport 2005; 2009; Culbertson et al. 2012). In contrast, our study on morphophonological learning found that adults neither learned the variability veridically nor regularized the language. Instead, they modulated the variant distribution towards harmony when the dominant input was disharmonic. This discrepancy between the current study and studies on morphosyntactic learning may be attributed to the different learning mechanisms involved in the two types of learning. In morphosyntactic learning, learners primarily identify the structural relationships among sentence components. For example, mass nouns belong to one class and count nouns to another, with each class taking a different determiner. Learners must identify which class of nouns (count vs. mass) takes which determiner. Variability was introduced by exposing participants to a mixture of sentences with and without determiners (Kam & Newport 2005). In our study on morphophonological learning, learners had to identify vowel harmony, a pattern that is not only related to the structural relationship among numeral modifiers and nouns but also to the phonetic grounding of the morphophonological pattern. We speculate that the learning mechanisms might differ in morphosyntactic and morphophonological learning due to the involvement of phonetic components in the latter. However, further exploration is needed to accurately identify the specific differences between the learning mechanisms of morphosyntactic and morphophonological variation learning.

One potential limitation of the current experiment is that the stem vowels consistently exhibited harmony, which could serve as an additional cue for learners to detect the harmony pattern. It is known that learners tend to acquire a harmony pattern when both stems and affixes follow harmony (Chong 2017). However, the presence of stem harmony alone cannot be attributed as the main factor contributing to the substantive bias observed in variable learning, as participants in the Control condition and the categorical training conditions did not show a bias towards harmony. Another limitation is that the alternation in variable learning involved only one suffix ([-me]∼[-mo]), while another suffix [-ge]/[-go] remained non-alternating. Natural languages are more complex, and when learners encounter a small number of suffixes undergoing harmony, they may resist acquiring the pattern (Baer-Henney et al. 2015; Finley 2021). Future studies could address this limitation by incorporating a larger number of suffixes, which may result in a stronger bias towards harmony in variable learning.

Abbreviations

CNL = a categorically natural language

CUL = a categorically unnatural language

VNL1 = a variably natural language with a [-back] suffixal vowel as the non-alternating suffix

VNL2 = a variably natural language with a [+back] suffixal vowel as the non-alternating suffix

VUL1 = a variably unnatural language with a [-back] suffixal vowel as non-alternating suffix

VUL2 = a variably unnatural language with a [+back] suffixal vowel as non-alternating suffix

ALL = artificial language learning

DAT = dative

Data availability

Dataset:

- Data file 1. Raw data related to the categorical learning conditions
- Data file 2. Raw data related to the variable learning conditions
- Data file 3. Raw data related to the control condition
- Data file 4. Code

Supplementary file:

- Supplementary Materials A. Stimuli in the training phase
- Supplementary Materials B. An example of visual stimuli
- Supplementary Materials C. Test items
- Supplementary Materials D. Questionnaire

DOI: https://osf.io/ue482/

Ethics and consent

The study was reviewed and approved by the Human Research Ethics Committee at the authors' institute. HREC's Reference Number: EA1812044. All participants provided their informed consent to participate in this study.

Acknowledgements

This paper has been significantly enhanced due to the valuable comments provided by Associate Editor Sara Finley, as well as the insightful feedback from anonymous reviewers. The research was enriched through the stimulating discussions with the members of the Language Development Lab at the University of Hong Kong. We would also like to express our gratitude to all the participants who generously contributed to this study.

Competing interests

The authors have no competing interests to declare.

References

Albright, Adam & Do, Youngah. 2017. A substantive bias for perceptually minimal alternations in Artificial Grammar learning. Presentation given at the *14th Old World Conference on Phonology*.

Baer-Henney, Dinah. 2015. *Learners' little helper: Strength and weakness of the substantive bias in phonological acquisition*. Doctoral dissertation. University of Potsdam.

Baer-Henney, Dinah & Kügler, Frank & van de Vijver, Ruben. 2015. The interaction of language-specific and universal factors during the acquisition of morphophonemic alternations with exceptions. *Cognitive Science* 39(7). 1537–1569. DOI: https://doi.org/10.1111/cogs.12209

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Boersma, Paul & Weenink, David. 2017. *Praat: Doing phonetics by computer (Version 6.0.49)*. Retrieved October 1, 2022, from http://www.praat.org/

Beguš, Gašpar. 2018. *Unnatural phonology: A synchrony-diachrony interface approach*. Doctoral Dissertation, Harvard University, MA.

Chambers, Kyle E. & Onishi, Kristine H. & Fisher, Cynthia. 2010. A vowel is a vowel: Generalizing newly learned phonotactic constraints to new contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36(3). 821–828. DOI: https://doi.org/10.1037/a0018991

Chong, Junxiang. 2017. *On the relation between phonotactic learning and alternation learning*. Doctoral dissertation, University of California at Los Angeles.

Clements, George. Nickerson. 1976. The autosegmental treatment of vowel harmony. In Dressier, Wolfgang U. & Pfeiffer, Oskar E. (eds.), *Phonologica*. 111–119. Innsbruck: Institut fur Sprachwissenschaft.

Cristià, Alejandrina & Seidl, Amanda. 2008. Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development* 4(3). 203–227. DOI: https://doi.org/10.1080/15475440802143109

Culbertson, Jennifer & Newport, Elissa L. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139. 71–82. DOI: https://doi.org/10.1016/j.cognition.2015.02.007

Culbertson, Jennifer & Smolensky, Paul & Legendre, Géraldine. 2012. Learning biases predict a word order universal. *Cognition* 122(3). 306–329. DOI: https://doi.org/10.1016/j.cognition.2011.10.017

Do, Youngah. 2013. *Biased learning of phonological alternations*. Doctoral dissertation. Massachusetts Institute of technology.

Do, Youngah & Havenhill, Jonathan. 2021. Production and Substantive Bias in Phonological Learning. In Bennett, Ryan & Bibbs, Richard & Brinkerhoff, Mykel & Kaplan, Max & Rich, Stephanie & Handel, Nicholas & Cavallaro, Maya (eds.), Proceedings of the 2020 Annual Meeting on Phonology. Washington, DC: Linguistic Society of America. DOI: https://doi.org/10.3765/amp.v9i0.4925

Do, Youngah & Mooney, Shannon. 2021. Variation awaiting bias: Substantively biased learning of vowel harmony variation. *Journal of Child Language* 49(2). 397–407. Cambridge University Press. DOI: https://doi.org/10.1017/S0305000920000719

Do, Youngah & Zsiga, Elizabeth & Havenhill, Jonathan. 2016. Naturalness and frequency in implicit phonological learning. Talk presented at the *90th Annual Meeting of the Linguistic Society of America*. Washington, D.C.

Finley, Sara. 2008. *Formal and cognitive restrictions on vowel harmony*. Doctoral dissertation. Johns Hopkins University. DOI: https://doi.org/doi:10.7282/T3K64H05

Finley, Sara. 2012. Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes* 27(10). 1550–1562. DOI: https://doi.org/10.1080/01690965.2012.660168

Finley, Sara. 2021. *Learning exceptions in phonological alternations. Language and Speech* 64(4). 991–1017. DOI: https://doi.org/10.1177/0023830920978679

Finley, Sara & Badecker, William. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61(3). 423–437. DOI: https://doi.org/10.1016/j.jml.2009.05.002

Gafos, Adamantios & Dye, Amanda. 2011. Vowel harmony: Opaque and transparent vowels. In Oostendorp, Marc van & Ewen, Colin J. & Hume, Elizabeth & Rice, Keren (eds.), *The Blackwell Companion to Phonology*. John Wiley & Sons, Ltd. DOI: https://doi.org/10.1002/9781444335262.wbctp0091

Glewwe, Eleanor. 2019. *Bias in phonotactic learning: Experimental studies of phonotactic implicationals*. Doctoral Dissertation, UCLA, CA.

Glewwe, Eleanor & Zymet, Jesse & Adams, Jocob & Jacobson, Rachel & Yates, Anthony & Zeng, Ann & Daland, Robert. 2018. Substantive bias and the acquisition of final (de)voicing patterns. Presentation given at the *92nd Annual Meeting of the Linguistic Society of America*. Salt Lake City.

Gordon, Matthew. 2016. *Phonological typology*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199669004.001.0001

Hayes, Bruce & Londe, Zsuzsa Cziráky. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23(1). 59–104. DOI: https://doi.org/10.1017/S0952675706000765

Hayes, Bruce & Siptár, Péter & Zuraw, Kie & Londe, Zsuzsa. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4). 822–863. DOI: https://doi.org/10.1353/lan.0.0169

Hochmann, Jean-Rémy & Carey, Susan & Mehler, Jacques. 2018. Infants learn a rule predicated on the relation same but fail to simultaneously learn a rule predicated on the relation different. *Cognition* 177. 49–57. DOI: https://doi.org/10.1016/j.cognition.2018.04.005

Ito, Junko. 1984. Melodic dissimilation in Ainu. *Linguistic Inquiry* 15. 505–513.

Kam, Carla & Newport, Elissa. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1(2). 151–195. DOI: https://doi.org/10.1207/s15473341lld0102_3

Kam, Carla & Newport, Elissa. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1). 30–66. DOI: https://doi.org/10.1016/j.cogpsych.2009.01.001

Kaun, Abigail Rhoades. 2004. The typology of rounding harmony. In Hayes, Bruce & Kirchner, Robert & Steriade, Donca (eds.), *Phonetically Based Phonology*. 87–116. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486401.004

Kenesei, Istvan & Vago, Robert M. & Fenyvesi, Anna. 2002. *Hungarian*. Routledge. DOI: https://doi.org/10.4324/9780203192238

Kuo, Li-Jen. 2009. The role of natural class features in the acquisition of phonotactic regularities. *Journal of Psycholinguistic Research* 38(2). 129–150. DOI: https://doi.org/10.1007/s10936-008-9098-7

Lenth, Russell V. & Buerkner, Paul & Herve, Maxime & Love, Jonathan & Singmann, Henrik. 2020. emmeans: Estimated marginal means, aka least-squares means (R package version 1.4.4). Retrieved October 1, 2022, from https://cran.r-project.org/web/packages/emmeans/index.html

Linker, Wendy. 1982. Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study. *UCLA Working Papers in Phonetics* 56. UCLA, Los Angeles.

Lynch, John. 2003. Low vowel dissimilation in Vanuatu languages. *Oceanic Linguistics* 42. 359–406. DOI: https://doi.org/10.1353/ol.2003.0025

Lysvik, Julian K. 2018. An Artificial Language Learning experiment finds no bias against word-final voicing. Poster presented at the *Twenty-Sixth Manchester Phonology Meeting*. Manchester.

Lysvik, Julian K. 2020. *Where does naturalness in phonology come from? Insights from artificial language learning*. Doctoral dissertation. University of Oslo.

Mailhot, Frédéric. 2013. Modeling the emergence of vowel harmony through iterated learning. In Yu, Alan (eds.), *Origins of sound change: approaches to phonologization*. 247–261. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199573745.003.0012

Martin, Alexander. 2017. *Biases in phonological processing and learning*. Doctoral dissertation. École Normale Supérieure.

Martin, Alexander & Peperkamp, Sharon. 2020. Phonetically natural rules benefit from a learning bias: a re-examination of vowel harmony and disharmony. *Phonology* 37. 65–90. DOI: https://doi.org/10.1017/S0952675720000044

Martin, Alexander & White, James. 2021. Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry* 52(1). 227–239. DOI: https://doi.org/10.1162/ling_a_00375

Mooney, Shannon & Do, Youngah. 2018. Learners change artificial languages to constrain free variation in line with typological principles. *15th Old World Conference on Phonology (OCP15)*. London, United Kingdom.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25(1). 83–127. DOI: https://doi.org/10.1017/S0952675708001413

Moreton, Elliott & Pater, Joe. 2012a. Structure and substance in artificial-phonology learning, part I: Structure. *Language and Linguistics Compass* 6(11). 686–701. DOI: https://doi.org/10.1002/lnc3.363

Moreton, Elliott & Pater, Joe. 2012b. Structure and substance in artificial-phonology learning. part II: Substance. *Language and Linguistics Compass* 6(11). 702–718. DOI: https://doi.org/10.1002/lnc3.366

Ohala, John J. 1994. Towards a universal, phonetically-based, theory of vowel harmony. In *Proceedings of the 3rd International Conference on Spoken Language Processing*. DOI: https://doi.org/10.21437/ICSLP.1994-113

Ott, Susan. 2011. *Feld–fällt–fehlt: Untersuchungen zur Phonologie-Morphosyntax-Schnittstelle bei Kindern und Erwachsenen*. Doctoral dissertation. Potsdam: University of Potsdam.

Pavlovia. 2021. Pavlovia: Where behaviour is studied. *Open Science Tools Ltd.* Retrieved October 1, 2021, from https://pavlovia.org/.

Peirce, Jonathan & Gray, Jeremy R. & Simpson, Sol & MacAskill, Michael & Höchenberger, Richard & Sogo, Hiroyuki & Kastman, Erik & Lindeløv, Jonas. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods.* DOI: https://doi.org/10.3758/s13428-018-01193-y

Peperkamp, Sharon & Dupoux, Emmanuel. 2007. Learning the mapping from surface to underlying representations in an artificial language. In Cole, Jennifer & Hualde, José I. (eds.), *Laboratory Phonology* 9. 315–338. Berlin: Mouton de Gruyter.

Peperkamp, Sharon & Skoruppa, Katrin & Dupoux, Emmanuel. 2006. The role of phonetic naturalness in phonological rule acquisition. In Bamman, David & Magnitskaia, Tatiana & Zoller, Colleen (eds.), *Papers from the 30th Boston University Conference on Language Development* (BUCLD 30). 464–475. Somerville, MA: Cascadilla Press.

Pycha, Anne & Nowak, Pawel & Shin, Eurie & Shosted, Ryan. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In Tsujimura, Mimu & Garding, Gina (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics* (WCCFL 22). 101–114. Cambridge, MA: Cascadilla Press.

Qualtrics. 2021. Qualtrics survey software. Qualtrics, Provo, Utah, USA. Retrieved October 1, 2021, from https://www.qualtrics.com/.

R Core Team. 2018. R: A language and environment for statistical computing. Retrieved October 20, 2021, from https://www.r-project.org/.

Saffran, Jenny R. & Thiessen, Erik D. 2003. Pattern induction by infant language learners. *Developmental Psychology* 39(3). 484–494. DOI: https://doi.org/10.1037/0012-1649.39.3.484

Schuler, Kathryn D. & Yang, Charles & Newport, Elissa L. 2016. Testing the tolerance principle: Children form productive rules when it is more computationally efficient to do so. In Papafragou, Anna & Grodner, Daniel & Mirman, Daniel & Trueswell, John (eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. 2321–2326. Austin, TX: Cognitive Science Society.

Seidl, Amanda & Buckley, Eugene. 2005. On the learning of arbitrary phonological rules. *Language Learning and Development* 1(3 & 4). 289–316. DOI: https://doi.org/10.1080/15475441.2005.9671950

Shapp, Allison. 2012. Substantive bias in the learning of harmony patterns. Qualifying Paper. New York University.

Singleton, Jenny L. & Newport, Elissa L. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407. DOI: https://doi.org/10.1016/j.cogpsych.2004.05.001

Skoruppa, Katrin & Lambrechts, Anna & Peperkamp, Sharon. 2009. The role of phonetic distance in the acquisition of phonological alternations. In Lima, Suzi & Mullin, Kevin & Smith, Brian (eds.), *Proceedings of the 39th Meeting of the North-East Linguistics Society* (NELS). Amherst, MA: Graduate Linguistics Students Association.

Skoruppa, Katrin & Peperkamp, Sharon. 2011. Adaptation to novel accents: Feature-based learning in context-sensitive phonological regularities. *Cognitive Science* 35(2). 348–366. DOI: https://doi.org/10.1111/j.1551-6709.2010.01152.x

Sporepedia. 2009. Sporepedia. Retrieved September 10, 2021, from https://www.spore.com/sporepedia#qry=ftr-creature%3Apg-140%3Aview-newest.

Steele, Ariana & Denby, Thomas & Chan, Chun & Goldrick, Matthew. 2015. Learning non-native phonotactic constraints over the web. In The Scottish Consortium for ICPhS 2015 (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow.

van de Vijver, Ruben & Baer-Henney, Dinah. 2014. Developing biases. *Frontiers in Psychology* 5. 1–8. DOI: https://doi.org/10.3389/fpsyg.2014.00634

White, James. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130(1). 96–115. DOI: https://doi.org/10.1016/j.cognition.2013.09.008

White, James & Kager, René & Linzen, Tal & Markopoulos, Giorgos & Martin, Alexander & Nevins, Andrew & Peperkamp, Sharon & Polgárdi, Krisztina & Topintzi, Nina & van de Vijver, Ruben. 2018. Preference for locality is affected by the prefix/suffix asymmetry: Evidence from artificial language learning. In Hucklebridge, Sherry & Nelson, Max (eds.), *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society* 3. 207–220. GLSA

Wilson, Colin. 2003. Experimental investigation of phonological naturalness. *WCCFL 22 Proceedings*. 533–546.

Wilson, Colin. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982. DOI: https://doi.org/10.1207/s15516709cog0000_89

Zhang, Jie & Lai, Yuwen. 2010. Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* 27(1). 153–201. DOI: https://doi.org/10.1017/S0952675710000060