The Obligatory Contour Principle Effects in Phonological Learning

By

Shuxiao Gong
M.A., The University of Kansas, 2019
M.A., University College London, 2017
B.A., Beijing Language and Culture University, 2015

Submitted to the graduate degree program in Linguistics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Jie Zhang

Robert Fiorentino

Allard Jongman

Yan Li

Joan Sereno

Date Defended: 10 May 2022

The dissertation committee for Shuxiao Gong certifies that this is the
approved version of the following dissertation:

The Obligatory Contour Principle Effects in Phonological Learning

Chair: Jie Zhang

Date Approved: 10 May 2022

Abstract

Understanding how native speakers acquire the phonological patterns in their language is a key task for the field of phonology. Numerous studies have suggested that phonological learning is a biased process: certain phonological patterns are easily accessed and learned by the speakers, while others show acquisition difficulties. These differences in pattern learnability can partially explain the observed phonological typology: patterns that are easier to learn occur more frequently in the world's languages, whereas patterns that show learning disadvantages may occur less frequently. Therefore, an important aspect of understanding phonological learning and typology is to understand the nature of these learning biases.

In this dissertation, I approach the issue of phonological learning bias by focusing on the phenomenon of the Obligatory Contour Principle (OCP), i.e., the avoidance of adjacent similar units in the lexicon. Although this pattern has been attested in a wide range of languages (e.g., Cantonese, Dutch, English, Mandarin, Muna, Quechua, among many others) and supported by psycholinguistic experiments, how the OCP impacts phonological learning has yet to be explored in detail. Using a series of artificial grammar learning experiments, I present evidence that the OCP plays an active role in phonological learning, in that similarity avoidance patterns are easier to learn compared to other patterns. Specifically, an OCP-based phonotactic pattern was better learned than a complexity-matching consonant major place harmony phonotactic pattern, and an arbitrary control pattern (Experiment 1). Similarly, an OCP-triggered alternation pattern was better learned than an arbitrarily conditioned alternation pattern (Experiment 2). Experiment 3 further showed that an OCP effect in stem phonotactics could aid the learning of an OCP-triggered alternation.

Typological asymmetries can be observed within the OCP patterns. For example, OCP effects in the labial place of articulation are stronger than other consonant places, and identical sequences sometimes can be exempted from the co-occurrence restriction. Two further artificial grammar learning experiments were conducted to investigate if these asymmetries can be explained by learning biases as well. While a learning bias favoring OCP-Labial was found (Experiment 4), no learning advantage for the identity avoidance was reproduced in a lab setting (Experiment 5).

Based on these experimental results, I argue that phonological learning is a biased process: speakers are equipped with a synchronous learning bias that causes the similarity avoidance patterns to receive an advantage during phonological learning. The typological frequency of the OCP patterns is thus explained by this bias. By incorporating the similarity avoidance into the set of possible biases, this dissertation provides further insights into the nature of phonological learning bias, phonological acquisition, language change, and typology.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This dissertation could not have been completed without the support from numerous wonderful people. I can only try my best to make a full list of you.

First, my deepest gratitude goes to Jie Zhang, my advisor at the University of Kansas. I admire your exceptional knowledge in both theoretical and experimental phonology. Your ideas and critiques are the foundation of this dissertation. Thank you for giving me an opportunity to do a PhD in linguistics here at Kansas five years ago. Five years later, I am proud to say that I am a phonologist trained by you.

Thanks to my other committee members, Robert Fiorentino, Allard Jongman, Yan Li, and Joan Sereno, for valuable comments that helped to improve the quality of this dissertation.

More generally, during my graduate study at KU, I benefited from many other people. Thanks to my teachers: Caitlin Coughlin, Philip Duncan, Alison Gabriele, Utako Minai, Annie Tremblay, as well as my committee members. Thank you for providing me with background in other fields of linguistics and useful experimental and statistical skills for conducting research. Thanks to my linguistics schoolmates: Ruoxi Cheng, Sheyenne Fishero, Hyoju Kim, Jinmyung Lee, Chang Liu, Charlie Redmon, Seulgi Shin, Wenting Tang, Tingting Wang, Delaney Wilson, Xiao Yang, Yuyu Zeng, and many others. Thank you for your companion, friendship, and exchanges of ideas. I am grateful to Yan Li, who taught me how to become a good Chinese language teacher. I really enjoyed working as a Chinese instructor over the past three years.

Outside KU, many others have also assisted my journey in linguistics. Special thanks to Shanti Ulfsbjorninn, who sparked my interest in phonology and persuaded me to pursue an academic career. I don’t know what my life would be like if you did not recommend Gussenhoven and

Jacobs' 'Understanding Phonology' to me at that class break, when you caught me reading a phonetics textbook. Everything started after I got that book. I cannot be grateful enough to San Duanmu, who was patient enough to respond to a random noob undergraduate's emails and kept nourishing my interests in Chinese phonology by giving feedback on my drafts. My core concepts of phonology were heavily influenced by Professor Duanmu's work. Thanks to Michael Proctor at Macquarie University, the instructor of the first official linguistics class I took. I am grateful to your teaching and mentoring. I would also like to thank James White, my advisor at University College London, who first introduced the artificial grammar learning paradigm to me. Thanks to my other teachers at UCL, John Harris and Andrew Nevins, for providing me with a solid background in various phonological theories. I had a fruitful year in the UK. Thanks to James Myers, for insightful discussions about Chinese linguistics and data exchanges. Thanks to Fei Chen, Yan Feng, Gang Peng, and Hao Zhang at SIAT, for familiarizing me with experimental linguistics and tone perception.

My life in America has become unforgettable because of the following people. Thanks to Haseop, Micah, and Xingzhe, for being the best roommates. Thanks to Ben, Mei, Siqi, Wanbin, and Zhenyu, for weekly dinner parties. Thanks to Ruotong, for hosting me in New York City to have the most spectacular vacations. What we had in NYC was definitely epic. Thanks to all the Chinese Meteorology Enthusiasts fellows, for providing 24/7 online chatting support and cutting-edge news reports. Thanks to Jianguo, for being my personal financial advisor. Thanks to Jingyu and Weikun, for hosting me in Washington D.C. Thanks to LCEC fellows, for companionship and food. Thanks to Runsen, for teaching me how to stay sane and for interdisciplinary collaborations. The path I have chosen is not very wide, but it is also not very narrow. Yet somehow, you all managed to polish it up for me.

I extend special thanks to two of my friends. Chulong (龙龙), thank you for being my best friend, mentor, research assistant, librarian, role model, and life coach. Your advice and encouragement helped me to survive in academia. I don't remember how many times you cheered me up and saved me from depression. Yilei (小明), thank you for being my confidant and listening to all my secrets. You are the only person in this world I can be 100% honest with. Therefore, each of our weekly hours-long talks became psychoanalysis, and I felt much relieved for simply letting out my true emotions.

Lastly, most sincere gratitude to my grandma (奶奶), Guangyu Zhao, who single-handedly raised me up and provided me with endless love that my parents could not offer. You taught me what is love, and why is it meaningful to live in these unnatural times. Thank you so much for your contribution and sacrifice to our family, and especially, to your grandson (棒棒). This dissertation is dedicated to you.

The Obligatory Contour Principle Effects in Phonological Learning

1. Introduction

Understanding the nature of speakers' phonological knowledge and the process of phonological learning is a challenging endeavor. Due to its abstract, implicit, and unconscious nature, phonological knowledge can only be indirectly examined by investigating speakers' output, such as data elicited from naturalistic experiments or corpus data. The usage aspects of these data, such as frequency and task-specific behaviors, also make them difficult to interpret. In addition, the input presented to language learners consists of continuous strings of speech signals full of acoustic variation. Learners need to retrieve meaningful units and combine them into hierarchical linguistic structures from this noisy input. Finally, phonological learning in experimental settings is not the same as natural language acquisition; therefore, the unavoidable task-specific meta-linguistic strategies participants bring into the experiments may also add additional noise to the data we collect.

This dissertation aims to provide more insights on the nature of phonological knowledge and its acquisition. To acquire the phonological knowledge of a language, learners are presented with the lexicon of the language and need to make statistical generalizations from this input. However, the learning outcome does not always match with what is evidenced in the lexicon. Results from wug-tests and non-word judgment tasks have shown that this phonological learning is a biased process (Becker et al., 2011; Hayes et al., 2009; Hayes & White, 2013, 2015; Moreton, 2008; White, 2017). In this type of studies, researchers first use a corpus to find out the statistical patterns of the lexicon, and then design novel items to see how speakers will generalize these statistical patterns to unseen forms. The results demonstrate that speakers often make generalizations that are inconsistent with their language experience, that is, speakers may fail to acquire certain patterns in the lexicon (underlearning) or make generalizations that are absent in the input (overlearning). For example, the Turkish lexicon shows that a stem-final consonant voicing alternation is conditioned by the place of the consonant, word length, and the quality of the preceding vowel. However, Becker et al.'s (2011) wug-test results demonstrated that only the consonant place and word length factors influenced speakers' alternation behaviors, while the vowel quality factor had little to no effects. These results indicate that the effect of preceding vowel quality on Turkish voicing alternation represented in the lexicon failed to be picked up by the speakers — an example of underlearning. More specifically, phonological patterns supported by phonetic motivations or conforming to certain complexity-based restrictions and cognitive predispositions may be easier to learn than other patterns. Consequently, the biased learning among different patterns leads to differences in their typological frequency. Patterns favored by the learning biases have an advantage in acquisition and are more likely to be innovated and preserved during the evolution of the lexicon. Patterns that are not supported by learning biases, on the other hand, may not be able to survive from generation to generation. Specifically, in this study, I examine how similarity avoidance biases the learning of phonotactic patterns and phonological alternations. Many of the world's languages exhibit such similarity avoidance effect in their phonotactics. For example, consonants with the same place of articulation occur much less frequently within a root compared with chance level, and this dispreference for similarity within a root has been known as the Obligatory Contour Principle (OCP) effect (McCarthy, 1986).

The causes of the typological prevalence of the OCP effects are the focus of the current study. Phonological typology is determined by various types of cognitive predispositions and/or phonetic principles, known as learning biases (Moreton, 2008; Moreton & Pater, 2012a, 2012b). Patterns that are favored by these biases receive an advantage in phonological acquisition and are thus more likely to occur in language and survive generations of language change. Nevertheless, the nature of these biases has yet to be investigated in greater details. Since the OCP effects have been attested in many of the world's languages (e.g., Arabic, Dutch, English, Hebrew, Muna, Quechua, among many others), it is hypothesized that there is a learning bias behind that favors the acquisition of the OCP patterns, and this study is conducted to test this hypothesis.

In addition to comparing speakers' behaviors in wug-tests and non-word judgment tests with phonological modeling results, artificial grammar learning (AGL) is another useful tool to study the learning biases speakers bring to the task during phonological acquisition. Using AGL experiments, I present evidence to show that OCP patterns are favored by learners during synchronic phonological learning. Specifically, learners who are trained with an OCP-based phonotactic pattern can better internalize such pattern compared to other structurally parallel phonotactic patterns or more complex, arbitrarily determined patterns (Experiment 1). Moreover, participants also find an OCP-triggered phonological alternation easier to learn than other arbitrarily conditioned alternations (Experiment 2), and they can utilize the OCP in stem phonotactics to facilitate the learning of an OCP-triggered alternation (Experiment 3). Finally, Experiments 4 and 5 show that some of the typological asymmetries within OCP patterns can be attributed to learning biases as well.

Based on these experimental results, I propose that the OCP can function as a learning bias during synchronic phonological acquisition to cause the widely attested similarity avoidance patterns found in various world languages. Thus, this dissertation makes a contribution on the nature of phonological learning biases and factors that shape phonological typology.

The remainder of this chapter first lays out the details of various types of learning biases reported in the literature, including structural bias, substantive bias, channel bias, and analytic bias. Their relations, effects on learning, and how the OCP can be accounted for by them, will also be discussed. Then, a review of the artificial grammar learning paradigm is provided.

1.1. Structural Biases

Many researchers agree that phonological learning is a biased process. What remains controversial is which types of biases are effective during phonological acquisition and how to account for these biases in phonological theory. Previous AGL studies in phonological acquisition have suggested two types of biases: structural bias and substantive bias (Moreton & Pater, 2012a, 2012b). A structural bias favors the learning of formally simpler patterns over patterns with high formal complexity, while a substantive bias favors the learning of phonetically grounded or typologically frequent patterns over the phonetically arbitrary ones. Although these two biases can both affect natural language typology, their effects on phonological learning in the laboratory are not always equivalent. AGL studies that directly compare the learning of formally simple patterns with complex ones nearly always find an advantage for the simple ones, but studies that compare the learning of phonetically motivated or typologically frequent patterns with unmotivated ones do not always exhibit an advantage for the former (Moreton & Pater, 2012a, 2012b). Because structural bias effects occur more reliably than substantive bias, it has been proposed that substantive biases, if they exist at all, should be weaker than structural biases (Pycha et al., 2003).

Structural bias can be defined in several ways. The first is the number of relevant features involved. These biases may reflect domain-general cognitive ability and are not necessarily restricted to language learning (Moreton & Pater, 2012a). A group of sounds can be categorized by phonological features, just as objects can be categorized by a set of domain-general attributes.

Shepard et al. (1961) illustrate this feature number complexity using the following example. Suppose that there are 8 geometric figures in total, namely 1) large black circle, 2) small black circle, 3) large black triangle, 4) small black triangle, 5) large white circle, 6) small white circle, 7) large white triangle, and 8) small white triangle. Three binary features ([±size], [±color], [±shape]) can be used to describe these figures. Furthermore, these figures can be partitioned into two groups with equal numbers (i.e., 4 figures in each group) in six ways.

In Type I, only one feature ([+color]) is needed to define the selected group (indicated by the boxed figures). For Type II, two features, color and shape, are needed to define the selected group: [+color, −shape] (black circles) and [−color, +shape] (white triangles). From Type III to Type V, three features are needed, but with certain subsets of values used less. For example, the selected group in Type III can be described as [−size, +color, −shape] (small black circle), [−size, +color, +shape] (small black triangle), and [−color, +shape] (white triangles). The subset, two white triangles, can be described by two features only, so the feature [±size] is used less in this case. Types III, IV, and V all have the same degree of formal complexity. Finally, in Type VI, all three features are fully required to define the selected group.

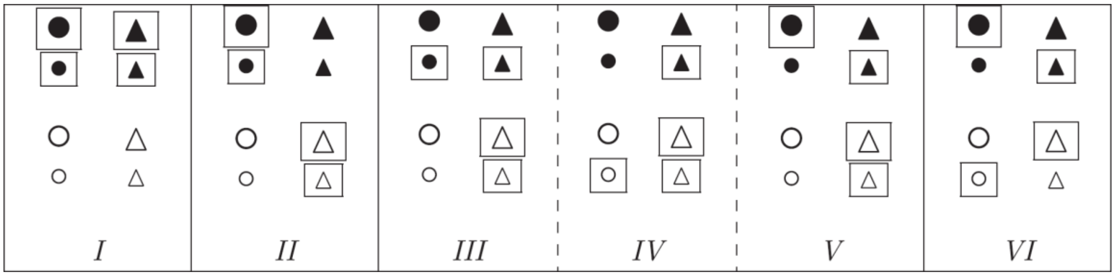


Figure 1 Six possible equal partitions of eight geometric figures defined by binary features of color, shape, and size. Boxes indicate the arbitrary positive class. (Adapted from Moreton & Pater (2012a).)

Psychological experiments have been extensively conducted to test how participants internalize these six concepts. The main finding is that the learning difficulty increases along with the number of features involved (e.g., Shepard et al., 1961): the difficulty hierarchy is Type I < Type II < Types III, IV, V < Type VI. The same mechanism also works for phonological patterns. For example, Skoruppa & Peperkamp (2011) taught French listeners artificial variants of French with different phonotactic patterns on vowels: vowel harmony, vowel disharmony, and a mixed pattern of harmony and disharmony. The mixed pattern had higher structural complexity than the other two patterns, and the results suggested a structural bias effect: the more complex pattern was not learned as well as the other two patterns that referred to only one feature ([back]). In sum, AGL experiments in phonology have shown that Type I systems are better learned than Type II systems (Cristià & Seidl, 2008; Pycha et al., 2003; Saffran & Thiessen, 2003), and also Type II systems are better learned than Type VI systems (Kuo, 2009). In sum, a wide variety of phonological learning AGL experiments shows that a pattern with a single feature is easier to learn than ones with more features.

Formal complexity can also be determined by certain syntagmatic relations among features. For example, assimilation and dissimilation patterns are widely observed in natural languages, because dependencies between the realizations of the same feature may receive a learning advantage. AGL studies have found that patterns based on featural agreement are better learned than those cross-tier agreement patterns. Moreton (2008, 2012) report that participants' learning performance is better when the pattern involves height agreement between vowels, or voicing agreement between consonants, compared with patterns where the dependency is established on the height of the vowel and the voicing of the consonant. These findings indicate that patterns concerning two instances of the same feature are easier to learn than those concerning instances of two different features. Another formal complexity factor is the locality effect, that is, a dependency is more learnable if the distance between the two sounds is shorter than a long-distance dependency. This structural bias is also attested in AGL phonology experiments. Finley (2011) trained participants to learn both long-distance and short-distance sibilant harmony patterns. The results showed that if participants learned the long-distance sibilant harmony in the exposure, they were more likely to generalize this sibilant harmony to unseen cases where the agreement was shorter in distance. However, if they learned the harmony pattern in short distance, they were not able to generalize to the unseen long-distance cases. This learning asymmetry suggests that the default setting for a dependency pattern should be local, and dependencies based on long-distance usually imply short-distance dependencies. This also explains why the OCP is a co-occurrence restriction on adjacent elements.

Finally, a pattern's computational properties can also play a role in defining its formal complexity. For example, a constraint which bans the repetition of certain segments in a word can be viewed as using variables to put restrictions on identical sounds, and this constraint will be computationally different from constraints that do not refer to variable relations (Berent et al., 2012). Another kind of computational property is the distinction between subregular and non-subregular patterns. Subregular patterns are those based on 1) strictly local (*xy, local dependencies), 2) strictly piecewise (*x...y, long-distance dependencies), and 3) tier-based strictly local (a strictly local pattern operated on a phonological tier) (Heinz, 2010). An example of a non-subregular pattern is the so-called first-last restrictions, where the dependencies are established on the first and last sounds of a word. AGL experiments have revealed that only the subregular phonotactic patterns can be reliably learned in the lab, while the non-subregular first-last patterns are not learnable (Lai, 2015).

If we want to test the validity of the OCP in phonological learning, we may start by comparing it with the acquisition of a formally more complex pattern. However, as mentioned earlier, such structural biases may originate from domain-general cognitive ability rather than reflect domain-specific linguistic competence. Therefore, although it still can provide insights on the general mechanisms of human cognitive ability, from a phonological perspective, the role of structural bias in phonological acquisition is considered less ‘interesting’, because we cannot guarantee that the learning asymmetry revealed by the experiments is genuinely due to linguistic factors. For example, in our case, if the two patterns in question, one OCP-based pattern and one arbitrary pattern, differ in formal complexity, and if the results turn out to suggest that the OCP pattern is learned better, then we are not in a position to conclude whether this learning asymmetry is solely based on a domain-general bias about human cognitive ability or a bias that specifically regulates linguistic acquisition.

Another type of learning bias that has been defined to play a role in phonological acquisition is substantive bias. Because the content of this type of bias is largely based on phonetic principles, it is assumed to be domain-specific and reflects true linguistic competence. This bias is reviewed in the next section.

1.2. Substantive Biases

Apart from formal complexity, phonological patterns can differ in terms of other features, and the most important one is whether the pattern can be accounted for by phonetic substance. Patterns that are matched in formal complexity can have very different typological frequencies depending on how these patterns conform to certain phonetic explanations. For example, both vowel-height harmony and consonant-continuancy harmony refer to only one feature, and the dependencies are established based on featural agreement, yet the former is much more commonly attested in the world’s languages than the latter (Rose & Walker, 2004). In terms of OCP, although OCP-Place has been found in many languages as reviewed earlier, consonant major place harmony is non-attested and thus a typological gap (Hansson, 2010). Usually, phonetically natural patterns are caused by phonetic precursors. For instance, the high frequency of vowel-height harmony is due to phonetic vowel-to-vowel co-articulation, while there is no such precursor available for two consonants in a word. Similarly, OCP-Place recurs in natural languages because it is supported by various perceptual and production motivations, as reviewed in §2, but major place harmony has no resort to phonetic explanations.

A phonetically natural pattern is a pattern that is 1) typologically common and 2) phonetically motivated (Hayes & White, 2013). From these two factors, we hypothesize that typologically common patterns are easier to learn than typologically less frequent patterns, and phonetically motivated patterns are easier to learn than phonetically unmotivated patterns. These hypotheses can be subsumed under the term ‘substantive bias’, another type of inductive bias that can affect phonological learning and finally shapes the typology (Moreton & Pater, 2012b). This study tests if there is a learning asymmetry between OCP-Place and consonant place harmony in the acquisition of phonotactic patterns. Crucially, these two patterns match in terms of formal complexity. If it turns out that the phonetically natural OCP pattern is better learned than consonant place harmony, it will serve as evidence for a substantive bias in phonological learning.

Several AGL studies have reported the substantive bias effect in experimental work. For example, Wilson (2006) compared the learning of mid-vowel-triggered velar palatalization on the preceding consonant and that of high-vowel-triggered palatalization. In typology, velar palatalization before mid vowels asymmetrically implies velar palatalization before high vowel. The experimental results are in line with typology, since participants in the mid-vowel training groups can extend the palatalization to high vowel environments, whereas the high-vowel training group is more conservative and only applies this alternation to pre-high-vowel contexts in the testing phase.

In another example, Finley (2012b) tested the substantive bias effect in the learning of vowel rounding harmony. Cross-linguistically, mid vowels are better triggers of rounding harmony than high vowels, and this is backed by phonetic motivations: because high round vowels are perceptually ‘rounder’ than mid round vowels, rounding harmony can better enhance the perceptibility of the round feature between two mid vowels than it does for high vowels (Kaun, 2004). Finley trained the participants to learn a vowel harmony pattern where the suffix vowels agree with the rounding of the stem vowels. In one condition, the vowel harmony is instantiated by mid vowels, and in the other by high vowels. The results indicate that participants in the mid vowel condition learned the rounding harmony better, because they could extend this pattern to novel forms. However, participants in the high vowel condition failed to learn the rounding harmony pattern and their performance was not better than the control group. This study clearly shows that the phonetically natural harmony patterns are easier to learn during artificial grammar learning.

White (2013) conducted a series of AGL experiments on saltation patterns to support for the substantive bias effect. Saltation means that two dissimilar sounds alternate with each other while other relatively more similar sounds do not and are ‘jumped over’. For example, in a hypothetical language, if [p] alternates with [v] while [b] and [f] also exist in the segment inventory, this would be unnatural because the more similar sound pairs [p]~[b] and [p]~[f] do not involve in alternations but the less similar pair [p]~[v] does. White’s experiments show that when participants are trained on patterns that are potentially saltatory, they tend to behave in a way to avoid this saltatory system. And when they are trained on an explicitly saltatory system, they underlearn this pattern compared to non-saltatory systems. Hence, a substantive bias must be in effect to guide the participants to lean towards the phonetically natural patterns during learning.

Nevertheless, Moreton & Pater’s (2012b) literature review suggests that substantive bias is less reliable than structural bias, because there are many studies that also fail to support substantive bias by reporting equal learning effects on phonological patterns that differ only in phonetic naturalness. For example, Pycha et al.’s (2003) study reviewed above finds that vowel frontness harmony and disharmony patterns are learned the same by the participants, and only the formally more complex pattern is learned worse. These findings lead Pycha et al. to conclude that the effects of structural bias are stronger than substantive bias in phonological learning, if substantive bias exists at all.

Recall the first two conditions of the Skoruppa & Peperkamp (2011) study. Two artificial variants of French were created to feature vowel harmony and vowel disharmony. In other words, in Harmonic French, all vowels in words were modified to agree in frontness, while in Disharmonic French, all vowels in words were set to differentiate in frontness. Since vowel harmony is typologically more frequent than vowel disharmony, a hypothesis based on substantive bias would predict that the learning of Harmonic French to be more successful than Disharmonic French. Yet participants learned the two variants equally well, as indicated by their ability to identify words that belonged to the variant they were exposed to. This experiment thus provides no evidence for a substantive bias[1].

---

[1] However, in a more recent study, Martin & Peperkamp (2020) also compared the learning of vowel harmony vs. disharmony patterns in AGL tasks and found a salient substantive bias effect: the vowel harmony pattern was better learned than vowel disharmony. They mentioned that these results might be due to a larger sample size of their exposure phase. In addition, their stimuli were recorded by a French native speaker while their participants were English native speakers. The use of non-native stimuli may cause more phonetic processing of the items and then increase the likelihood of a substantive bias effect.

Some studies have commented that the effects of the substantive bias depend on the type of phonetic naturalness. Glewwe (2019) conducted a series of artificial grammar learning experiments on consonant voicing. Her results suggested that the substantive bias effect only showed up in the learning of position-based voicing patterns (the voicing difference in word initial vs. final positions), whereas the substantive bias effect was missing in voicing-based voicing patterns (final devoicing vs. final voicing). She attributed these conditional emergences of the substantive bias effect to the distinction between perception vs. production-driven phonetic naturalness. Because the position-based pattern is motivated by perception (voicing distinction is more perceptible at onset positions than coda positions), it is more likely to receive a substantive bias effect during learning, while the voicing-direction-based pattern has a source in production (aerodynamically, voiced obstruents are harder to produce at word edges than voiceless obstruents) and is thus not subject to the substantive bias. Another factor that may condition the emergence of substantive bias effect is the mode of the experimental method. It has been found that perception-only AGL tasks sometimes fail to produce a substantive bias effect (Y.-L. Lin, 2019; Nevins, 2010; Zhang & Lai, 2010), but AGL experiments that involve production tasks consistently report such an effect (Moreton, 2008, 2012; White, 2014; Wilson, 2006).

Another concern on the reliability of substantive bias comes from its confound with structural bias. In some studies that claim to have discovered a substantive bias effect, they in fact fail to control for formal complexity among the patterns they test on. For example, Wilson (2003) trained the participants to learn an allomorph alternation between suffixes [-la] and [-na] based on the second consonant of CVCV stems. In one condition, the alternation is built on nasal assimilation: if the second consonant is nasal, the [-na] suffix is used, and otherwise [-la] is used. In the other random condition, [-na] is used when the second stem consonant is dorsal, and otherwise [-la] is used. The results indicate that the nasal assimilation pattern is better learned than the other random pattern, and this is interpreted as a substantive bias effect since nasal assimilation is phonetically more natural than the random agreement pattern. However, these two patterns also differ in formal complexity. The nasal assimilation pattern is based on featural agreement of one single feature [nasal], whereas the random pattern relates to the dependency between two features: a place feature [dorsal] and a manner feature [nasal]. Therefore, the learning advantage found by the AGL experiment can also attribute to a structural bias effect. Similarly, Baer-Henney & van de Vijver (2012) compare the learnability of two suffix allomorph alternation patterns; one based on vowel backness harmony, and one based on the tenseness of the stem vowel. The vowel harmony pattern is considered natural, where the vowel in the suffix agrees with the backness of the stem vowel ([-y] with stem front vowels and [-u] with stem back vowels). The tenseness pattern is considered as phonetically unmotivated, in which [-y] is paired with tense vowels and [-u] with lax vowels. Although native speakers of German learn the vowel harmony pattern better, this pattern, just as in Wilson (2003), is also structurally simpler than the tenseness pattern. The harmony pattern refers to instances of only one feature [back], but the tenseness pattern necessarily relates to two features: [tense] and [back]. Thus, the natural pattern is also formally simpler, and we cannot conclude that the learning advantage is solely due to substantive bias.

If we attempt to uncover a domain-specific substantive effect in phonological acquisition, the caveat mentioned above requires us to compare language patterns that truly differ in phonetic substance by adequately controlling formal complexity. This dissertation serves this goal by investigating the learning asymmetries between OCP-Place and consonant place harmony in Experiment 1 and between OCP-Labial and OCP-Dorsal in Experiment 4. The investigated patterns in these two experiments match in formal complexity and only differ in terms of phonetic motivation. If such asymmetries are discovered, it will serve as solid evidence for the role of substantive bias in phonological learning.

1.3. Channel Bias and Analytic Bias

With regard to the factors that shape the typological asymmetries among phonological patterns, there are two general theories. The first one is Phonetically Based Phonology (Hayes et al., 2004), which argues that phonology is equipped with phonetic substance. Speakers have detailed perceptual and articulatory knowledge, and this knowledge is part of their phonological grammar. This grammatical system enriched with phonetic substance can aid the acquisition of the patterns that conform to phonetic principles; thus the phonetically natural patterns are better learned and more easily to survive during generations than the unnatural patterns. In this view, phonetic substance exists in the synchronic acquisition mechanism, which biases the learning process and causes typological asymmetries. An alternative theory, known as Evolutionary Phonology (Blevins, 2004), claims that phonetic substance influences typology via diachronic sound changes, but does not play a role in synchronic grammar. In the channel of language transmission, which happens through speech perception and production, due to the phonetic substance (perceptibility and articulatory difficulty), some sound structures suffer more easily from misperception and misproduction. These mistakes serve as the phonetic precursors for sound changes, and these diachronic sound changes lead to the present phonological typology. For example, due to articulatory difficulty, voiced obstruents at word-final positions are more difficult to produce. Therefore, speakers may occasionally mispronounce voiced obstruents as voiceless in these positions, causing listeners to misperceive them as voiceless sounds and interpret them as underlyingly voiceless. A final devoicing sound change thus begins and the phonological grammar of this language is influence by the phonetic substance (articulatory difficulty of word-final voiced obstruents). Since the phonetic precursor is universal, similar sound change may happen cross-linguistically, resulting in the high typological frequency of final devoicing. These ‘phonetically systematic errors in transmission’ are known as the channel bias (Moreton, 2008). Crucially, this theory presumes that phonetic substance can only indirectly affect sound patterns via diachronic sound changes, but does not play a role in speakers’ synchronic phonological grammar.

Evolutionary Phonology believes that phonological typology is shaped by the channel biases only. Phonetically Based Phonology, however, argues that analytic bias, “cognitive predispositions making learners more receptive to some patterns than others”, is also effective in shaping the typology (Moreton, 2008, p. 83). This analytic bias represents the detailed phonetic knowledge speakers possess in their synchronic grammatical computational systems, which guides them to better learn the phonetically natural patterns. Indeed, the effects of channel and analytic biases on typology may be overlapping and hard to tease apart. Take the final devoicing pattern as an example again. It could be the result of a diachronic sound change instigated by channel bias, or it could be that speakers have the phonetic knowledge in their synchronic grammar to favor the learning of final devoicing, representing the analytic bias effect. Similarly, the typological frequency of OCP-Place can be attributed to the channel bias factors such as perceptual confusability of adjacent similar sounds and their articulatory difficulty, or we can view it as the result of analytic bias that favors the acquisition of OCP-conforming patterns.

By controlling the strength of the phonetic precursors, Moreton (2008) used the typology of vowel height interaction (HH) and vowel height – consonant voicing interaction (HV) to argue that analytic bias had an independent effect on shaping typology. Typologically, HH patterns outnumber HV patterns, and phonetically, the precursors of both patterns are of the same amount, and even HV pattern shows a more robust precursor. Using AGL experiments, he compared the learning of an HH pattern and an HV pattern, and the results suggested that even though channel bias did not favor the HH pattern, it was still better learned than the HV pattern. This study showed that channel bias alone cannot determine all the typological asymmetries. Because if so, we would predict that the phonetically more grounded HV pattern should be more frequent in the world's languages and better learned during AGL tasks. Moreton then argued that an analytic bias that favors within-tier dependencies (vowel-to-vowel) was in effect to aid the learning of HH patterns.

1.4. Artificial Grammar Learning Experiments

The OCP effect in phonological knowledge has been previously examined by non-word judgment tasks (Berent & Shimron, 1997; Frisch & Zawaydeh, 2001; Gong & Zhang, 2021) and lexical decision tasks (Berent et al., 2001; Kager & Shatzman, 2007). This study employs artificial grammar learning experiments to investigate the learning of OCP-based phonological patterns. AGL has been widely adopted to test the acquisition of phonological patterns and what biases learners bring to the task during that acquisition (Finley, 2011, 2012b; Finley & Badecker, 2009; Glewwe, 2019; Gomez & Gerken, 1999; Y.-L. Lin, 2019; Moreton, 2008, 2012; Moreton & Pater, 2012a, 2012b; Pycha et al., 2003; Reber, 1967; White, 2013; Wilson, 2003, 2006). In this paradigm, participants are exposed to an artificial miniature language encoded with certain phonological patterns absent from their native language during the learning phase. Then in the test phase, participants are evaluated on how well they have internalized the patterns, for example, by asking them to judge whether a novel item belongs to the language they have just been exposed to.

There are several reasons to use artificial grammar to investigate learning biases. First, testing speakers on artificial materials has the advantage of controlling for potential influence caused by lexical frequency, pragmatics, and diachronic change in real language, so we can reduce the impact of learners' prior knowledge during the learning phase, and monitor learning in real time. Second, AGL experiments can allow us to design and manipulate different types of phonological patterns we are interested in, and can directly compare the learning outcome of these patterns (Gomez & Gerken, 1999). Natural language materials usually do not offer such flexibility, since it is rare that all the patterns we want to examine occur simultaneously in one single natural language. Therefore, AGL can provide more direct insights on the acquisition process of linguistic patterns, and can reveal linguistically universal tendencies by soliciting performance data from the participants (Nevins, 2009).

Numerous studies that incorporated artificial grammar learning tasks have suggested that this paradigm is a useful tool for investigating phonological learning processes, including the acquisition of phonotactics, alternations, and their interactions. Previous studies on the acquisition of phonotactics have shown than by the age of 9 months, infants have gained phonotactic knowledge by showing preferences for words that conform to the phonotactic regularities of their native language (Jusczyk & Luce, 1994). Artificial grammar learning experiments have shown that both infants (Chambers et al., 2003; Cristià & Seidl, 2008) and adults (Linzen & Gallagher, 2017; Onishi et al., 2002) are capable of learning phonotactic patterns from short exposure to data. Therefore, both infants and adult learners are able to extract phonotactic patterns from the statistical properties of the lexicon in natural and artificial languages. In addition to the two studies reviewed above, evidence for the successful learning of alternations that are absent in participants' native language abounds in AGL experiments (Finley, 2011, 2012b; Finley & Badecker, 2009; Alexander Martin & Peperkamp, 2020; Myers & Padgett, 2014; White & Sundara, 2014), and such learning is guided by various types of biases. For the interaction between phonotactics and alternation learning, Chong (2021) compared the alternation learning when there is a match or mismatch in the stem phonotactics. The results support such a link between phonotactics and alternation: if the stems obey vowel harmony, it will facilitate the learning of a vowel-harmony-based alternation rule in the suffix, compared to the case when stem phonotactic does not show vowel harmony.

In many AGL studies, two artificial patterns are assigned for participants to learn. After some amount of training on these patterns, the learning outcome of these two patterns will be compared to determine whether there are learning asymmetries between the two. These asymmetries indicate the learning biases learners bring to the learning task (Moreton, 2008; Pycha et al., 2003). For instance, Pycha et al. (2003) trained English speakers to learn a series of suffixation patterns. The suffix, being either [-ɛk] or [-ʌk], was attached to CVC stems based on the frontness of the stem vowel. Three allomorph selection patterns were created: 1) the frontness of the suffix vowel agreed with the stem vowel, 2) the frontness of the suffix vowel disagreed with the stem vowel, and 3) the frontness agreement between suffix and stem vowels was arbitrary and could not be summarized by a single feature. Critically, the formulation of pattern 3 was more complex than the other two patterns, and indeed, the grammaticality judgments in the test phase were more accurate for the agreement and disagreement patterns than for the arbitrary pattern. The asymmetrical learning behaviors among the three conditions indicate that a bias is in effect so that patterns that depend on more features are more difficult to learn.

In another common AGL paradigm, certain cases are withheld in the training; later in the test phase, we can see if the participants will respond to the withheld cases in a biased way. This is sometimes known as the 'poverty of the stimulus' design. For instance, Wilson (2006) designed a language game in which [k g] palatalized to [tʃ dʒ] in certain vowel contexts (before [i e]). English speakers were trained on this alternation and then tested on a mixture of old and novel prompts to measure their palatalization rates in different conditions. In one condition, palatalization only occurred before the vowel [i], and in the other condition only before [e]. In other words, in the training phase of each condition, a crucial vowel context was withheld. The test phase then tested how learners would respond to these unseen cases. The results showed that when trained on the pre-[e] condition, participants were able to extend the palatalization pattern to the novel pre-[i] environments. Such extension did not happen for the participants trained in the pre-[i] condition, where they did not palatalize as much before [e] in testing. Wilson (2006) argued that the different behaviors in the two conditions were phonetically motivated since velar stops and palatoalveolar affricates are more similar before high vowels. The typological consequence is that palatalization is more likely to happen when the following vowel is higher; and palatalization before [e] implies palatalization before [i], but not vice versa. The more successful extension of velar palatalization from pre-[e] to pre-[i] environments than the other way again suggested that phonological learning is a biased process.

To summarize, artificial grammar learning experiments allow us to control the structure of the language materials participants receive, so that we can monitor the learning process that occurs independent of secondary factors such as usage frequency in natural language settings. By comparing the learning outcomes of different patterns, and by looking at how learners generalize to unseen environments and how they go beyond the data they have, we obtain synchronic behavioral data on how speakers internalize different phonological patterns and shed light on how learning processes are potentially biased.

1.5. Summary

In this chapter, I have introduced what phonological learning likely involves, factors that may influence this process, and why it is important to investigate phonological learning (for example, because it can help us understand how phonological typology might take shape). I discussed the properties of the assorted types of learning biases, including structural, substantive, channel, and analytic biases. These labels of learning biases, sometimes, are overlapping and difficult to tease apart. This study is conducted to provide more insights on the nature of these learning biases, how speakers' phonological acquisition is influenced by them, and how the learning biases interact with each other. The experimental paradigm I use in this dissertation to investigate phonological acquisition is artificial grammar learning, a paradigm that has been shown to approximate phonological learning. The next chapter narrows down our discussion of phonological learning to a specific phenomenon: the Obligatory Contour Principle - the focus of this dissertation.

2. The Obligatory Contour Principle

2.1. OCP Effects in Phonotactics

Numerous languages restrict the co-occurrence of adjacent similar segments within a morpheme. This type of phonotactic restrictions has been described by the Obligatory Contour Principle (OCP), which requires segments in a morpheme to disagree in phonological features (Frisch et al., 2004; McCarthy, 1986; Yip, 1988). This principle was originally formulated to describe the tonal disharmony patterns observed in West-African languages (Leben, 1973). The OCP was later extended to account for co-occurrence restrictions on segmental features such as the avoidance of homorganic consonants in the roots of Semitic languages like Arabic (McCarthy, 1986) and Hebrew (Berent & Shimron, 1997). In Arabic, verbal roots consist of a sequence of consonants and actual words are formed by inserting vowels around the root consonants. For example, the verbal root /k t b/ can form words like katab-a 'he wrote', kutib-a 'it was written', etc. with interleaving vowels. Essentially, consonants with the same place of articulation tend not to co-occur in the same root in Arabic. That is, the observed roots with homorganic consonant pairs are fewer than expected if all consonants occurred randomly. Although on the surface, these consonants are interleaved with vowels, they are adjacent on the consonantal tier, where the OCP takes effect. Another interesting pattern is that this autosegmental OCP effect does not behave in an all-or-nothing manner. Instead, the stringency of the OCP depends on the similarity between the targets: the more features the two consonants share, the stronger the co-occurrence restriction is enforced on them (Frisch et al., 2004; Pierrehumbert, 1993). The strength of a phonotactic restriction can be measured by the Observed/Expected value, which is the ratio between the ‘observed’ total number of a certain sequence in the lexicon and the ‘expected’ number of such sequence that would occur if all sounds appeared at chance level. An O/E ratio below 1 means that the target sequence occurs less frequently than expected and such sequence is underrepresented in the lexicon, and an O/E ratio above 1 indicates that the target sequence occurs more frequently than expected and is overrepresented. The smaller the value, the stronger the co-occurrence restriction is reflected in the lexicon. Previous studies have shown that in many languages that exhibit OCP-place effects, the O/E ratio of a consonant pair is negatively correlated with its similarity; that is, the more similar two consonants are, the less likely they would co-occur in a morpheme (Coetzee & Pater, 2008; Frisch et al., 2004). Such gradient OCP effects are observed in many genetically and geographically unrelated languages, including Dutch (Kager & Shatzman, 2007), English (Berkley, 1994), Muna (Coetzee & Pater, 2008), Quechua (Gallagher, 2010), among others. Since similarity avoidance is widely attested in typology, some scholars believe that the OCP effect in the lexicon is a statistical universal (Mayer et al., 2010; Pozdniakov & Segerer, 2007).

2.2. OCP Effects in Alternation

In addition to acting as a morpheme structure constraint that regulates the phonotactics of lexical items, the OCP also plays an active role in dynamic phonological processes. McCarthy (1986) reports that in many languages, syncope rules (vowel deletion in between two consonants) are blocked if the consonants on both sides of the deletion are identical. For example, in Afar, unstressed vowels are often deleted before stressed syllables (a-c). But exceptions occur when the surrounding consonants are identical (d-f). This repetition blocking process is known as antigemination.

Afar vowel syncope

a.	gut**u**ca	gutce	push / he pushed
b.	bar**i**say	barse	let him teach / he taught
c.	dig**i**bte	digbe	she / I married
d.	xar**a**rte	xar**a**re	she / he burned
e.	dan**a**nte	dan**a**ne	she / he hurt
f.	wal**a**lta	wal**a**la	she / he hurt

Yip (1988) examined the OCP effects in a dozen of typologically distinct languages and concluded that the OCP can act as a morpheme structure constraint, a rule blocker, and a rule trigger. For example, in English, onset consonant clusters disfavor homorganic sequences: [bl], [gl] are attested yet [dl] is missing. This can be understood as an OCP-based morpheme structure constraint for onsets. A spirantization rule which changes [t] to [s] when followed by [i] is blocked when the output results in a [ss] sequence (pira[t]e → pira[s]y but hones[t] → hones[t]y). This is an example of the OCP as a rule blocker. Finally, a schwa is inserted between two coronal sibilants during plural formation (do[g] → do[gz], but ro[z]e → ro[zəz]). This epenthesis rule can be analyzed as triggered by the OCP violation of two adjacent sibilants.

Another example comes from Cantonese. Not only does the Cantonese lexicon exhibit a series of OCP-based phonotactic constraints, some phonological rules of language games Cantonese are motivated by the OCP (Yip, 1982). Cantonese has a C_1VC_2 syllable structure. The language game La-mi first reduplicates the base syllable, then replaces the C_1 with [l] in the first syllable and the V with [i] in the second syllable; e.g., the base [jɐt] will be turned into [lɐt jit]. When the base syllable ends with a labial consonant, an additional alternation is observed, such that the labial coda in the second syllable will dissimilate to a coronal (e.g., [sɐ**p**] → [lɐp si**t**]). It is assumed that this dissimilation rule is triggered by the OCP constraint of Cantonese that forbids two labials to co-occur in the same morpheme (Bauer & Benedict, 1997). Moreover, such labial dissimilation happened in the diachronic sound change of Cantonese as well. Historically, the labial codas [p m] changed into [t n] if the syllable onset was also a labial (e.g., *biam → pin, but *liam → lim) (Yue-Hashimoto, 1972). This is why modern Cantonese has a synchronic co-occurrence constraint on two labials in the onset and coda positions, respectively, with only a few exceptions in loanwords and onomatopoetic words.

2.3. The Perception and Production Bases of the OCP

Speakers' tacit knowledge of the OCP has been widely shown in behavioral experiments as well. Berent & Shimron (1997) conducted a nonword acceptability judgment experiment on native Hebrew speakers. Nonwords either violating the OCP or not were presented to the participants, and they were asked to rate how possible the nonwords sounded like Hebrew words on a three-level scale (most possible, least possible, in between). The results showed that OCP-violating items were consistently rated as less likely than non-violating items. Following this study, Berent et al. (2001) conducted a lexical decision task and found that nonwords violating the OCP were rejected faster than non-violating ones by Hebrew speakers. Crucially, the test items were controlled for lexical statistics such as neighborhood density and transitional probabilities, so that the difference in reaction time was indeed due to the OCP violation. In addition to Hebrew, psycholinguistic experiments provided converging evidence on the role of the OCP in various other languages, such as Arabic (Frisch & Zawaydeh, 2001), Dutch (Kager & Shatzman, 2007), and English (Coetzee, 2008). These experiments support the contention that the OCP is a part of the speakers' phonological knowledge, and that this knowledge guides language processing.

Since the OCP effect is found in the phonotactics and phonological processes of a wide range of languages, many functional explanations have been proposed to account for it. Ohala's (1981, 1993) hypercorrection theory states that dissimilar sound changes will take place when listeners assume that phonologically intended repeated features are coarticulatory and then cancel them out in perception. For example, an on-going sound change in Cantonese drops the labialized velar onsets' labial glide [w] before back rounded vowels, but this change does not apply in front of unrounded vowels. What is happening here is that the intended labial feature of the labialized velars is mistakenly perceived as coarticulatory rounding from the following rounded vowels. Therefore, listeners mistreat the intended labialized velars as plain velars followed by a rounded vowel. Boersma (2000) argued that the OCP is the result of a tendency to merge two acoustically similar sounds as a single percept during the perception of surface auditory forms. At the juncture of two similar sounds, it is difficult for the perceptual system to determine when each sound starts and ends. The perceptual system therefore prefers a lexicon that underrepresents adjacent similar units as such a lexicon reduces the likelihood of mapping two similar units onto separate underlying forms and hence facilitates lexical access. In other words, adjacent repetitive occurrences of phonological features tend to be misperceived as single instances of these features (Gallagher, 2010). Moreover, Frisch (2004) argued that repeated similar segments are avoided because the perception of the second sound in such sequences is easily distorted. Word perception involves serial encoding of phonemes. This encoding process requires segments to be activated and deactivated in turn. To facilitate this encoding, after one segment is activated and encoded in the perception, it will immediately be deactivated to get ready for the activation of the next segment. Therefore, if two identical segments occur adjacently, the inhibition and activation of the same unit may overlap so the encoding of the second unit is disrupted. These perception-based theories are supported by experimental evidence. For example, when presented in speech-spectrum noise, CVC syllables whose initial and final consonants share the same place of articulation were identified less accurately than those with different places (Woods et al., 2010).

The OCP has also been argued to have a production basis, as sequential encoding is at play in production planning as well. Dell et al. (1997) argued that there is a turn-off function that deactivates each gesture after it is articulated. When two similar segments are placed in close proximity, they will show competition in articulatory planning because the deactivation mechanism slows down the planning of the latter segment. Moreover, similar segments are co-activated via their shared features. Hence, during articulatory encoding, the similar but unintended counterpart segment is more likely to be mistakenly selected (Dell, 1986). This increased difficulty of production planning of nearly identical segments induces more speech errors in utterances containing adjacent similar segments such as sound misorderings (Dell, 1984). Complete segmental identity may also elicit this production difficulty, and thus they are exceptionally vulnerable to speech errors and repairs in production (Walter, 2007). Results from production studies suggested that vowels in the middle of two consonants with the same place gesture are longer than those between two non-identical gestures, because repeating gross motor gestures is more difficult, and thus requires more time, than switching between different gestures (Walter, 2007). This production-based account for similar-consonant cooccurrence restrictions is called the biomechanical repetition avoidance hypothesis. Repetition of gestures is articulatorily difficult, because it involves sustained activity without a resting period for the involved articulators and requires rapid reversal of an articulator’s trajectory. Meanwhile, Walter (2007) found that when vowels occur in between two consonants that share the same place of articulation, these vowels tend to be articulated with longer durations compared to vowels flanked by two heterorganic consonants, because identical gestures may cause overlapping coarticulations, which costs extra transition time for resolving these overlaps. Therefore, the consonant co-occurrence restrictions on similar places of articulation are caused by an articulatory pressure.

In sum, these functional pressures make adjacent similar units vulnerable to misperception and misproduction. These phonetic groundings represent the channel bias effect on phonological typology. Over time, these OCP-based transmission errors make listeners more likely to perceive adjacent similar sounds as distinct units, and eventually shape the lexicon to favor similarity avoidance via diachronic language changes. Novel forms that conform to the OCP are more likely to enter the lexicon, and OCP-violating items are more likely to be discarded because they are harder to process (Andrew Martin, 2007). As a result, the synchronic lexicon exhibits a statistical pattern of similarity avoidance. Whether the pressures from the OCP are also synchronic is a question that remains unanswered. AGL experiments can help us examine if there is a synchronic learning bias favoring the OCP patterns, so that they are more likely to occur in languages.

2.4. Place Asymmetry

The strength of the OCP effect often varies in different consonant places. In Cantonese, the co-occurrence restrictions predominantly target the Labial feature, while the OCP effect on other places is scant (Do & Lai, 2020; Yip, 1988). For the languages that show gradient OCP effect, by comparing the O/E ratios of the consonant pairs with different places of articulation, oftentimes we can see that the co-occurrence restriction on two labial consonants is the strongest, followed by dorsal and coronal. This tendency is true for Arabic (Frisch et al., 2004), Dutch (Kager & Shatzman, 2007), and English (Berkley, 1994). Typological surveys on the OCP patterns of the world's languages also suggest that OCP-Labial effects are stronger than the co-occurrence restrictions on other places of articulation (Graff, 2012; Pozdniakov & Segerer, 2007; Walter, 2007). For example, Graff (2012) examined the lexicons of 60 languages from 25 major language families, 29 of them are identified with OCP-Place effects. For these 29 languages that exhibit co-occurrence restrictions on place of articulation, 25 of them show OCP-Labial effects, and in 13 of them, the OCP-Labial effects are significantly stronger than the effects of other place co-occurrences. OCP-Coronal effects are also found in 25 languages, yet the coronal co-occurrence restrictions are either significantly weaker or only as strong as the labial co-occurrence effects. OCP-Dorsal effects are the least stable among the major places; only 14 languages have them, and they are always weaker than OCP-Labial effects. These patterns suggest that OCP-Labial is typologically more common than OCP restrictions on other places.

This typological advantage of OCP-Labial is likely rooted in perception. Woods et al. (2010) collected identification data of CVC tokens presented in speech-spectrum noise. The results showed that when the two consonants differed in place of articulation, the words were identified more accurately than words with two consonants sharing the same place. Importantly, the confusion effect was stronger when the two consonants were labial: based on Graff's (2012) reanalysis of the primary data in Woods et al.'s (2010) identification experiments, CVC words with two labials were more likely to be misidentified than words with consonants sharing other features, such as coronal and dorsal. In terms of acoustics, labial consonants have lower amplitude than other places of articulation, because there is little or no oral cavity in front of the constriction to amplify the burst of the labial sounds (Ohala, 1996). Woods et al.'s (2010) explained their identification results by the weaker noise signaling the presence of two labials than that signaling the presence of two coronals or two dorsals.

2.5. Identity Exemption

Another characteristic among the OCP patterns in the world's languages is that identical consonants sometimes are exempted from the OCP-based co-occurrence restrictions. The Semitic languages mentioned above, in fact, allow identical consonants to co-occur in stem-final positions, so that stems like samam 'poison' are not uncommon in the Arabic lexicon. On the surface, these examples look like violations of the OCP. The repeated consonants in stem-final position are analyzed as being derived from an underlying single consonant which surfaces as two identical ones via rightward spreading (Berent et al., 2001; McCarthy, 1986). Moreover, in Muna, roots with homorganic consonants are underrepresented in the lexicon, showing an OCP-place effect, yet identical consonants can co-occur freely within a root. Among the 18 identical consonant pairs in this language, 15 of them have an O/E ratio above 1 (Coetzee & Pater, 2008). Zuraw (2002) has proposed that there is a phonological tendency to promote reduplication-like structures among words in the lexicon. For example, the word 'persevere' is often misspelled as 'perse**r**vere', so that the vowels of the first two syllables will agree. In addition, the tendency for consonants within a root to assimilate is formally described as 'agreement by correspondence'. The phonetically similar sounds within a word tend to establish a correspondence and then undergo assimilation processes in certain features (Rose & Walker, 2004). As a result, similar sounds have a tendency to be alternated into identical pairs. It is possible that the Muna pattern is a reflection of these principles. The exemption of identity patterns thus may result from the special algebraic dependency between two identical units, in which the two units are referenced by variables, instead of being represented by feature matrices. In this case, the functional difficulty of processing adjacent similar units can be avoided by a variable repetition structure, and this option is only available for total identical units, not for non-identical similar units. If identical sounds are represented by features as well, we would expect identity to result in more serious violations of the OCP compared to mere similarity, since two identical sounds share more features than two non-identical similar sounds. Non-word judgment and lexical decision tasks have shown that Hebrew speakers are aware of the phonotactic constraint on identical root consonants at stem-initial positions and can extend this pattern to novel lexical items (Berent et al., 2001; Berent & Shimron, 1997) and these empirical results are better explained by a model that treats identity as repetitions of a variable which can generalize to unbounded sets of elements, but not as sequences of sounds with the same features which only work for this particular set of elements (Berent et al., 2012).

The identity exemption effect causes a typological asymmetry between similarity and identity avoidance. Languages that show similarity avoidance may also show identity avoidance, whereas no language, to my knowledge, only shows identity avoidance but does not show similarity avoidance. This is similar to the situation in the place asymmetry mentioned above, where OCP-Labial languages may also exhibit OCP-Dorsal effects, but very few languages only show OCP-Dorsal but do not have OCP-Labial.

The identity exemption phenomenon may indicate that adjacent identical sounds can be represented as a repeated variable and are less likely to cause perceptual confusion, hence they are allowed to co-occur in the lexicon. However, results from perception experiments do not support this hypothesis. For example, the perception of laryngeal features such as ejective is equivalent among identical consonant pairs and non-identical similar ones. That is, [k’aki] and [k’ak’i] are as difficult to discriminate as [k’api] and [k’ap’i], even though the latter pair contains consonants with different places (Gallagher, 2010). This example shows that the perception of the ejective feature is the same across identical and non-identical consonant pairs. In a monosyllabic word identification experiment, CVC words were presented to the participants in spectrum noise. The results showed that the misperception rates were not higher when the two consonants in the CVC stimuli were identical, compared to two consonants sharing place, manner, or voicing but non-identical (Woods et al., 2010).

To summarize, even though algebraic relations may be established among identical segments, and this relation can be a part of speakers’ phonotactic knowledge (Berent et al., 2012; Gallagher, 2016), this does not imply that identical segments in speech sequences will cause less perceptual confusion than non-identical similar segments. There is no identity exemption in perception: words with identical consonants are not exempted from dissimilatory misperceptions. The existence of identity exemption in some languages like Arabic and Muna is then likely due to a language-specific lexical effect which attempts to maximize the perceptual distance among all words in the lexicon (Graff, 2012). By examining the lexicons in his corpus, he found a correlation between the underattestedness of similarity avoidance and the overattestedness of identical consonant pairs in the lexicon. That is, if the OCP effect in the lexicon is stronger, this language is more likely to exhibit identity exemption. For example, the less likely non-identical labials are to co-occur, the more likely identical labials are to co-occur within words. He further argued that this correlation is driven by the pressure to maximize the overall featural distance among words in the lexicon. Because words like [pVb], [pVf] are underattested, words with identical consonants [pVp] can stand out to better distinguish with words like [pVt], [pVk], and thus overall the words in the lexicon are perceptually more distinct from each other.

Another theory for the special case of identical segments is proposed by Zuraw's (2002) formal account of Aggressive Reduplication. Using a series of active phonological processes in Tagalog, she argues for a markedness constraint that enforces different substrings of a word to be identical. This theory is applied to explain the identity exemption effect in Peruvian Aymara, where a morpheme cannot contain more than one ejective consonant, unless these two ejectives are identical. One piece of evidence that this pattern may have originated from Aggressive Reduplication is that, in these words with multiple ejectives, the vowels directly after them are more likely to be identical than chance, indicating a tendency that a CV sequence is partially reduplicated. The identity exemption in this language can thus be understood as a pressure to have identical substrings in words. Zuraw (2002) further argued that the rankings of these Aggressive Reduplication constraints that enforce piecewise identity are language-specific; hence some languages may tolerate identical segments more, while other languages may still treat identity as a special case of similarity and avoid them in the lexicon.

The exemption of the co-occurrence of identical segments is also predicted by another formal account based on long-distance agreement by correspondence (Rose & Walker, 2004). Consonant harmony patterns, in this analysis, are treated as feature agreement among phonetically similar sounds. These phonetically similar sounds tend to establish a correspondence within a word and undergo agreement in certain features, such as nasality, laterality, voicing, etc. Psycholinguistic evidence has shown that speakers tend to assimilate similar sounds in production and adjacent similar but different sounds are articulatorily difficult to produce. Long-distance agreement by correspondence is thus established to facilitate production. This process can be viewed as the phonologization of the tendency to match certain features among a group of similar sounds. In this case, similar but non-identical consonants may participate in a correspondence and fulfill featural agreement, so that similar consonants might be turned into identical ones.

2.6. OCP in Artificial Grammar Learning

The OCP biasing effect in AGL has been attested in some previous studies as well. For example, Boll-Avetisyan & Kager (2014) created artificial languages composed of concatenated CV syllables without pauses. The shape of such speech stream followed the pattern of ...$P_1P_2TP_1P_2T$..., where P stands for a CV syllable with a labial onset and T for a CV syllable with a coronal onset. In other words, a coronal-initial syllable was always followed by two labial-initial syllables. These languages were presented to Dutch native speakers and their non-word acceptability judgments indicated that PTP sequences were preferred over PPT or TPP sequences. They interpreted these results as an OCP bias effect: since the Dutch lexicon shows an OCP-labial effect (words that contain two labial consonants are underrepresented in Dutch), speakers generalize this knowledge onto this task and thus more likely to put a word boundary in between two labial-initial syllables.

Brohan (2014) designed a series of experiments to test whether OCP-based consonant co-occurrence restrictions can be learned in an AGL setting. Participants were trained with CVCV items (C = {p b f v t d s z k g} and V = {i e a o u}) and later were asked to offer acceptability ratings on a novel set of items to see if they sounded like words that could belong to the language they heard in the exposure. In Experiment 1, there were two experimental conditions: for the OCP condition, all training items did not contain consonant pairs with homorganic places (e.g., *pVbV, *tVsV, *kVgV); and for the arbitrary condition, all training items did not allow labials and coronals to cooccur (e.g., *pVtV, *sVfV). The arbitrary phonotactic rule was selected so that the number of illegal consonant pairs was similar with the number for the OCP condition. The test items contained familiar words that already occurred in the exposure, and novel words that either conformed to the phonotactics reflected in the exposure or not. Statistical analysis showed that the acceptability difference between legal and illegal test items was significantly larger in the OCP condition than in the arbitrary condition, suggesting that participants in an AGL task could better learn an OCP-based phonotactic restriction as compared to learning an arbitrary phonotactic restriction. In Experiment 2, he further tested whether the training of an OCP pattern realized by two places could generalize to the third place, but the results were only marginally significant and inconclusive. Altogether, this study provides evidence that an OCP-based phonotactic pattern can be better learned than an arbitrary pattern.

Chen (2022) conducted an AGL study on the learnability of two types of tonal co-occurrences: terminal-node based tonal OCP and unit-node based tonal OCP. The former assumes that contour tones can be decomposed into sequences of level tones, and any adjacent same level tones are prohibited (e.g., *H-H, *LH-H, *HL-L, etc.). The latter considers an additional unit node for contour tones, and tonal co-occurrence operates on this unit node (e.g., *H-H, *HL-HL, *LH-LH). The results of the two AGL experiments indicate that only the OCP-Terminal constraints had reliable effects on the responses in the test phases, and the OCP-Unit constraints only had marginal effects and were mostly due to explicit training knowledge. In sum, this study provides evidence for the learnability of OCP patterns, when the target sequences are lexical tones.

Moreover, phonotactic learning can also happen when learners are not directly presented with negative evidence for the missing pattern; that is, when they are only exposed to legal patterns, but never explicitly told the missing patterns are ill-formed. This is particularly relevant when the generalization of the missing pattern can be more easily made than the generalization for the legal pattern. For example, participants were able to induce a vowel rounding disharmony regulation from merely being exposed to stems with disagreeing round features (Skoruppa & Peperkamp, 2011). These types of evidence suggest that OCP patterns, which are also based on feature disagreement, should also be learnable in an AGL setting.

3. Current Study

3.1. Goals and Research Questions of the Dissertation

Understanding the acquisition of phonological patterns is a formidable task. Most of the phonological knowledge is unconscious and implicit, so the data we collect from corpora or from experiments can only serve as indirect inferences of this implicit knowledge. Moreover, any type of experiments, no matter how carefully designed, is unlikely to avoid the effects of task-specific meta-linguistic strategies on participants' behaviors. This adds additional noise to our data and makes the extraction of linguistic knowledge even more challenging. The goal of this dissertation is to gain a better understanding of how speakers learn the phonological patterns of their language. After we gain more insights on the nature of phonological acquisition and how this process may be biased, it further sheds light on why certain patterns are better learned and are more likely to survive during generations of language change, and eventually how phonological typology is influenced by various learning biases. Due to the complex nature of phonological acquisition, I focus on one specific phenomenon as a test case. Typological observations suggest that similarity avoidance occurs in a wide range of languages, and this phenomenon is backed by solid perception and production motivations. The current study aims to investigate the role of the OCP in phonological learning, in particular, how the learning process may be biased by similarity avoidance, and how this bias can potentially influence typology.

Numerous studies have demonstrated that phonological learning is a biased process because the learning outcome does not always match what is available in the linguistic input. Speakers may fail to internalize some statistical patterns in the lexicon (underlearning) or make certain assumptions which are absent in the input (overlearning). These disparities between the language input and learning outcome indicate that the learning is skewed by various types of biases (Becker et al., 2011; Hayes et al., 2009; Hayes & White, 2013). Discovering and understanding the nature of these learning biases can help us to better determine the factors that shape phonological typology. In addition, most studies on learning biases used vowel harmony or obstruent voicing alternation as the target patterns, and the focus has been on agreement patterns. Few studies have investigated the effect of disagreement patterns such as the OCP in phonological learning. The implication of studying disagreement patterns in phonological learning is significant, as the evidence available in the learning data may be less obvious. The current study thus offers a comprehensive examination of how similarity avoidance biases the acquisition of phonotactic patterns, alternations, and the typological asymmetries within the OCP.

3.1.1. The Learning of Phonotactics

The phonotactics of a language dictates what sound sequences are permissible in that language. Infants as young as 9 months old have developed phonotactic knowledge of their native language (Jusczyk & Luce, 1994), and non-word acceptability judgment tasks have shown that speakers can utilize their phonotactic knowledge to make gradient word-likeness judgments on various types of non-words (Gong & Zhang, 2021; Hayes & White, 2013). The phonotactic patterns that have been investigated in AGL studies contained a wide range of phenomena, including vowel harmony

(Skoruppa & Peperkamp, 2011), final devoicing (Myers & Padgett, 2014), and other feature-based generalizations (e.g., Linzen & Gallagher, 2017). However, few studies have investigated the OCP effects on phonotactic learning using this methodology. Therefore, this dissertation starts from examining the learnability of OCP-based phonotactic patterns compared to other patterns, since the OCP effect is most commonly realized as morpheme structure constraints on similar sounds in the target language's phonotactics.

3.1.2. The Learning of Alternations

In addition to phonotactics, speakers must also learn the phonological alternations in their native language. Phonological alternations refer to the phenomenon that certain morphemes vary in pronunciation depending on the contexts they reside in. For example, the English past tense morpheme is pronounced as a voiceless [t] when preceded by a voiceless obstruent, whereas it surfaces as [d] after a voiced sound. Wug tests have shown that such knowledge about alternations is an active part of the phonological capability as the same alternations will also apply when speakers are presented with novel stems (Berko, 1958). As reviewed earlier, the OCP may also trigger or block phonological alternations, and thus we are interested in whether OCP-triggered alternations are easier to learn than other arbitrarily conditioned alternation patterns.

Phonotactic regulations and phonological alternations can often be motivated by the same set of constraints. Morphophonological alternations, in many cases, seem to avoid the same structures as the language's static phonotactics may want to prevent. For example, in English, the plural suffix alternates between [-z] and [-s] to agree with the voicing of the preceding consonant (e.g., do[gz] and ca[ts]). Yet at the same time, the phonotactic grammar of English has also clearly stated that consonants with disagreeing voicing cannot co-occur in the syllable coda position (*gs# and *tz#). In traditional rule-based phonology, the phonotactic restrictions are captured by Morpheme

Structure Constraints, which are a property of the language's lexicon. Oftentimes, morphophonological rules apply to satisfy these same Morpheme Structure Constraints. In our case, the English voicing alternation rule seems to do the same thing as the Morpheme Structure Constraints by repeating that there must be no consonant clusters that disagree in voicing. This results in a representational redundancy: the generalization (e.g., voicing agreement) is represented twice; once in the lexicon, and once in the phonological grammar. This is also known as the Duplication Problem (Kenstowicz & Kisseberth, 1977) and later served as a motivation for constraint-based phonology (Prince & Smolensky, 1993). In a constraint-based framework such as Optimality Theory, the voicing agreement in static phonotactics and suffix alternations in English can be covered under the same principle: a constraint penalizing voicing disagreement.

Although the OCP effect is mostly manifested in consonant co-occurrence in the world's languages, OCP-based alternations are indeed reported in diachronic sound changes and synchronic language games, such as in Cantonese (Yip, 1988). Zymet (2020) also reports an OCP-based vowel alternation in which the passive imperative suffix /-u/ of Malagasy undergoes backness dissimilation and changes into [-i] if the stem also contains a [u] vowel. Therefore, the learnability of OCP-based alternations is worth investigating.

In AGL studies that investigated the acquisition of alternations, most of them chose vowel alternations as the target, especially patterns triggered by vowel harmony (Baer-Henney & van de Vijver, 2012; Chong, 2021; Finley, 2012b; Alexander Martin & Peperkamp, 2020; Pycha et al., 2003). For studies that focused on consonants, the alternations are either triggered by sibilant harmony (Finley, 2011), velar palatalization (Wilson, 2006), or conditioned by perceptual similarity (White, 2013). No previous studies have specifically looked at the acquisition of OCP-based alternations. Therefore, this dissertation also compares the acquisition of OCP-based alternations with other arbitrarily conditioned alternations.

3.1.3. The Phonotactics-Alternation Link

As mentioned earlier, since static phonotactics and active alternations can often be described by the same set of constraints (the Duplication Problem), it is hypothesized that the acquisition of static phonotactics and alternations is subject to the same synchronic grammatical mechanism (Hayes, 2004; Tesar & Prince, 2003, 2007). Phonotactic regulations not only account for the systematic gaps in languages, but also function as the trigger of alternations. In addition, while infants as young as 10 months old can show sensitivities to the phonotactic patterns of their native languages (Jusczyk & Luce, 1994), children acquire phonological alternations at a much later age (Sundara et al., 2021). Since phonotactic knowledge is acquired earlier than alternation processes, a further hypothesis that phonotactic knowledge can aid the learning of alternations (Hayes, 2004; Prince & Tesar, 2004; Tesar & Prince, 2007).

Few empirical studies have directly tested this hypothesis, because in natural languages, it is hard to find phonological alternations which differ only in terms of their phonotactic motivations (Pater & Tessier, 2006). Some AGL experiments have been conducted to shed light on this issue, and the results are inconclusive. Pater & Tessier (2006) recruited native English speakers to compare the learnability of two alternations that differed in phonotactic support from English phonotactics. In one language, a C~Ø alternation was conditioned by the tenseness of the stem vowel: stems ending with a lax vowel underwent a stop insertion (e.g., /kɛ/ → [kɛt]) while stems ending with a tense vowel remained the same (e.g., /blej/ → [blej]). This alternation was consistent with English phonotactics, where lax vowels cannot occur in open syllables. In another language, the same C~Ø alternation was conditioned by the frontness of the stem vowel: stems with a front vowel underwent stop insertion (e.g., /lij/ → [lijt]) while back vowels did not (e.g., /vuw/ → [vuw]). This pattern was not motivated by English phonotactics. The results showed that, indeed, the phonotactically supported alternation was better learned. However, since the design of this experiment was based on L1 English phonotactics, it is possible that participants performed better in the first language simply because it aligned with their L1 knowledge; thus the learning asymmetry might not be genuinely due to the aid from phonotactics, but a bias against an unfamiliar pattern.

Chong (2021) avoided potential L1 influence by testing vowel harmony patterns on native English speakers, because vowel harmony is missing from the English lexicon. He tested the learnability of an alternation pattern based on vowel rounding harmony. In the artificial language Chong constructed, the plural suffix had two allomorph [-mi] and [-mu], and the allomorph selection was based on the rounding of the final vowel in the stem. Across different experimental conditions, the manipulation was how much vowel harmony was attested in the stems: the vowels in the CVCV stems either all agreed in rounding, or 75% agreed, or had no evidence for vowel harmony. In training, participants were exposed to both singular and plural forms, and in testing, they were asked to choose the correct plural form for novel singular stems. The results indicated that the rounding alternation in the suffix was better learned if the stems also exhibited vowel harmony, suggesting that stem phonotactics aided the acquisition of alternations. However, when the plural forms were removed from the exposure phase and participants were tested on the rounding alternation purely based on the characteristics of the stem phonotactics, they failed to extend the vowel harmony in the stems to unseen plural alternations.

Do & Yeung (2021) replicated Chong's (2021) experiment on native Cantonese speakers but did not find such a facilitative effect of stem phonotactics on the learning of the alternation. Regardless of whether Cantonese speakers were trained on stem vowel harmony or not, their behaviors in the alternation test were not significantly different. They argued that this null result was due to the L1 of the participants. English is a language that has evidence for phonotactically aided alternations (e.g., the voicing agreement mentioned earlier), yet Cantonese is a language that lacks such mechanism in the phonological system. They proposed that the link between phonotactics and alternation is language specific: if a speaker's native language is equipped with phonotactically motivated alternations, then a bias becomes available to help the learning of alternations backed by phonotactics; if the evidence is missing, phonotactics and alternation will be treated as separate components in the grammar. In the latter scenario, whether the two components agree or disagree in terms of phonotactic motivation becomes irrelevant, and that explains why Cantonese speakers did not respond differently in learning phonotactics-alternation matching or mismatching patterns.

Using the OCP as a test case, this dissertation further investigates such phonotactics-alternation link by looking at whether an OCP-based phonotactic pattern can facilitate the learning of an OCP-triggered alternation rule.

3.1.4. Typological Asymmetries within OCP

Finally, typological asymmetries within the OCP patterns, such as the labial place advantage and identity exemption, suggest that learning biases may be effective in favoring the learning of certain specific OCP patterns than others. For the place asymmetry effect, typology indicates that OCP-Labial occurs more often than other places of articulation (Graff, 2012), and this asymmetry is backed by perceptual and acoustic groundings (Ohala, 1996; Woods et al., 2010). Therefore, a learning bias favoring the OCP-Labial should be effective during phonological acquisition. Specifically, when participants are trained with OCP evidence from other places of articulation only, they should more readily generalize this OCP pattern to the labial place. But when they are provided with OCP evidence from the labial place only, the speakers will be less likely to generalize the OCP effect to other places. Another asymmetry within the OCP patterns is identity exemption. Typology suggests that identity avoidance asymmetrically implies similarity avoidance, so identity avoidance should serve as stronger evidence for the OCP effect than similarity avoidance. Consequently, when speakers are presented with an identity avoidance pattern only, they should be more likely to generalize the OCP effect to other similarity avoidance situations. However, this prediction from the typology, as argued by Graff (2012), is not supported by perceptual grounding. Identical segments, in perception, do not cause less misperception than non-identical similar segments; and identity exemption is a language-specific lexical effect. Due to the lack of perceptual grounding, the bias against identity avoidance may be absent in AGL. In addition, if identity exemption pattern is due to the structure of the lexicon, the miniature language setup of AGL experiments may not be sufficient to establish clear lexical patterns, because lexical patterns can only be gleaned from a larger input that includes sufficient type and token frequency information. This dissertation investigates the nature of the learning biases within various types of OCP effects.

3.1.5. Potential L1 Influence on Artificial Grammar Learning

The artificial language introduced to the participants in an AGL task is analogous to a second language to be acquired. Therefore, participants' L1 knowledge may transfer to the artificial language (L2) and influence how it is learned. Indeed, previous AGL studies have shown that a particular L1 may yield different learning outcomes on different artificial languages (Onnis & Thiessen, 2013). To control for this potential confound, the target patterns carried by the artificial language should be absent from participants' L1. Otherwise, we cannot distinguish whether the preference of the pattern is truly due to acquisition biases or simply L1 transfer effects.

The current study recruited Mandarin native speakers as the participants of the AGL tasks. This is because, as opposed to many other languages, the consonantal OCP-Place effects beyond monosyllables[2] are absent in the Mandarin lexicon. In a lexicon containing 266,642 word tokens (Li & Xu, 2001), CVC sequences where the two consonants are homorganic are neither over- nor under-represented (Boll-Avetisyan, 2012). Therefore, Mandarin native speakers serve as ideal participants for AGL experiments that investigate the consonantal OCP effects across syllables.

3.1.6. Research Questions

To summarize, in order to better understand whether the OCP can bias phonological learning, and if so, how; this dissertation asks the following five research questions: 1) Is an OCP-based phonotactic pattern easier to learn than other phonotactic patterns? 2) Is an OCP-triggered alternation easier to learn than an arbitrarily conditioned alternation? 3) Does an OCP pattern in stem phonotactics facilitate the learning of an OCP-triggered alternation? 4) Does the training in a non-labial OCP pattern lead to greater generalization to other places compared to the training in an OCP-Labial pattern? 5) Does the training in an identity avoidance pattern lead to greater generalization to other non-identical OCP cases or does the training in a non-identical similarity avoidance pattern generalize to identity avoidance?

3.2. Outline of the Dissertation

Corresponding to the research questions raised in §3.1.5, a series of artificial grammar learning experiments is designed to investigate the role of Obligatory Contour Principle in phonological learning. Experiments 1 and 2 test whether static OCP phonotactic restrictions and OCP-based

---

[2] Within a syllable, co-occurrence restrictions based on the OCP have been proposed (Gong & Zhang, 2021; Yi & Duanmu, 2015). But the majority of the proposed phonotactic constraints within a syllable are related to co-occurrence restrictions between consonants and vowels and between vocoids.

alternations are easier to learn than other arbitrary patterns. This is to provide baseline evidence that similarity avoidance is privileged in phonological learning, and that the AGL paradigm can be used to investigate this privilege. The dissertation further tests whether consonant place disagreement in the stems can aid the learning of an OCP-Place-triggered alternation (Experiment 3). Finally, the dissertation will examine whether the typological asymmetries within OCP (place asymmetries and identity exemption) have bases in phonological acquisition. Experiment 4 tests whether the training of the OCP in one place will be equally likely to transfer into another place, and Experiment 5 examines if the training in identity avoidance patterns extends to (non-identical) similarity avoidance patterns. Finally, theoretical implications of the current study will be discussed in the concluding remarks.

4. Experiment 1: Phonotactics

The current study starts with an experiment that investigates the learning of phonotactic patterns. Previous artificial grammar learning experiments have shown that both children and adults can pick up the statistical regularities in the input and acquire many different types of linguistic structures through this statistical learning (Newport, 2016). The current experiment was conducted to compare the learning of an OCP-based pattern with other types of phonotactic patterns.

The first type of phonotactic pattern to be compared with the OCP is an arbitrarily selected pattern that is formally more complex than the OCP-based pattern. While the OCP pattern can be described by only one feature (*$[\text{place}_i][\text{place}_i]$), the arbitrary pattern necessarily requires more than one feature to be represented. As reviewed in §1.1, most of the AGL studies on phonotactics agree that structurally simpler patterns are better learned than patterns referring to more features. The arbitrary pattern in this experiment is designed to investigate this structural bias effect.

The second pattern to be compared with the OCP is a consonant major place harmony pattern. The opposite pattern of the OCP-Place – major place harmony – is missing in natural languages. Based on Clements & Hume's (1995) feature geometry on the internal featural organization of sounds, the place of articulation of a consonant can be categorized into three levels: labial, coronal, and dorsal, which are known as the consonant major places. Therefore, consonant major place harmony refers to the agreement of these major place features within a root. Although consonant harmony effects are attested in some languages, such as the sibilant harmony in Samala and nasal consonant harmony in Yaka, consonant major place agreement is a typological gap among all languages (Hansson, 2010; Rose & Walker, 2004). The absence of consonant major place harmony has been a puzzle for phonological theory. But one possible explanation is that because major place features involve separate articulators, there are no phonetic precursors for coarticulation among different place gestures; as a result, the retention of consonants' major place features is preferred (Rose & Walker, 2004). The comparison of the learning between these two patterns would reflect a substantive bias effect, because the OCP pattern is backed by functional explanations and widely attested in typology, while the major place harmony does not have phonetic support and is a typological gap. As reviewed before, whether the effects of substantive bias can be instantiated in AGL is controversial (Glewwe, 2019; Moreton & Pater, 2012b). But this effect often shows up in natural phonotactic learning. For example, Hayes & White (2013) found that for the phonotactic constraints with equal importance, the phonetically natural ones, for example, ones that conform to the Sonority Sequencing Principle or homorganicity / heterogenicity, had more effect on speakers' nonword judgments, while the unnatural ones had little to no effect. Also, the two patterns in the current experiment, OCP-Place and place harmony, match in terms of structural complexity: both only referring to one feature (place) but differ in directions (a dissimilation pattern *[place$_i$][place$_i$] vs. an assimilation pattern [place$_i$][place$_i$]). The matching complexity ensures that if any learning asymmetry emerges from these two patterns, it would serve as strong evidence to support the substantive bias effect, and we can conclude that the learning advantage of the OCP patterns reflects a domain-specific linguistic effect.

In sum, the current experiment compares the acquisition of an OCP-based phonotactic pattern with a consonant major place harmony pattern and an arbitrary pattern by examining participants' behaviors in the learning of the three artificial languages. In the OCP language, the consonants in the words all had disagreeing places of articulation; in the harmony language, all consonants agreed in terms of place of articulation; and in the arbitrary language, the distribution of the consonants followed an arbitrarily determined complex pattern and was difficult for participants to discern. Participants were randomly assigned to learn either of the three languages. After hearing all the words in the exposure phase, participants offered acceptability judgments for novel words that are distinct from the exposure words.

4.1. Participants

Participants were 90 native Mandarin speakers (32 male, 58 female, mean age = 19.89, age SD = 1.66) without any speech or hearing problems. They were randomly assigned to the three conditions of this experiment: OCP, Harmony, and Arbitrary. Monetary reimbursement was provided for participating in this experiment. Participants were recruited via online social media (WeChat groups, etc.). Most of the participants were college students studying in mainland China.

4.2. Materials

All stimuli in the exposure and test phases were CVCV non-words of Mandarin, where C was drawn from the set of [p^h p f k^h k x] and V from [a ei ou u]. Stimuli were created from all possible combinations of these consonants and vowels excluding identical pairs of the vowels. All CV combinations except [k^{h}ei] and [pou] were existing syllables in Mandarin, although [k^{h}ei] appears in some colloquial forms. All vowels in the stimuli carried the falling tone, because two consecutive falling tones is the most frequent tonal sequence in Mandarin disyllabic words[3] (I. Lin, 2017). This means that some CV combinations were tonal gaps of Mandarin (e.g., [fou3] with the low-dipping tone is a word 'to deny' in Mandarin, but [fou4] with the falling tone is missing). Moreover, resulting CVCV combinations that formed existing Mandarin words (e.g., [fu4fei4], 'to pay') were excluded from the experimental stimuli.

The consonants in the artificial languages did not contain coronal sounds, so that in all 36 possible consonant pairs, half of them agreed in the place of articulation and half did not agree. In the exposure phase of the OCP condition, the stimuli consisted of the 18 disagreeing consonant pairs only (e.g., p^h...k, x...p, etc.), and in the Harmony condition, the stimuli consisted of the 18 agreeing pairs only (e.g., p^h...p, x...k, etc.). For the Arbitrary condition, 9 agreeing (p^h...f, p...p^h, p...p, f...p^h, f...p, k^h...x, k...k^h, k...k, k...x) and 9 disagreeing (p^h...k^h, p^h...x, p...k^h, f...k^h, k^h...p^h, k^h...p, k...p, x...p^h, x...f) consonant pairs were arbitrarily selected to appear in the exposure stimuli. Therefore, consonant pairs in the OCP and Harmony conditions abided by the generalizations that they do or do not agree in place, but no clear phonotactic generalization on consonants can be made in the Arbitrary condition.

In the exposure phase, each consonant pair was assigned five vowel pairs, giving 90 exposure stimuli for each condition. The 90 stimuli in each condition were presented to the participants in a pseudo-randomized order. The test phase included 36 words that were different from any of the

---

[3] Adjacent repeated tones seem to be a violation of the OCP. However, if we analyze the falling tone, a contour tone, as a sequence of level tones (high plus low) (Duanmu, 1994), then two adjacent falling tones do not violate the OCP.

stimuli in the exposure (i.e., novel vowel pairs) with all 36 possible consonant pairs. Half of these test words' consonant pairs were attested in the exposure and the other half were not, while exactly which pairs were attested depended on the experimental condition. Test words were presented in a pseudo-randomized order as well. The test words were the same across the three experimental conditions. The complete lists of all exposure and test stimuli are provided in Appendix 1. All stimuli are transcribed in Pinyin in the appendices.

Words in the artificial languages, as in all other languages created in this study, were recorded by a male native Mandarin speaker. The recording session was conducted in a quiet place using a condenser microphone. Other than applying amplitude normalization to 70 dB using Praat (Boersma & Weenink, 2017), the recorded stimuli were not modified in any other way.

4.3. Procedures

All experiments in the current study were conducted using Gorilla, a web-based platform for running online experiments (Anwyl-Irvine et al., 2020). Before the experiment, participants were presented with an Information Statement and consent was requested to participant in this study. At the beginning of the experiment, participants were told that they will be learning a novel language in the exposure phase and then decide whether some new words sound like they could belong to this novel language they have been listening to in the exposure phase. For each trial in the exposure phase, participants heard the stimulus and were encouraged to read it out. Then they clicked on a button to move on to the next trial. The next-trial button would not appear until the stimulus finished playing. In the test phase, trials were presented in a similar way, except that after hearing the stimulus in each trial, participants were asked to make a binary acceptability judgment by clicking either the 'yes' or 'no' button on the screen. No orthographic forms were provided throughout this experiment. Participants' binary decisions were recorded for analysis.

Previous studies have shown that individual's language aptitude such as executive functions can modulate the performance of learning artificial languages (Kapa & Colombo, 2014). To explore the potential influence of between-subject differences on artificial grammar learning behavior, this experiment, and all other experiments of this study, included a Chinese version of the LexTALE vocabulary test (Chan & Chang, 2018). This is a character-based vocabulary test which contains 60 real Chinese characters and 30 nonce characters with varying orthographic plausibility, i.e., some nonce characters conform to Chinese orthographic rules and some do not by having radicals appearing in an incorrect position. The 90 test characters were presented to the participants in a pseudo-randomized order, and they were asked to judge whether the character they saw was an existing Chinese character. Participants' accuracy scores were recorded for analysis.

After finishing the main experiment and the vocabulary test, participants were redirected to fill out a questionnaire about their basic demographic information and language background. The whole experiment lasted for around 15 minutes.

4.4. Predictions

First, I expect a main effect of consonant pair attestedness on the acceptability rates of the test stimuli, since it represents the effect of training: participants in all three conditions should be more likely to accept novel words with consonant pairs they have heard in the exposure. More importantly, I predict that there should be a significant interaction between condition and attestedness, because this would indicate that the training effect (the acceptability difference between attested and unattested test words) is different across the three experimental conditions. Based on the widely reported complexity effect in phonotactic learning, I first predict that the training effect should be weaker in the Arbitrary condition compared to the other two structurally simpler phonotactic patterns. With regard to the difference between the OCP and Harmony conditions, it is likely that these two patterns are learned equally well because they match in complexity. However, the typology indicates that consonant major harmony is almost non-existent in world languages; therefore, a substantive bias effect may also play a role and renders the OCP pattern more learnable than the Harmony pattern.

4.5. Data Analysis

Participants' acceptability choices were coded as a binary dependent variable of the statistic model (yes was coded as 1, and no was coded as 0). The two independent variables were experimental condition (OCP vs. Harmony vs. Arbitrary) and the attestedness of the consonant pair in the test items (Attested vs. Non-Attested). The data were analyzed by mixed-effects logistic regression models using the *lme4* package (Bates et al., 2015) of the R software (R Core Team, 2018). The fixed effect structure included the two independent variables and their interaction. The random effect structure was determined by backwards stepwise model comparisons. We started with the full random effect structure with ransom intercepts and slopes for participants and items. Each random effect factor was taken out one by one; the simpler model was then compared with the more complex model using the likelihood ratio test, and the random factor was excluded if it did not significantly improve the model.

4.6. Results

Among the 90 participants, 23 of them (9 from the OCP condition, 7 from the Harmony condition, and 7 from the Arbitrary condition) gave a 'yes' response to all test words. Their results were excluded from the data analysis because they might not have learned any pattern or simply did not pay attention to the experiment.

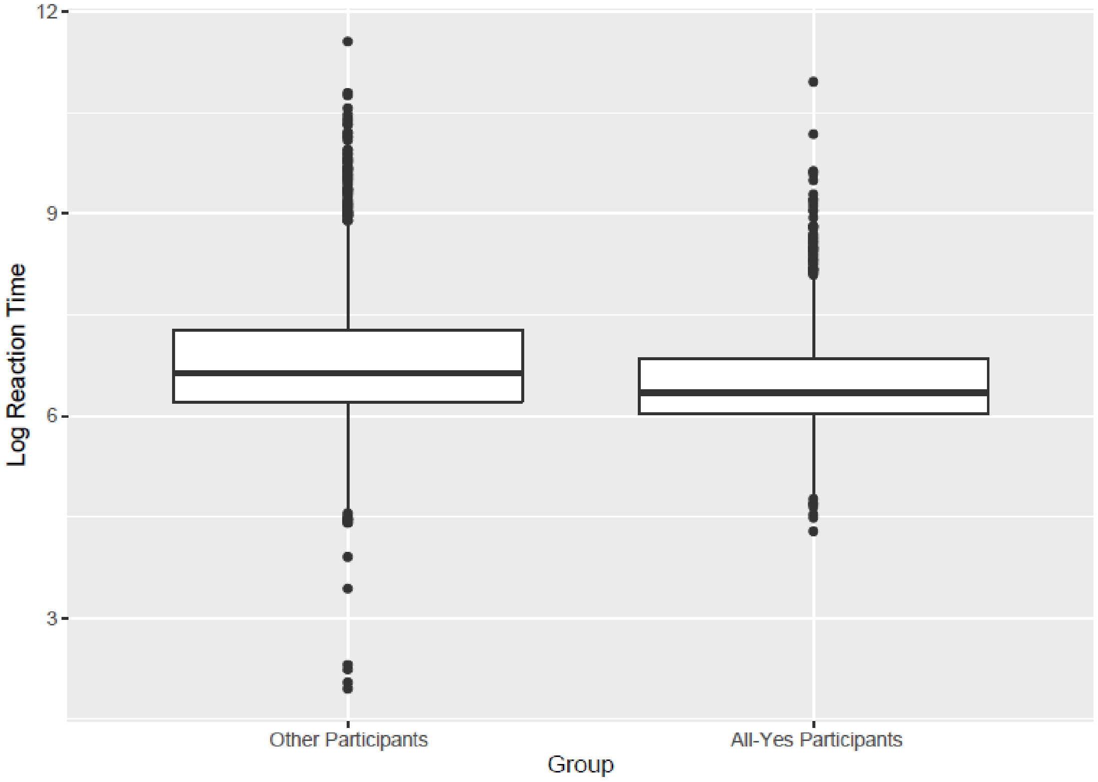


Figure 2 The distributions of log-transformed reaction time among 'all-yes' participants and other participants.

The distributions of log-transformed reaction time for all the test trials across participants who gave 'yes' responses to all trials (mean = 6.535, SD = 0.865) and other remaining participants (mean = 6.824, SD = 0.958) are shown in Figure 2. A two-sample t-test between the log reaction times on each test trial of these 23 participants and the remaining 67 participants showed that these all-yes participants spent significantly less time on doing the test phase tasks ($t = 7.90$, $p < .0001$), which further indicates that they might not have engaged with the task enough.

The mean endorsement rates (percentage to give a 'yes' response for the test words) of each participant across different conditions and different consonant pair attestedness are shown in Figure 3.

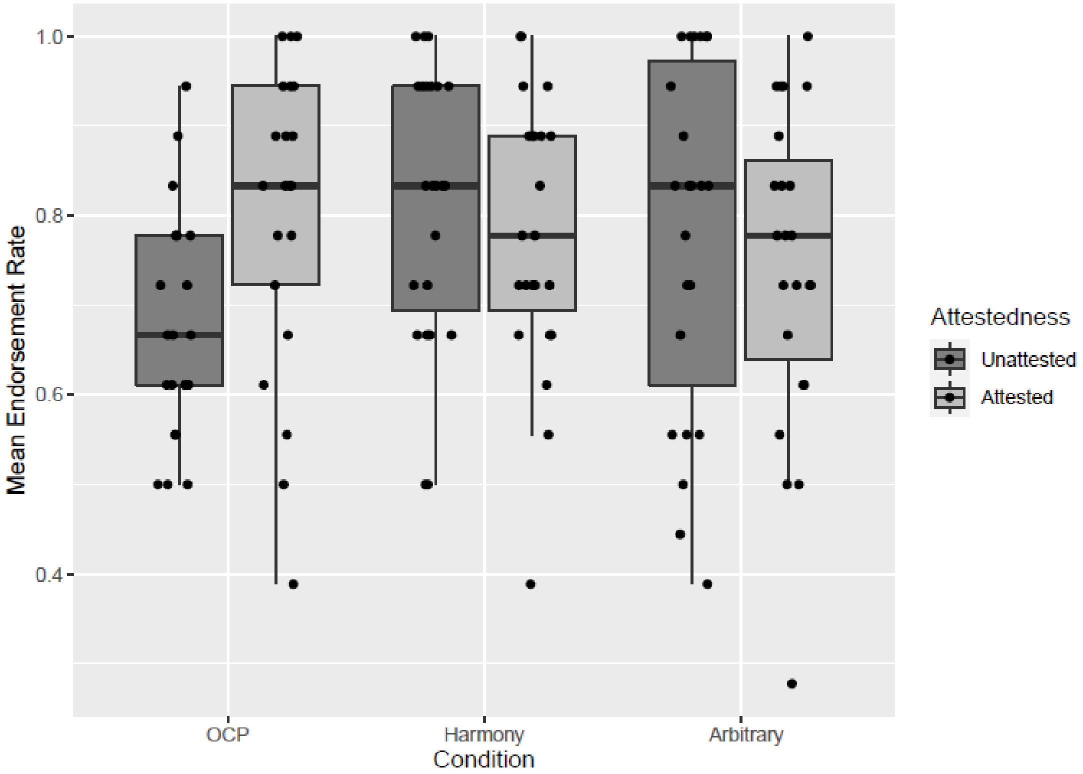


Figure 3 Proportion of 'yes' responses to the test stimuli in Experiment 1 grouped by condition and consonant pair attestedness. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

Logistic mixed-effects models were fitted to participants yes/no acceptability responses (yes = 1, no = 0) on the test words using the *lme4* package (Bates et al., 2015) in the R software (R Core Team, 2018). Fixed effects included a three-level factor experimental condition (OCP, Harmony, and Arbitrary), a binary factor consonant pair attestedness (whether the consonant pair occurs in the exposed language), and their interactions. For the random effects to be included, I started with the most complex structure, but the model failed to converge. I then subtracted random slopes from the model and the maximal model to converge was the one with random intercepts for participant and test word, and random slopes for attestedness by participant and by test item. Random slopes were not attempted for participant because participants in this design were nested within conditions. The maximal model's parameter estimates are shown in Table 1. The baseline for attestedness is unattested and the baseline for the experimental condition is Arbitrary. The *p*-values were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

	Estimate	Std Error	*t* value	*p* value
(Intercept)	1.4960	0.2518	5.9408	0.0000***
Attested	-0.2955	0.2713	-1.0891	0.2761
OCP	-0.5871	0.2982	-1.9689	0.0490*
Harmony	0.2748	0.3582	0.7670	0.4431
Attested:OCP	1.1089	0.3220	3.4439	0.0006***
Attested:Harmony	0.0005	0.3881	0.0012	0.9990

Table 1 The maximal model for endorsement rates.

The significant effect of OCP indicates that for the unattested items, participants in the OCP condition gave significantly fewer 'yes' responses than participants in the Arbitrary condition. More importantly, there is a significant interaction between item attestedness and experimental condition. The acceptability difference between the two types of test items was significantly higher in the OCP condition than in the Arbitrary condition, and the difference was not significantly different among participants in the Harmony condition and the Arbitrary condition. We therefore conclude that the OCP pattern was better learned than the Harmony and the Arbitrary patterns.

4.7. Discussion

The results of Experiment 1 suggest that phonotactic learning is guided by a structural bias, because the phonotactic pattern in the OCP condition was better learned than the more complex pattern as encoded in the Arbitrary condition, which echoes with most previous AGL studies that structurally simpler patterns are easier to learn (Kuo, 2009; Pycha et al., 2003). The OCP-Place constraint represented in the OCP condition can be summarized as *[place$_i$][place$_i$], which only refers to one feature, yet the pattern in the Arbitrary condition must refer to more than one feature to be formally described.

The learning asymmetry between the OCP and the Harmony conditions also reveals a substantive bias effect. The generalizations for the phonotactic patterns in the OCP and Harmony conditions only refer to one feature: *[place$_i$][place$_i$] and [place$_i$][place$_i$], respectively; thus these two patterns match in terms of complexity. By comparing the outcomes of these two patterns, we also found a learning asymmetry that the OCP pattern is better learned than the Harmony pattern. This constitutes strong evidence for a substantive bias effect in phonotactic learning, because the difference between these two patterns only lies in phonetic substance and typological attestedness: OCP-Place is motivated by perceptual and production factors and is widely attested in the world’s languages, while place harmony is a typological gap. The learnability difference between these two conditions also indicates that there is a learning bias favoring an OCP-based phonotactic pattern, because the reverse pattern – major place harmony – was not properly learned in the experiment.

The improper learning of the Harmony pattern is particularly interesting, because long-distance consonant harmony patterns like sibilant harmony can be successfully learned in AGL studies (Finley, 2011, 2012a). Meanwhile, arbitrary and typologically unattested sound patterns are also shown to be learnable (Koo & Callahan, 2012; Seidl & Buckley, 2005). For example, vowel disharmony is rarely attested in world languages but it is learnable in Pycha et al.’s (2003) AGL experiments. For the studies that tested arbitrary phonotactic patterns, e.g. Koo & Callahan (2012), they often target consonant co-occurrence of a single pair (e.g., [m] always occurs after [s]), so the learning might be quite specific. The remaining question is that, given that neither vowel disharmony nor consonant place harmony is attested in typology, and neither has clear phonetic motivation, why the current experiment showed that place harmony is unlearnable, while vowel disharmony was found to be learnable (Pycha et al., 2003; Skoruppa & Peperkamp, 2011)? My hypothesis here is that consonant major place harmony is a systematic typological gap, and a substantive bias in the acquisition mechanism impedes the proper learning of such a pattern. In addition, under closer scrutiny, when extending the stimulus length to four syllables, Martin & White (2021) reported that a general vowel harmony pattern across the whole quadrisyllabic word ([+F][+F][+F][+F] or [-F][-F][-F][-F]) was better learned than a general disharmony pattern ([+F][-F][+F][-F] or [-F][+F][-F][+F]). Martin & Peperkamp (2020) used French accented stimuli to test the learnability of vowel harmony and disharmony patterns on English native speakers, and the results showed that vowel harmony was more easily learned than disharmony. They argued that this finding is due to the more phonetic processing of the materials when participants were presented with sounds from a foreign language. Therefore, it is possible that the proper learning of vowel disharmony patterns in earlier studies (Pycha et al., 2003; Skoruppa & Peperkamp, 2011) is due to alternative strategies such as memorizing specific vowel dependencies (e.g., suffix [-u] always occurs after a preceding [e]), but in fact, vowel disharmony still receives a learning disadvantage, due to its typological infrequency. Finally, it should be noted that the bias against these typologically rare patterns is a ‘soft’ bias, i.e., it is not the case that consonant place harmony or vowel disharmony is completely unlearnable. Compared to other typologically attested and phonetically natural patterns, they are in a disadvantaged position to be acquired, but if given enough exposure, these patterns may still show learning effects, as many studies have indeed reported the successful learning of arbitrary phonological patterns.

Since the current experiment used a between-subject design, there is a concern that any learning difference could result from between-subject variations. To mitigate this concern, I added participants' vocabulary test scores as a measure for individual differences into the mixed-effects model to see if this factor can significantly improve the model (Kapa & Colombo, 2014). The mean vocabulary test score for the participants was 79.07 (sd = 2.69) in the OCP condition, 73.23 (sd = 8.49) in the Harmony condition, and 76.67 (sd = 3.84) in the Arbitrary condition. The results showed that vocabulary scores did not have an effect on participants' yes/no responses. By comparing the model with and without the vocabulary score predictor using the chi-squared test, the model with such predictor was not a significantly better one ($\chi^2 = 2.0995$, $p = .1473$).

As a further proof-of-concept for the between-subject design, for the 60 participants in the OCP and Arbitrary conditions, I invited them back to take the other condition that they did not participate in after a one-month interval in order to create a within-subject design for these two conditions. Among the 30 participants who participated in the OCP condition first, 23 of them came back for the Arbitrary condition; among the 30 participants who participated in the Arbitrary condition first, 25 came back for the OCP condition. For the combined dataset, a participant's data were not considered if the participant attended only one session, or if the participant offered 'yes' responses to all of the test trials in any one session. Consequently, 34 participants' data remained, among whom 17 participants finished the OCP condition first and the other 17 finished the Control condition first.

For this dataset, a logistic mixed-effects model was built using consonant pair attestedness and experimental condition to predict participants' yes/no responses. The model summary and the boxplot in Figure 4 indicate that there is still a significant interaction between attestedness and condition, showing that the OCP pattern was better learned than the Arbitrary pattern. Furthermore, the order in which participants took the two sessions (OCP first or Arbitrary first) did not affect the learning outcome, because adding the session order predictor did not significantly improve the model (χ^2 = 1.1475, p = .2841). This within-subject verification further confirms that the learnability difference across conditions observed in the main experimental sessions was not due to individual propensities in the AGL tasks. Notice that in this dataset, participants' behaviors in the Arbitrary condition showed a different pattern compared to the full dataset: test items with attested consonant pairs received numerically higher ratings in this restricted dataset, while the full dataset showed the reverse pattern. However, according to the regression models, in which the baseline for the experimental condition was Arbitrary, the main effect of consonant pair attestedness in neither dataset is not significant, so we cannot conclude that participants in the two datasets truly behaved differently in the Arbitrary condition.

	Estimate	Std Error	*t* value	*p* value
(Intercept)	1.2803	0.2280	5.6142	0.0000***
Attested	0.1900	0.2297	0.8272	0.4081
OCP	-0.4176	0.1679	-2.4871	0.0129*
Attested:OCP	0.4645	0.2040	2.2772	0.0228*

Table 2 Model for endorsement rates based on the verification experiment dataset

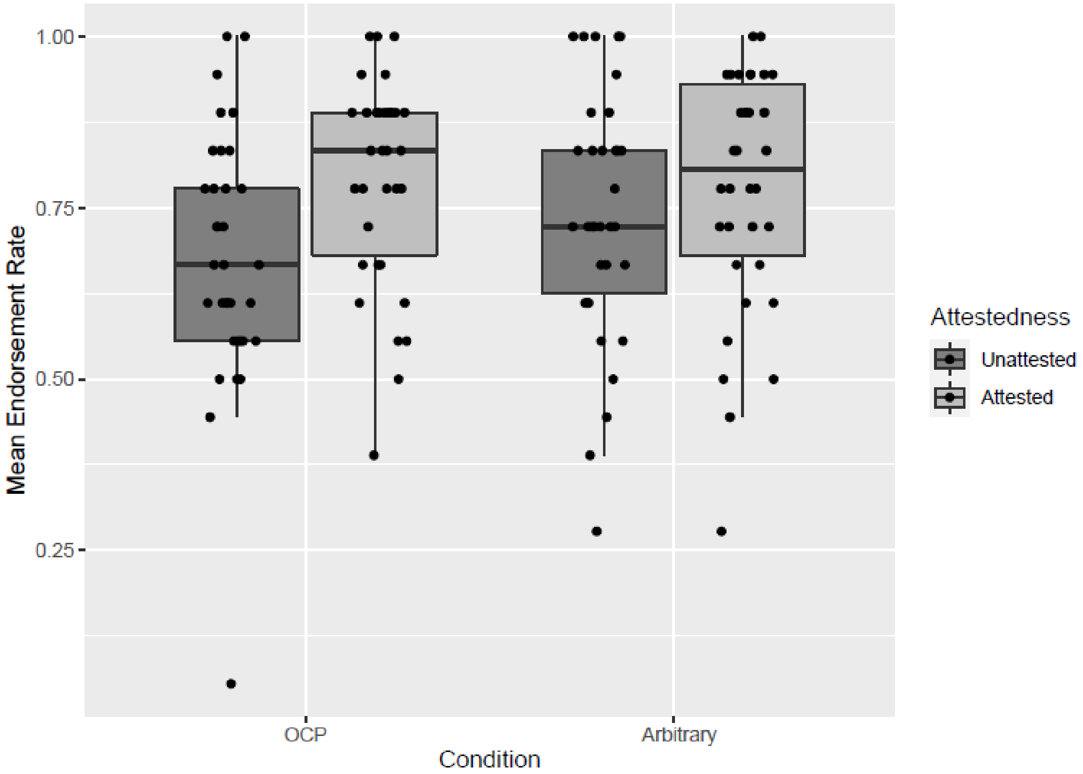


Figure 4 Proportion of ‘yes’ responses to the test stimuli in the verification experiment grouped by condition and consonant pair attestedness. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

The results of these within-subject verification processes provide evidence that the between-subject design is a valid design for AGL experiments that compare the learning of different patterns. Hence, the following experiments in this study all adopted the between-subject design.

I also considered a number of nuisance factors that may have had an influence on participants’ acceptability judgments in the test phase. First, participants might judge the two Mandarin segmental gaps [k^hei4] and [pou4] (mentioned in Section 4.2) as less acceptable than other existing syllables. I added a binary factor that distinguishes the test stimuli based on whether they contain one of these two segmental gaps into the model shown in Table 1, and maximum likelihood ratio test suggested that this factor did not significantly improve the model (χ^2 = 0.0113, p = .9152). This means that participants' behaviors on test items containing these segmental gaps were not different from other items. Second, although the vowels in the stimuli were randomly assigned, all participants were trained on the same set of stimuli in each condition. That means the vowel distributions for each participant were also fixed, so that they may have accidentally picked up on vowel co-occurrence patterns as well. If certain vowel pairs occurred less frequently in the training, would that affect participants' acceptability judgment in the test phase? To rule out this possibility, I looked at the vowel combinations in the training to see if there were discernable patterns, and if so, whether these patterns had an effect on participants' judgments. For example, there were altogether 12 possible vowel pairs (16 pairs based on factorial combinations among [a, ei, ou, u] minus 4 identical pairs) for the CVCV test words, and there were 90 items in the exposure phase. On average, each vowel pair should occur 90 ÷ 12 = 7.5 times, yet the vowel combination [...ou...ei] only occurred three times in the exposure items of the OCP condition. If vowel statistics influenced acceptability, we may expect that test items carrying this vowel combination (e.g., [koupei]) would receive lower ratings than other items in the OCP condition. Adding a factor that distinguished test items with the [...ou...ei] vowel pair from others did not significantly improve the statistical model for acceptability in the OCP condition (χ^2 = 0, p = 1). Third, participants may have accidentally noticed the similarity between the two vowels in the exposure items and make generalizations. I used the feature matrix from Hayes (2009) to define the five monophthong vowels used in the experiment and counted how many features were different among all the vowel pairs. When the comparison was between a monophthong and a diphthong, the monophthong was repeated twice.

	high	low	front	back	round
i	+	-	+	-	-
u	+	-	-	+	+
e	-	-	+	-	-
o	-	-	-	+	+
a	-	+	-	-	-

Table 3 Feature definition of the monophthong vowels used in the experimental stimuli

For example, for the two vowels [a] and [u], they differ in [high], [low], [back], and [round], so their distance was counted as 4. For the vowel [a] and the diphthong [ei], [a] and [e] differ in [low] and [front], [a] and [i] differ in [high], [low], and [front], so the distance between [a] and [ei] was counted as (2 + 3) / 2 = 2.5. Using this method, I calculated the vowel distance for all exposure items, and crucially, the mean vowel distances across the three artificial languages in the experiment are the same. Adding this vowel distance factor did not significantly improve the model for acceptability ($\chi^2 = 0.5821$, $p = .446$).

The main findings in Experiment 1 are that speakers learned an OCP-based phonotactic pattern better than an arbitrarily designed consonant co-occurrence pattern. Meanwhile, the learning of the OCP pattern was also significantly better than that of a consonant harmony pattern which had the same formal complexity. These results thus provide strong evidence for a structural bias and a substantive bias in phonotactic learning. While most AGL experiments have shown a structural bias effect, the substantive bias does not always emerge in these studies (Glewwe, 2019; Moreton & Pater, 2012b). This experiment hence contributes an argument for the existence of substantive bias in that the phonetically motivated OCP pattern is better learned than its structural counterpart: consonant major place harmony, which is generally missing in real languages. Meanwhile, I have conducted additional analyses to ensure that the observed learnability differences are not due to between-subject variations or other accidental distributional patterns in the exposure phases.

A remaining issue from previous discussion is why the current experiment found a substantive bias effect, while other studies (such as Skoruppa & Peperkamp (2011) on vowel harmony vs. disharmony and Glewwe (2019) on the consonant place contrasts in word-initial vs. final positions) failed to show such an effect. This may not be due to the amount of training, as Skoruppa & Peperkamp used 304 target words in their exposure phases, which highly exceeds the stimuli number of the current experiment (90). It is likely that the difference lies in the nature of the patterns being investigated. The motivation for vowel harmony is mostly production-based: vowel-to-vowel co-articulation serves as the phonetic precursor for vowel features to agree in production. On the other hand, the OCP pattern investigated here is both perceptually and productively motivated, as reviewed in §2.3, and thus it is more likely for a substantive bias effect to emerge (Glewwe, 2019).

5. Experiment 2: Alternation

Experiment 1 provided evidence for the learnability of OCP-based phonotactic patterns, it is natural to ask the question: are OCP-based alternation patterns also easier to learn? The current experiment is designed to test whether a phonological alternation triggered by the OCP is better learned than an alternation happened on an arbitrarily selected set of consonant pairs. Participants were asked to learn a simple language game where labial consonants became coronal in certain contexts (Gallagher, 2013; Wilson, 2006). In one language, the alternation happened when the consonant pairs of the stimuli agreed in place of articulation, that is, the second labial consonant changed into coronal to avoid an OCP-labial violation. The alternation did not apply when the consonant pairs disagreed in place. This is referred to as the OCP condition. In a second language, such alternation happened in 9 arbitrarily chosen consonant pairs, while the remaining pairs did not exhibit alternations. This is referred to as the Arbitrary condition. Participants were randomly assigned to either of the two conditions. After going through the language game training, they were presented with novel prompts to test if they would apply the just-learned alternation to them.

5.1. Participants

Participants were 60 native Mandarin speakers (26 male, 34 female, mean age = 20.6, age SD = 1.96) without any speech or hearing problems. They were randomly assigned to the two conditions of this experiment: OCP and Arbitrary. Monetary reimbursement was provided for participating in this experiment. Participants were recruited via online social media (WeChat groups, etc.). Most of the participants were college students studying in mainland China. Because the condition was randomly assigned by the Gorilla platform, there was an imbalance in the final dataset between the number of participants in the two conditions: 26 for OCP and 34 for Arbitrary.

5.2. Materials

The stimuli for this experiment were $C_1V_1C_2V_2$ Mandarin non-words, where C_1 and C_2 were all possible combinations of [p^h p f k^h k x] in C_1 and [p^h p f] in C_2, and V_1 and V_2 were all non-identical pairs of the vowels [a ei ou u]. In addition, filler items were created for the test phase, which took the same vowel patterns and all combinations of [tʂ ʂ] in C_1 and [p^h p f] in C_2. C_2 contained only labial sounds, no dorsals, because it was the consonant that underwent the alternation in the language game.

There were altogether 18 possible consonant pairs; 9 of them agreed in place of articulation (both sounds are labial), and 9 had disagreeing places. In the OCP condition, a prompt whose consonants were from the 9 agreeing pairs were paired with a response in which the second consonant changed to coronal (e.g., prompt [p^{h}ap^{h}u], response [p^{h}at^{h}u]), whereas a prompt whose consonants were from the 9 disagreeing pairs were paired with an identical response (e.g., prompt [k^{h}ap^{h}u], response [k^{h}ap^{h}u]). Therefore, this condition represented an alternation pattern similar to the labial dissimilation process reported in the historical sound change of Cantonese, where two labial consonants are not allowed to co-occur within a stem and the second one changes into a coronal (Yip, 1988). In the Arbitrary condition, prompts carrying 9 arbitrarily chosen consonant pairs underwent the alternation (4 agreeing pairs: p^h...p, p^h...f, p...f, f...p^h, and 5 disagreeing pairs: k^h...p^h, k^h...p, k...f, x...p^h, x...p), whereas prompts with the remaining 9 pairs did not alternate in the responses. For the 9 prompts that were arbitrarily selected to be the alternating prompts, the C$_2$ [p^h p f] each occurred three times, while the occurrence of C$_1$ was not controlled. Table 4 illustrates the stimuli of the two conditions using V as the placeholder for vowels. The 18 consonant pairs combined with 2 vowel combinations, giving rise to 36 prompts for the exposure phase in each condition. The complete lists of all exposure and test stimuli are provided in Appendix 2.

OCP				Arbitrary			
alternating		non-alternating		alternating		non-alternating	
prompt	response	prompt	response	prompt	response	prompt	response
p^{h}Vp^{h}V	p^{h}Vt^{h}V	k^{h}Vp^{h}V	k^{h}Vp^{h}V	p^{h}VpV	p^{h}VtV	p^{h}VpV	p^{h}VpV
p^{h}VpV	p^{h}VtV	k^{h}VpV	k^{h}VpV	p^{h}VfV	p^{h}VsV	pVp^{h}V	pVp^{h}V
p^{h}VfV	p^{h}VsV	k^{h}VfV	k^{h}VfV	pVfV	pVsV	pVpV	pVpV
pVp^{h}V	pVt^{h}V	kVp^{h}V	kVp^{h}V	fVp^{h}V	fVt^{h}V	fVpV	fVpV
pVpV	pVtV	kVpV	kVpV	k^{h}Vp^{h}V	k^{h}Vt^{h}V	fVfV	fVfV
pVfV	pVsV	kVfV	kVfV	k^{h}VpV	k^{h}VtV	k^{h}VfV	k^{h}VfV

fVpʰV	fVtʰV	xVpʰV	xVpʰV	kVfV	kVsV	kVpʰV	kVpʰV
fVpV	fVtV	xVpV	xVpV	xVpʰV	xVtʰV	kVpV	kVpV
fVfV	fVsV	xVfV	xVfV	xVpV	xVtV	xVfV	xVfV

Table 4 Stimuli for the OCP and Arbitrary experimental conditions

The test items consisted of 18 novel prompts with all 18 consonant pairs and 6 fillers beginning with either [tʂ] or [ʂ]. Although they contained the same consonant pairs as the prompts in the exposure phase, all prompts in the test phase had novel vowel pairs and thus participants had not heard them before. The test prompts were the same across the two experimental conditions.

All experimental stimuli were recorded by a male native Mandarin speaker. The recording session was conducted in a quiet place using a condenser microphone. Other than applying amplitude normalization to 70 dB using Praat (Boersma & Weenink, 2017), the recorded stimuli were not modified in any other way.

5.3. Procedures

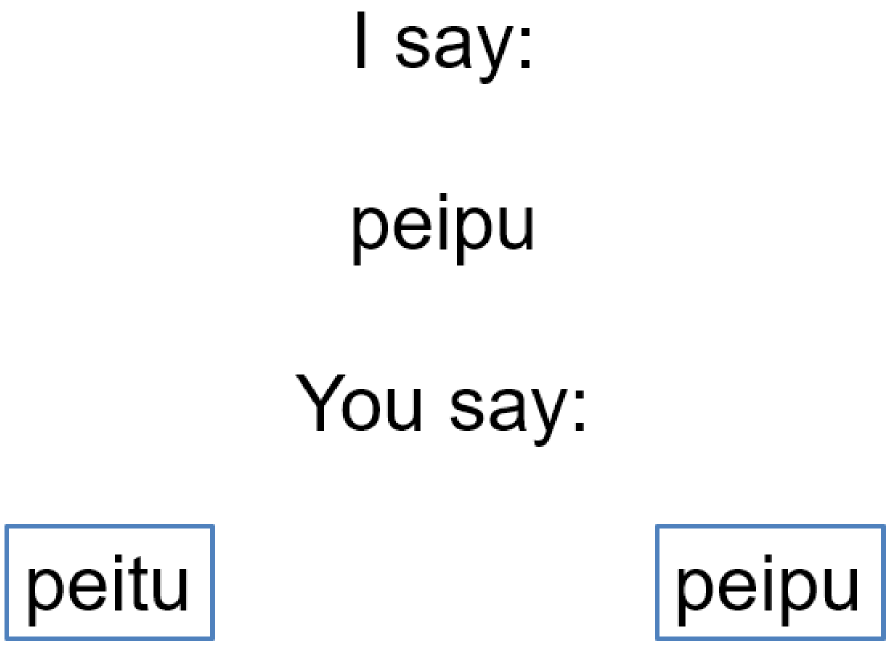
The language game design of this experiment followed that of Wilson (2006) and Gallagher (2013). Participants were presented with a prompt and had to provide a response. In the exposure phase, prompts and responses were presented to the participants under the frame of ‘I say X, you say Y’. In the test phase, participants were presented with prompts only and were asked to provide a response.

The exposure phase of both experimental conditions contained 36 prompt-response pairs. Half of them exhibited alternation and half did not, so that the overall alternation rate was 50% in both conditions. Exactly which prompts alternated depended on the condition. These 36 pairs were repeated four times in four pseudo-random orders for a total of 144 trials. In each trial, the text ‘I say X’, in which X stood for the Pinyin[4] orthographic form of the prompt, appeared on the screen, and then the sound of the prompt was played. After the auditory form of the prompt was displayed, a second line of text ‘You say Y’ appeared, in which Y stood for the Pinyin transcription of the response. Participants were then required to click a button to hear the response. After both the prompt and the response were played, a ‘continue’ button appeared which led the participants to the text trial. Participants were encouraged to read out the stimuli to help them better learn the pattern.

The test phase contained 24 items, including 18 novel prompts with all possible consonant pairs and 6 fillers beginning with [tʂ] or [ʂ] that did not occur in the exposure phase. Similar to the exposure, in each test trial, participants first saw the text ‘I say X’ and heard the sound of the prompt. After that, the text ‘You say’ appeared, followed by the sound of two choices. One of the choices was an alternating response of the test prompt (the labial consonant in C_2 changing into a coronal), and the other one was the same as the prompt (non-alternating). The prompt and the non-alternating response were acoustically identical. The Pinyin transcriptions of the two choices appeared in two text boxes, and participants were asked to click on one of the boxes to make a two-alternative forced-choice as the response to the prompt word. Whether the alternating choice occurred on the left or the right of the screen was counterbalanced. After making a choice, participants were brought to the next test trial. A screenshot of the test phase is given in Figure 5.

---

[4] Pinyin is the official romanization writing system of Mandarin Chinese. This system is taught at elementary schools as a way to transcribe the pronunciation of Chinese characters. Therefore, most native Mandarin speakers are familiar with this system and are able to use it to annotate the sounds in a syllable, at least for all existing Mandarin syllables.



I say:

peipu

You say:

Figure 5 Screenshot of a testing trial for Experiment 2. Text instruction is translated from Mandarin into English for illustration.

The experiment was run using the online experiment platform Gorilla (Anwyl-Irvine et al., 2020). After finishing the main experiment, participants were directed to complete the Chinese vocabulary test and the language background questionnaire. The whole experiment lasted for around 30 minutes.

5.4. Predictions

It is expected that prompt type will have a main effect on participants' alternating behaviors. In both conditions, participants should be more likely to alternate the prompts if the corresponding consonant pairs also alternate during the exposure. Crucially, I predict a significant interaction between condition and prompt type, because the OCP pattern should result in a stronger learning effect compared to the Arbitrary pattern. This finding will indicate that OCP-triggered alternations receive a learning advantage compared to other arbitrarily conditioned phonological alternations.

5.5. Data Analysis

Participants' choices were coded as a binary dependent variable of the statistic model (choosing the alternating response was coded as 1, and choosing the non-alternating response was coded as 0). The two independent variables were experimental condition (OCP vs. Arbitrary) and prompt type, i.e., the alternating behavior of the target consonant in the training (Alternating vs. Non-Alternating). The data were analyzed by mixed-effects logistic regression models using the *lme4* package (Bates et al., 2015) of the R software (R Core Team, 2018). The fixed effect structure included the two independent variables and their interaction. The random effect structure was determined by backwards stepwise model comparisons.

5.6. Results

Six participants' results were removed from the analysis because they either gave alternating responses to all the test prompts (1 from the OCP condition, 2 from the Arbitrary condition) or gave non-alternating responses to all (2 from the OCP condition and 1 from the Arbitrary condition), because their skewed performance might indicate that they were not learning any pattern or were not paying attention[5]. Therefore, the remaining dataset for analysis contained data from 23 participants in the OCP condition and 31 participants in the Arbitrary condition. Responses to the filler items were not analyzed.

---

[5] However, according to the distribution of the log reaction time on all the test trials for these skewed participants (mean = 7.417, sd = 0.574), they did not show faster responses compared to the remaining participants (mean = 7.347, sd = 0.954).

The comparison between the participants' responses on the two types of prompts in the two experimental conditions is shown in Figure 6.

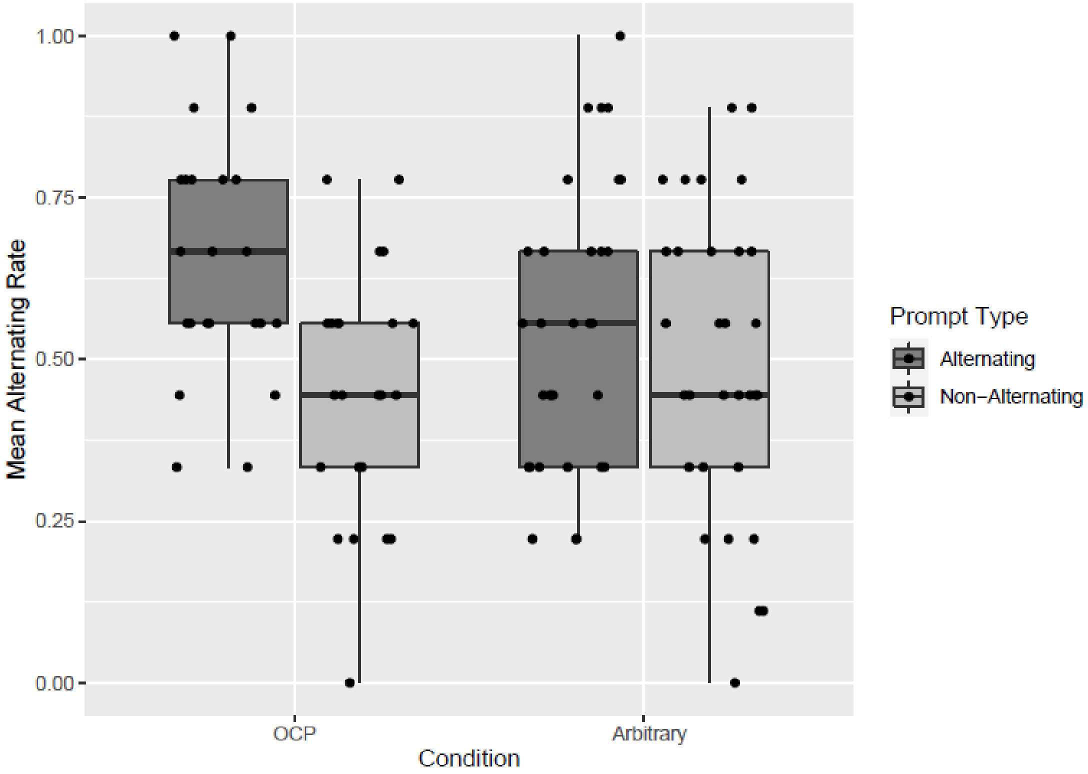


Figure 6 Proportion of alternating responses to the test prompts grouped by condition and prompt type. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

A logistic mixed-effects model was fitted to participants' alternating/non-alternating responses on the test prompts using the *lme4* package (Bates et al., 2015) in the R software (R Core Team, 2018). Fixed effects included experimental condition (OCP vs. Arbitrary), prompt type (whether the prompt alternated in the exposure phase), and their interaction. Random effects included random intercepts for participant and test prompt, and a random slope for prompt type by participant. Other effects were not included since this was the maximal model that converged. The maximal model's parameter estimates are shown in Table 5. The baseline for prompt type is non-alternating and the baseline for the experimental condition is Arbitrary. The *p*-values were obtained using the *lmerTest* package (Kuznetsova et al., 2017). There is a marginally significant effect of the interaction between prompt type and condition.

	Estimate	Std Error	*t* value	*p* value
(Intercept)	-0.0402	0.1915	-0.2100	0.8337
Alternating	0.2868	0.2218	1.2928	0.1961
OCP	-0.1893	0.2728	-0.6940	0.4877
Alternating:OCP	0.6469	0.3350	1.9309	0.0535

Table 5 The maximal model for alternating rates.

The model summary indicates that participants were more likely to give an alternating response for the alternating prompts, yet this effect is not significant in the Arbitrary condition. The lack of a prompt type effect indicates that participants did not properly learn the alternation pattern encoded in the Arbitrary condition. After switching the baseline for the experimental condition to OCP, the effect of prompt type was significant ($t = 3.788$, $p = .0002$). Crucially, the marginally-significant interaction between prompt type and condition results from the larger alternating rate difference between the two prompt types in the OCP condition than in the Arbitrary condition, as shown in the boxplots.

5.7. Discussion

The results of Experiment 2 demonstrate that participants were able to learn the OCP-Labial-based alternation pattern, but did not succeed in learning the arbitrarily conditioned alternation pattern. The better learning of the OCP pattern indicates that OCP-based alternation has an advantage in phonological acquisition compared to an arbitrary pattern. This result also suggests that alternation patterns based on consonant co-occurrence are learnable by Mandarin speakers in AGL tasks; this lays the foundation for the following experiments of the current study: other OCP-related learnability issues will be investigated by comparing the learning outcome of various OCP-triggered alternations, such as OCP-Labial vs. OCP-Dorsal, and identity avoidance vs. similarity (but non-identity) avoidance.

These results can be interpreted as an effect of a structural bias. The alternation pattern in the OCP condition can be summarized as the avoidance of two consonants that agree in place, whereas the arbitrary pattern will necessarily involve more than one feature to describe. Hence, the OCP-based pattern in structurally simpler and that is why it is better learned compared to the Arbitrary pattern. Whether this learning advantage can also be accounted for by a substantive bias effect is an interesting question. Unlike in Experiment 1, the current experiment did not include a condition that is comparable in complexity with the OCP condition. Because the alternation implemented in this experiment was a Labial to Coronal alternation, if we were to set up a mirror pattern where the alternation is conditioned by consonant place harmony (changing the place of the second consonant into Coronal if the two consonants do not match in place, e.g., [kVpV] → [kVtV], while [pVpV] stays the same), then we would end up with a situation in which the alternation is not motivated by any principle. After changing the second consonant into Coronal, the resulting word would still violate place harmony (kVtV). Therefore, if we really wanted to test alternations conforming to consonant harmony, the design would essentially look very different from what was adopted in Experiment 2. At the very least, the alternation rule to be learned ‘changing the C2 into a Coronal’ would need to be different.

Previous AGL studies suggest that tasks involving the learning of two alternation rules with matching complexity have often reported a substantive bias effect (Finley, 2011, 2012b; Alexander

Martin & Peperkamp, 2020; White, 2013; Wilson, 2006) (but not all of them, see for example Pycha et al. (2003) who did not find a substantive bias effect). Due to this complication as well as the findings from Experiment 1 that suggest the relevance of the OCP in learning beyond complexity, I leave the investigation of this issue (e.g., by looking at the learnability between an OCP-based alternation pattern and an alternation that turns differing consonant pairs into major place harmony pairs) to future research.

Coetzee's (2009) AGL experiment on the learning of an alternation pattern found a bias against giving an alternating response in the test phase; in other words, the default option when seeing a novel prompt is to not alternate. Other AGL experiments on the acquisition of alternation patterns reported similar biasing effects (Gallagher, 2013; Wilson, 2006). For example, the overall alternation rate (across all conditions) in Gallagher's (2013) experiment was about one-third, despite the fact that alternation rate in the training phase was 50%. This non-alternating default preference was not found in the current experiment, as the overall alternation rate was 53.2%. This may have resulted from the nature of the patterns involved: the OCP-based alternation patterns in the current task may be more subtle than the patterns used in the other studies (e.g., identity-triggered alternation in Gallagher (2013)), hence overall participants found the current task more challenging and relied more on guessing in giving responses. L1 influence might be another possibility. With a few exceptions of tone sandhi rules and limited vowel alternations triggered by the diminutive suffixation process, Mandarin is generally considered as a language that lacks systematic phonological alternation. Because of this, the default non-alternating bias might not be available for Mandarin speakers, as the whole alternation mechanism is missing from Mandarin speakers' phonological grammar. As a result, in an AGL task, Mandarin speakers are free from the non-alternating bias and are more likely to match the alternation rate provided in the training when tested by the novel prompts.

The potential effects of nuisance factors, such as the missing syllables [$\mathrm{k^h}$ei4] and [pou4], and the vowel distributions, were also examined. Whether the test items contained the syllable [$\mathrm{k^h}$ei4] or [pou4] did not significantly improve the model on alternation responses (χ^2 = 0.2415, p = .6231), and neither did the distances between the two vowels in the test items (χ^2 = 0.3688, p = .5437). Individual differences on the participants' vocabulary test scores again did not affect their learning behaviors. The mean vocabulary test score for the participants was 74.27 (sd = 5.44) in the OCP condition, and 73.91 (sd = 5.59) in the Arbitrary condition. Adding participants' vocabulary test scores as a factor into the statistical model did not result in significant improvement of the model (χ^2 = 1.1828, p = .2768).

6. Experiment 3: Phonotactics and Alternation

Experiments 1 and 2 of this dissertation have provided evidence for the advantage of learning an OCP-based phonotactic pattern and an OCP-triggered alternation over other arbitrary patterns. They provide an apt backdrop for the investigation of whether an OCP effect in stem phonotactics can aid the acquisition of an OCP-triggered phonological alternation. To this end, Experiment 3 was designed to test the phonotactics-alternation link using OCP as the target phonological phenomenon and native Mandarin speakers as the subject group. This experiment aims to contribute more empirical data to the relation between phonotactics and alternation in phonological learning by testing a consonantal pattern that has not been examined before. By testing native Mandarin speakers, who also have little evidence for the link between phonotactics and alternation due to the paucity of alternation in Mandarin, the experiment also provides a further test to the language-specificity hypothesis of Do & Yeung (2021) on this relation in phonological learning.

Specifically, Experiment 3 is designed to compare the learning of an OCP-triggered alternation when there is a match or a mismatch in stem phonotactics. In both conditions of this experiment, there was an exceptionless alternation where the consonant in the plural suffix [-ma ~ -na] alternated to disagree with the place of the second consonant of the singular stem. This alternation was motivated by the OCP-labial-triggered place dissimilation. Participants in both conditions were trained on the same alternation with the same amount of evidence. The two conditions differed in whether the stem phonotactics supported this alternation. In the OCP condition, all stems had disagreeing consonant pairs, which matched with the alternation pattern in the plural suffix; in the Control condition, agreeing and disagreeing consonant pairs were equally likely to occur in the stems, resulting in a mismatch between alternation and stem phonotactics. In the exposure phase, participants listened to the singular stems and the suffixed plural forms separately, together with visual cues for all the stimuli. In the test phase, participants were presented with novel singular stems and were asked to choose their correct plural forms (using either the [-ma] or [-na] allomorph). This design followed Chong's (2021) and Do & Yeung's (2021) AGL experiments on how phonotactics and alternation interact in vowel harmony.

6.1. Participants

Participants were 60 native Mandarin speakers (22 male, 38 female, mean age = 23.1, age SD = 5.21) without any speech or hearing problems. They were randomly assigned to the two conditions of this experiment: 30 participants for the OCP condition and 30 participants for the Control condition. Monetary reimbursement was provided for participating in this experiment. Participants were recruited via online social media (WeChat groups, etc.).

6.2. Materials

The stimuli for the artificial languages in the two experimental conditions were drawn from CVCV stems where C and V were all possible combinations of [p^h p f k^h k x] and [a ei ou u], except for identical vowel pairs. In addition, each stem was paired with a plural form by adding a suffix that had two allomorphs [-ma] and [-na]. The place of the consonant in the plural suffix always disagreed with the second consonant in the stem, such that [-ma] was used when the second stem consonant is [k^h], [k], or [x] and [-na] was used when it was [p^h], [p], or [f]. Again, this alternation was built on the labial place dissimilation in Cantonese.

There were altogether 36 possible consonant pairs, 18 of which agreed in place of articulation and the other 18 did not agree. The stems in the OCP condition contained only the 18 disagreeing consonant pairs, so that all stems abided by OCP-Place. They were combined with two distinct vowel pairs, giving rise to 36 singular stems. Stems plus their plural suffixed forms added up to 72 stimuli for the OCP condition. For the Control condition, 36 stems contained all possible 36 consonant pairs, meaning that there were no phonotactic patterns in consonant co-occurrence. These stems were paired with plural forms, giving rise to also 72 stimuli.

Because the second stem consonant was 50% labial and 50% dorsal, the occurring frequencies of the two allomorphs [-ma] and [-na] were both 50%. Thus, both conditions had the same amount of evidence for the place alternation rule. All training and test stimuli can be found in Appendix 3.

The test stimuli contained 36 novel stems and their two suffixed forms (either taking the [-ma] or [-na] allomorph). These novel stems were all different from the stems in the training and had only the 18 disagreeing consonant pairs. These 18 disagreeing pairs were combined with two distinct vowel pairs, which gave rise to the 36 test stems. The reason why only disagreeing consonant pairs were used was because in the OCP condition, participants were only trained in stems with disagreeing pairs. Both conditions used the same test stimuli. This design may introduce a bias on participants' answer choices, as the training and testing items always match in stem phonotactics. However, in a very similar study about the link between phonotactics and alternation, Do & Yeung (2021) presented both types of test items: whether they comply with the stem phonotactic of the training items or not. The results suggested that participants showed similar behaviors no matter the training and testing items matched in phonotactics or not.

Half of the test stems had labials as the second consonant (e.g., [k^{h}eip^{h}ou]). Therefore, for the two suffixed choices in the testing, one violated the OCP-Place ([k^{h}ei**p^h**ou**m**a]) and the other one did not ([k^{h}ei**p^h**ou**n**a]). For these stimuli, there is a motivation to pick the [-na] suffix. For the other half of the test stems, they had dorsals as the second consonant (e.g., [p^{h}ak^{h}ou]). Despite the fact that in training, all stems whose second consonant was a dorsal were paired with a [-ma] suffix, in the test phase, when participants saw these dorsal-ending stems and the two corresponding choices, neither suffixed form violated the OCP-Place ([p^{h}a**k^h**ou**m**a] and [p^{h}a**k^h**ou**n**a]). Even though the consonant sequence [...k^h...n...] was never attested in the exposure phase, because it does not violate the OCP, there is no harm for participants to pick the [-na] allomorph. In the data analysis, the place of the second stem consonant was entered as a factor, because participants may exhibit different behaviors on these two types of test stems.

6.3. Procedures

The experiment consisted of an exposure phase and a test phase. At the beginning, participants were told that they were going to learn words from a foreign language. In the exposure phase, each trial was either a singular stem or a plural form. Singular stems were paired with images that displayed a single object, while the plural forms were paired with images showing two of the same objects. Participants heard the stimulus word and the Pinyin transcription also appeared on the screen. The singular and plural forms for the same object were never presented in adjacent trials, so there was no direct instruction on plural marking or the alternation rule. After listening to the stimulus, participants clicked the 'continue' button to move on to the text trial. The 72 stimuli were repeated three times, and each time, the stimuli presented in a pseudo-randomized order. In each trial, participants were only allowed to hear the stimulus once.

The test phase included a wug-test on novel singular stems. Participants heard a novel stem and saw a corresponding image. They then heard two possible plural forms taking either the [-ma] or [-na] suffix. The Pinyin orthographic forms of the two choices and the plural image were also displayed. Participants were asked to pick the correct plural form they thought should be. Screenshots of a test trial can be found in Figure 7. The presentation order of the two possible plural forms was counterbalanced such that each of the [-ma]-ending and the [-na]-ending forms was presented first 50% of the time.

Figure 7a Screenshot of a testing trial of Experiment 3 (part 1).

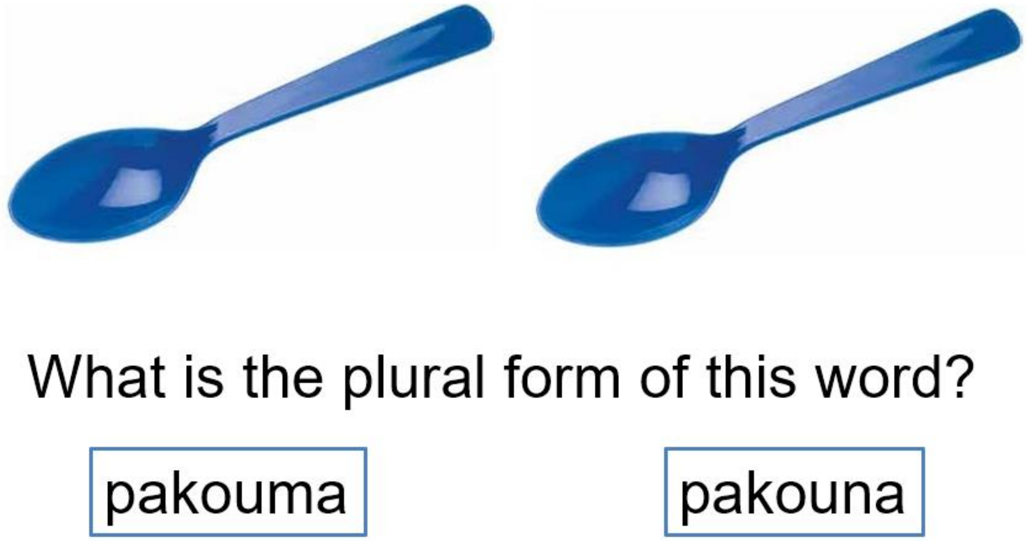

What is the plural form of this word?

Figure 7b Screenshot of a testing trial of Experiment (part 2). Text instruction is translated from Mandarin into English for illustration.

The experiment was run using the online experiment platform Gorilla (Anwyl-Irvine et al., 2020). After finishing the main experiment, participants were directed to complete the Chinese vocabulary test and the language background questionnaire. The whole experiment lasted for around 30 minutes.

6.4. Predictions

Based on Chong's (2021) findings, we expect that stem phonotactic can aid the learning of the alternation in the suffix, and participants will be more likely to choose the [-na] suffix to avoid an OCP violation in the OCP condition than in the Control condition. However, Do & Yeung's (2021) study found that when participants' L1, such as Cantonese, shows little evidence for the phonotactics-alternation link, they may not use stem phonotactics to help them learn the alternation patterns. Given that Mandarin is similar to Cantonese in lacking phonotactically triggered segmental alternations, Experiment 3 is also likely to return a null result, i.e., the place disagreement phonotactic pattern in the stems of the OCP condition may not be able to help the participants to learn an OCP-based alternation in the plural suffixes.

I also predict that the allomorph selection behaviors will be different between labial-final stems ([k^{h}eip^{h}ou]) and dorsal-final stems ([p^{h}ak^{h}ou]), because for the dorsal-final stems, neither choice in the test phase violates the OCP. In other words, there is no harm for participants to choose the [-na] suffix for these dorsal-ending stems, even though in exposure they never hear such combinations. As a result, the allomorph choice between the OCP and Control conditions for the dorsal-final stems might not be as different, since the stem phonotactics (OCP-Place) is not helpful in guiding the participants to choose [-ma] or [-na] for dorsal-ending stems.

6.5. Data Analysis

Participants' suffixation choices were coded as a binary dependent variable of the statistic model (choosing the [-na] suffix was coded as 1, and choosing the [-ma] suffix was coded as 0). The two independent variables were experimental condition (OCP vs. Control) and the place of the second consonant of the test stems (Labial vs. Dorsal). The data were analyzed using mixed-effects logistic regression models using the *lme4* package (Bates et al., 2015) of the R software (R Core Team, 2018). The fixed effect structure included the two independent variables and their interaction. The random effect structure was determined by backwards stepwise model comparisons.

6.6. Results

Participants' rates of choosing the alternated plural forms (i.e., rates of choosing the [-na] allomorph, assuming [-ma] form is the default unalternated form) in the two experimental conditions are shown in Figure 8. We are mainly interested in whether participants in the OCP condition could benefit from the stem phonotactics and thus were more likely to choose the [-na] suffix for the labial-ending stems in their plural form responses.

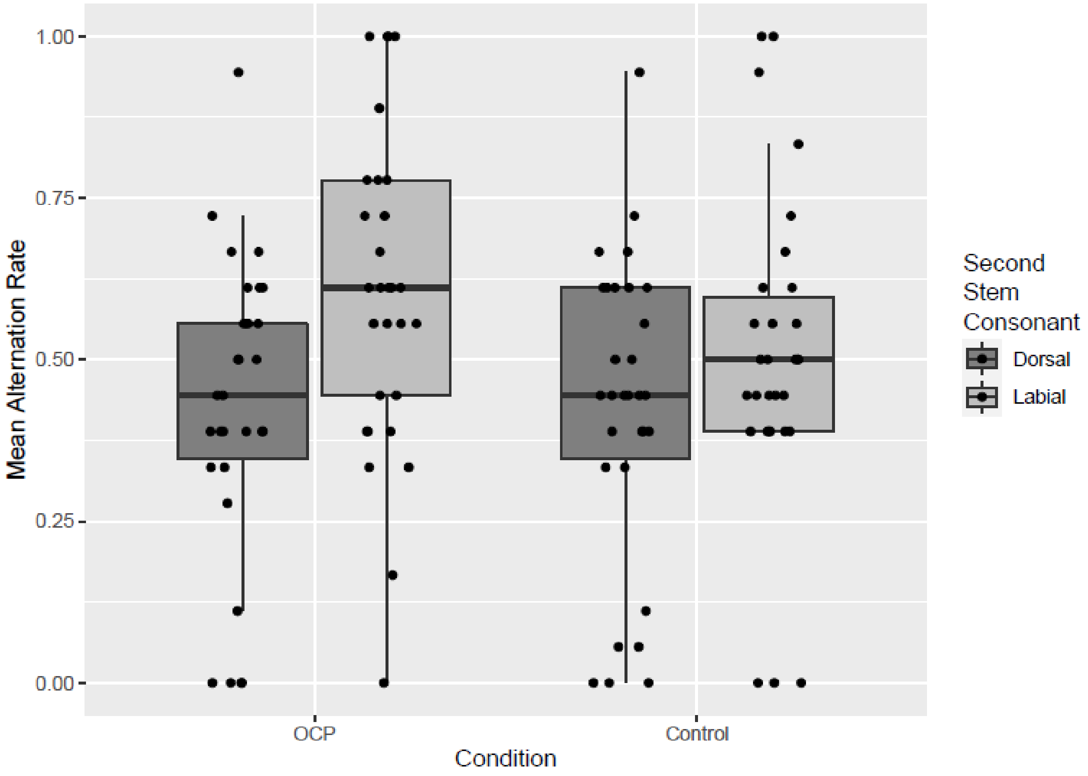


Figure 8 Proportion of alternating responses to the test stems grouped by condition and the place of the second stem consonant. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

The rate of choosing the alternated form (the [-na] suffixed form) was analyzed using a logistic mixed-effects model with experimental condition and the place of the second stem consonant as categorical independent factors. The model also contained by-subject random intercepts, and this was the maximal model that converged. Introducing by-item intercepts into the model caused a singular fit problem, because there was not enough additional item-level variation to warrant adding an additional item-level random effect to explain all the observed variation. The model summary is provided in Table 6. The *p*-values were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

	Estimate	Std Error	*t* value	*p* value
(Intercept)	-0.3763	0.1748	-2.1532	0.0313*
OCP	0.0446	0.2458	0.1816	0.8559
Labial	0.3631	0.1302	2.7894	0.0053**
OCP:Labial	0.4570	0.1838	2.4861	0.0129*

Table 6 The maximal model for alternation rates.

The baseline for this model is Dorsal for consonant place and Control for experimental condition. The effect of experimental condition was not significant, which is in line with the prediction: for the dorsal-ending stems, neither [-ma] nor [-na] suffix violated the OCP, so the stem phonotactics may not affect participants' choices, leading to a lack of experimental condition effect for the dorsal-ending stems. However, there was a significant interaction between the two factors, indicating that for the labial-ending stems, participants gave significantly more [-na] suffix choices in the OCP condition than in the Control condition. Switching the baseline for the consonant place factor from Dorsal to Labial gave rise to a significant effect for the experimental condition (t = 2.039, p = 0.041). Further post-hoc tests showed that for labial-ending stems, participants in the OCP condition gave significantly above-chance rates of choosing the alternated [-na] suffix (t = 2.809, p = 0.005), indicating a learning effect of this OCP-based alternation. Meanwhile, for the labial-ending stems in the Control condition, participants' [-na] suffix choices were not significantly different from chance (t = -0.076, p = 0.940). These results suggest that stem phonotactics did have an effect on the learning of the suffix alternation, at least for the labial-ending stems: if the stems all had disagreeing consonant pairs, then participants were more likely to choose a plural suffix that also had disagreeing place with the nearest consonant.

6.7. Discussion

The results of Experiment 3 suggest that an OCP-based alternation is easier to learn when the generalizations in stem phonotactics agree with the dynamic generalizations about the alterations in the suffix. Recall that in both conditions, the allomorph selection was based on the exact same principle: the consonant in the suffix always disagreed in place with the second consonant of the stem. Participants in the Control condition, where the stem consonant pairs provided no clue to the place disagreement, failed to learn this OCP-conditioned alternation across the morpheme boundary, despite the exceptionless evidence for this alternation in the exposure phase. Participants in the OCP condition, however, seemed to be able to resort to the place disagreement pattern in the stem phonotactics and apply this knowledge to the learning of the alternation. The alternation learning is made easier if the stem phonotactics supports the generalizations about the alternation. Therefore, the results provide evidence for the hypothesis that stem phonotactics can aid the learning of alternations that are triggered by the same set of principles (Chong, 2017, 2021; Pater & Tessier, 2006).

The pattern in the Control condition is an example of the Derived Environment Effects (Kiparsky, 1993), where the OCP-triggered alternation only happens in suffixation (derived environment), while the stems in the exposure contain violations of the OCP. The results of Experiment 3 show that when stem phonotactics and alternation do not match, as in the Control condition, the alternation is not properly learned, compared to situations where phonotactics and alternation go hand in hand (as in the OCP condition). This provides evidence that alternations involving Derived Environment Effects may be more difficult to learn than patterns where stem phonotactics and alternations match (Chong, 2021). Moreover, if we conduct closer inquiries on the Derived Environment Effects patterns, we may question whether they are a valid cross-linguistic generalization. Finnish Assibilation is a textbook example of the Derived Environment Effects (Kiparsky, 1993). This rule changes an underlying /t/ into /s/ before /i/, but it only applies across morpheme boundary and does not apply to /ti/ sequences within the roots. For example, when the third-person-singular-present affix /-i/ is attached to the root verb /tilat/ ‘to order’, Assibilation will convert the heteromorphemic /t-i/ sequence into /s-i/, while the root-internal /ti/ sequence remains intact (/tilat + -i/ → /tilas-i/, */silas-i/). However, a closer examination of Finnish reveals that Assibilation is actually morphologically and lexically conditioned and cannot be simply predicted by morphological derivedness (Anttila, 2006), because some /i/-initial suffixes will not trigger Assibilation when attached to root-final /t/ (/tunte + isi/ → /tunt-isi/, */tuns-isi/). If Derived Environment Effects are subject to such lexical exceptions, it raises the question of how general these effects truly are. If we view Derived Environment Effects as a special case of morphologically conditioned phonological alternation, then we may have an answer to their acquisition disadvantage. Compared to alternations that only refer to phonological information, morphologically conditioned alternations additionally require information from morphology, which may lead to higher complexity and higher learning difficulties.

Because neither Mandarin nor Cantonese has phonotactically aided alternations in their language systems, one may expect to see the Mandarin participants in Experiment 3 behave more similarly to Do & Yeung’s (2021) Cantonese participants and hence not show a learning difference between the two experimental conditions. In addition, the designs of the current experiment and Do & Yeung’s are also similar in terms of number of stimuli in the exposure phase (36 singular stems in the current study vs. 32 in Do & Yeung (2021)) and the duration of the training. The current experiment, however, did not replicate Do & Yeung’s null result. It is possible that the OCP-Labial generalization used in the current study is a more salient pattern than the vowel harmony generalizations used in previous studies, so that the Mandarin participants can learn the consonant co-occurrence patterns better and later extend them to alternations in suffixation. Another possibility is that Mandarin in fact contains a small number of phonotactically motivated alternations. For example, the Tone 3 Sandhi rule, where a Tone 3 becomes Tone 2 when it precedes another Tone 3, is a suprasegmental alternation that avoids two adjacent Tone 3s. Moreover, the diminutive suffix [-ɚ] also triggers a series of alternations such as coda deletion, vowel nasalization, and schwa insertion, etc. These adaptations can be viewed as being motivated to satisfy a fixed CGVX syllable structure, which only allows two timing slots for the rimes (Duanmu, 2007). These alternations may provide Mandarin speakers with clues to match stem phonotactics with alternations, while this mechanism is missing in Cantonese. I leave the exploration of this disagreement of findings for future research.

Experiment 3 did not include a well-formedness rating task for the test stems to see if participants had learned the phonotactic regulations in the exposure. Based on the results of Experiment 1 where participants trained on consonant disagreement showed significantly more endorsements on the novel disagreeing words than agreeing words, I posit that participants in the OCP condition of Experiment 3 were able to detect the consonant dissimilation pattern and internalize it.

The potential effects of nuisance factors, such as the missing syllables [k^hei4] and [pou4], and the vowel distributions, were examined. Whether the test items contained the syllable [k^hei4] or [pou4] did not significantly improve the model on allomorph choices (χ^2 = 0.782, p = .3765), and neither did the distances between the two vowels in the test items (χ^2 = 1.6614, p = .1974). Individual differences among the participants on their vocabulary test scores, again, did not affect their alternation learning outcome. The mean vocabulary test score for the participants in the OCP condition was 73.10 (sd = 7.75), and 74.20 (sd = 7.27) for the Control condition. Model comparison demonstrates that adding participants' vocabulary test scores as a factor into the model did not result in a significant improvement of the model (χ^2 = 0.16, p = .689).

7. Experiment 4: Place Asymmetry

The previous experiments have provided a solid foundation for the learnability of OCP patterns compared to other arbitrary patterns, and OCP in stem phonotactics can aid the acquisition of an OCP-triggered alternation. The following two experiments examine the typological distributions within the OCP to see if these detailed typological asymmetries can be accounted for by learning biases as well.

Experiment 4 was designed to test if there is a learning asymmetry between an OCP-Labial-triggered alternation pattern and an OCP-Dorsal-triggered one. Previous studies have shown that multiple learning biases can shape phonological typology by skewing the acquisition of different patterns (Moreton, 2008). Using the AGL paradigm, Experiment 4 was conducted to examine if the typological asymmetry between OCP-Labial and OCP-Dorsal is the result of a learning bias favoring the acquisition of labial co-occurrence than dorsal. Similar to Experiment 2, participants were asked to learn a language game in which alternations happened on the second consonant of the prompts in the artificial language.

Experiment 4 contained two conditions. In the Labial condition, participants were trained on CVCV prompts with two labial consonants or with two disagreeing consonants, and the second consonant in the prompt changed to coronal in the response if the two stem consonants were labial. Importantly, participants in this condition were not presented with any stimuli that contained two dorsal consonants. In other words, the crucial information about whether the alternation applied in stems with two dorsal consonants was withheld (a poverty of the stimulus design). Later in the test phase, participants were tested on whether they will extend the alternation to the withheld contexts, i.e., whether training in OCP-Labial could generalize to a novel OCP-Dorsal pattern. On the contrary, in the Dorsal condition, participants were trained on prompts with two dorsals, which underwent a similar alternation so that the second dorsal consonant changed to coronal, and prompts with disagreeing places, which remained the same in the responses. In this condition, prompts with two labial consonants were withheld from the exposure. Participants' readiness to generalize the alternation to the withheld contexts was compared across the two conditions. If the observed OCP-Labial privilege over OCP-Dorsal is shaped by a learning bias, we would expect that training in the OCP-Dorsal pattern is more likely to generalize to an OCP-Labial context. Since typologically, if a language sees co-occurrence restrictions in two dorsal consonants, it implies that labial co-occurrence is highly likely to be found in the lexicon as well. However, if a language has labial co-occurrence, it does not imply that dorsal co-occurrence is also present. Therefore, training in dorsal dissimilation licenses the inference to alternations triggered by labial dissimilation (and potentially other places as well), but training in labial dissimilation does not license the inference to alternations triggered by dorsal dissimilation.

7.1. Participants

Participants were 60 native Mandarin speakers (17 male, 43 female, mean age = 21.0, age SD = 5.84) without any speech or hearing problems. They were randomly assigned to the two conditions of this experiment: 30 participants for the Labial condition and 30 participants for the Dorsal condition. Monetary reimbursement was provided for participating in this experiment. Participants were recruited via online social media (WeChat groups, etc.). Most of the participants were college students in mainland China.

7.2. Materials

The stimuli for this experiment were drawn from CVCV non-words, where Cs were all possible combinations of [p^h p f k^h k x], and Vs were all non-identical pairs of the vowels [a ei ou u]. Based on the consonant place of articulation, there were altogether 36 possible consonant pairs: 9 labial agreeing pairs which contained two labials, 9 dorsal agreeing pairs which contained two dorsals, and 18 disagreeing pairs. In the Labial condition, a prompt whose consonants were from the 9 labial agreeing pairs were paired with a response in which the second consonant changed to coronal (e.g., prompt [p^hap^hu], response [p^hat^hu]), whereas a prompt whose consonants were from the 18 disagreeing pairs were paired with an identical response (e.g., prompt [p^hak^hu], response [p^hak^hu]). Prompts with dorsal agreeing pairs did not occur in this condition. In the Dorsal condition, prompts with any of the 9 dorsal agreeing pairs were paired with a response in which the second consonant changed to coronal (e.g., prompt [k^hak^hu], response [k^hat^hu]), whereas prompts with the 18 disagreeing pairs did not alternate in the responses. The 9 agreeing consonant pairs in each condition combined with 4 vowel combinations, and the 18 disagreeing pairs were assigned 2 vowel combinations each. This gave rise to 72 prompts for the exposure phase in each condition, and the alternation happened in 50% of the prompts in both conditions. Importantly, in the Labial condition, consonant pairs with two dorsal consonants were withheld during exposure, while in the Dorsal condition, consonant pairs with two labial consonants were withheld.

The test items consisted of 36 novel prompts with all 36 consonant pairs. Each consonant pair was assigned one vowel combination. All prompts in the test phase had novel vowel pairs and thus participants did not hear them before. The same set of test prompts was used in both experimental conditions. Appendix 4 presents all exposure and test stimuli for Experiment 4.

7.3. Procedures

Experiment 4 followed the procedures adopted in Experiment 2. In the exposure phase of both conditions, 72 prompt-response pairs were presented using the ‘I say X, you say Y’ framework and were repeated twice in two distinct pseudo-randomized orders, giving 144 exposure trials. In the test phase, the 36 test prompts were presented in a pseudo-randomized order. In each test trial, participants first saw the text ‘I say X’ and heard the sound of the prompt. After that, the text ‘You say’ appeared, followed by the sound of two choices. One of the choices was an alternating response of the test prompt (the labial consonant in C_2 changing into a coronal for the Labial condition, and the dorsal consonant in C_2 changing into a coronal for the Dorsal condition), and the other one was the same as the prompt (non-alternating). The Pinyin transcriptions of the two choices appeared in two text boxes, and participants were asked to click on one of the boxes to make a two-alternative forced-choice as the response to the prompt word. Whether the alternating choice occurred on the left or the right of the screen was counterbalanced. After making a choice, participants were brought to the next test trial.

The experiment was run using the online experiment platform Gorilla (Anwyl-Irvine et al., 2020). After finishing the main experiment, participants were directed to complete the Chinese vocabulary test and the language background questionnaire. The whole experiment lasted for around 30 minutes.

7.4. Predictions

The key issue addressed by this experiment is whether participants’ willingness to generalize the alternation to the withheld contexts will be different depending on the training conditions. According to the typological asymmetry that OCP-Labial is more well-attested than OCP-Dorsal, we hypothesized that language learners have a learning bias that favors the OCP-Labial pattern. In typology, the existence of OCP-Dorsal in the lexicon implies the existence of OCP-Labial, but not vice versa. Therefore, we hypothesized that training in OCP-Dorsal would serve as stronger evidence for OCP-triggered alternation. In other words, we expect that the training of a dorsal-to-coronal alternation triggered by the OCP will lead to more generalization to the labial agreeing contexts compared to the labial training condition. In terms of statistical analysis, this should be reflected by a significant interaction between experimental condition (Labial vs. Dorsal) and the prompt type (whether the consonant pair occurred in the exposure phase or not). On the other hand, if the learners were not equipped with such learning bias, the degree of generalization should be equal across the two conditions.

7.5. Data Analysis

Participants' choices were coded as a binary dependent variable of the statistic model (choosing the alternating response was coded as 1, and choosing the non-alternating response was coded as 0). The two independent variables were experimental condition (Labial vs. Dorsal) and prompt type, i.e., whether the consonant pair of the test prompt occurred in the training (Exposure vs. Novel vs. Dissimilar). The data were analyzed by mixed-effects logistic regression models using the *lme4* package (Bates et al., 2015) of the R software (R Core Team, 2018). The fixed effect structure included the two independent variables and their interaction. The random effect structure was determined by backwards stepwise model comparisons.

7.6. Results

Participants' responses were coded as alternated or non-alternated. One participant from the Labial training condition gave non-alternating choices to all 36 test prompts, and two participants, one from Labial training and one from Dorsal training, gave alternating responses to all 36 test prompts. Their data were excluded from further analysis[6]. The remaining dataset thus included 28 participants in the Labial condition and 29 participants in the Dorsal condition.

The mean rates of alternated choices for each prompt type in each experimental condition are shown in Figure 9. There are three types of prompts: 18 prompts with two place disagreeing consonants (Dissimilar), 9 prompts with two labial consonants (Labial), and 9 prompts with two dorsal consonants (Dorsal). Recall that the typology of OCP-Place patterns indicates that OCP-Dorsal implies OCP-Labial in most of the attested languages. If this typological asymmetry is caused by phonological learning biases, we would expect that participants should be more likely to generalize the OCP-triggered alternation from dorsal agreeing to labial agreeing in the Dorsal training condition, while in the Labial training condition, they should be less likely to generalize this alternation from labial agreeing to dorsal agreeing.

---

[6] However, according to the distribution of the log reaction time on all the test trials for these skewed participants (mean = 7.348, sd = 0.472), they did not show faster responses compared to the remaining participants (mean = 7.278, sd = 0.955).

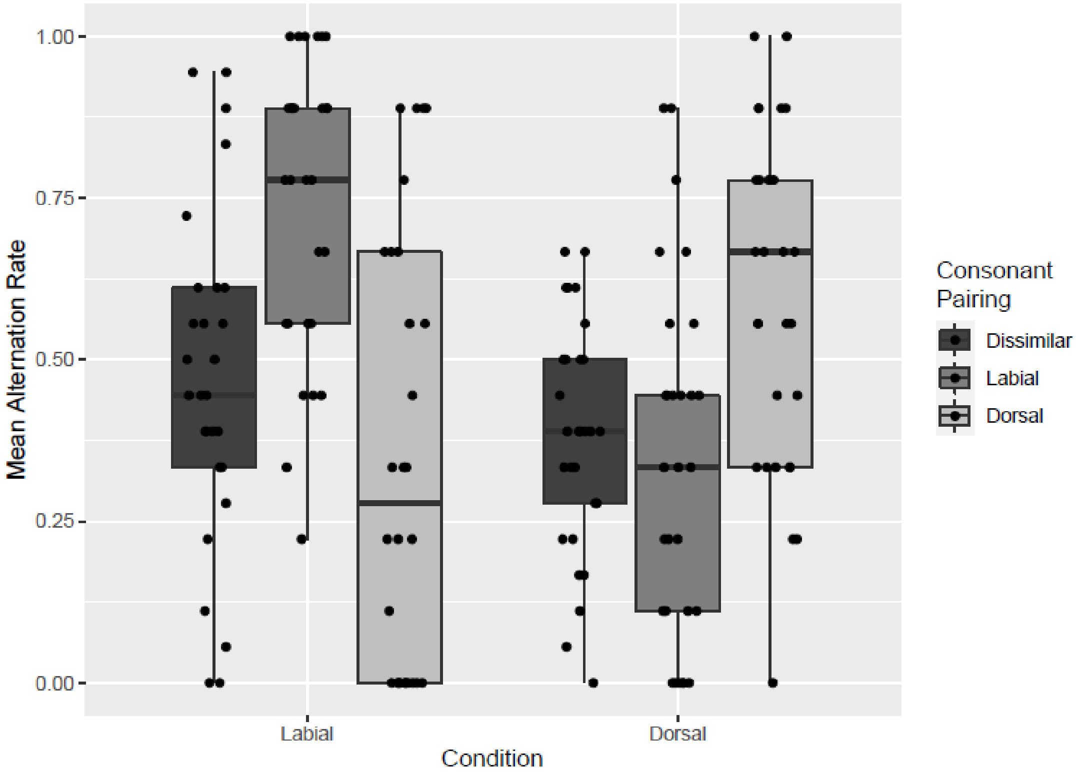


Figure 9 Proportion of alternating responses to the test prompts grouped by condition and prompt type. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

A logistic mixed-effects regression model on the alternation responses was built using experimental condition (Labial vs. Dorsal) and prompt type (Exposure, Novel, Dissimilar) as the fixed-effects predictors. Note that the prompt type had three levels and was coded differently from what is seen in Figure 9. A prompt was coded as Exposure if the consonant pair it carried occurred in the training phase, and prompts having the withheld consonant pairs were coded as Novel. In other words, labial agreeing test prompts were coded as Exposure for participants in the Labial training condition and as Novel for participants in the Dorsal training condition. Correspondingly, dorsal agreeing test prompts were coded as Novel for participants in the Labial training condition and as Exposure for participants in the Dorsal training condition. Finally, all prompts with disagreeing consonant pairs were coded as Dissimilar, and there were no specific predictions for these test prompts. Random factors included random intercepts for participants and test prompts. This was the maximal model that converged. Adding any other random slopes caused the model to fail to converge. The model summary is provided in Table 7. The *p*-values were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

	Estimate	Std Error	*t* value	*p* value
(Intercept)	1.2414	0.2214	5.6072	< 0.0001***
Dorsal	-0.8021	0.3012	-2.6633	0.0077**
Novel	-1.9711	0.2261	-8.7194	< 0.0001***
Dissimilar	-1.3736	0.1933	-7.1068	< 0.0001***
Dorsal:Novel	0.7837	0.3215	2.4379	0.0148*
Dorsal:Dissimilar	0.3885	0.2555	1.5208	0.1283

Table 7 The maximal model for alternation rates.

The baseline for experimental condition was set as Labial and the baseline for prompt type was Exposure. The effect of the experimental condition suggests that participants in the Labial condition gave more alternating responses to the exposure test prompts than the participants in the Dorsal condition. The effect of prompt type suggests that the alternating rate for the exposure test prompts in the Labial condition was significantly higher than the other two prompt types, illustrating a learning effect. Crucially, there was a significant interaction between condition and prompt type: the alternation rate difference between Exposure and Novel test prompts in the Labial condition was significantly larger than that in the Dorsal condition, suggesting that participants were more reluctant to generalize the OCP-based alternation to unseen environments in the Labial condition than in the Dorsal condition. This is in line with our hypothesis that participants exposed to OCP-Dorsal would show greater generalization towards OCP-Labial than vice versa.

By-participant analysis showed that in the Labial training condition, 21 out of 28 participants exhibited a greater alternation rate in the labial agreeing prompts than in the dorsal agreeing prompts. Two Labial condition participants mastered the rule completely and gave all labial agreeing prompts an alternation response and all dorsal agreeing prompts a non-alternation response. However, in the Dorsal condition, only 19 out of 29 participants showed a greater alternation rate in the dorsal agreeing prompts than in the labial agreeing prompts, and the differences in the alternation rates were smaller than those in the Labial condition.

7.7. Discussion

In this experiment, during the exposure phase, participants were exposed to incomplete miniature languages, only seeing the OCP-triggered alternations affecting prompts with two labials or only seeing the alternations affecting prompts with two dorsals. Later in the test phase, participants were tested on the environments they had not seen. The setup of the two experimental conditions (the two miniature languages) was symmetrical, yet we found asymmetrical results: participants in the Dorsal condition were more ready to generalize to the unseen environments compared to participants in the Labial condition.

The results of Experiment 4 support the hypothesis that the typological asymmetry among OCP-Labial and OCP-Dorsal patterns may be attributed to a phonological learning bias. In both conditions, participants were conservative about generalizing the alternation to unseen environments. However, due to this OCP-Labial learning bias, participants were more cautious about generalizing the OCP-triggered alternation from the labial agreeing to dorsal agreeing environments, while they were less cautious to do so in the opposite direction. This is a case of the poverty of the stimulus effect: in the absence of evidence about whether the OCP-triggered alternation should generalize to the other withheld environments, participants still adopted a bias to favor the learning of the OCP-Labial based pattern and generalized less in this condition. This indicates that OCP-Labial patterns received an advantage in phonological learning. Through the development of the lexicon in generations, this learning advantage may accumulate and cause OCP-Labial patterns to be more likely to occur, compared to OCP patterns based on other places of articulation. These findings also support the position that the learning bias is part of speakers' synchronic phonological knowledge, since it can guide speakers' learning behaviors in AGL tasks. Therefore, language typology is not merely the results of historical sound changes (Blevins, 2004); synchronic learning biases are also at play in shaping the typology by favoring certain phonetically principled patterns, in our case, the OCP-Labial pattern.

I also interpret this learning bias as a substantive bias. Due to the symmetrical design of the two exposure languages, the structural complexity of the two OCP-triggered alternation patterns was the same, yet the learning outcome was still different across the two conditions. This asymmetry thus results from a substantive bias that favors OCP-Labial patterns. The origin of this substantive bias can be traced back to perceptual reasons. Woods et al.'s (2010) CVC word identification results found that multiple labial consonants had the strongest negative effect on identification. CVC words with two labials were more likely to be misperceived than CVC words with consonants sharing other features, such as dorsal and coronal. Acoustically, labials have lower amplitude than coronals and dorsals because there is little or no oral cavity in front of the constriction, so the noise in the labial sounds cannot be amplified as much as other consonant places (Ohala, 1996). This also explains why there are more misperceptions in CVC words with two labials compared to CVC words with two coronals or two dorsals, because the noise signaling the presences of two labials is weaker. This experiment decided to focus on the place asymmetry between labial and dorsal OCP because of the labial-to-coronal alternation attested in the historical sound change of Cantonese. It is also interesting to see if the learning asymmetry still holds between labial and coronal places. Since the typology suggests that OCP-Coronal is at the weakest among all places of articulation, would this asymmetry be even stronger if we compare the learning of OCP-Labial and OCP-Coronal?

The potential effects of nuisance factors, such as the missing syllables [k^{h}ei4] and [pou4], and the vowel distributions, were examined. Whether the test items contained the syllable [k^{h}ei4] or [pou4] did not significantly improve the model on participants' alternation choices ($\chi^2 = 0.5119$, $p = .4743$), and neither did the distances between the two vowels in the test items ($\chi^2 = 0.7294$, $p = .3931$).

8. Experiment 5: Identity

Another characteristic of the attested OCP patterns is the exemption of identical consonants from the co-occurrence restriction. For languages that exhibit the consonant place OCP effect in their lexicon, identical consonants may be excused from the restriction and can freely co-occur or even be over-represented in the lexicon (e.g., Muna (Coetzee & Pater, 2008)).

Experiment 5 tests whether this identity exemption effect results from a learning bias that favors identity patterns. Again, using the language game framework and the poverty of the stimulus design adopted in Experiment 4, this experiment compares participants' generalization behaviors on an identity-avoidance-triggered alternation and a similarity-avoidance-triggered alternation. In the Identity condition, participants learned a labial-to-coronal alternation when the two consonants in the CVCV prompts were identical. The prompts remained the same in responses when the two consonants had disagreeing places. Crucially, prompts with agreeing but non-identical consonant pairs were withheld in this condition; thus the evidence for this OCP-based alternation was based on identity avoidance. In the Similarity condition, participants learned the same labial-to-coronal alternation from hearing prompts and responses. Now the alternation happened only when the two consonants agreed in place and were not identical; prompts with identical consonants were withheld from the exposure phase. In this case, the alternation was conditioned by (non-identical) similarity avoidance.

In the typology of OCP patterns in the world languages, as far as I know, no language only bans the co-occurrence of identical segments while shows no OCP effect on similar-but-non-identical segments. On the contrary, as reviewed earlier, many languages that show OCP effects will tolerate the co-occurrence of identical segments. Therefore, based on this typological asymmetry, we expect that training in identity avoidance will serve as stronger evidence for the OCP-triggered alternation, hence will lead to greater generalization to the withheld contexts compared to training in similarity avoidance.

On the other hand, because perceptual studies do not support that identical segments will be pardoned in misperception, it has been argued that identity exemption is merely a language specific lexical effect. Therefore, the typological asymmetry mentioned above might not be due to a learning bias that favors identical segments. Considering that structurally, identity is an extreme case of similarity, it is also reasonable to assume that training in similarity avoidance will more easily generalize to identity avoidance, because, for example, an OCP constraint like *[place$_i$][place$_i$] is not only applicable to similar consonants that share the place feature (pVb), but also to identical consonants (pVp). When the training is based on similarity avoidance, participants receive more robust evidence for this *[place$_i$][place$_i$] constraint from multiple similar consonant pairs. However, when the training is based on identity avoidance, in addition to the *[place$_i$][place$_i$] generalization, participants may also learn a more specific subset constraint *Identity and hence more reluctant to generalize the OCP-based alternation into contexts with similarity. If this is the case, then we expect that training in similarity avoidance will serve as stronger evidence for the target alternation rule, hence will lead to greater generalization to the withheld contexts compared to training in identity avoidance.

8.1. Participants

Participants were 60 native Mandarin speakers (16 male, 44 female, mean age = 19.57, age SD = 1.48) without any speech or hearing problems. They were randomly assigned to the two conditions of this experiment: Similarity and Identity. Because the condition assignment conducted by the Gorilla platform was random, there was an imbalance in the final dataset between the number of participants in the two conditions: 26 for Similarity and 34 for Identity. Monetary reimbursement was provided for participating in this experiment. Participants were recruited via online social media (WeChat groups, etc.). Most of the participants were college students in mainland China.

8.2. Materials

The stimuli for this experiment were drawn from $C_1V_1C_2V_2$ non-words, where C_1 and C_2 were all possible combinations of [p^h p f k^h k x] in C_1 and [p^h p f] in C_2, and V_1 and V_2 were all non-identical pairs of the vowels [a ei ou u]. In terms of consonant place of articulation, there were altogether 18 consonant pairs: 3 identical pairs, 6 non-identical agreeing pairs, and 9 disagreeing pairs. In the Identity condition, a prompt whose consonants were from the 3 identical pairs was paired with a response in which the second consonant changed to coronal (e.g., prompt [p^{h}ap^{h}u], response [p^hat^{h}u]), whereas a prompt whose consonants were from the 9 disagreeing pairs was paired with an identical response (e.g., prompt [k^hap^{h}u], response [k^hap^{h}u]). Prompts with non-identical agreeing pairs did not occur in this condition. In the Similarity condition, prompts the 6 non-identical agreeing pairs (p^h...p, p^h...f, p...p^h, p...f, f...p^h, f...p) were paired with a response in which the second consonant changed to coronal (e.g., prompt [p^haf^{h}u], response [p^hasu]), whereas prompts with the 9 disagreeing pairs did not alternate in the responses. For the 3 alternating consonant pairs in the Identity condition, they combined with 6 vowel combinations each. For the 6 alternating consonant pairs in the Similarity condition, they combined with 3 vowel pairs each. Finally, for the 9 disagreeing pairs in both conditions, they were assigned 2 vowel combinations each. This gave rise to 36 prompts for the exposure phase in each condition, and the alternation happened in 50% of the prompts in both conditions.

The test items consisted of 36 novel prompts with all 18 consonant pairs. Each consonant pair was assigned two vowel combinations. All prompts in the test phase had novel vowel pairs and thus participants would not have heard them before. The same set of test prompts was used in both experimental conditions. Appendix 5 presents all exposure and test stimuli for Experiment 5.

8.3. Procedures

Experiment 5 followed the same procedures adopted in Experiments 2 and 4. In the exposure phase of both conditions, 36 prompt-response pairs were presented using the ‘I say X, you say Y’ framework and were repeated four times in four distinct pseudo-randomized orders, giving 144 exposure trials. In the test phase, the 36 test prompts were presented in a pseudo-randomized order. In each test trial, participants first saw the text ‘I say X’ and heard the sound of the prompt. After that, the text ‘You say’ appeared, followed by the sound of two choices. One of the choices was an alternating response of the test prompt (the labial consonant in C_2 changing into a coronal), and the other one was the same as the prompt (non-alternating). The Pinyin transcriptions of the two choices appeared in two text boxes, and participants were asked to click on one of the boxes to make a two-alternative forced-choice as the response to the prompt word. Whether the alternating choice occurred on the left or the right of the screen was counterbalanced. After making a choice, participants were brought to the next test trial.

The experiment was run using the online experiment platform Gorilla (Anwyl-Irvine et al., 2020). After finishing the main experiment, participants were directed to complete the Chinese vocabulary test and the language background questionnaire. The whole experiment lasted for around 30 minutes.

8.4. Predictions

This experiment investigates the extent to which the generalization from the training contexts to the withheld contexts of the alternation rule is different between the two experimental conditions. Two possible predictions are considered. First, from phonological typology we learn that the OCP-place effect may be missing among identical consonant co-occurrences. In other words, for languages that show the OCP effect in the lexicon, identity avoidance implies (non-identical) similarity avoidance, while not all languages with similarity avoidance also show identity avoidance. Based on this typological asymmetry, we would predict that identity avoidance will lead to more generalization to the unseen similarity cases, and identity avoidance serves as stronger evidence for the alternation. Second, since the generalization in the Identity condition (*Identity) is a proper subset of the pattern encoded in the Similarity condition ($*[\text{place}_i][\text{place}_i]$), $*[\text{place}_i][\text{place}_i]$ automatically generalizes to apply in identical consonant pairs, but identity avoidance is just a special case of similarity avoidance and may not generalize from this subset pattern to its superset. This subset-superset relation between the two patterns predicts that being exposed in the more general pattern, the similarity avoidance pattern, may serve as stronger evidence for the alternation rule compared to the identity avoidance pattern. Therefore, the typology and the structural relation between the two exposure patterns make different predictions on the direction of generalization.

In terms of statistical analysis, these predicted generalization differences should be reflected by a significant interaction between experimental condition (Similarity vs. Identity) and the prompt type (whether the consonant pair occurred in the exposure phase or not). Yet the different predictions above dictate different direction of this interaction term.

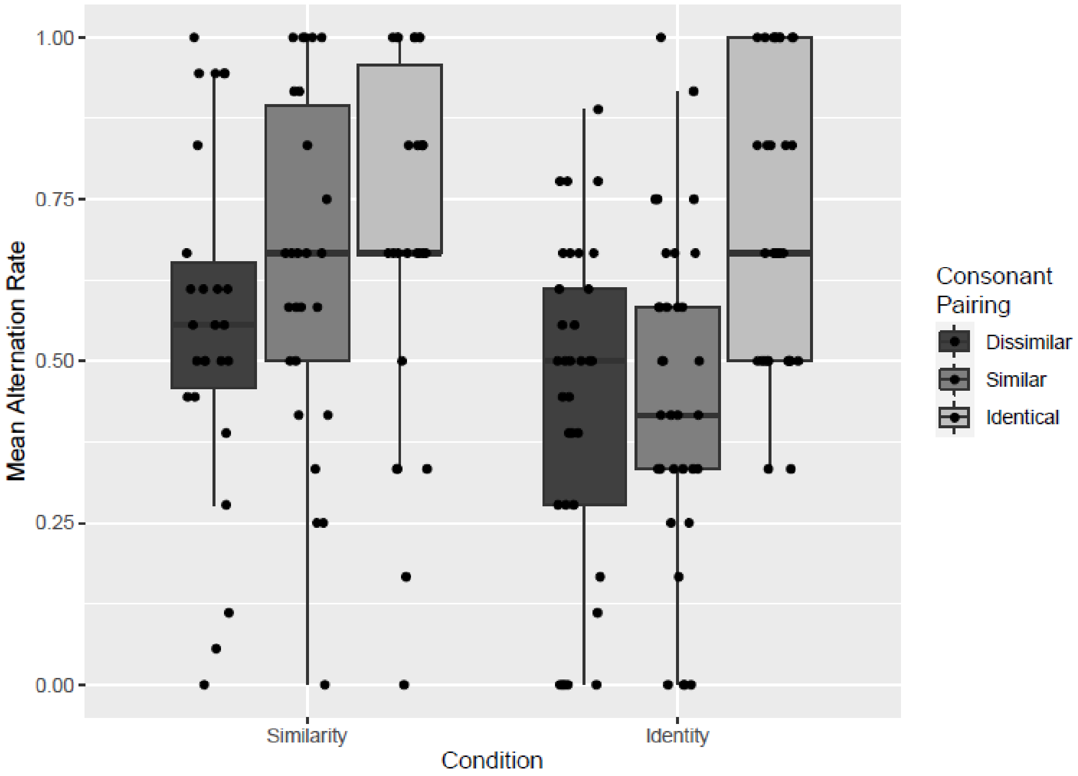
8.5. Data Analysis

Participants' choices were coded as a binary dependent variable of the statistic model (choosing the alternating response was coded as 1, and choosing the non-alternating response was coded as 0). The two independent variables were experimental condition (Similarity vs. Identity) and prompt type, i.e., whether the consonant pair of the test prompt occurred in the training (Exposure vs. Novel vs. Dissimilar). The data were analyzed by mixed-effects logistic regression models using the *lme4* package (Bates et al., 2015) of the R software (R Core Team, 2018). The fixed effect structure included the two independent variables and their interaction. The random effect structure was determined by backwards stepwise model comparisons.

8.6. Results

Participants' responses were coded as alternated or non-alternated. One participant from the Identity training condition gave alternating choices to all 36 test prompts. Her data were excluded from further analysis[7]. The remaining dataset thus included 26 participants in the Similarity condition and 33 participants in the Identity condition.

The mean rates of the alternated choice for each prompt type in each experimental condition are shown in Figure 10. There are three types of prompts: 18 prompts with two place disagreeing consonants (Dissimilar), 12 prompts with two consonants agreeing in place but not identical (Similar), and 6 prompts with two identical consonants (Identical). Recall that the prediction from phonological typology states that the generalization of the alternation rule to the withheld environments should be stronger in the Identity condition, while the prediction from structural bias says that the generalization should be stronger in the Similarity condition.



---

[7] However, according to the distribution of the log reaction time on all the test trials for the skewed participant (mean = 7.363, sd = 0.631), they did not show faster responses compared to the remaining participants (mean = 7.170, sd = 0.974).

Figure 10 Proportion of alternating responses to the test prompts grouped by condition and prompt type. Boxes indicate the range between the first and third quartiles. Whiskers delimit the minimum and maximum data points, excluding any outliers.

A logistic mixed-effects regression model on the alternation responses was built using experimental condition (Similarity vs. Identity) and prompt type (Exposure, Novel, Dissimilar) as the fixed-effects predictors. Note that the prompt type had three levels and was coded differently from what is seen in Figure 10. A prompt was coded as Exposure if the consonant pair it carried occurred in the training phase, and prompts having the withheld consonant pairs were coded as Novel. In other words, similar agreeing test prompts were coded as Exposure for participants in the Similarity training condition and as Novel for participants in the Identity training condition. Correspondingly, test prompts with two identical consonants were coded as Novel for participants in the Similarity training condition and as Exposure for participants in the Identity training condition. Finally, all prompts with disagreeing consonant pairs were coded as Dissimilar, and there are no specific predictions about these dissimilar prompts. Random factors included random intercepts for participants and test prompts. This was the maximal model that converged. Adding any other random slopes caused the model fail to converge. The model summary is provided in Table 8. The *p*-values were obtained using the *lmerTest* package (Kuznetsova et al., 2017).

	Estimate	Std Error	*t* value	*p* value
(Intercept)	0.7826	0.2564	3.0523	0.0023**
Identity	0.2644	0.3683	0.7179	0.4728
Novel	0.3133	0.2582	1.2137	0.2249
Dissimilar	-0.4380	0.1869	-2.3428	0.0191*
Identity:Novel	-1.5886	0.3764	-4.2201	<.0001***
Identity:Dissimilar	-0.9128	0.2771	-3.2945	0.0010***

Table 8 The maximal model for alternation rates.

The baseline for experimental condition was set as Similarity and the baseline for prompt type was Exposure. The main effect of prompt type suggests that the alternating rate for the exposure test prompts in the Similarity condition was significantly higher than the dissimilar prompts, illustrating a learning effect. Crucially, there was a significant interaction between condition and prompt type. The alternation rate for the Novel prompts in the Similarity condition was higher than the Exposure prompts, while for the Identity condition, the alternation rate for the novel prompts was lower than the Exposure prompts. The alternation rate differences between prompt types were thus significantly different across the two experimental conditions. These results show that participants were highly prone to extend the OCP-based alternation to unseen identical environments when they were trained on similarity avoidance, whereas they were reluctant to do so when they were trained on identity avoidance. Because identity avoidance is a subset pattern of similarity avoidance, participants trained in similarity avoidance may naturally extend to the subset identity avoidance pattern, while the other direction is less likely to happen. This corroborates with the structural bias view.

By-participant analysis showed that in the Similarity training condition, only 11 out of 26 participants exhibited a greater alternation rate in prompts with two similar consonants (Exposure) than in prompts with two identical consonants (Novel). However, in the Identity condition, 25 out of 33 participants showed a greater alternation rate in prompts with two similar consonants (Exposure) than in prompts with two identical consonants (Novel). And the alternation rate differences between the two prompt types were bigger in the Identity condition.

8.7. Discussion

Experiment 5 tested participants' willingness to generalize to novel contexts from similarity avoidance to identity avoidance as well as the opposite direction, from identity avoidance to similarity avoidance. The typology of OCP patterns predicts that identity training should lead to more generalization, because some languages with OCP effects tolerate identical sequences and OCP on identity implies OCP on similar segments but not vice versa. However, the results showed that similarity avoidance training led to more generalization to identity avoidance, which contradicted with the typological observation. I offer some preliminary explanations to why identity exemption is found in typology but does not show up in a laboratory setting.

First, it is noteworthy that identity avoidance is a subset of the more general similarity avoidance pattern, that is, *Identity is a special case of *[place$_i$][place$_i$]. In the Identity condition of the experiment, participants were trained to alternate when the two consonants in the prompts were identical. Two possible types of generalizations can be learned from this pattern: 1) the more specific *Identity, and 2) the more general *[place$_i$][place$_i$]. Then, for the test prompts with two identical consonants, identity avoidance originates from the additive effects of both generalizations, whereas test prompts with two similar consonants can only benefit from one generalization (the more general *[place$_i$][place$_i$]). Therefore, the alternation rate is higher for the identical test prompts than the similar test prompts because the former is influenced by both generalizations. In the Similarity condition, participants were trained on similar-but-non-identical prompts, while the evidence about identity avoidance was withheld, which means that participants could only acquire one generalization, *[place$_i$][place$_i$]. Crucially, this generalization is effective in both identical and similar test prompts, and the effect should be equal between the two types of prompts. This is indeed what we found: the regression analysis in Table 8 suggests that the alternations rates between Exposure and Novel prompts in the Similarity condition are not significantly different from each other. The outcome is that similarity avoidance effectively generalizes to identity avoidance, but the generalization from identity avoidance to similarity avoidance is weak.

Second, recall that there were altogether 9 possible place-agreeing consonant pairs. In the training of the Identity condition, participants were only trained on 3 identical pairs out of all 9 possible ones: they needed to alternate the second consonant if the two consonants in the prompts were the same. Nevertheless, for the Similarity condition, the alternation rule was carried by 6 similar-but-non-identical consonant pairs in the training, which means that participants in this condition were exposed to quantitatively more diverse environments for this alternation rule, and hence the learning was more efficient in the Similarity condition due to this broader evidence available in the exposure phase.

Third, as Graff (2012) has proposed, identity exemption is not supported by perceptual grounding. For example, CVC words with two identical consonants in spectrum noise were as likely to be misidentified as words with non-identical consonants sharing other features (Woods et al., 2010). Unlike the asymmetrical implicational relation between OCP-Labial vs. OCP-Dorsal, which is backed by solid perceptual and acoustic factors, the asymmetrical relation between similarity and identity avoidance in the typology lacks such grounding. Therefore, the observed typology of identity exemption might just be a language-specific lexical effect that is adopted to maximize the perceptual distinction among all the lexical entries in the lexicon. Some languages happen to choose identity exemption to achieve this goal, and this choice is not supported by a substantive bias which favors the co-occurrence of two identical segments. If the observed identity exemption effect is caused by the structure of the lexicon, then the minimal design of the AGL training materials may not be sufficient to allow these lexical trends to be acquired. Moreover, the typology of identity exemption is not clear compared to the OCP-Place asymmetry: we do not know exactly how many languages with similarity avoidance tolerate two identical segments, neither do we know whether there is a language that only bans identity but allows similarity. If the typology about identity itself is inconclusive, this could be the reason why the results of AGL tasks do not match with it.

Another possibility is that the identity exemption effect found in languages like Muna is representationally different from the identity effect in the experiment. In the experiment, the six consonants involved were viewed as collections of feature bundles. Therefore, two identical consonants constituted a more serious violation of the OCP compared to similar consonants, because the former had more featural overlaps than the latter. A closer look at the data of previous experiments can further support this argument. Recall that in the OCP condition of Experiment 1, participants were trained with CVCV words whose two consonants disagreed in place. If we look at the binary acceptability data of all the test items with agreeing consonant pairs (items whose consonant pairs were not attested in the training), we can see that items with two identical consonants received significantly lower acceptability than items with two similar-but-non-identical consonants (t = -5.502, p < .0001). This indicates that the learning effect of the OCP pattern is stronger in the identity situation, because identity avoidance is a special and a more typical case of similarity avoidance, and training in a general similarity avoidance pattern (OCP-Place) can efficiently generalize to the specific case of identity avoidance. In real languages that exhibit identity exemption, the segments themselves may be viewed as representational primitives. The correspondence between two identical segments is no longer a special case of the OCP but is encoded as multiple instantiations of the same unit, hence they are protected from the OCP and are allowed to co-occur.

The results of Experiment 5 also demonstrate that phonotactic constraints based on variables have a special status in phonological grammar. Some phonological constraints are algebraic, in that they operate on variables that stand for a whole class of items, rather than specific instances. The training on similarity avoidance could extend to identical consonant pairs, which means that participants learned a constraint based on variables *[place$_i$][place$_i$], not just the specific co-occurrence restrictions on the 12 similar-but-non-identical consonant pairs. The OCP constraints can be analyzed as piecemeal co-occurrence restrictions on certain segments (e.g., OCP-Labial = *[p^h...p^h], *[p^h...p], *[p...f], etc.) or constraints operated on variables (e.g., OCP-Labial = *[Labial][Labial], which extends to all labial sounds). There is good reason to believe that the variable analysis is the preferred one. Berent et al. (2012) compared the prediction accuracies of these two analyses on speakers' non-word judgments and found that the variable account offered a better match with speakers' behaviors. Native speakers of Hebrew can generalize a co-occurrence constraint on two repeated consonants within a root to sounds that do not exist in Hebrew, suggesting that this constraint on identical sounds must be represented as a constraint on repeating variables, not just a series of piecemeal constraints that only account for the sounds existing in Hebrew. Gallagher (2016) also reported that variable representations have an advantage over piecemeal representations. There are two phonotactic patterns in Quechua which can be summarized as 1) ejective co-occurrence *[+constricted glottis][+constricted glottis] and 2) ejective ordering *[−cont, −son, −constricted glottis][+constricted glottis]. Production and perception experiments showed that the first phonotactic pattern (ejective co-occurrence) was phonologically more 'active' in native speakers' phonotactic knowledge, because ejective co-occurrence exhibited consistent misperception across different environments and caused more phonological repairs in production repetition tasks. The results of Experiment 5 support that OCP constraints operate on variables. Since the co-occurrence restriction on the same place (*[place$_i$][place$_i$]) can be generalized as an algebraic constraint on the repeated place feature, it explains why in the Similarity condition, the training on similarity avoidance can generalize to identity avoidance. If this co-occurrence restriction is represented as piecemeal constraints which account for each consonant pair individually (*p...p^h, *p...f, etc.), we would not predict that the training will generalize to other consonant pairs, identical or non-identical.

The current results also support that speakers favor simpler and more general patterns when facing multiple generalizations. Durvasula & Liter (2020) asked participants to learn a phonotactic pattern where consonants had both voicing and continuancy agreements. Multiple generalizations can be made from this pattern, including simple ones like voicing harmony, continuancy harmony, and more complicated ones like voicing plus continuancy harmony. They prepared test stimuli which only conformed to voicing harmony, only to continuancy harmony, and to both voicing and continuancy harmony. The results showed that the acceptability of fully conforming items could be analyzed as simple additive effects of the acceptability in voicing harmony only and in continuancy harmony only. In other words, conforming to both voicing and continuancy did not result in extra acceptability boost, which means that participants were not learning the more complicated voicing plus continuancy harmony pattern. In the Similarity condition of Experiment 5, for example, participants were also facing two possible generalizations: 1) a co-occurrence constraint on two consonants with the same place, and 2) the same co-occurrence constraint plus an identity exemption rule, which allows two identical consonants to occur in a word. Structurally, the former one is simpler, and thus participants learned this single *[place$_i$][place$_i$] constraint only and extended that to the unseen identical consonant pairs.

The potential effects of nuisance factors, such as the missing syllables [k^{h}ei4] and [pou4], and the vowel distributions, were examined. Whether the test items contained the syllable [k^{h}ei4] or [pou4] could not significantly improve the model on participants' alternation choices ($\chi^2 = 0.0109$, $p = .9167$), and neither did the distances between the two vowels in the test items ($\chi^2 = 3.2546$, $p = .0712$).

9. Concluding Remarks

9.1. The Roles of Channel Bias and Analytic Bias

Channel bias refers to the systematic perception and production errors during language transmission. Evolutionary Phonology believes that channel bias is the only factor that influences phonological typology via sound changes triggered by these phonetic predispositions. On the other hand, Phonetically Based Phonology additionally posits the role of analytic bias in shaping phonological typology. Speakers possess detailed phonetic or structural knowledge that can guide synchronic phonological acquisition.

Most of the results in the current study can be interpreted as channel bias effects. For example, in Experiment 1, an OCP phonotactic pattern was easier to learn than a place harmony pattern and an arbitrary pattern, because the disagreeing consonant pairs were more clearly perceived than agreeing consonant pairs (cf. Woods et al.'s (2010) CVC identification results). However, the poverty of the stimulus design of Experiments 4 and 5 argues for an analytic bias effect. Recall that in these two experiments, certain contexts were withheld in the training, so the participants would not know the phonetic precursors of the withheld items. For example, in the Labial condition of Experiment 4, participants were only exposed to labial-agreeing consonant pairs (and also dissimilar pairs), so they had no idea about the perceptibility of dorsal-agreeing pairs. Later in the test phase, participants' responses on dorsal-agreeing items would not be affected by the precursor strength of these items. Participants' asymmetrical behaviors in contexts they did not receive any training are an argument for analytic bias.

Nevertheless, the experiments in this study were not designed to control the precursor strength in different patterns of the experimental conditions, so the results could not tell us whether the OCP learning advantage is simply due to perceptual difficulty or is part of speakers' synchronic grammatical system. My interpretation is that both channel and analytic biases may contribute to this OCP advantage and bring about its typological well-attestedness.

9.2. Implications for AGL as a Method for Phonological Learning

A major concern about using AGL to investigate phonological learning is how it compares with real language acquisition. It may not represent first language acquisition, and the short exposure phase in these tasks may not necessarily lead to a long-lasting first language phonological pattern (Moreton & Pater, 2012b). It seems like what is learned from AGL is mostly short-term memory, yet studies have shown that the learned knowledge can become long-term memory if overnight sleep is involved. For example, Martin & Peperkamp (2020) reported that an alternation based on vowel harmony was better learned if the training sessions were separated by an overnight sleep. Ettlinger et al. (2016) argued that AGL is equivalent to second language learning, especially when there was a semantic learning component involved. However, for most phonological patterns implemented in AGL, they appear to be easier to learn than the patterns in real second language (Moreton & Pater, 2012b). Perhaps due to the limited variability of the training items, as discussed in §3.1.4, AGL may be more explicit than real second language learning, and participants' strategies can also influence the results. Since AGL does not completely imitate real language acquisition, this may help us to understand why some typological asymmetries are successfully reproduced (e.g., the OCP-Place asymmetry), while other asymmetries are not (e.g., the identity exemption). The identity avoidance pattern carried by the artificial grammars in Experiment 5 is explicitly a subset pattern of similarity avoidance; while in real languages, identity avoidance is established as statistical tendencies and may have exceptions. These lexical variations may be acquired by a different mechanism than what is adopted in AGL.

Related to participants' individual strategies, many AGL studies have shown that there are noticeable individual differences in the learning outcome and learning strategies of AGL experiments (e.g., Berent, 2013; Finley, 2011; Y.-L. Lin, 2019; Zimmerer et al., 2011). Kapa & Colombo (2014) also found that executive functions such as the inhibitory control ability can affect participants' AGL performance. In the discussion sections of each experiment, I offered some preliminary descriptions about participants' individual behaviors. Some of them could figure out and master the intended pattern, while others were more uncertain about their training. Further analysis showed that the amount of time spent during exposure and the reaction time on the test trials cannot predict these individual differences. Although individual difference in AGL is an interesting topic and is an understudied area, this dissertation was not designed to study its role in AGL tasks. Future studies should be conducted to further investigate this effect.

9.3. Implications for Phonological Theories

This dissertation has provided evidence for the role of the Obligatory Contour Principle in phonological learning; hence we have obtained a better understanding of the nature of various learning biases speakers implement during acquisition. Moreover, the learnability difference across the OCP and Harmony conditions in Experiment 1 suggests that this bias is a domain-specific linguistic substantive bias, not just a domain-general cognitive predisposition. Therefore, the current study adds another piece of evidence for the role of substantive bias in phonological learning.

During typological survey, we find that some phonological patterns occur over and over again, some other patterns occur less frequently, and some logically possible patterns are completely unattested in language. In the case of similarity avoidance, because it is supported by perceptual and articulatory grounding, lexical items that conform to the OCP are easier to process and produce. As a result, these forms will be favored in language acquisition, lexical borrowing, and coining of new words. Over the course of many generations, the OCP patterns can better survive waves of language changes and become typologically well-attested (Frisch et al., 2004).

As mentioned earlier, whether these pressures from phonetics are purely instantiated by diachronic sound changes, or synchronic predispositions also play a role, is a question under debate (Evolutionary Phonology vs. Phonetically Based Phonology). By examining phonological acquisition using the AGL paradigm, the current study has shown that synchronic learning biases are also one of the factors that help to shape phonological typology, hence the theory of Phonetically Based Phonology is supported. Due to this synchronically biased learning, children may acquire a linguistic system slightly different from adult’s grammar (Kiparsky, 1982), and the synchronic learning biases may influence this process by facilitating the learning of certain phonetically grounded (substantive bias) or structurally simple (structural bias) patterns and hindering the learning of unmotivated patterns.

The main contribution of the current study is that similarity avoidance (OCP) belongs to the set of such synchronic learning biases. By favoring the learning of OCP-based patterns, similarity avoidance can be easily and efficiently learned through generations, and thus OCP is a typologically frequent pattern. Notice that this argument does not preclude the effects of diachronic sound changes on phonological typology. Both synchronic and diachronic factors can, and should, contribute to the typological asymmetries in the world's languages.

The effects of learning biases can be assessed by computational modeling. Since the input of the AGL experiments is fixed, we can feed the learning materials to phonological learning models such as the MaxEnt framework (Hayes & Wilson, 2008) to obtain predictions about speakers' behaviors based on these inputs. Then, we can bias the learning algorithm to favor the OCP-based patterns, to see if the biased predictions can better match with participants' learning outcome (Gong & Zhang, 2021; White, 2017; Wilson, 2006; Zhang et al., 2011). I leave this modeling work for future inquiries.

9.4. Summary of the Dissertation

Understanding the nature of phonological learning is a key issue in the field of phonology. Previous literature has shown that phonological learning is guided by various types of biases that render certain patterns easier to learn and thus typologically more common among languages. The similarity avoidance effects in phonology have been argued to universally exist in the world's languages (Suzuki, 1998; Walter, 2007). Experimental evidence has also suggested that similarity avoidance plays an active role in speakers' phonological knowledge (Berent et al., 2001; Berent & Shimron, 1997; Berkley, 1994; Frisch & Zawaydeh, 2001). In recent phonological research, artificial grammar learning paradigm has often been used to test what biases speakers bring in during phonological learning (Moreton, 2008; Moreton & Pater, 2012a, 2012b; Wilson, 2006). This paradigm allows for the flexibility to test linguistic patterns that are difficult to tap into in natural languages. Focusing on the typologically wide-attested OCP effect, using the artificial grammar learning paradigm, this dissertation investigates whether phonological learning is biased by similarity avoidance, such that consonant major place disagreement patterns receive a learning advantage and thus recur in the world's languages.

I first present evidence to show that an OCP-based phonotactic constraint that prohibits within-stem similar consonant place co-occurrence is more learnable than a) a consonant major place harmony pattern, and b) an arbitrary, more complex phonotactic pattern (Experiment 1). Moreover, I show that an OCP-motivated alternation pattern is more learnable than an arbitrarily conditioned alternation (Experiment 2). These two experiments lay out the foundation that an OCP bias exists in both phonotactic and alternation learning. Based on these findings, a natural follow-up question would be the interaction between phonotactics and alternation, which received vast attention since the early traditional generative phonology (the Duplication Problem, Kenstowicz & Kisseberth (1977)), and later motivated the constraint-based phonological framework (Optimality Theory, Prince & Smolensky (1993)). This interaction is supported by the finding that OCP effect in stem phonotactics can aid the learning of an OCP-motivated alternation (Experiment 3). After the learning advantage of the OCP-based patterns has been explored by these experiments, I further investigate whether two typological asymmetries within the OCP patterns – strength variation among consonant places and identity exemption – have their sources in learning biases as well. The results suggest that the place asymmetry can be accounted for by a learning bias favoring OCP-Labial (Experiment 4), while the identity exemption effect is not reproduced in the laboratory setting (Experiment 5).

This dissertation also shows that the learning advantage of the OCP patterns is not solely due to the domain-general structural bias which favors formally simpler patterns. Results from Experiments 1 and 4 suggest that learning asymmetry still emerges when the two patterns match in formal complexity (OCP-Place vs. place harmony, and OCP-Labial vs. OCP-Dorsal). Therefore, we can conclude that this OCP learning bias belongs to domain-specific linguistic knowledge.

It should be noted that the native language of the participants in the current study – Mandarin, does not show OCP-Place effects among the consonants in multisyllabic words (Boll-Avetisyan, 2012; Graff, 2012). Therefore, participants' behaviors in the current tasks cannot be attributed to prior L1 knowledge, so the learning advantage of the OCP patterns provides strong evidence for the synchronic learning biases in phonological acquisition and the universality of such biases.

Together, this dissertation contributes to our understanding of the nature of the OCP and how this principle may serve as a learning bias in phonological acquisition and how it can influence phonological typology. In a broader sense, using OCP as a test case, this dissertation provides further evidence to support that phonological learning is a biased process. We thus moved one more step forward in understanding the mechanisms of how speakers internalize the linguistic input they receive and how they use that knowledge to produce language.

References

Anttila, Arto. (2006). Variation and opacity. *Natural Language & Linguistic Theory*, *24*(4).

Anwyl-Irvine, Alexander L., Massonnié, Jessica, Flitton, Adam, Kirkham, Natasha, & Evershed, Jo K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407.

Baer-Henney, Dinah, & van de Vijver, Ruben. (2012). On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology*, *3*(2), 221–249.

Bates, Douglas, Mächler, Martin, Bolker, Benjamin M., & Walker, Steven C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*.

Bauer, Robert S., & Benedict, Paul K. (1997). Modern Cantonese phonology. In *Trends in linguistics Studies and monographs* (Issue 102). Mouton de Gruyter.

Becker, Michael, Ketrez, Nihan, & Nevins, Andrew. (2011). the Surfeit of the Stimulus: Analytic Biases Filter Lexical Statistics in Turkish Laryngeal Alternations. *Language*, *87*(1), 84–125.

Berent, Iris. (2013). *The phonological mind*. Cambridge University Press.

Berent, Iris, & Shimron, Joseph. (1997). The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition*, *64*, 39–72.

Berent, Iris, Shimron, Joseph, & Vaknin, Vered. (2001). Phonological constraints on reading: evidence from the obligatory contour principle. *Journal of Memory and Language*, *44*(4), 644–665.

Berent, Iris, Wilson, Colin, Bemis, Douglas K., & Marcus, Gary F. (2012). On the Role of Variables in Phonology: Remarks on Hayes and Wilson 2008. *Linguistic Inquiry*, *43*(1), 97–119.

Berkley, Deborah. (1994). The OCP and gradient data. *Studies in the Linguistic Sciences*, *24,* 59–72.

Berko, Jean. (1958). A child's learning of English morphology. *Word*, *14*(November), 150–77.

Blevins, Juliette. (2004). *Evolutionary Phonology*. Cambridge University Press.

Boersma, Paul. (2000). The OCP in the perception grammar. *Rutgers Optimality Archive*, *435*, 1–52.

Boersma, Paul, & Weenink, David. (2017). *Praat: doing phonetics by computer* (6.0.33).

Boll-Avetisyan, Natalie. (2012). *Phonotactics and its acquisition, representation, and use: An experimental-phonological study*. LOT Publications.

Chambers, Kyle E., Onishi, Kristine H., & Fisher, Cynthia. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69–B77.

Chan, I. Lei, & Chang, Charles B. (2018). LEXTALE_CH: A quick, character-based proficiency test for Mandarin Chinese. *Proceedings of the Annual Boston University Conference on Language Development*.

Chen, Tsung-ying. (2022). On the learnability of level-based and unit-based tonal OCP constraints: An artificial grammar learning study. *Glossa*.

Chong, Adam Junxiang. (2017). *On the relation between phonotactic learning and alternation learning (PhD thesis)*. University of California, Los Angeles.

Chong, Adam Junxiang. (2021). The effect of phonotactics on alternation learning. *Language*, *97*(2), 213–244.

Clements, George N., & Hume, Elizabeth V. (1995). The internal organization of speech sounds. In *The handbook of phonological theory* (Issue 1).

Coetzee, Andries W. (2008). Grammaticality and ungrammaticality in phonology. *Language*, *84*(2), 218–257.

Coetzee, Andries W. (2009). Learning lexical indexation. *Phonology*, *26*(1), 109–145.

Coetzee, Andries W., & Pater, Joe. (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory*, *26*(2), 289–337.

Cristià, Alejandrina, & Seidl, Amanda. (2008). Is Infants' Learning of Sound Patterns Constrained by Phonological Features? *Language Learning and Development*, *4*(3).

Dell, Gary S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(2), 222–233.

Dell, Gary S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*.

Dell, Gary S., Burger, Lisa K., & Svec, William R. (1997). Language Production and Serial Order : A Functional Analysis and a Model. *Psychological Review*, *104*(1), 123–147.

Do, Youngah, & Lai, Ryan Ka Yau. (2020). Incorporating tone in the modelling of wordlikeness judgements. *Phonology*, *37*(4), 577–615.

Do, Youngah, & Yeung, Ping Hei. (2021). Evidence against a link between learning phonotactics and learning phonological alternations. *Linguistics Vanguard*, *7*(1), 839–878.

Duanmu, San. (1994). Against contour tone units. *Linguistic Inquiry*, *25*(4), 555–608.

Duanmu, San. (2007). *The Phonology of Standard Chinese* (2nd ed.). Oxford University Press.

Durvasula, Karthik, & Liter, Adam. (2020). There is a simplicity bias when generalising from ambiguous data. *Phonology*, *37*(2), 177–213.

Ettlinger, Marc, Morgan-Short, Kara, Faretta-Stutenberg, Mandy, & Wong, Patrick C. M. (2016). The relationship between artificial and second language learning. *Cognitive Science*, *40*(4).

Finley, Sara. (2011). The privileged status of locality in consonant harmony. *Journal of Memory and Language*.

Finley, Sara. (2012a). Testing the limits of long-distance learning: Learning beyond a three-segment window. *Cognitive Science*, *36*(4).

Finley, Sara. (2012b). Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes*.

Finley, Sara, & Badecker, William. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437.

Frisch, Stefan A. (2004). Language processing and segmental OCP effects. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 346–371). Cambridge University Press.

Frisch, Stefan A., Pierrehumbert, Janet B., & Broe, Michael B. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, *22*(1), 179–228.

Frisch, Stefan A., & Zawaydeh, Bushra Adnan. (2001). The psychological reality of OCP - place in Arabic. *Language*, *77*(1), 91–106.

Gallagher, Gillian. (2010). Perceptual distinctness and long-distance laryngeal restrictions. *Phonology*, *27*(3), 435–480.

Gallagher, Gillian. (2013). Learning the identity effect as an artificial language: bias and generalisation. *Phonology*, *30*(2), 253–295.

Gallagher, Gillian. (2016). Asymmetries in the representation of categorical phonotactics. *Language*, *92*(3), 557–590.

Glewwe, Eleanor. (2019). *Bias in Phonotactic Learning: Experimental Studies of Phonotactic Implicationals (PhD thesis)*. UCLA.

Gomez, Rebecca L., & Gerken, Louann. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Gong, Shuxiao, & Zhang, Jie. (2021). Modelling Mandarin speakers' phonotactic knowledge. *Phonology*, *38*(2), 241–275.

Graff, Peter. (2012). *Communicative efficiency in the lexicon (PhD thesis)*. MIT.

Hansson, Gunnar Ólafur. (2010). *Consonant harmony: long-distance interaction in phonology*. University of California Press.

Hayes, Bruce. (2004). Phonological acquisition in Optimality Theory: The early stages. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in Phonological Acquisition* (pp. 158–203). Cambridge University Press.

Hayes, Bruce. (2009). *Introductory Phonology*. John Wiley & Sons, Inc.

Hayes, Bruce, Kirchner, Robert, & Steriade, Donca. (2004). *Phonetically Based Phonology* (B. Hayes, R. Kirchner, & D. Steriade (eds.)). Cambridge University Press.

Hayes, Bruce, Siptar, Peter, Zuraw, Kie, & Londe, Zsuzsa. (2009). Natural and Unnatural Constraints in Hungarian Vowel Harmony. *Language*, *85*(4), 822–863.

Hayes, Bruce, & White, James. (2013). Phonological naturalness and phonotactic learning. *Linguistic Inquiry, 44*(1), 45–75.

Hayes, Bruce, & White, James. (2015). Saltation and the P-map. *Phonology*, *32*(2), 267–302.

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

Heinz, Jeffrey. (2010). Learning long-distance phonotactics. *Linguistic Inquiry, 41*(4), 623–661.

Jusczyk, Peter W., & Luce, Paul A. (1994). Infants′ sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*(5), 630–645.

Kager, René, & Shatzman, Keren. (2007). Phonological constraints in speech processing. *Linguistics in the Netherlands*, *24*(1), 99–111.

Kapa, Leah L., & Colombo, John. (2014). Executive function predicts artificial language learning. *Journal of Memory and Language*, *76*.

Kaun, Abigail R. (2004). The typology of rounding harmony. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically Based Phonology*. Cambridge University Press.

Kenstowicz, Michael John, & Kisseberth, Charles Wayne. (1977). *Topics in phonological theory*. Academic Press.

Kiparsky, Paul. (1982). *Explanation in phonology*. Foris Publications.

Kiparsky, Paul. (1993). Blocking in nonderived environments. In S. Hargus & Ellen M. Kaisse (Eds.), *Phonetics and phonology 4: Studies in lexical phonology* (pp. 277–313). Academic Press.

Koo, Hahn, & Callahan, Lydia. (2012). Tier-adjacency is not a necessary condition for learning phonotactic dependencies. *Language and Cognitive Processes*, *27*(10).

Kuo, Li Jen. (2009). The role of natural class features in the acquisition of phonotactic regularities. *Journal of Psycholinguistic Research*, *38*(2), 129–150.

Kuznetsova, Alexandra, Brockhoff, Per B., & Christensen, Rune H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13).

Lai, Regine. (2015). Learnable vs. Unlearnable Harmony Patterns. *Linguistic Inquiry, 46*(3), 425–451.

Leben, William. (1973). *Suprasegmental Phonology (PhD thesis)*. MIT.

Li, A., & Xu, B. (2001). *Chinese Annotated Dialogue and Conversation Corpus*. Institute of Linguistics, Chinese Academy of Social Sciences.

Lin, Isabelle. (2017). *Tone sequences in lexical processing of Beijing Mandarin (MA thesis)*. University of California, Los Angeles.

Lin, Yu-Leng. (2019). *Substantive bias and natural classes: An empirical approach* (C.-R. Huang & R. Guo (eds.)). Peking University Press and Springer Nature Singapore.

Linzen, Tal, & Gallagher, Gillian. (2017). Rapid generalization in phonotactic learning. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1), 1–32.

Martin, Alexander, & Peperkamp, Sharon. (2020). Phonetically natural rules benefit from a learning bias: A re-examination of vowel harmony and disharmony. *Phonology*.

Martin, Alexander, & White, James. (2021). Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry*, *52*(1), 227–239.

Martin, Andrew. (2007). *The Evolving Lexicon (PhD thesis)*. University of California, Los Angeles.

Mayer, Thomas, Rohrdantz, Christian, Plank, Frans, Bak, Peter, Butt, Miriam, & Keim, Daniel A. (2010). Consonant co-occurrence in stems across languages: Automatic analysis and visualization of a phonotactic constraint. *Proceedings of the ACL 2010 Workshop on NLP and Linguistics: Finding the Common Ground*.

McCarthy, John J. (1986). OCP Effects: Gemination and antigemination. *Linguistic Inquiry*, *17*(2), 207–263.

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

Moreton, Elliott. (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language*, *67*(1), 165–183.

Moreton, Elliott, & Pater, Joe. (2012a). Structure and Substance in Artificial-Phonology Learning, Part I: Structure. *Language and Linguistics Compass*, *6*(11), 686–701.

Moreton, Elliott, & Pater, Joe. (2012b). Structure and Substance in Artificial-Phonology Learning, Part II: Substance. *Language and Linguistics Compass*, *6*(11), 702–718.

Myers, Scott, & Padgett, Jaye. (2014). Domain generalisation in artificial language learning. *Phonology*, *31*, 399–433.

Nevins, Andrew. (2009). On formal universals in phonology. In *Behavioral and Brain Sciences* (Vol. 32, Issue 5). https://doi.org/10.1017/S0140525X09990537

Nevins, Andrew. (2010). *Locality in Vowel Harmony*. The MIT Press.

Newport, Elissa L. (2016). Statistical language learning: computational, maturational, and linguistic constraints. *Language and Cognition*, *8*(3), 447–461.

Ohala, John J. (1981). The listener as a source of sound change. In *Papers from the parasession on language and behavior Chicago Linguistic Society* (pp. 178–203).

Ohala, John J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical Linguistics: Problems and Perspectives* (p. 237−278). Longman Academic.

Ohala, John J. (1996). Speech perception is hearing sounds, not tongues. *The Journal of the Acoustical Society of America*, *99*(3).

Onishi, Kristine H., Chambers, Kyle E., & Fisher, Cynthia. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1).

Onnis, Luca, & Thiessen, Erik. (2013). Language experience changes subsequent learning. *Cognition*, *126*(2).

Pater, Joe, & Tessier, Anne Michelle. (2006). L1 phonotactic knowledge and the L2 acquisition of alternations. In *Inquiries in Linguistic Development: In Honor of Lydia White*.

Pierrehumbert, Janet. (1993). Dissimilarity in the Arabic verbal roots. In *Proceedings of the 23rd meeting of the {North East Linguistic Society}* (pp. 367–381).

Pozdniakov, Konstantin, & Segerer, Guillaume. (2007). Similar Place Avoidance: A statistical universal. *Linguistic Typology*, *11*(2), 307–348.

Prince, Alan, & Smolensky, Paul. (1993). *Optimality Theory: constraint interaction in generative grammar*. Rutgers University, New Brunswick and University of Colorado, Boulder.

Prince, Alan, & Tesar, Bruce. (2004). Learning phonotactic distributions. In *Constraints in Phonological Acquisition*.

Pycha, Anne, Nowak, Pawel, Shin, Eurie, & Shosted, Ryan. (2003). Phonological rule-learing and its implications for a theory of vowel harmony. *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL 22), January 2003*, 101–114.

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reber, Arthur S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6).

Rose, Sharon, & Walker, Rachel. (2004). A typology of consonant agreement as correspondence. *Language*, *80*(3).

Saffran, Jenny R., & Thiessen, Erik D. (2003). Pattern Induction by Infant Language Learners. *Developmental Psychology*, *39*(3).

Seidl, Amanda, & Buckley, Eugene. (2005). On the Learning of Arbitrary Phonological Rules. *Language Learning and Development*, *1*(3–4), 289–316.

Shepard, Roger N., Hovland, Carl I., & Jenkins, Herbert M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13).

Skoruppa, Katrin, & Peperkamp, Sharon. (2011). Adaptation to Novel Accents: Feature-Based Learning of Context-Sensitive Phonological Regularities. *Cognitive Science*, *35*, 348–366.

Sundara, Megha, White, James, Kim, Yun Jung, & Chong, Adam J. (2021). Stem similarity modulates infants' acquisition of phonological alternations. *Cognition*, *209*.

Suzuki, Keiichiro. (1998). *A typological investigation of dissimilation (PhD Thesis)*. The University of Arizona.

Tesar, Bruce, & Prince, Alan. (2003). Using phonotactics to learn phonological alternations. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 241–269.

Tesar, Bruce, & Prince, Alan. (2007). Using phonotactics to learn phonological alternations. *Proceedings of the Thirty-Ninth Conference of the Chicago Linguistics Society*, 209–237.

Walter, Mary. (2007). *Repetition avoidance in human language (PhD Thesis)*. MIT.

White, James. (2013). *Bias in phonological learning: evidence from saltation*. University of California, Los Angeles.

White, James. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*(1).

White, James. (2017). Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language*, *93*(1), 1–36.

White, James, & Sundara, Megha. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, *133*(1), 85–90.

Wilson, Colin. (2003). Experimental Investigation of Phonological Naturalness. *West Coast Conference on Formal Linguistics 22 (WCCFL22)*, 101–114.

Wilson, Colin. (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science*, *30*(5), 945–982.

Woods, David. L., Yund, E. William, Herron, Timothy J., & Cruadhlaoich, Matthew A. I. Ua. (2010). Consonant identification in consonant-vowel-consonant syllables in speech-spectrum noise. *The Journal of the Acoustical Society of America*, *127*(3), 1609–1623.

Yi, Li, & Duanmu, San. (2015). Phonemes, features, and syllables: Converting onset and rime inventories to consonants and vowels. *Language and Linguistics*, *16*(6), 819–842.

Yip, Moira. (1982). Reduplication and CV Skeleta in Chinese Secret Languages. *Linguistic Inquiry*, *13*(4), 637–661.

Yip, Moira. (1988). The obligatory contour principle and phonological rules : A loss of identity. *Linguistic Inquiry*, *19*(1), 65–100.

Yue-Hashimoto, Anne Oi-Kan. (1972). *Studies in Yue Dialects 1: Phonology of Cantonese*. Cambridge University Press.

Zhang, Jie, & Lai, Yuwen. (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology*, *27*, 153–201.

Zhang, Jie, Lai, Yuwen, & Sailor, Craig. (2011). Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua*, *121*(2), 181–206.

Zimmerer, Vitor C., Cowell, Patricia E., & Varley, Rosemary A. (2011). Individual behavior in learning of an artificial grammar. *Memory and Cognition*, *39*(3), 491–501.

Zuraw, Kie. (2002). Aggressive reduplication. *Phonology*, *19*(3), 395–439.

Zymet, Jesse. (2020). Malagasy OCP Targets a Single Affix: Implications for Morphosyntactic Generalization in Learning. *Linguistic Inquiry*, *51*(3), 624–634.

Appendix 1: Stimuli for Experiment 1

All stimuli are transcribed in Pinyin. The correspondence between Pinyin and IPA is as follows: p = [p^h], b = [p], k = [k^h], g = [k], h = [x]. All other symbols are the same in both systems.

Exposure

OCP	Harmony	Arbitrary	OCP	Harmony	Arbitrary	OCP	Harmony	Arbitrary
pukou	pupou	pukou	feigu	feibu	hupei	fagu	fabou	heipou
pouka	poupa	pouka	fahei	fafu	hafu	fouga	fouba	houpu
paku	papu	paku	fuhei	fufa	hufei	fugei	fubei	hapei
peikou	peipou	peikou	feihou	feifou	houfa	heipu	heiku	boubu
pakei	papei	pakei	fuhou	fufou	houfu	heipou	heikou	babu
peigou	peibou	peihou	fouha	foufa	hafei	houpu	huka	beibou
puga	puba	puha	koupu	kouku	foupu	hapei	hakei	buba
pagu	pabu	pahou	kupa	kuka	feipou	hupei	hukei	boubei
pouga	pouba	puhou	keipou	keikou	foupa	habei	hagei	poufu
pugou	pubou	peiha	keipa	keika	fapu	hubou	hukou	peifou
peihou	peifou	bukei	koupei	koukei	fupou	heibu	heigu	peifa
puha	pufa	beiku	kouba	kouga	kahei	habou	hagou	poufa
pahou	pafou	bukou	kabu	kagu	kahu	heiba	heiga	pafu
puhou	pufou	bakei	koubu	kougu	kuhou	hafu	hahu	gugei
peiha	peifa	beikou	kabou	kagou	kahou	hufei	huhei	gouga
bukei	bupei	fakou	koubei	kagei	kouhei	houfa	houha	guga
beiku	beipu	foukei	kafei	kahei	bupei	houfu	heiha	geiga
bukou	bupou	feiku	keifou	keihou	bupou	hafei	hahei	geigu
bakei	bapei	fouka	keifu	kahou	beipa			
beikou	beipou	fukei	koufu	kouhu	bapou			
bagou	babou	koupu	kufa	kuha	bapei			
bagei	babei	kupa	geipu	geiku	gukei			
bugei	bubei	keipou	gupou	gukou	gouku			
buga	buba	keipa	geipa	geika	gakei			
bouga	bouba	koupei	gapu	gaku	gouka			
bahou	bafou	kouba	goupa	gouka	geikou			
buhou	bufou	kabu	gabei	gagei	fuba			
bouha	boufa	koubu	geibou	gouga	fouba			
beihu	beifu	kabou	goubu	gougu	feibou			
bouhu	bafei	koubei	gabou	gagou	feiba			
fakou	fapou	gabei	gubei	gugei	fabou			
foukei	foupei	geibou	gafou	gahou	geihou			
feiku	feipa	goubu	goufu	gouhu	geiha			
fouka	foupa	gabou	gafei	gahei	gahou			
fukei	fupei	gubei	goufa	gouha	guha			
fuga	fuba	heipu	geifa	geiha	gouha			

Test

papou	feipu	gapei
peibu	fabu	goubei
poufei	fafei	goufei
peiku	feikou	gakou
peigu	feigou	geigou
pahei	fouhei	gahu
bapu	kapu	houpei
bubou	kubou	heibou
boufu	kufei	hafou
bouku	kukei	houku
beigou	kougei	houga
bahei	keihu	houhu

Appendix 2: Stimuli for Experiment 2

Exposure

OCP				Arbitrary			
alternating		non-alternating		alternating		non-alternating	
prompt	response	prompt	response	prompt	response	prompt	response
papu	patu	kupou	kupou	fapei	fatei	papu	papu
papou	patou	kapu	kapu	foupa	fouta	papou	papou
puba	puda	kouba	kouba	kupou	kutou	bapou	bapou
pabei	padei	keiba	keiba	kapu	katu	beipa	beipa
peifu	peisu	kafu	kafu	hupa	huta	boubu	boubu
poufei	pousei	koufa	koufa	houpei	houtei	babou	babou
bapou	batou	gapou	gapou	puba	puda	feiba	feiba
beipa	beita	goupei	goupei	pabei	padei	fabu	fabu
boubu	boudu	gubei	gubei	kouba	kouda	fufa	fufa
babou	badou	geiba	geiba	keiba	keida	foufei	foufei
bufa	busa	gafei	gafei	heibu	heidu	kafu	kafu
boufei	bousei	gafou	gafou	huba	huda	koufa	koufa
fapei	fatei	hupa	hupa	peifu	peisu	gapou	gapou
foupa	fouta	houpei	houpei	poufei	pousei	goupei	goupei
feiba	feida	heibu	heibu	bufa	busa	gubei	gubei
fabu	fadu	huba	huba	boufei	bousei	geiba	geiba
fufa	fusa	hufei	hufei	gafei	gasei	hufei	hufei
foufei	fousei	hafou	hafou	gafou	gasou	hafou	hafou

Test Prompts

OCP			Arbitrary		
alternating	non-alternating	filler	alternating	non-alternating	filler
peipu	koupu	shoufei	feipa	peipu	shoufei
pubei	kuba	zhapou	koupu	boupa	zhapou
pufa	keifa	shoufa	hupei	bubou	shoufa
boupa	gapei	zhoubu	pubei	feibou	zhoubu
bubou	gabou	sheibu	kuba	foufu	sheibu
boufu	gufa	zhupa	houbu	keifa	zhupa
feipa	hupei		pufa	gapei	
feibou	houbu		boufu	gabou	
foufu	hafei		gufa	hafei	

Appendix 3: Stimuli for Experiment 3

Exposure				Test		
OCP		Control				
stem	suffixed	stem	suffixed	stem	[-ma] suffixed	[-na] suffixed
peika	peikama	poupu	poupuna	pakou	pakouma	pakouna
pukei	pukeima	peibou	peibouna	puga	pugama	pugana
pagou	pagouma	pafei	pafeina	peiha	peihama	peihana
pougei	pougeima	poukei	poukeima	boukei	boukeima	boukeina
peihou	peihouma	peiga	peigama	beigou	beigouma	beigouna
peihu	peihuma	pahou	pahouma	buhei	buheima	buheina
buka	bukama	boupa	boupana	fouku	foukuma	foukuna
bukou	bukouma	babu	babuna	feigou	feigouma	feigouna
bougu	bouguma	bufa	bufana	fahu	fahuma	fahuna
bagei	bageima	bukei	bukeima	keipou	keipouma	keipouna
bouha	bouhama	beigu	beiguma	keibu	keibuma	keibuna
bahu	bahuma	buhou	buhouma	koufu	koufuma	koufuna
feika	feikama	foupu	foupuna	gupou	gupouma	gupouna
fukei	fukeima	feibou	feibouna	gubei	gubeima	gubeina
fagou	fagouma	fafei	fafeina	goufu	goufuma	goufuna
fougei	fougeima	fakou	fakouma	hapou	hapouma	hapouna
feihou	feihouma	fuga	fugama	heibu	heibuma	heibuna
feihu	feihuma	feiha	feihama	heifu	heifuma	heifuna
kupa	kupana	koupa	koupana	peiku	peikuma	peikuna
kupou	kupouna	kabu	kabuna	pugou	pugouma	pugouna
koubu	koubuna	kufei	kufeina	pouhu	pouhuma	pouhuna
kabei	kabeina	kukei	kukeima	baku	bakuma	bakuna
koufa	koufana	keigu	keiguma	bagou	bagouma	bagouna
kafu	kafuna	kuhou	kuhouma	beiha	beihama	beihana
geipa	geipana	goupu	goupuna	fakei	fakeima	fakeina
gupei	gupeina	geibou	geibouna	fugou	fugouma	fugouna
gabou	gabouna	gafei	gafeina	fahei	faheima	faheina
goubei	goubeina	goukei	goukeima	kapei	kapeima	kapeina
geifou	geifouna	gagou	gagouma	kuba	kubama	kubana
geifu	geifuna	geiha	geihama	keifou	keifouma	keifouna
hupa	hupana	heipa	heipana	geipu	geipuma	geipuna
hupou	hupouna	habu	habuna	gabu	gabuma	gabuna
houbu	houbuna	hufa	hufana	gufou	gufouma	gufouna
habei	habeina	hukei	hukeima	hupei	hupeima	hupeina
houfa	houfana	heigu	heiguma	habou	habouma	habouna
hafu	hafuna	huhou	huhouma	houfei	houfeima	houfeina

Appendix 4: Stimuli for Experiment 4

Exposure

Labial				Dorsal			
alternating		non-alternating		alternating		non-alternating	
prompt	response	prompt	response	prompt	response	prompt	response
pupa	puta	pakou	pakou	kuka	kuta	pakou	pakou
pupei	putei	peigou	peigou	kakou	katou	peigou	peigou
peibu	peidu	puhei	puhei	kugou	kudou	puhei	puhei
pabei	padei	beika	beika	kagou	kadou	beika	beika
peifa	peisa	bagei	bagei	kouhu	kousu	bagei	bagei
pufa	pusa	beiha	beiha	kahu	kasu	beiha	beiha
beipou	beitou	fouka	fouka	gouka	gouta	fouka	fouka
beipu	beitu	fougei	fougei	guka	guta	fougei	fougei
boubu	boudu	fahou	fahou	gougei	goudei	fahou	fahou
bouba	bouda	kapei	kapei	gagou	gadou	kapei	kapei
beifa	beisa	kuba	kuba	guhei	gusei	kuba	kuba
boufei	bousei	keifa	keifa	gouhei	gousei	keifa	keifa
feipou	feitou	goupu	goupu	haku	hatu	goupu	goupu
fapou	fatou	guba	guba	hakei	hatei	guba	guba
fuba	fuda	gafu	gafu	heigou	heidou	gafu	gafu
foubu	foudu	houpei	houpei	heigu	heidu	houpei	houpei
fafu	fasu	habei	habei	houha	housa	habei	habei
fafou	fasou	hafou	hafou	huhei	husei	hafou	hafou
poupa	pouta	pouka	pouka	kukou	kutou	pouka	pouka
papei	patei	pugou	pugou	kaku	katu	pugou	pugou
poubu	poudu	pouhu	pouhu	keigou	keidou	pouhu	pouhu
pabu	padu	bukou	bukou	kagei	kadei	bukou	bukou
poufu	pousu	bagou	bagou	keihou	keisou	bagou	bagou
poufa	pousa	buha	buha	kahou	kasou	buha	buha
beipa	beita	fakou	fakou	gouku	goutu	fakou	fakou
bupa	buta	fuga	fuga	gaku	gatu	fuga	fuga
beibou	beidou	fouhu	fouhu	gougu	goudu	fouhu	fouhu
babei	badei	keipou	keipou	gugei	gudei	keipou	keipou
beifu	beisu	keibu	keibu	gouha	gousa	keibu	keibu
bufei	busei	kafu	kafu	guha	gusa	kafu	kafu
foupei	foutei	geipou	geipou	houku	houtu	geipou	geipou
fapei	fatei	gabu	gabu	heiku	heitu	gabu	gabu
feibou	feidou	geifu	geifu	huga	huda	geifu	geifu
fabei	fadei	hupei	hupei	heiga	heida	hupei	hupei
fufou	fusou	habu	habu	heiha	heisa	habu	habu
foufa	fousa	houfei	houfei	hahou	hasou	houfei	houfei

Test Prompts

Labial			Dorsal		
alternating	withheld	non-alternating	alternating	withheld	non-alternating
papou	kukei	pukei	kukei	papou	pukei
pouba	kougu	peigu	kougu	pouba	peigu
pufou	kuhou	pahou	kuhou	pufou	pahou
boupa	gukei	bakei	gukei	boupa	bakei
bubou	gugou	beigou	gugou	bubou	beigou
bafei	gouhu	bahu	gouhu	bafei	bahu
fupa	huka	fouku	huka	fupa	fouku
fouba	hougu	feiga	hougu	fouba	feiga
feifa	huhou	fuhou	huhou	feifa	fuhou
		kupei			kupei
		kabei			kabei
		kufou			kufou
		gapei			gapei
		geibu			geibu
		geifa			geifa
		hupou			hupou
		hubei			hubei
		heifou			heifou

Appendix 5: Stimuli for Experiment 5

Exposure

Identity				Similarity			
alternating		non-alternating		alternating		non-alternating	
prompt	response	prompt	response	prompt	response	prompt	response
peipa	peita	kupou	kupou	bapu	batu	kupou	kupou
poupa	pouta	kapu	kapu	beipou	beitou	kapu	kapu
poupu	poutu	kouba	kouba	bapei	batei	kouba	kouba
peipu	peitu	keiba	keiba	feipou	feitou	keiba	keiba
papu	patu	kafu	kafu	fapei	fatei	kafu	kafu
pupei	putei	koufa	koufa	foupei	foutei	koufa	koufa
buba	buda	gapou	gapou	peibu	peidu	gapou	gapou
boubu	boudu	goupei	goupei	poubei	poudei	goupei	goupei
bubou	budou	gubei	gubei	pabu	padu	gubei	gubei
beibou	beidou	geiba	geiba	foubu	foudu	geiba	geiba
babou	badou	gafei	gafei	fubei	fudei	gafei	gafei
beibu	beidu	gafou	gafou	fabu	fadu	gafou	gafou
fafou	fasou	hupa	hupa	poufu	pousu	hupa	hupa
feifou	feisou	houpei	houpei	pafei	pasei	houpei	houpei
fufa	fusa	heibu	heibu	pufa	pusa	heibu	heibu
foufu	fousu	huba	huba	beifu	beisu	huba	huba
fafu	fasu	hufei	hufei	boufei	bousei	hufei	hufei
fafei	fasei	hafou	hafou	bufei	busei	hafou	hafou

Test Prompts

Identity			Similarity		
alternating	withheld	non-alternating	alternating	withheld	non-alternating
pupa	pubei	koupu	pubei	pupa	koupu
boubei	peifa	kuba	peifa	boubei	kuba
feifu	boupei	keifa	boupei	foufei	keifa
poupei	boufu	gapei	boufu	poupei	gapei
babu	foupu	gabou	foupu	babu	gabou
fufou	feibou	gufa	feibou	fufou	gufa
	pabei	hupei	pabei		hupei
	pafu	houbu	pafu		houbu
	boupa	hafei	boupa		hafei
	beifa	keipa	beifa		keipa
	fupa	kubou	fupa		kubou
	fabei	koufu	fabei		koufu
		goupu			goupu
		geibu			geibu
		geifa			geifa
		hapei			hapei
		habei			habei
		houfu			houfu