Sonority Effects and Learning Bias in Nasal Harmony

by

Yu-Leng Lin

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Linguistics
University of Toronto

Sonority Effects and Learning Biases on Nasal Harmony

Yu-Leng Lin

Doctor of Philosophy

Department of Linguistics

University of Toronto

2016

An important question in linguistics involves the nature of apparent substantive biases. Biases are often claimed to be universal based on typological evidence. This thesis tests for a substantive bias, the proposed universal implicational nasalized segment hierarchy in vowel-consonant nasal harmony, using an Artificial Grammar paradigm. In particular, I address whether a pattern that is predicted by the implicational hierarchy is in fact easier to learn than one that is not predicted or that is indeterminate with regard to predictions. I use a grammaticality judgment wug test paradigm to investigate whether it is easier to make a generalization when a more marked blocker (more sonorant segment) or target (less sonorant segment) is presented during an exposure phase and a less marked blocker (less sonorant segment) or target (more sonorant segment) in the test phase than vice versa. I call this the sonority hierarchy type prediction.

In addition to testing the predictions on the basis of the hierarchy, I also test predictions based on natural classes. The natural class hypothesis predicts that a grammar is more learnable if a new segment (a segment introduced in the test phase but not present in the exposure phase) is of the same natural class as an old segment (a segment introduced in the exposure phase).

The experiment was run with speakers of Min (Taiwan Southern Min), a language with no apparent evidence for sonority classes, using a method based on that of Wilson

(2006). Experiments were carried out that allow both the sonority hierarchy type and the natural class hypotheses to be tested, taking individual differences (learner types) into account. The results show that both the sonority hierarchy and natural classes play a role, supporting the claim that it is easier to learn a grammar that exhibits a substantive bias than one that does not. In conclusion, this thesis suggests that the implicational nasalized segment hierarchy is testable and learnable in artificial grammar learning to some extent and natural classes are psychologically real and actively used by participants in nasal harmony.

Acknowledgments

I dedicate this work to my beloved family, my impenetrable support, with deepest gratitude:
My parents Cheng-Chia Lin and Li-Chin Yang,
elder brother Yu-Hsien Lin, and younger sister Yu-Ting Lin

If life consists of a series of journeys, getting a PhD has definitely been a long road for me. There is a saying from Taiwan, "a scholar needs to learn how to enjoy solitude. I thought I was well prepared to embrace this fate before stepping onto Canadian soil, given the fact that I had already pursued my bachelor and master's degrees and lived by myself for years. Quickly, I learned I was totally wrong when my mom left the day before our departmental welcome party. I locked myself out from people for my first semester. I could have been alone for my entire PhD life, but luckily Dr. Christopher Spahr (cohort buddy, Thanksgiving Sushi Potluck founder) rescued me by bringing me to the parties held by departmental friends from the second semester onward. That is the time I began to earn friendships.

My dissertation would not have been possible without my committee. My supervisor Prof. Keren Rice (Thanks!) is a role model for me who is knowledgeable in phonology and field methods. She has dedicated herself to helping me in any way she could (even during the weekends!), always believes in me, and reminds me to defend myself without compromising my thoughts. Prof. Yoonjung Kang (고맙습니다!), also my generals paper supervisor, is always helpful and kind to me, offering me valuable feedback in terms of theory, experimental design and stats. Prof. Alexei Kochetov (спасибо!) reviewed my dissertation carefully and gave me wonderful feedback, especially in terms of phonetics. Prof. Philip Monahan (Thanks!) provided great views in terms of speech processing and cognitive science, and Peter Jurgec (hvala!) provided great views in terms of typology – both pushing my dissertation forward in excellent ways. To my external, Prof. Gunnar Hansson (Þakka þér!), an expert in artificial phonological grammar learning, thank you for a thorough review and insightful feedback, and for highlighting and really seeing the value and contributions in my dissertation.

I extend my thanks to the great support I have received by our department (e.g., Dr. Margaret 'Meg' Grant, Prof. Daphna Heller, Prof. Elizabeth Cowper, Prof. Alana Johns, Prof. Elan Dresher, Prof. Diane Massam, Prof. Nicholas Welch, Mary Hsu, Jill Given-King). Special thanks are directed to people who have provided valuable feedback on this work: Prof. Chandan Narayan, Prof. Sara Finley, Jeffrey Lamontagne, Prof. Alan Yu, Prof. Paul Boersma, Prof. Rachel Walker, Prof. Ron Smyth, and Prof. Harald Baayen. Thanks also for the audience's feedback at UofT phonology-phonetics group, MOT-2013, CRC-2013, 21st mfm, LAGB-2014, and AMP-2015.

Here is a list of names I consider to be my Torontonian family (if I may): Prof. Keren Rice (mentor who is always patient and friendly to every academic and non-academic question I have. I couldn't imagine working on my dissertation without her being my supervisor. Thanks for hosting a post-defense party in honor of ME), Iryna Osadcha (cohort buddy who always checks up on my mental state, my personal cheer squad), Dr. Sarah Clarke (my rock climbing buddy who loves and cares about me so much, my secret project contributor), Prof. Ailís Cournane (mentor who shows me how to better teach, how to better write and how to cope with some issues, my secret project contributor), Prof. Jessamyn Schertz (academic mentor who helps me with bits and pieces here and there), Rebecca "Becky" Tollan (grading buddy who always believes in me, reminding me that I can make it to the defense), and Jada Fung (play-harder leader and mental support). Here is another list of names I consider to be my friends: Élodie Thomas (cohort buddy), Radu Craioveanu (cohort buddy, tech support, phonetics consultant, now school bus companion), Ross Godfrey (cohort buddy, p-side alliance), Dr. Marisa Brook (cohort buddy), Dr. Colin Gorrie (work buddy), Dan Milway (office mate), Joanna Chociej (teaching mate), Tomohiro Yokoyama (cheesecake master), Emily Clare (always a sweetheart, writing me an awesome praat script), Mercedeh Mohaghegh (psycholinguistics alliance and zumba buddy), Alex Motut (FitBit companion and teaching consultant), Dr. Derek Denis (stats consultant), Julianne Doner (teaching mate who offered me her mom's home-made turkey lunch for thanksgiving), Dr. Kenji Oda (scholar sharing his tips and experience), Na-Young Ryu (great cook of Korean food), Emilia Melara (workout buddy), Jessica Yeung (language exchange companion), Kaz Bamba (後輩), Maida Percival (thanks for being my sister's language exchange companion), Shayna Gardiner (FitBit companion), Clarissa Forbes (CUPE steward and my lovely grader), Michael Schwan (loveable crazy guy), James Smith (quantitative methods companion), and Dr. Safieh Moghaddam (smiley lady).

I also have personal armed forces: Prof. Shih-Hung Lin (my Number 1 Supporter), Prof. Jiun-Shiung Wu (undergrad supervisor who introduced me to the field of linguistics), Prof. James Myers (MA supervisor who led me to the world of artificial phonological grammar learning), the sedate Meng-Ying Lin, the keen social worker Deng-Min Chuang, the productive scholar Yu-Te Huang, the smart Chi-Fan Freya Chen, the out-going Yung-Shan Sandy Chiang, the considerate Yung-Yi Amy Chiang, the kind-hearted Yi-Chun Jean Lin, the vigorous Chang-Hua Carrie Chen, and the Pisces sweet girl Feng Cao.

Please accept my sincere apology if I miss anyone here.

My floating feature of [+Dr.], after five years, finally docks into its Dr. node!

Dr. Yu-Leng Lin
Toronto, Canada
September 17th, 2016

List of Tables

List of Figures

Chapter 1 Introduction

In recent years, questions have been raised in linguistics about just what properties of grammar are universal and what emerges through exposure to language (Blevins 2004, 2006, Moreton 2008, Botma 2011, Mielke 2011, Parker 2008, Parker 2011, among others). In my thesis, I address a particular aspect of this issue, focusing on the susceptibility of segments of different types to nasalization. More particularly, the study examines nasal harmony. Nasal harmony is cross-linguistically common (a process where a nasal trigger spreads the feature [nasal] rightward or leftward, stopping when it encounters a blocker, /nawaka/→[nãw̃ãka], see Walker 2011), but it is not understood whether nasal harmony can be reflected in artificial grammar learning.

The susceptibility of segments of different types to nasalization in nasal harmony is generally considered to be related to sonority, with a more sonorant segment more likely to be nasalized, and with a less sonorant segment being nasalized in nasal harmony only if more sonorant segments are (Cohn 1990, 1993, Piggott 1992, Piggott & van der Hulst 1997, Boersma 1998, Walker 1998, 2000). This can be formulated as the nasalized segment hierarchy, vowels > glides > liquids > fricatives > obstruent stops proposed by Walker (2011), concentrating on the relationship between targets and blockers of nasal harmony. This thesis takes on this question.

This study addresses the following questions, contributing to a better understanding of the interaction between nasal harmony and learnability: (1) whether a pattern that is predicted by the nasalized segment hierarchy is easier to learn than one that is not predicted or that is indeterminate with regard to predictions in artificial grammar learning, (2) whether natural classes are psychologically real and actively used by participants, (3) how to improve the methodology to detect the nasalized segment hierarchy.

1.1 Universal implicational nasalized segment hierarchy

Walker (2011) proposes a universal implicational nasalized segment hierarchy based on evidence from typological frequency, as in (1).

(1) Nasalized segment hierarchy (from most nasalizable to least nasalizable)
vowels > glides > liquids > fricatives > obstruent stops

She draws evidence for this hierarchy from nasal harmony systems, arguing that in a language, if a more marked manner class blocks harmony (with vowels the least likely to be blockers), so do the less marked classes (obstruent stops are the most likely to be blockers). Furthermore, if a more marked class is a target, so is a less marked class.

1.2 Artificial grammar learning paradigm as the methodology

This thesis tests for the universal nature of the nasalized segment hierarchy through artificial grammar experiments. In addition to testing for the nasalized segment hierarchy, another goal of this work is to improve the methodology of artificial grammar (AG) learning experiments (e.g., Reber 1967, Gómez & Gerken 1999, Wilson 2003, see detailed discussion in Chapter 3). Typical artificial grammar learning includes two phases: an exposure phase and a following test phase. In the exposure phase, participants are exposed to stimuli which have been generated with a certain grammar; then in the test phase, participants are tested on their ability to distinguish novel stimuli that did not occur in the exposure phase. Artificial grammar learning is of value because it can control 'prior learning' in the exposure phase (which is impossible for natural languages), and watch learning in 'real time', and it is relatively easier than with natural languages to pinpoint the exact factors relevant to the participants' discriminations in AG than in natural languages (cf. Gómez & Gerken 1999). Above all, it is claimed that linguistically universal tendencies can be uncovered through artificial grammar research (Nevins 2009).

Artificial grammar learning has been used as a methodology to test markedness and implicational universals by a number of researchers, including Wilson (2006) (palatalization), Moreton (2008) (voicing, vowel height), 2012 (voicing, vowel backness), Lin (2010) (nasal harmony), Finley & Badecker (2009) (rounding and backness harmony), 2010 (vowel harmony), and Finley (2011a) (sibilant harmony). Moreton & Pater (2012a,b), in an overview article, address both the strengths and the limits of the artificial grammar paradigm.

Moreton & Pater (2012b) note that in artificial grammar studies on phonetic naturalness that investigate substantive universals, the results suggest that both phonetically natural and phonetically unnatural patterns are learnable, which challenges the premise of the artificial grammar learning paradigm that phonetically natural patterns are easier to learn than phonetically unnatural ones (i.e., that there is substantive bias). Many are still trying to find evidence to support the premise of substantive bias. Wilson (2006), for example, succeeded in finding an implicational universal between vowel height and velar palatalization, but he failed to find an effect concerning voicing and velar palatalization. Both of these relationships have been claimed as implicational universals, involving phonetic naturalness. Palatalization is found with high front vowels: if it is triggered by mid front vowels, it is also triggered by high front vowels. Similarly, if palatalization targets a voiced velar stop (/g/) in a language, it should also targets a voiceless velar stop (/k/). Moreton & Pater (2012b) suggest that the lack of effect is not because this asymmetrical relation of phonetic naturalness is wrong, it is because, as Pycha et al. (2003) argue, substantive bias (phonetic naturalness) is weaker than formal bias (e.g., the number of phonological features). The most robust evidence so far for substantive bias found in artificial grammar learning is vowel harmony.

1.3 Outline of the current study

The current study aims to test for a substantive bias, namely the proposed universal implicational nasalized segment hierarchy in vowel-consonant nasal harmony (Walker 2011), an innovative use of artificial grammar learning (see Moreton & Pater 2012a,b, discussed in Sections 3.2 and 3.4). In particular, I address whether a pattern that is predicted by the implicational hierarchy is in fact easier to learn than one that is not predicted or that is indeterminate. I use a grammaticality judgment wug test paradigm to investigate whether it is easier to make a generalization when a more marked blocker (more sonorant segment) is presented during exposure and a less marked blocker (less sonorant segment) in the test than vice versa[1]. This work incorporates the positives of the methodology of Wilson (2006) to enhance the possibility of finding an implicational universal of phonetic naturalness, since as discussed earlier, Wilson succeeded in testing an implicational universal of the relationship between palatalization and vowel height.

Chapters 2 and 3 review the typology of vowel-consonant nasal harmony, the representation of [nasal], and artificial grammar learning, all of which serve as background to the current experimental design. Chapter 4 presents Experiment 1, examining the relationship between stops and fricatives in terms of their patterning with respect to nasal harmony. Chapters 5 and 6 introduce follow-up experiments. Chapter 7 concludes and presents questions for future work.

[1] Note that I call the experimental task a 'wug test' simply because the stimuli consist of singular and plural forms. It might more appropriately be called a lexical decision task.

Chapter 2 Vowel-consonant nasal harmony

In this chapter, I introduce the typology of vowel-consonant nasal harmony to show what patterns look like in real languages, supporting the nasalized segment hierarchy introduced in Chapter 1. I briefly address two important factors in nasal harmony, opacity and transparency, and then introduce different analyses of how nasal harmony is triggered and represented in the generative phonology literature.

This chapter serves as background for the artificial grammar experiments to show that my experiments are designed based on real language data[2]. Note that even though it is difficult to find nasal flow on obstruents (e.g., /k, s, h/ used for the current study) phonetically and it is likely impossible for them to be nasalized acoustically or articulatorily (see Boersma 1998, 2003 for discussion), throughout this thesis for my stimuli I use the phonological form (e.g., /mãw̃ẽs̃ã/, /mãw̃ẽh̃ã/, /mãw̃ẽk̃ã/) instead of the phonetic form (e.g., [mãw̃ẽsã], [mãw̃ẽhã], [mãw̃ẽkã]) to indicate nasal harmony[3]. This notation should be interpreted as meaning that these consonants are transparent or undergoers to nasal harmony. In my experiments, participants were choosing not so much whether a segment such as /k/, /s/, or /h/ is a blocker or a target, but whether it is an opaque non-target (I call these blockers) or a transparent non-target (I call these targets).

2.1 Typology of nasal vowel-consonant harmony

Nasal harmony is a phenomenon where the feature [nasal] spreads to include a larger domain than a single segment. Piggott (1992), in an influential article on nasal harmony, classifies vowel-consonant nasal harmony into two types, Type A: vowel-consonant harmony with opaque segments, and Type B: vowel-consonant harmony with transparent segments. Type A involves nasal harmony *targeting vowels and certain consonants* (i.e., targets), but being *blocked by other segments* (i.e., blockers/opaque segments). For instance, in (2a), harmony targets glides but not fricatives, and fricatives block harmony. In Type B (see 2b), the transparent segments do not bear nasalization themselves, but they do not stop nasal spreading from transmitting beyond them (i.e., transparent segments).

---

[2] Some people are concerned that participants would treat phonological artificial grammar learning like a non-linguistic game such as solving a puzzle, so that they would not use their phonological knowledge to learn the artificial grammar. I will discuss this concern in more detail in Chapter 3.

[3] Boersma (2003) indicates that in generative phonology, nasal harmony can be dealt with in three processes, underlying form, phonological form, and phonetic form (e.g., /mawesa/→/mãw̃ẽs̃ã/→[mãw̃ẽs̃ã]).

(2) a. Type A b. Type B

Specifically, Type A nasal harmony, where fricatives and stops are blockers, and glides and vowels are targets, can be found in Ijo (Ijoid), Urhobo (Edoid), and Narum (Autronesian). Consider the data in Ijo in (3), with leftward spreading.

(3) Type A nasal harmony in Ijo (Walker 1998: 31–32, Williamson 1965, 1969, 1987).

(a) /izoŋgo/	→	[izõŋgo]	'jug'
(b) /abanu/	→	[abãnu]	'loft'
(c) /sɔɾɔ̃/	→	[sɔ̃ɾ̃ɔ̃]	'five'
(d) /waĩ/	→	[w̃ãĩ]	'prepare sugarcane'

Type B harmony, where fricatives are transparent segments, is found in Mòbà (Benue-Congo) and Kaingang (Macro-Gê)[4]. Consider the data in the Mòbà dialect of Yoruba in (4), with leftward spreading.

(4) Type B nasal harmony in Mòbà (Piggott 2003: 379, Ajíbóyè 2002)

(a) /isĩ/	→	[ĩsĩ]	'worship'
(b) /uwã/	→	[ũw̃ã]	'lie'

All researchers who have worked in this area note two types of nasal harmony, although they account for them in different ways (cf. Piggott & van der Hulst 1997, Piggott 2003, Walker 2011, Boersma 2003). In Section 2.3, I summarize Type A harmony since the Type A pattern is the focus of the thesis. The purpose of this section is to understand the interaction between nasal harmony,

[4] Nasal harmony only occurs within a single syllable in Kaingang (see Walker 2011, Piggott & van der Hulst 1997: 101).

sonority and phonological classes and to give an overall sense of what kind of segments usually serve as triggers, targets, and blockers.

2.2 Nasal vowel-consonant harmony with opaque segments

In Type A nasal harmony, a trigger can either be an underlying nasal stop or a nasalized vowel. The direction of spreading can either be rightward or leftward, but rightward spreading is more common than leftward. Stops are always blockers.

Type A nasal harmony is summarized in Table 1 (modified from Piggott 2003, Walker 2011). Possible targets and blockers are shown in the columns, and the rows indicate the different harmony types that have been identified. The shaded cells represent neutral (non-nasalizable) and opaque segments (blockers).

Table 1. Type A: vowel-consonant harmony with opaque segments

	Vowels	Glides	Liquids		Fricatives	Stops	Languages
A1	target	blocker	blocker		blocker	blocker	Sundanese, Mixtec
A2	target	target	blocker		blocker	blocker	Malay, Warao, Capanhua, Arabela
Between A2 and A3	target	target	target (tap)	blocker (trill)	blocker	blocker	Epena Pedee[5]
A3	target	target	target		blocker	blocker	Ijo, Urhobo, Narum
A4	target	target	target		target	blocker	Applecross Gaelic

In the A1 pattern, vowels are targets, but consonants are blockers. The languages Sundanese (Austronesian) and Mixtec (Oto Manguean) belong to this category (e.g, Sundanese: /ɲãĩãñ/ 'wet', /mãwur/ 'spread', /mõlok/ 'stare', /ŋãtur/ 'arrange', Piggott 2003: 377, Robins 1957).

In the A2 pattern, vowels and glides are targets, and the rest of the consonants are blockers. Malay (Austronesian), Warao (isolate), and Capanahua (Panoan) belong to A2 (e.g., Malay: /mẽw̃ãh̃/ 'be luxurious', /mə̃laran/ 'forbid', /mãkan/ 'eat', Piggott 2003: 377, Onn 1976).

[5] Harms's (1985: 16) analysis hypothesizes that /s/ blocks rightward Type A nasal harmony but Harms (1994: 6) indicates that /s/ could be a blocker as in /ˈmĩə̃su/ 'spear'. In this thesis, I follow Harms (1994), treating /s/ as a blocker (see discussion in Walker 2011: 1841–1842).

Walker (2011) mentions that Epena Pedee (Choco) shows a pattern where vowels, glides, and lateral taps are targets, but the trill /r/ is a blocker, so I put this language in a category that falls between A2 and A3 (e.g., Epena: /nãw̃ẽ/ 'mother', /kʰãˈɲãɾ̃ã/ 'than', /ˈtũːra/ 'pelican', /ˈh̃õᵐpʰe/ 'fish', Walker 2011: 1842, Harms 1985, 1994).

In the A3 pattern, vowels, glides, and liquids are targets, but obstruents are blockers. Ijo (Ijoid), Urhobo (Edoid), and Narum (Autronesian) belong to A3. The current thesis is particularly interested in cases where vowels and glides are targets, while fricatives and stops are blockers, so I illustrate with Ijo as an example. Piggott (2003) assumes /n/ is in the underlying representation, serving as a trigger. The trigger spreads leftward to vowels (/a e i o ɔ/, glides (/w, j/) and liquids (/r/), but spreading is halted by fricatives (/z/) and stops (/k/). Walker (2011: 1842) also mentions that taps and laterals are targets in this language (cf. Williamson 1965, 1969, 1987). (5) illustrates Kolokuma Ijo nasalization. (5a-c) show vowels, glides, and /r/ as target, while (5e) shows that a fricative blocks harmony[6].

(5) Kolokuma Ijo nasalization (Piggott 2003: 377, Williamson 1965)

a.	/bein/	→ [bẽĩ(n)]	'be full'
b.	/owein/	→ [õw̃ẽĩ(n)]	'bite'
c.	/yarin/	→ [ỹãɾ̃ĩ(n)]	'shake'
d.	/kɔrɔŋgbɔː/	→ [kɔ̃ɾ̃ɔ̃ŋgbɔː]	'thin'
e.	/izoŋgo/	→ [izõŋgo]	'mosquito'

In the A4 pattern, sonorants and fricatives are targets, and only stops are blockers. Applecross Gaelic (Celtic) belongs to A4 (e.g., Applecross Gaelic (bidirectional spreading): /**kʰ**ɔ̃ĩs̃**p**axk/ 'wasp' /s**t**ɾ̃ãĩːɣ̃/ 'string', /ʃ̃ɛ̃nɛ̃ːṽãɾ̃/ 'grandmother', /mãːh̃ãɾ̃/ 'mother', Piggott 2003: 377, Ternes 1973, 2006).

[6] The blocking by fricatives suggests that, if the hierarchy is correct, stops should also block harmony. I was not able to find examples with stops in the appropriate position to confirm this.

2.3 Predictions of the nasalized segment hierarchy

Type A nasal harmony meets the predictions of the nasalized segment hierarchy – if, for instance, a glide is a blocker, it is predicted that sounds of lower sonority (liquids, fricatives, stops) are also blockers. If a stop is a blocker, we cannot predict the patterning of glides or other more sonorant segments. That is, if a language has glides as blockers, then sounds of lower sonority must be blockers as well. No predictions are possible based on blocking by less sonorant sounds – the fact that a stop is a blocker reveals nothing about the patterning of glides, for instance.

The predictions of Type A nasal harmony are summarized in (6). These form the basis for the experiments reported in this thesis.

(6) Implicational universals:

Blocking	more sonorant blocker	→	less sonorant blocker
Targeting	more sonorant target	←	less sonorant target

Chapter 3 Introduction to experimental methods: artificial grammar learning

Artificial Grammar (AG) is extensively used (e.g., Reber 1967, 1976, 1989, Gómez & Gerken 1999, Wilson 2003, 2006, Finley 2008, 2011, 2012, 2013, 2015, Finley & Badecker 2010, Moreton 2008, Lin 2009, 2010, 2011, 2012, 2013, 2014, 2016, Moreton & Pater 2012a,b, Nevins 2013, McMullin & Hansson 2014) and affords the potential to understand whether learning biases are universal, as discussed in Chapter 1. Artificial Grammar is of value in studying biases for several reasons. First, it is possible in AG but not possible with real languages to control prior learning in the exposure phase and to observe learning in real time (cf. Gómez & Gerken 1999). For instance, AG can minimize external factors such as lexical frequency that affect judgment about real words. Second, it is possible with AG to control exactly factors with respect to participants' judgment and discrimination, something that is not possible with real languages (cf. Gómez & Gerken 1999). Specifically, this paradigm allows us to design an artificial language that obeys a certain phonological pattern (Gómez & Gerken 1999), and it is possible to manipulate and compare types of phonological patterns (Nevins 2009, Finley 2011a). Third, artificial grammar work can reveal linguistically universal tendencies (cf. Nevins 2009). Lin (2012) points out that though people argue that performance data might not reflect competence (UG) (see Newton & de Villiers 2007, Chomsky 2007, Fodor & Garrett 1966), we cannot deny that people exhibit their knowledge through performance. Performance data (e.g., artificial grammar learning) can give indirect clues about competence (see the discussion about recursion in artificial grammar learning in Lobina 2011). We need to design and look at our data carefully, and try to tease various factors apart to see which factors affect the results.

Chapter 2 reviewed the relevant literature on nasal harmony. This chapter introduces the literature on artificial grammar that forms the foundation for this study. I will also discuss caveats and controversies of adopting an artificial grammar learning paradigm. Since the current study tackles one kind of implicational universal, the nasalized segment hierarchy, I will discuss artificial grammar learning research which tests implicational universals. In addition to implicational universals (which represent one kind of substantive bias), I will also present formal complexity bias (the number of features positively correlated with difficulty of learnability) as it is often addressed in the literature on artificial grammar learning. The structure of this chapter is as follows.

Section 3.1 introduces the premise of the artificial grammar learning paradigm. Section 3.2 presents why artificial grammar learning is argued to shed light on natural markedness (based on typological frequency and phonetic naturalness) and substantive bias, including discussion of Wilson (2006), Berent (2013), and Finley (2012). Section 3.3 summarizes the artificial grammar learning studies that test whether cross-linguistic implicational universals can be reflected in a lab setting using an artificial grammar paradigm, including Wilson (2006: palatalization) (substantive bias), Finley (2011a: sibilant harmony) (formal complexity bias), Finley (2012: round harmony) (substantive bias). Section 3.3 also summarizes the studies related to the sonority hierarchy, nasal assimilation and dissimilation, including Berent (2013: sonority hierarchy), and Wilson (2003) (nasal assimilation and dissimilation). Section 3.4 summarizes artificial grammar research concerning formal complexity bias (Moreton & Pater 2012a). Section 3.5 concludes this chapter.

3.1 Assumptions of artificial grammar learning: poverty of the stimulus

Hayes (2009) note that many universal grammar experiments in phonology depend on the poverty-of-the-stimulus argument (Chomsky 1980: 34). The logic is that experiments will manifest Universal Grammar (UG) if they demonstrate that participants show consistent learning behavior that could not be inferred from their own language background or from information provided during the experiments (eg., feedback). The artificial grammar paradigm is one type of experiment designed to get at UG.

Wilson (2003) discusses the structure of experiments that are based on this assumption. In particular, as mentioned in Chapter 1, the typical phonological artificial grammar learning paradigm consists of an exposure phase and a test phase. In the exposure phase, participants are exposed to stimuli which have been generated using a particular phonological pattern. Then in the test phase, participants are tested on novel stimuli[7]. During the test phase, if participants consistently show learning in certain directions and not in others, a possible inference is that UG principles are guiding the process of generalization, because the stimuli of the experiment alone were too impoverished to have determined the participants' responses. Less direct forms of poverty-of-the-stimulus experiments like those reported by Wilson (2003) and Moreton (2008) use more than one artificial phonological grammar, and compare the learnability of the different

[7] Note that all artificial grammar work cited in this chapter presupposes that natural classes and features exist (e.g., /p, t, k/ are voiceless stops; /p, t, k/ share features [-sonorant, -continuant, -voice]).

artificial languages. In this case, the poverty-of-the-stimulus argument is based on the assumption that if one artificial language is learnable and another is not, or if one artificial language is learned better than another artificial language, then it must be that UG principles underlie this learning difference (this kind of experiment is less direct because it 'compares' the learnability among two or more artificial languages instead of looking at one artificial language directly).

There are two versions of UG proposed in the literature, strong and weak versions (Wilson 2003, Hayes et al. 2009). The strong version of UG is that it is a rigid system permitting certain grammars and forbidding others. That is, any grammatical rule or constraint not countenanced under UG cannot be learned (Hayes et al 2009: 3). The weaker version of UG is that it consists of a set of biases: aspects of the system of mental representation and computation cause processes with certain properties to be easier to learn and/or represent, but do not exclude processes that fail to have such properties (Wilson 2003: 102). Wilson (2003) and Moreton (2008) assume a weaker version of UG. Even though certain patterns are rare or unattested, they could nevertheless be learnable.

This thesis adopts artificial grammar learning based on the assumption of the poverty of the stimulus, but is open to the possibility that participants' own language knowledge would influence the learning. Less direct forms of poverty-of-the-stimulus experiments are conducted since I compare the learnability of two different artificial grammars, one predicted by the nasalized segment hierarchy, and the other not, hypothesizing that the former would be learned better than the latter. The current study leans toward a weaker version of UG: any pattern could be learnable even though it is not predicted by UG. Nevertheless, there is a tendency that a pattern predicted by UG would be learned more successfully, in less time than a pattern that is not. In particular, I use an artificial phonological grammar paradigm as a tool to test the learnability of nasal harmony, to see whether humans' learning of nasal harmony matches the predictions of natural markedness, and to determine whether the artificial grammar paradigm is appropriate to test this kind of natural markedness and implicational universals.

3.2 Natural markedness and substantive bias

The premise of AG that certain properties are easier to learn/represent than others (i.e., a weaker version of UG) is based on natural markedness[8]. This is also termed “frequency markedness” (Bybee 2001: 202, Rice 2007: 81). Natural markedness refers to the phonetic basis of an opposition (Anderson 1985, Rice 2007: 80); for instance, voiceless obstruents are considered to be less marked than voiced obstruents based on factors such as frequency and implication. Factors such as typological frequency and phonetic naturalness serve as a basis to define what is more marked and what is less marked. Experimental hypotheses are defined in terms of natural markedness: (a) a typologically common/frequent pattern is easier to learn than a typologically uncommon/less frequent pattern, (b) a phonetically natural/motivated pattern is easier to learn than a phonetically unnatural/unmotivated pattern[9].

Both hypotheses (a) and (b) concerning phonetic naturalness can be subsumed under the term ‘substantive bias’ (a more phonetically unnatural/unmotivated pattern is harder to learn than a more phonetically natural/motivated pattern) (Moreton & Pater 2012a,b). Substantive bias, the focus of the current study[10], concerns “phonetic” naturalness, so it is domain-specific (specific to language) rather than domain-general (not specific to language but part of general cognitive abilities). Formal complexity bias, on the other hand, (a pattern with more features is harder to learn than a pattern with fewer features) is domain-general and non-linguistic rather than domain-specific (see discussion in Section 3.4). Moreton & Pater (2012b) argue that if phonologists can find any artificial language learning evidence for substantive bias, then we can be sure that this is

[8] Note that we cannot directly test UG, all we can test is natural markedness.

[9] In addition to natural markedness, there is a second type of markedness, called structural markedness by Bybee (2001), where evidence comes from phonological processes. Specifically, Rice (2007) notes that marked features often show the following patterns: they are susceptible to neutralization, they are unlikely to be epenthetic, they trigger assimilation, and they are retained in coalescence and deletion.

[10] There is a large literature on artificial grammar learning (Reber 1967, 1976, 1989, Wilson 2003, 2006, Finley 2008, 2011, 2012, 2013, Moreton 2008, 2010, 2012, Lin 2009, 2010, 2011, 2013, Nevins 2010, 2013, Albright & Do 2013, Boersma et al. 2013, among others). In this section, I review only Moreton & Pater (2012a) in detail for three reasons. First, Moreton & Pater specifically discuss artificial “phonological” grammar learning, which is exactly what I use this paradigm for. Second, they are concerned with both substantive bias (phonetic naturalness) and formal bias (feature contrasts). Third, they summarize the strengths and weaknesses of this paradigm in testing phonological patterns. Moreton & Pater (2012a,b) review the literature on artificial-phonology learning, and point out that phonetic substance bias and formal complexity bias are two factors which might reveal something about natural-language phonology. On the one hand, if a learner has a phonetic substance bias, s/he would acquire phonetically motivated patterns better than phonetically arbitrary ones. On the other hand, if a learner has a formal complexity bias, s/he would acquire simpler patterns faster or better than complex ones. Both cases are based on the assumption that the training data presents patterns well, participants perceive the patterns correctly, and experimenters control for other factors.

domain-specific and is about natural-language phonology rather than simply non-linguistic games. However, so far the literature on artificial learning has shown that formal complexity bias is stronger than substantive bias in a lab setting. The majority of artificial language work involving substantive bias fails to find any significant effects. Such research confirms Pycha et al.'s (2003) claim that substantive bias is weaker than formal complexity bias. The only strong evidence so far for substantive bias is that vowel agreement/harmony is learned more easily than consonant agreement/harmony (cf. Toro et. al 2008a,b, Moreton 2012)[11], which is consistent with natural-language phonology in that dependencies between non-adjacent consonants are rarer than ones between non-adjacent vowels (cf. Gafos 1996: 7–8, Baković 2000: 5–6, Hansson 2001: 1–2). This also supports the generalization that vowel harmony is more common than consonant harmony (Moreton 2008). From a perspective of processing between vowels and consonants, Finley (2011b) also points out that generalizing to new vowels (vowel harmony) is easier than generalizing to new consonants[12].

While some literature shows that the learning of an AG matches natural markedness/substantive bias (see Wilson 2003, Moreton 2008), some literature shows an inconsistency between learnability and natural markedness/substantive bias. These issues are discussed in this section.

3.2.1 Implicational universals in substantive bias

Artificial grammar work which touches on cross-linguistic typology (implicational universals) with respect to substantive bias has not found a robust and consistent effect. For example, Wilson (2006) investigates two kinds of implicational universals: (1) palatalization before more back vowels implies palatalization before more front ones; (2) palatalization of voiced velars implies palatalization of voiceless ones. He found that velar palatalization before /e/ implies velar palatalization before /i/, which corresponds to the prediction of (1). However, he found that the palatalization of /g/ was significantly more than that of /k/, which contradicts the prediction of

[11] The strong evidence of substantive bias is in the sense that the experimental results of vowel harmony in the artificial grammar literature are consistent, whereas the experimental results of consonant harmony in the artificial grammar literature are sometimes inconsistent.
[12] Finley (2011b) drew this conclusion based on the studies about vowels (Chambers et al. 2010, Finley & Badecker 2009, Skoruppa & Peperkamp 2011) and studies about consonants (Peperkamp et al. 2006, Peperkamp & Dupoux 2007).

(2). Cross-linguistic typology suggests that rounding harmony in mid vowels implies rounding harmony in high vowels, but Finley & Badecker (2009, Exp. 3) did not find this asymmetrical implication in their experimental work.

Thus sometimes proposed implicational universals in substantive bias fail to be replicated in experimental work. But does this mean that an artificial grammar learning paradigm is not an appropriate tool to test natural markedness/substantive bias? I will discuss this question in Sections 3.2.2 and 3.2.3.

3.2.2 Statistical tendencies

Berent (2013) and Finley (2012) evaluate the positives of using artificial grammar learning as a paradigm. Berent (2013: 133) regards typology as E-language instead of I-language (Chomsky 1972). I-language reflects the internal structure of grammar, while E-language reflects external non-grammatical factors (e.g., historical, functional, social factors). Specifically, several claim that typological universals do not equate with grammatical universals (i.e., Universal Grammar). Typological universals reflect statistical tendencies, not absolute stipulations (Berent 2013: 147). This can explain why sometimes artificial grammar learning experiments fail to find the expected trend, namely that a typologically frequent pattern is easier to learn than a typologically less frequent pattern. That is, the absence of absolute universals does not falsify the claim that grammatical phonological universals exist (Berent 2013: 133).

3.2.3 Robustness and learnability

In addition to the disassociation between typological frequency and learnability, there is also a mismatch between markedness and phonetic naturalness. Although marked structures might be more perceptible and robust than less marked ones, this does not necessarily entail that the former's phonetic encoding is equally easy to learn (Berent 2013: 163). Despite these two mismatches, Berent (2013: 150) argues that it remains valid to do experiments, since by comparing typological trends with individual participant's behavior, we can seek a common universal source that molds both of them.

Finley (2012) states that it is impossible to look at all possible languages in order to give evidence for crosslinguistic universal tendencies, but artificial grammar learning can show the psychological reality of universal tendencies without running experiments in all languages. In addition, Finley (2012) cites Gerken & Bollt (2008), who show that younger infants (7.5 months) make generalizations based on unnatural stress patterns, whereas older infants (9 months) make generalizations depending on natural stress patterns, suggesting that the phonetic biases develop over time. If phonetic biases are molded by experience (Kuhl et al. 1992, Kuhl 2001, Gerken & Bollt 2008), adult data might reflect the most accurate learning biases and might tell us something more about first language learning, since adult phonetic knowledge is fully developed, and adults have more exposure to language than infants.

3.2.4 Summary

In sum, Section 3.2 introduced the literature using the artificial phonological grammar paradigm as a tool to test learnability, including Wilson (2003), Wilson (2006), Moreton (2008), and Hayes et al. (2009), presenting the rationale and premise behind artificial grammar learning. The sections add the debate concerning the mismatch between learnability and natural markedness (Moreton 2008). At the same time, the sections also summarize the values of artificial grammar learning (Finley 2012, Berent 2013).

3.3 Artificial grammar studies in substance bias

Although some artificial grammar research fails to show evidence for implicational universals in substantive bias, there are many artificial grammar experiments that succeed in finding positive results to support implicational universals in substance bias.

In this section I summarize the research that is related to implicational universals, the sonority hierarchy, and harmony/assimilation.

3.3.1 Implicational universals involving a substance bias: palatalization

In this section, I review Wilson (2006) carefully, since he used artificial phonology learning to test an implicational universal. Specifically, his implicational universal is as follows: velar stops becoming palatal affricates before less front vowels implies velar stops becoming palatal affricates before more front vowels, but not vice versa[13]. My experimental logic is the same as his: the more marked pattern during an exposure phase implies the less marked pattern during a test phase. For this reason, I treat Wilson's work as a model for mine. Note, however, that he examines a single feature [-back], whereas I investigate a single feature [nasal] with respect to the sonority hierarchy and the relationship between natural classes involving two features, [sonorant] and [continuant].

3.3.1.1 *Fronting effect*

Wilson (2006) chose velar palatalization as a focus because the substantive bias toward fronting contexts is solid in terms of articulation, acoustics, perception and phonology. For articulation, Wilson cites Keating & Lahiri 1993, Ladefoged 2001, and Butcher & Tabain 2004 for evidence that the articulation of /k/ and /g/ is more forward on the palate when immediately preceding front vowels like /i/ and /e/ than when preceding back vowels like /a/. This fronting effect makes /ki/, /ke/, /gi/ and /ge/ articulatorily more similar to palato-alveolar affricates /tʃi/, /tʃe/, /dʒi/ and /dʒe/.

From an acoustic perspective, the fronting effect can cause velars to have a higher peak in the spectrum of the consonant release, since velars before front vowels have relatively smaller resonant cavities, and hence higher frequency peaks (Keating & Lahiri 1993, Guion 1996, 1998, Butcher & Tabain 2004).

From the perspective of perception, Wilson cites an identification task by Guion (1996, 1998) which examines confusion rates in perception. Her results show that the rate of [ki] mistakenly identified as [tʃi] is higher than the rate of [ka] mistakenly identified as [tʃa]. A similar trend is also found for the voiced counterpart. That is, the rate by which [gi] is mistakenly identified as [dʒi] is higher than the rate by which [ga] is mistakenly identified as [dʒa].

From the perspective of phonology, Ohala (1992), Guion (1996, 1998), Bateman (2007), and Kochetov (2011) investigate languages that have velar palatalization or have undergone a sound change of velar palatalization. The typological facts show that if a language has velar stops becoming palatal affricates before less front vowels then the language also has velar stops becoming palatal affricates before more front vowels, but not vice versa. This implicational

[13] Wilson also looked at an implicational universal involving voicing: voiced velar palatalization implies voiceless velar palatalization, but he failed to detect any effect. The design was similar to the vowel-context one, so I will not review it here.

universal is tested by Wilson's (2006) experiments. The following presents Wilson's experimental design.

3.3.1.2 *Experimental design*

Wilson used stimuli of the form $C_1V_1C_2V_2$. V_2 was always /ə/. C_2 was selected from English consonants /p b k g m n f v θ ð s z tʃ dʒ l r w/. C_1 was selected from /p b **k g tʃ dʒ**/, and V_1 was selected from /i e a/. The critical words/items involve /k g/ and /tʃ dʒ/. /p b/ were fillers.

As noted in Section 3.3.1.1, velars /k, g/ often palatalize before front vowels (vowel frontness). I follow Wilson's terms and label vowels in terms of vowel height instead of vowel frontness as high front vowels are more likely to be palatalization triggers than mid vowels are.

The Vowel-height condition involves a between-participant design. In the High-vowel condition (Group 1), during the exposure phase, critical items contained only the high vowel /i/ as V_1 after velars; see Table 2. The number in parentheses indicates the number of pairs.

The experiments included practice, exposure, break, and test. In the practice, two pairs were presented. One was /balə/ …/balə/ (no alternation). The other was /gibə/…/dʒibə/ (High-vowel condition).

During the exposure phase, 32 pairs were presented. The logic of the experiment is that high-vowel contexts (palatalization before /i/) in the exposure phase give no indication about whether during the test, velars before a mid vowel /e/ should be palatalized.

Table 2. Exposure trials for High-vowel condition (Group 1)

Condition	Trial type	
High	kiC_2V_2 **tʃ**iC_2V_2 (4)	giC_2V_2 **dʒ**iC_2V_2 (4)
Both	kaC_2V_2 kaC_2V_2 (3)	gaC_2V_2 gaC_2V_2 (3)
	piC_2V_2 piC_2V_2 (3)	biC_2V_2 biC_2V_2 (3)
	peC_2V_2 peC_2V_2 (3)	beC_2V_2 beC_2V_2 (3)
	paC_2V_2 paC_2V_2 (3)	baC_2V_2 baC_2V_2 (3)

In the Mid-vowel condition (Group 2), during the exposure phase, critical items contained only the mid vowel /e/ after velars; see Table 3.

Table 3. Exposure trials for Mid-vowel condition (Group 2)

Condition	Trial type (number)	
Mid	keC$_2$V$_2$ **tʃ**eC$_2$V$_2$ (4)	geC$_2$V$_2$ **dʒ**eC$_2$V$_2$ (4)
Both	kaC$_2$V$_2$ kaC$_2$V$_2$ (3)	gaC$_2$V$_2$ gaC$_2$V$_2$ (3)
	piC$_2$V$_2$ piC$_2$V$_2$ (3)	biC$_2$V$_2$ biC$_2$V$_2$ (3)
	peC$_2$V$_2$ peC$_2$V$_2$ (3)	beC$_2$V$_2$ beC$_2$V$_2$ (3)
	paC$_2$V$_2$ paC$_2$V$_2$ (3)	baC$_2$V$_2$ baC$_2$V$_2$ (3)

Again, the experiments included practice, exposure, break, and test. In the practice, two pairs were presented. One was /balə/ …/balə/ (no alternation). The other practice pair was /gebə/…/dʒebə/ (Mid-vowel condition).

During the exposure phase, 32 pairs were presented. The logic of the experiment is that mid vowel contexts (palatalization before /e/) in the exposure phase would help participants generalize from mid-vowel to high-vowel contexts that velars before a high vowel /i/ should be palatalized.

During the test phase (see Table 4), the same 80 pairs were presented to both groups. But note that for the Mid-vowel condition (Group 2), velar stops combined with high vowels would be 'novel' test trials, while the ones combined with mid vowels would be exposure trials (presented in the exposure before). The reverse held for the High-vowel condition (Group 1). Note also that in the test phase, velars before /a/ did not undergo palatalization, only velars before mid vowel /e/ and high vowel /i/ did.

Table 4. Test trials for both conditions (presented to both Groups 1 and 2)

Critical trial type		Filler trial type	
ki C$_2$V$_2$ **tʃ**iC$_2$V (8)	giC$_2$V$_2$ **dʒ**iC$_2$V$_2$ (8)	piC$_2$V$_2$ piC$_2$V$_2$ (6)	biC$_2$V$_2$ biC$_2$V$_2$ (6)
ke C$_2$V$_2$ **tʃ**eC$_2$V (8)	geC$_2$V$_2$ **dʒ**eC$_2$V$_2$ (8)	peC$_2$V$_2$ peC$_2$V$_2$ (6)	beC$_2$V$_2$ beC$_2$V$_2$ (6)
ka C$_2$V$_2$ kaC$_2$V (6)	gaC$_2$V$_2$ gaC$_2$V$_2$ (6)	paC$_2$V$_2$ paC$_2$V$_2$ (6)	baC$_2$V$_2$ baC$_2$V$_2$ (6)

The prediction is that if participants were exposed to the Mid-vowel context during the exposure phase, then they would be likely to palatalize velar consonants in the High-vowel context.

However, if participants were exposed to the High-vowel context during the exposure phase, they would randomly palatalize velar consonants in the Mid-vowel context.

3.3.1.3 *Procedure*

Before the experiment started, participants were told that they would learn a new spoken language game (a way of pronouncing certain words). During the practice, participants listened to two practice trials. During the exposure phase, the text 'I say …' appeared on the screen, and then the first member of a stimulus pair was played (e.g., ki C_2V_2). The other text 'you say …' appeared, and then the second member of a stimulus pair was played (e.g., tʃiC_2V). Participants needed to repeat the second member, and were told that repetition could assist them in learning the game. During the break, they did math problems using a pencil and paper. During the test, the procedure was almost the same as the exposure phase, except that they were asked to generate the second member of a stimulus pair.

3.3.1.4 *Results and analysis*

Recall that the implicational universal of velar palatalization is that the mid-vowel context implies the high-vowel context. If participants have a substantive bias, then it is expected to find that the rate of generalization (the rate of correct palatalization of critical test items) for the Mid-vowel contexts implies generalization to the High-vowel context.

In order to test this, Wilson compared the generalization rate for the exposure/old trials (presented in the exposure phase) and the generalization rate for the novel trials within each condition. A repeated-measure analysis of variance (ANOVA) was used. A crucial interaction is the one between *condition* (High-vowel condition vs. Mid-vowel condition) and *vowel context* (exposure/old trials vs. novel trials during the test), which was significant. Planned post hoc paired t-tests also confirmed this. Specifically, for the High-vowel condition (Group 1, cf. materials in Table 2 and Table 4), the palatalization rate of exposure context (old trials, /i/ pairs) was significantly higher than that of novel context (new items, /e/ pairs), suggesting that participants were not able to generalize from high-vowel contexts to mid-vowel contexts. On the other hand, for the Mid-vowel condition (Group 2, cf. materials in Table 3 and Table 4), the palatalization rate of old test items was not significantly higher than that for novel test items, suggesting that participants did generalize from mid-vowel contexts to high-vowel contexts.

This section summarized Wilson (2006) in detail to give exposure to what substantively biased artificial phonological grammar work is. In Chapter 4, I turn to my research on nasal harmony, and incorporate the positives of Wilson (2006) into my design.

3.3.2 Implicational universals involving a formal complexity bias: sibilant harmony

In this section, I review Finley (2011a), who investigates one kind of implicational universal concerning sibilant harmony (coronal assimilation), namely 'local' sibilant harmony where stem and suffix agree in [+strident] without intervening consonants, and 'non-local' sibilant harmony where stem and suffix agree in [+strident] with intervening consonants. Typologically, local sibilant harmony is less marked, and non-local sibilant harmony is more marked. That is, non-local sibilant harmony implies local sibilant harmony, but not vice versa (implicational universal). The substantive bias in this case is based on the fact that sibilant harmony is related to phonetic substance, coronal ([+strident]) assimilation. However, locality should not be counted as substance bias. Locality fits the formal complexity bias better (see more detail about contiguity-similarity tradeoff in Section 3.4.1.1). Specifically, a typological asymmetry in a local vs. non-local relationship should be categorized as an implicational universal involving a formal complexity bias (i.e., formal structure).

For local sibilant harmony, the structure of the stimuli was CV**ʃ**V-**ʃ**u or CV**s**V-**s**u. For non-local sibilant harmony, the structure was **ʃ**VCV-**ʃ**u or sVCV-**s**u. Non-sibilant consonants were selected from /p, t, k, b, d, g/. Vowels were drawn from /a, i, e, o, u/. There were two experiments. For the first experiment, 10 participants were exposed to local sibilant harmony during exposure (24 items), and were tested on non-local sibilant harmony (36 forced-choice pairs between harmonic and disharmonic stimuli). The opposite learning order was used for the second experiment. The stimuli were naturally produced by an adult female native speaker of English. The grammar in the second experiment is expected to be learned better than that of the first experiment if the implicational universal holds.

The between-participant factor was *Training* with two levels (i.e., *trained condition*: one group of participants were trained with a harmony pattern in the exposure phase vs. *control condition*: the other group was not trained where only stems were presented without suffixes, so that they would not have exposure to any information about harmony at all). The within-participant factor was *Test Items* with three levels (i.e, *Old Stems*, *New Stems*, *New Suffix* (non-local harmony for the first group, local harmony for the second group)). The data was analyzed using ANOVA.

For the first experiment (exposure to local harmony, see Figure 1), a main effect of *Training* was significant (mean: 0.80 for the *trained condition*, 0.47 for the *control*), suggesting that training did influence participants' learning compared to the control. A main effect of *Test Items* was significant. There was a significant difference between *New Stem* and non-local harmony (*New Suffix*) for the *trained condition*. To directly test generalizability, one sample *t*-tests were conducted to compare each test condition, with 50% as chance level. For the trained condition, the *Old Stem* was above chance, as was the *New Stem*. But the non-local sibilant harmony (*New Suffix*) was not above chance, suggesting that participants failed to generalize from local to non-local sibilant harmony. Though a main effect of *New Suffix* was not found, the accuracy rate, three out of ten participants, for non-local sibilant harmony (*New Stem*) was higher than 50%, suggesting that these three participants succeeded in generalizing from local to non-local sibilant harmony. This could mean that a claimed bias of locality might be questionable. But it could also mean that a claimed bias of locality stands for a "tendency" rather than an absolute trend, allowing for individual differences (corresponding to the argument of statistical tendencies proposed by Berent (2013) discussed in Section 3.2.2). In that case, we would expect the majority of participants would fail to generalize from local to non-local sibilant harmony, but a few participants (three participants for the first group) might actually succeed in doing so.

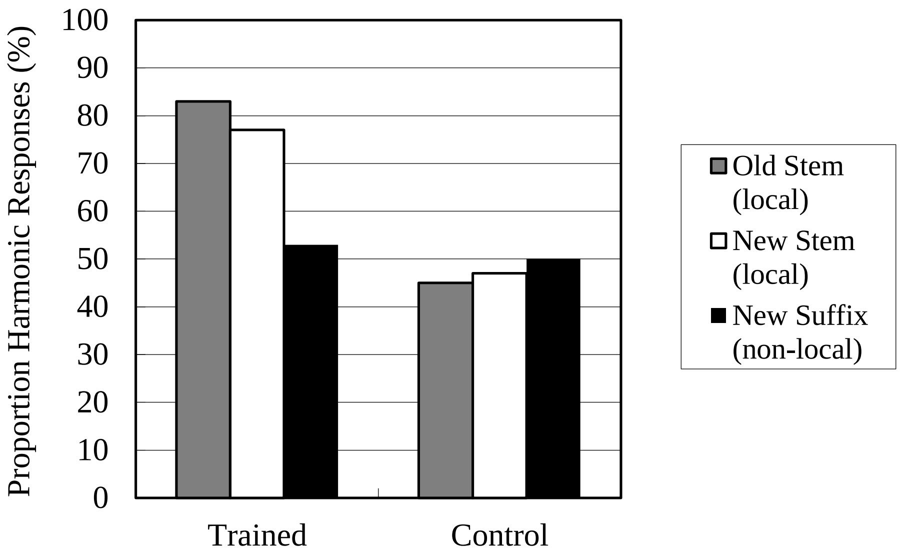


Figure 1. First experiment results: means (redrawn & modified from Finley 2011a: 79, Figure 1)

For the second experiment (see Figure 2), a significant effect of *Training* was also found. But no effect of *Test Items* was found, suggesting that there was no learning difference between *Old Stems*, *New Stems*, and *New Suffix*. A t-test was conducted to compare trained and control conditions for the local sibilant harmony (*New Suffix*), and it was significant, suggesting that participants generalized from a non-local pattern to local one. A one sample t-test showed that for the trained condition, *Old Stems, New Stems,* and *New Suffix* were all above chance, further confirming that participants generalized from the non-local to the local pattern. One caveat is that the accuracy rate for two of the participants for local sibilant harmony (*New Stem*) was less than 50%, suggesting that these two people failed to generalize from the non-local to the local pattern. However, these two participants might be able to learn the opposite learning order presented in the first experiment (i.e., from local to non-local pattern). That is, perhaps the two participants, like the three participants for the first experiment, were able to infer from the local to the non-local pattern, not the other way around (contradicting the prediction of the implicational universal). This is why they failed to learn well on *New Stem* since they were not exposed to the local pattern first, and then were tested on the non-local pattern. Again, it does not undermine the prediction of the implicational universal if we treat the prediction as a tendency (the "majority" of participants, would be able to generalize from non-local to local patterns) and allow individual variations.

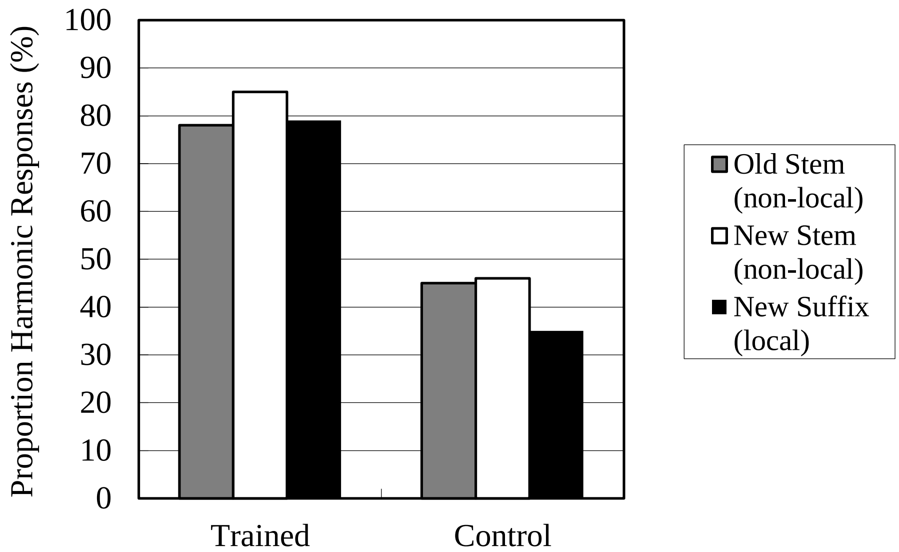


Figure 2. Second group results: means (redrawn & modified from Finley 2011a: 80, Figure 2)

In order to determine whether the results support the hypothesis of locality, a cross-experiment ANOVA was conducted to compare the trained condition for both the first and second experiments. No effect of *Training* was found, suggesting both orders were equally learnable. The effect of *Test Items* (old stems, new stems, new suffix) was significant, suggesting there was difference between old and new items. The generalizability between Experiments 1 and 2 was different, illustrated by the fact that the difference between New Stem and New Suffix was significant for the first experiment, but this difference was not significant for the second experiment.

Finley (2011a) is a successful case of adopting artificial grammar learning to test for an implicational universal involving sibilant harmony: local sibilant harmony implies long-distance sibilant harmony, but not vice versa. This study, like Wilson, is a foundation of the design adopted in Experiment 2 for the current study (the details are given in Chapter 5). Note that like Wilson, Finley's work examines a single feature ([-back] for Wilson 2006, [+strident] for Finley 2011a) rather than a hierarchy or the relationship between natural classes.

3.3.3 Implicational universals involving a substantive bias: round vowel harmony

In this section, I review Finley (2012), a study that provides a test of phonetic bias (substantive bias) in round vowel harmony. The current thesis also examines one kind of phonetic bias, nasal harmony. The typical pattern of round vowel harmony that Finley examines involves vowel harmony triggers: mid vowels are more likely to trigger round harmony than high vowels. One of the reasons for this typological asymmetry is that round harmony has greater perceptual benefits for mid vowels than for high vowels. Kaun (1995, 2004) discusses perceptual benefits about the relationship between vowel height and rounding, finding that high vowels are easier to identify as round than mid vowels are. Terbeek (1977) found that high and back vowels are more likely to be perceived as round than mid and front vowels based on a continuum of perceptibility data of round features created by multidimensional scaling. This work demonstrates that high vowels are relatively easy to perceive as round, and they are a relatively worse source trigger of round vowel harmony compared to mid vowels.

The hypothesis of Finley's AG experiment is as follows: participants should be able to generalize to novel items during the test phase if they are exposed to source triggers of mid vowels during the exposure phase, while participants would have difficulty making generalizations to novel items if they are exposed to high vowel triggers. This hypothesis is confirmed by her results.

The structure of the stimuli is CVCV-CV. There were two conditions, the Mid and High Vowel Triggers (Vowel Triggers). In the Mid Vowel Trigger condition, the two vowels in the stem were identical and mid (/e, o/, as in /b**e**d**e**/, and /g**o**b**o**/). In the High Vowel Trigger condition, the two vowels in the stem were the same and high (/i, u/, as in /b**i**d**i**/, and /g**u**b**u**/). Suffixes were of two types, and each had two allomorphs (-mi vs. -mu, and -ge vs. -go). Specifically, in the Mid Vowel Trigger condition, there were four vowel patterns in the exposure phase: (1) front suffix vowel, **ee-i** vs. **ee-e**, (2) back suffix vowel, **oo-u** vs. **oo-o**. In the High Vowel Trigger condition, there were also four vowel patterns: (1) front suffix vowel, **ii-i** vs. **ii-e**, (2) back suffix vowel, **uu-u** vs. **uu-o**. Stem consonants were selected from /b, d, g, p, t, k, m, n/. Each pattern had six tokens, yielding 24 stem + suffix pairs[14].

[14] The suffixed items were recorded by a male English speaker in three different forms (i.e., CVCVgə (the ə ensured that the final stem vowel was produced with minimal coarticulation), CVCV-gi, CVCV-gu). For instance, the test item /**bidigi**/ was created by cross-splicing the /bidi/ fragment from /**bidi**gə/, and the /gi/ fragment from /bidi**gi**/, while the test item /**bidigu**/ was created by cross-splicing the /bidi/ fragment from /**bidi**gə/, and the /gu/ fragment from /bidi**gu**/. In this case, the only difference between /bidigi/ and /bidigu/ was the suffix, it was either a harmonic suffix or a disharmonic suffix.

There were two types of control conditions. In the Control-Mid condition, during exposure, participants heard 12 harmonic unsuffixed stems (CVCV-) which were the same as the stems used in the Mid Vowel Trigger condition, and they were also exposed to 12 disharmonic unsuffixed stems (CVCV-). In the Control-High condition, during exposure, participants heard 12 harmonic stems identical to the ones used in the High Vowel Trigger condition, and the 12 disharmonic stems were the same as the ones used in the Control-Mid condition. Each exposure item was repeated five times and randomized for each cycle. Before the exposure phase began, participants were informed that they would learn a language which they had never heard before, but they did not have to memorize the items.

During the test phase, participants heard 24 forced-choice test items. Items consist of two words in a row, disharmonic stem-suffix vs. harmonic stem-suffix. They needed to choose which one was from the language they just learned.

A 3 by 2 ANOVA was conducted with Training (Mid Vowel Trigger, High Vowel Trigger, Control) as the between-participant factor, and Test Item (Old, New) as the within-participant factor, showing a main effect of condition with no effect of Test Item, and no interaction[15].

A marginally significant difference was found between the Mid Vowel Trigger and the High Vowel Trigger conditions. This marginal effect occurred because the Mid and High Vowel Trigger conditions for New test items showed a significant difference ($p < .05$, correction of multiple comparisons). Moreover, as shown in Figure 3, there was a significant difference between the Mid Vowel Trigger and Control-Mid conditions for new items (means: 0.67 vs. 0.47), but no significant difference was found between the High Vowel Trigger and Control-High conditions (0.55 vs. 0.54) for new items, suggesting that participants in the Mid Vowel Trigger condition could generalize to new novel items, but participants in the High Vowel Trigger condition failed to do so[16]. This supports Kaun (1995, 2004)'s view that mid vowels are better source triggers of round harmony than high vowels.

[15] Pair-wise comparisons showed a significant effect of training for the Mid Vowel Trigger condition (mean: 0.66) compared to the Control-Mid condition (0.51), but no significant effect of training for the High Vowel condition (0.58) was found when compared to Control-High condition (0.51).
[16] An AXB task (judging whether X is identical to A or B) was conducted after the main experiment, confirming that participants could distinguish different vowel height (criterion of distinguishing rate: 75%).

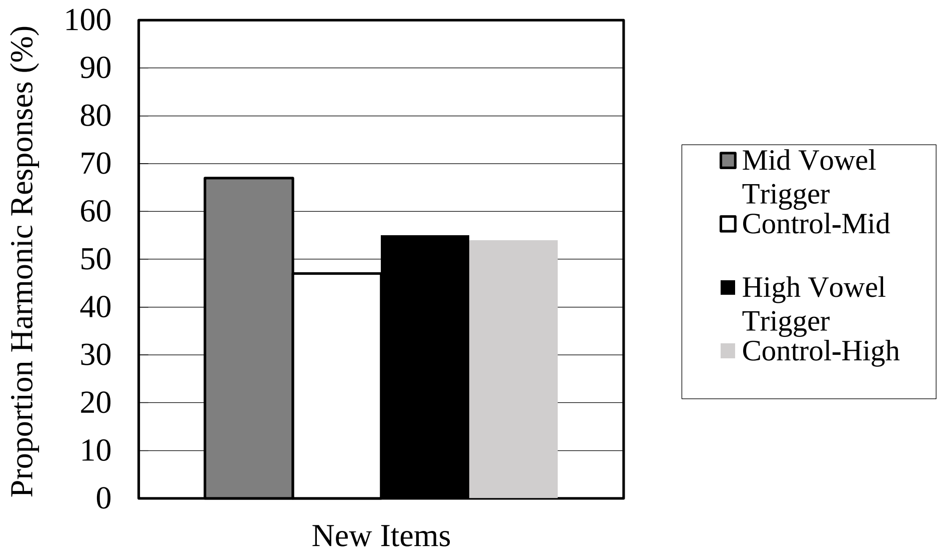


Figure 3. Four conditions for the new items

In summary, Finley (2012) successfully used the AG paradigm to test for an implicational universal involving round harmony: mid vowels are more likely to trigger round harmony than high vowels.

3.3.4 Sonority hierarchy

In this section, I review Berent's (2013) work on the sonority hierarchy as this hierarchy is relevant to my investigation. The work reported in this section assumes the phonological (i.e., universal and innate) nature of the hierarchy[17].

Berent (2013) asserts that grammatical constraints (Universal Grammar) are innate, with phonologically unmarked structure learned more readily than phonologically marked structure. Specifically, she studies word-initial onset clusters to examine the universal that all languages prefer onsets with larger increasing sonority distances, as shown in (7). That is, onsets with sonority rises are less phonologically marked than onsets with sonority plateaus, which are in turn less phonologically marked than onsets with sonority falls. More marked sonority distances imply less marked ones (cf. Greenberg 1978). Note also that I assume that the sonority hierarchy is phonological instead of phonetic.

(7) The preference for large sonority distances (Berent 2013: 168)

$\ldots\Delta s=2 \succ \Delta s=1 \succ \Delta s=0 \succ \Delta s=-1 \succ \Delta s=-2\ldots$

bl ≻ bn ≻ bd ≻ nb ≻ lb

Berent (2013) summarizes an experiment of Berent et al. (2007) that investigates whether English speakers are sensitive to sonority of unattested onsets in their native language. In the

[17] There is a debate about the status of the sonority hierarchy, with some arguing that it is phonological (e.g., Hooper 1976, Clements 1990, Blevins 1995, Zec 2007), and others arguing that it is a consequence of phonetic correlates, and epiphenomenal (e.g., Gordon 1999, Henke et al. 2012, Evans & Levinson 2009, Daland et al. 2011). One reason for considering the sonority hierarchy to be phonological is that there are some differences between languages in how particular segments interact in terms of sonority but the sonority hierarchy of different languages all follow the same tendency (least sonorant to more sonorant). For instance, one language might have a consonant sonority hierarchy as follows, from least to most sonorant: voiceless stops < voiced stops < voiceless fricatives < voiced fricatives < nasals < liquids < glides, while another language might have a relatively simplified consonant sonority hierarchy: stops < fricatives < nasals < liquids < glides (e.g., Prince & Smolensky 1993, 2004, de Lacy 2000, Smith 2002, Parker 2002, de Lacy 2006). Even within obstruents, Clements (1990) among others argues that languages can differ in whether obstruents are treated as a class or are more finely differentiated. In addition, certain segments may pattern as part of different classes– /v/ patterns as a fricative in some languages and as a sonorant in others. Such differences in sonority suggest a phonological account because such an account entails that the sonority hierarchy must be phonological rather than purely physical properties of speech sounds (phonetics) because sonority must be acquired and realized in language-specifically phonological patterning. In addition to patterning, sonority plays a similar role in various phonological phenomena cross-linguistically such as phonotactic constraints (e.g., consonant clusters, syllabification-vowel epenthesis, deletion, syllabic consonants, metathesis) and morphophonemic alternations (lenition, fortition) (see Parker 2002, Cser 2003, Honeybone 2008, Parker 2011).

experiment, participants were presented with three types of onsets, as shown in Table 5 (Berent 2013: 181).

Table 5. A sample of the materials from Berent et al. (2007)

	Monosyllables	Disyllables	Sonority distances
Small rise	bnif	bənif	Δs=1
Plateau	bdif	bədif	Δs=0
Fall	lbif	ləbif	Δs=-2

The first type of onset in Table 5 involves a stop-nasal sequence, the second type has a stop-stop sequence, and the third type has a liquid-stop sequence. In English, in general onsets with sonority distances that are more than 2 are allowed with one exception: s-initial onsets. If English speakers have knowledge about markedness of sonority distances, then participants will be most likely to misidentify the ill-formed monosyllable /lbif/ as the disyllable /ləbif/, and least likely to misidentify the ill-formed /bnif/ as /bənif/. That is, the likelihood of misidentification is predicted to be positively related to markedness: (1) strongest misidentification-/**lb**if/ (most marked), (2) moderate misidentification-/**bd**if/ (marked), (3) lowest misidentification-/**bn**if/ (least marked). In the experiments, participants were presented with one auditory stimulus at a time (either a monosyllable or a disyllable). Participants were asked to judge whether the word contained one or two syllables. Supporting the hypothesis, the results (see Figure 4) show that for the monosyllables, the likelihood of misidentification was the highest for the onsets with sonority falls (around 62%), with sonority plateaus in between (around 28%), and the lowest for the onsets with sonority rises (around 14%).

For disyllables, the response accuracy for /bənif/ was lower than for /ləbif/. Berent (2013) explains that this is not due to ill-formedness. Both /bənif/ and /ləbif/ are well-formed. It is also not due to the phonetic properties of stimuli. The duration of schwa for /bənif/ and /ləbif/ was controlled. Berent offers an alternative explanation, that the syllable count task is a forced choice. Since /lbif/ is a worse monosyllable than /bnif/, this makes /ləbif/ a relatively better disyllable than /bənif/. That is, when participants heard /bənif/ (Disyllables: “Fall” in Figure 4), they would more likely suspect that it might be a monosyllable compared to the scenario where participants heard /ləbif/ (Disyllables: “Rise” in Figure 4). This is why the accuracy of “Rise” for the disyllables condition (around 89%) is a bit lower than the accuracy of "Fall" for the disyllables condition (around 95%).

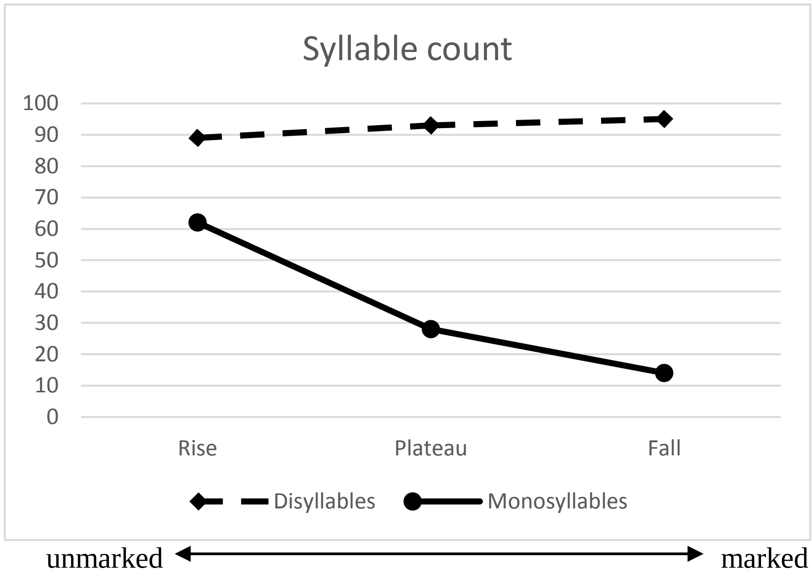


Figure 4. Response accuracy in the syllable count task (Redrawn from Figure 8.1, Berent et al. 2013: 182)

Ren et al. (2010) also looked at clusters, showing that Mandarin speakers favor /bl-/ (sonority rises) over /lb-/ (sonority falls), even though Mandarin does not have onset clusters, and has limited codas /n, ŋ, ɻ/. This suggests that Mandarin speakers are sensitive to onsets with different sonority distances, even though Mandarin bans onset clusters. Zhao & Berent (2011) further investigate this issue, finding that Mandarin speakers are sensitive to the sonority hierarchy of onset clusters: sonority rises ≻ sonority plateaus ≻ sonority falls. Korean, which also bans onset clusters, shows the same trend (cf. Berent et al. 2008).

This section summarized Berent's (2013) work that suggests that the sonority hierarchy represents a universal tendency in onsets. Even in languages like Mandarin and Korean which lack onset clusters in their phonological systems, speakers are sensitive to onsets with different sonority distances. This might provide evidence for learning biases in a language without consonant hierarchy effects. These learning biases match with the cross-linguistic typology that sonority rises in onsets: sonority falls imply sonority plateaus which in turn imply sonority rises.

3.3.5 Implicational universals: natural classes/features

In this section I summarize Wilson's (2003) work on nasal assimilation and dissimilation in Section 3.3.5.1, and Moreton's (2008) work on height-voice and voice-voice patterns in Section 3.3.5.2.

3.3.5.1 *Nasal assimilation and dissimilation*

Wilson (2003) examined nasal assimilation and dissimilation, using an artificial grammar learning paradigm to argue that certain types of processes have a 'privileged status' (Wilson 2003: 112). The following is a summary of the assimilation and dissimilation experiments. Note that the participants in these experiments were native speakers of American English.

For assimilation, Wilson examined the patterning of a suffix with two allomorphs (/-la/ and /-na/). In the first pattern (assimilation), the suffix is /-na/ when the preceding stem consonant is [+nasal], otherwise the default suffix is /-la/ ([-nasal]) (e.g., dumena vs. tukola). In the second pattern (dependency between [+nasal] and [+dorsal]), the suffix is /-na/ when the preceding stem consonant is [+dorsal], otherwise the default suffix is /-la/ ([-nasal]) (e.g., tukona vs. dumela). The first pattern involves typical consonant harmony (Rose & Walker 2002) and is attested ("natural" assimilation process). In contrast, the second pattern is arbitrary in the sense that there is no formal or substantive relation between the alternation in the suffix and the property of the stem conditioning the alternation. That is, there is no formal relation between the value of [dorsal] for the stem and the value of [nasal] for the suffix. Such an arbitrary pattern can exist in languages (Pierrehumbert 2002), compatible with the view of UG as a tendency rather than an absolute requirement, as discussed in Section 3.2, but it is typologically unusual.

To test the learnability of the assimilation pattern and the arbitrary pattern, Wilson conducted an artificial grammar learning experiment that contained these two patterns. In the exposure phase, 20 /CVCVCV/ items were presented. One group was exposed to consonant harmony (Group 1A), and the other group to the arbitrary pattern (Group 1B). The list was repeated twice, each time in a different random order. In the test phase, participants heard 80 /CVCVCV/ items arranged in two blocks of 40 each. Participants hit one button to respond "yes" (they remembered hearing an item in the exposure phase), and hit the other button to respond "no" (they did not remember hearing an item in the exposure phase). The test stimuli for the consonant harmony pattern (assimilation) are shown in Table 6 ((4) in Wilson 2003: 106). The grammaticality of the test stimuli for the arbitrary alternation (arbitrary pattern) was opposite to that of nasal assimilation except when the preceding stem consonant is [-nasal, -dorsal] (e.g., sutola, *sutona), so it is not shown here.

Table 6. Stimulus categories of the consonant harmony pattern in the test phase

		Grammaticality	
		grammatical	ungrammatical
Test stimuli	old	N=20 ex. dumena, sutola	N=20 ex. *dumela, *sutona
	new	N=20 ex. kinena, tagola	N=20 ex. *kinela, *tagona

The results show that Group 1A (trained on consonant harmony) accepted new grammatical items significantly more often than new ungrammatical ones, but Group B (trained on arbitrary alternation) did not show this significant effect. Since participants could generalize to new items, the results suggest that participants in Group 1A 'acquired' knowledge of the nasal assimilation rule in the exposure phase, and this knowledge made them reject ungrammatical items: they did not blindly accept ungrammatical items containing old stems. It also made them accept new grammatical items which conformed to the rule but were not presented in the exposure phase. By contrast, there was no evidence that participants in Group 1B acquired knowledge of the arbitrary rule.

Nevertheless, there is a possible confound in this assimilation experiment. Rose & Walker (2002) state that agreement in [sonorant], [continuant], and place features makes consonants susceptible to agree with other features. For the assimilation rule, the allomorph suffix /-na/ always agreed with the preceding stem consonant in [sonorant] and [continuant], and sometimes agreed in place. The arbitrary rule does not have this similarity property. If so, it is uncertain that the learning asymmetry is because of a cognitive bias about the stem-suffix relationship or because of the similarity property. Wilson therefore did another dissimilation experiment to rule out this similarity confound. That is, for nasal dissimilation, the suffix consonant /l/ in /-la/ and the stem [+nasal] consonant would not agree in values for [sonorant] and [continuant].

For the dissimilation experiment, Wilson also designed two patterns. In Group 2A (nasal dissimilation), the suffix was /-la/ if and only if the final stem consonant was [+nasal] (dumela vs. tukona), and in Group 2B (arbitrary dissimilation), the suffix was /-la/ if and only if the final stem consonant was [+dorsal] *(dumena vs. tukola)*. The procedure was the same as the first experiment. This experiment also showed that participants accepted new grammatical items significantly more often than new ungrammatical items, but no such effect was found in the arbitrary dissimilation pattern.

Having summarized the findings, I now examine the statistical analysis in detail. Wilson (2003) analyzed his data in two steps. At the first step, he used ANOVA to analyze the assimilation and dissimilation experiments separately. For the assimilation experiment, there were two within-participant factors (STEM TYPE: old vs. new, GRAMMATICALITY: grammatical vs. ungrammatical) and one between-participant factor (GROUP: Groups 1A vs. Group 1B). For the dissimilation experiment, one between-participant factor was GROUP (Group 2A vs. Group 2B), and two within-participant factors were the same as those of the assimilation experiment. For both experiments, no significant main effect of GROUP was found, implying that neither group made errors significantly more than the other.

At the second step, Wilson (2003) continued to use repeated-measures ANOVA to look at Group 1A and Group 1B separately. The same repeated-measures ANOVA was also adopted for Group 2A and Group 2B. For Group 1A, no significant main effect of STEM TYPE (old vs. new) was found. By contrast, the main effect of GRAMMATICALITY was significant: participants responded “yes” significantly more often to grammatical stimuli than to ungrammatical stimuli. The crucial pairwise comparison between new-grammatical and new-ungrammatical stimuli was also significant: new-grammatical stimuli were accepted more often than new-ungrammatical stimuli, suggesting that participants in Group 1A could generalize to new items and indeed learn this phonologically natural assimilation pattern. For Group 1B, a significant main effect of STEM TYPE was found: the stimuli with old stems were accepted significantly more often than the stimuli with new items. By contrast, the main effect of GRAMMATICALITY was not significant. The crucial pairwise comparison between new-grammatical and new-ungrammatical stimuli was not significant, suggesting that participants in Group 1B did not learn the arbitrary assimilation pattern. For Group 2A, the main effect of STEM TYPE was significant: the stimuli with old stems were accepted significantly more often than the stimuli with new stems. The main effect of

GRAMMATICALITY was also significant: the grammatical stimuli were accepted significantly more often than the ungrammatical ones. A crucial pairwise comparison between new-grammatical and new-ungrammatical stimuli was significant (new-grammatical stimuli > new-ungrammatical stimuli), demonstrating that participants in Group 2A really learned the phonologically natural dissimilation pattern. For Group 2B, the participants accepted the stimuli with old stems significantly more often than the stimuli with new items (main effect of STEM TYPE). No main effect of GRAMMATICALITY was found. The crucial pairwise comparison was not significant, showing that participants in Group 2B did not learn this arbitrary dissimilation pattern.

To summarize, Wilson (2003) examines nasal assimilation and dissimilation, and shows that it is easier to learn a natural pattern than an arbitrary pattern.

3.3.5.2 *Height-voice and voice-voice*

In this section, I summarize Moreton's (2008) work on height-voice and voice-voice dependencies. Moreton shows that sometimes typological frequency does not match learnability of artificial grammar. The following is a summary of the experiments and discussion of potential reasons for the mismatch between typological frequency and learnability.

In one of Moreton's (2008) experiments, participants learned two patterns, a height-voice pattern (a high vowel followed by a voiced consonant such as /p**ig**o/, or a non-high vowel followed by a voiceless consonant such as /p**ok**i/), and a voice-voice pattern (consonant agrees with another consonant in voicing such as /**k**i**k**i/ or /**g**i**b**u/). The results showed that that English speakers learned the voice-voice pattern better than the height-voice pattern. This cannot be explained by a typological asymmetry or the strength of the phonetic precursor as both patterns are typologically rare (see Hansson 2004, Rose & Walker 2004, Moreton 2008). Hence, Moreton (2008) concludes, Universal Grammar is set of learning biases that restricts participants from learning a typologically less frequent pattern and that helps participants learn a typologically more frequent pattern[18]. Since the results do not correspond to typological facts, according to Moreton (2008), some type of bias other than UG might favor single-feature dependencies ([voice]) over two-different-feature dependencies ([voice] and vowel height) or favor within-tier (within the consonant tier)

[18] This definition of UG is similar to substantive bias in the sense of typological frequency.

dependencies over between-tier (between the consonant tier and the vowel tier) dependencies. However, another alternative is possible: the height-voice pattern was not learned as well as the voice-voice pattern because vowel height and voicing are independent of one another, even though both patterns do exist. Some features are closely linked. For instance, the tonogenesis literature points to a relationship between voicing and tone (voiced and low tone vs. voiceless and high tone). If participants learn a tone-voice pattern (e.g., voiced + L tone; voiceless + H tone), we might find that the tone-voice pattern is learned better than the height-voice pattern, even though both patterns involve two features and both of them are typologically rare (see Moreton 2010). If an asymmetry between height-voice and tone-voice patterns was substantial, it could be attributable to vowel height and voicing being independent, while tone and voicing can interact. Specifically, voicing can have a co-articulatory effect on vowels as tone, but vowel height and voicing are unrelated. Whether two features are independent or interacting with each other would potentially influence learnability. The interactive relationship like vowels and tones might be easier to learn that the independent relationship between vowel height and voicing.

The above arguments about inconsistency between learnability and natural markedness indicate that artificial grammar learning does not necessarily match with typological frequency (i.e., natural markedness). If typological frequency is assumed to be explained well by UG in generative phonology (e.g., Chomsky & Halle 1968: 4, 251, 296–297, Archangeli & Pulleyblank 1994: 391–395, Clements & Hume 1995: 245, Steriade 2001: 235–237, Davidson et al. 2004, Hayes & Steriade 2004: 1–2, 6, Moreton 2008: 86), and if using an artificial grammar learning paradigm is claimed to reflect natural markedness, then we should not expect such inconsistencies between learnability and typological frequency. However, such mismatches between learnability, natural markedness, and UG do occur. How do we then know what direction to go? That is, is it necessarily the case that the more unmarked a pattern is, the easier it is to learn? It could be that there are things other than typological frequency that defines natural markedness. It seems that there is no consensus in artificial grammar work about what other things could define natural markedness in addition to typological markedness.

In brief, the above discussion suggests that typological frequency, phonetic naturalness, and other possible factors all play a role in learnability.

3.3.6 Summary

Section 3.3 summarizes the research with respect to substantive bias, including palatalization (Wilson 2006), sibilant harmony (Finley 2011a) (coronal assimilation), round harmony (Finley 2012), natural classes/features (Moreton 208), and nasal assimilation and dissimilation (Wilson 2003). Specifically, both artificial grammar studies involving implicational universals in substantive bias, including Wilson (2006), and Finley (2012), and artificial grammar studies involving implicational universals in formal complexity bias, including Finley 2011a (locality) serve as foundations for the current study in terms of experimental design and analysis.

3.4 Formal complexity bias

Recall that the literature on phonological learning posits two types of bias, substantive bias and formal complexity bias. In this section I discuss the latter, including (1) domain general: attribute-based object classification which involves the number of features, (2) contiguity-similarity tradeoff, (3) feature agreement, and (4) why domain-general knowledge could be active in natural language.

3.4.1 Domain-general: attribute-based object classification

In addition to substantive bias (typologically and phonetically motivated) accounts of learnability asymmetries with artificial grammars (cf. Sections 3.2 and 3.3), some research claims that these are not sufficient to account for all the results, and it has been proposed that feature count is also important: a simpler pattern (with fewer features) is easier to learn than a more complex pattern (with more features). A formal complexity bias, not a focus of the current study, involves the number of features or the relationship between features. It is worth reviewing this here because the evidence for formal complexity bias in learning is strong. Moreton & Pater (2012a) argue that formal complexity bias might be domain-general (i.e., is not restricted to language). Specifically, the number of phonological features is a type of formal complexity that is comparable to domain-general concepts such as attribute-based object classification. For instance, consider a situation with 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 features (i.e., size, color, shape) are involved with binary values (large – [+size] vs. small – [-size], black – [+color], white – [-color], triangle – [+shape], circle – [-shape]). These figures can be classified into two groups with equal numbers (i.e., 4 figures in each group) in six ways, as in Figure 5.

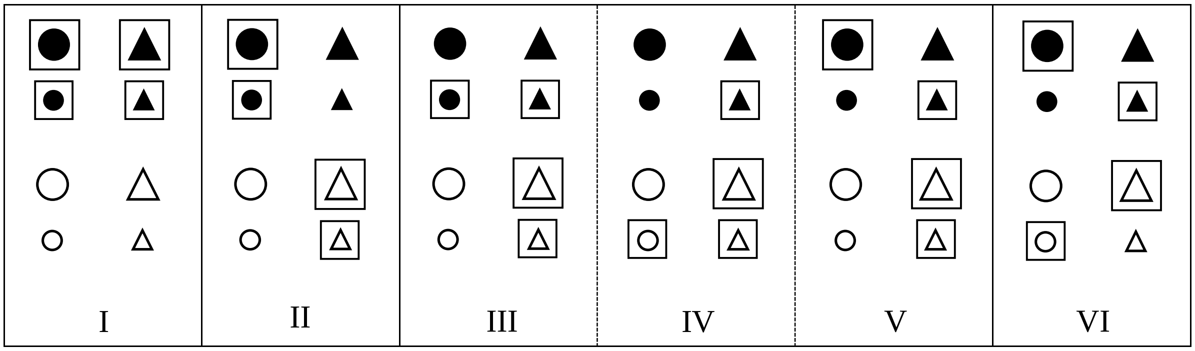


Figure 5. Six possible partitions defined by binary features (Redrawn from Figure (1), Moreton & Pater 2012a: 688)

In Type I, [+color] (black) figures are used to define one group (in boxes), and the other figures fall into a second group (group that is not boxed). The formal complexity is one (i.e., number of features). For Type II, two features, color and shape, are used to define one group: [+color, -shape] (black circles), [-color, +shape] (white triangles) (in boxes). The formal complexity is two. From Type III-Type V, three features are used, with certain subsets of values used less. The formal complexity for these three types is all three. Specifically, one group for Type III is [-size, +color, -shape] (small black circle), [-size, +color, +shape] (small black triangle), [-color, +shape] (white triangles). One group for Type IV is [-size, -color, -shape] (small white circle), [-size, +color, +shape] (small black triangle), [-color, +shape] (white triangles). One group for Type V is [+shape, +color, -shape] (big black circle), [-shape, +color, +shape] (small black triangle), [-color, -shape] (white triangle). Last, in Type VI, with a formal complexity of four, the four members of the group are defined as: [+size, +color, -shape] (big black circle), [-size, -color, -shape] (small white circle), [-size, +color, +shape] (small black triangle), [+size, -color, +shape] (big white triangle). In brief, psychologists have tested these six non-linguistic concepts (see Shepard et al. 1961, Neisser & Weene 1962, Nosofsky et al. 1994a,b, Feldman 2000, Love 2002, Smith et al. 2004, Moreton & Pater 2012a: 688), and found that the degree of formal complexity is: Type I < Type II < Type III, Type IV, Type V < Type VI.

Phonological experiments have tested Type I, Type II, and Type VI (Types I and II: Pycha et al. 2003, Saffran & Thiessen 2003, Cristiá & Seidl 2008, Types II and IV: Kuo 2009, among others), with the overall finding that systems with fewer features are learned better than systems with more features. In addition to number of features, formal complexity also involves relations between features such as the contiguity-similarity tradeoff [19] (Section 3.4.1.1), and feature agreement (Section 3.4.1.2).

3.4.1.1 *Contiguity-similarity tradeoff*

Formal complexity with respect to the contiguity-similarity tradeoff means that "phonological theory typically treats dependencies between adjacent elements as the normal case, excluding long-distance interactions unless the interacting segments share some property which is absent from intervening material" (Moreton & Pater 2012a: 693) (see also Jensen 1974, McCarthy 1981, Cole & Trigo 1988, Pierrehumbert 1993, Odden 1995, Gafos 1996, Hansson 2001, Frisch et al. 2004, Rose & Walker 2004, Heinz 2010). Moreton & Pater (2012a: 693–694) indicate that contiguity and similarity are expected to interact to facilitate learning, but no artificial grammar learning studies actually test this interaction. Artificial grammar studies often test formal complexity concerning contiguity, and this is relevant to the current thesis. Here I only discuss formal complexity with contiguity.

Formal complexity with respect to contiguity means that a dependency with a shorter distance should be learned better than one with a longer distance[20]. Finley (2011a) found that if participants learned sibilant harmony with a longer distance between the sibilants during the exposure phase, then they would be able to generalize to the one with a shorter distance, but not the other way (see Section 3.3.2). Along the same line, since speech continues forward (not backward), this implies that backward/leftward nasal harmony should be harder to learn than forward/rightward nasal harmony, confirmed by Nevins's (2013) findings that leftward (backward) nasal harmony implies rightward (forward) nasal harmony, but not vice versa (i.e., implicational universals)[21]. However, other works have not found the tendencies suggested by contiguity (cf.

[19] In general, dependencies between adjacent/contiguous segments are phonologically more natural than those between long-distance/non-contiguous segments, that is, local assimilation is phonologically more natural than long-distance assimilation.
[20] Moreton & Pater (2012a: 693–694) indicate that contiguity and similarity are expected to interact to facilitate learning, but no artificial grammar learning studies actually test this interaction.
[21] Whether all forward harmony is easier to learn than backward harmony is an empirical question.

Majerus et al. 2004, Warker & Dell 2006, Warker et. al 2008, Koo & Callahan 2011, Finley 2012, Moreton 2008, 2012).

3.4.1.2 *Feature agreement*

Wilson (2003) touches on one more aspect of formal complexity, feature agreement. For instance, he tested agreement or disagreement in [nasal] between stem and suffix. He also looked at the relationship between the features [dorsal] and [nasal], and [dorsal] and [lateral] with respect to agreement (see Section 3.3.5). The results show that participants learned the agreement or disagreement in one feature significantly better than control groups (without training). However, he did not find any significant learning effect on patterns involving two features. Moreton (2008) and Lin (2009) found that height agreement between vowels (one feature) was learned better than a dependency between the height of V_1 and voicing of C_2 (two features). Not specific to vowel height, voice agreement between consonants was learned significantly better than the dependency between the height of V_1 and voicing of C_2 (Moreton 2008), and the learning of backness agreement between vowels was significantly better than that of the dependency between backness of V_1 and voicing of C_2 (Moreton 2012).

While in the work noted above, patterns involving one feature were learned better than those involving two features, others have found no evidence to support the privilege of one feature over two features (e.g., Kuo 2009: place-place correlation vs. place-aspiration correlation; Seidl & Buckley 2005, Exp. 2: the agreement in labiality between C_1 and V_1 vs. the dependency between the labiality of C_1 and the vowel height of V_1).

3.4.2 Domain-specific: natural language

Hume & Johnson (2001) argue that domain-general mechanisms are active in natural-language phonology. Specifically, they argue that cognitive factors and generalizations are not only germane to linguistic category formation, but to category formation in general. If this is the case, then artificial grammar phonology concerning formal complexity is indicative of natural-language phonology. However, if natural-language phonology and artificial grammar language phonology do not share the same processes, then artificial grammar language phonology would be irrelevant to natural language.

In brief, in Sections 3.2 and 3.4, I discussed substantive bias (phonetic naturalness) and formal complexity bias, showing that substantive bias is relatively weaker than formal complexity bias and thus harder to uncover in artificial grammar learning. As mentioned in Chapter 1, one goal of the current study is to incorporate the positives of the methodology of Wilson (2006), enhancing the possibility to find an implicational universal of phonetic naturalness: whether it is easier to make a generalization when a more marked blocker is presented during exposure and a less marked blocker in test rather than vice versa.

3.5 Summary

In sum, this chapter summarizes a variety of works concerning natural markedness, formal complexity bias vs. substantive bias, implicational universals concerning palatalization, sibilant harmony and round harmony, the sonority hierarchy, and [nasal]-related work: nasal harmony and nasal assimilation/dissimilation. The following table summarizes the major work that I reviewed. Though these works tackle different issues and have different goals, they can be classified into two topics, substantive bias involving implicational universals, sonority hierarchy, and natural classes/features, and formal complexity bias. Since formal complexity bias involves features, sometimes substantive bias based on natural classes/features is not differentiable from formal complexity bias. Both could account for results. This is why I put Wilson (2003) and Moreton (2008) in both categories of "natural classes/features (substantive bias)" and of "formal complexity bias". I put Finley (2011a) in both "implicational universals in substantive bias" and "formal complexity bias" because, as discussed in Section 3.3.2, sibilant harmony is related to phonetic substance (coronal assimilation: [+strident]) and locality (formal complexity bias). One of the goals for the current study aims to investigate what matters in artificial grammar learning, especially for implicational universals, sonority hierarchy, and natural classes.

Table 7. Major work by topics[22]

Substantive bias	sonority hierarchy (substantive bias)	natural classes/ features (substantive bias)	formal complexity bias
Wilson (2006): palatalization (implicational universal) Finley (2011a): sibilant harmony Finley (2012): round harmony (implicational universal)	Berent (2013): onsets	Wilson (2003): assimilation/ dissimilation on [nasal] Moreton (2008): height-voice voice-voice	Wilson (2003): assimilation/dissimilation on [nasal] Moreton (2008): height-voice voice-voice Finley (2011a): locality (implicational universals)

As we will see in the following chapters, my study takes these works as a foundation, and builds on them in that it involves substantive bias between feature classes.

[22] Note that there are many phonological artificial grammar studies (foot-conditioned segmental phonotactis in Bennett 2012, saltation in White 2013, long-distance phonotactics in McMullin & Hansson 2014, among others), but here I only include the research that is most relevant to my thesis.

Chapter 4 Experiment 1

As discussed in Chapter 1, Walker (2011) proposes a universal implicational nasalized segment hierarchy based on evidence from typological frequency. This hierarchy is repeated in (8). Walker's claim is that the more sonorant a segment is, the likelier it is to get nasalized in nasal harmony. Walker argues that if a more marked blocker class blocks harmony (vowels are least marked targets, so the least likely to be blockers, and the most likely to be targets), so do the less marked blocker classes (stops are the most marked targets, so the most likely to be blockers, and the least likely to be targets). The universal implicational nasalized segment hierarchy can be framed into a concept schema as below. Recall that I use the term "targets" the segments involved that are not phonetically nasalized. If nasalization continues over a sound, I call it a target, but such sounds such as stops and fricatives might better be termed transparent.

(8) Vowel-consonant nasal harmony with opaque segments: a universal implicational nasalized segment hierarchy (Walker 2011)

Vowels > Glides > Liquids > Fricatives > Stops

more marked blocker ⇒ less marked blocker

less marked target ⇒ more marked target

In this work I address the question of whether a pattern that is predicted by this implicational universal is easier to learn than one that is not. In particular, I investigate if it is easier to make a generalization when a more marked phonological class is presented during training (i.e., old phonological class) and a less marked phonological class in testing (i.e., new phonological class) rather than vice versa. For example, the *blocking* of nasal harmony by vowels (more marked) implies the *blocking* by stops (less marked), but not vice versa. Further, the *targeting* of stops (more marked) implies of the *targeting* of vowels (less marked), but not vice versa.

These types of patterns are the topic of this study. In particular, the goal of the study is to look at: (1) the implicational relationship between the blocking by /s/ and /k/ (i.e., a more marked blocker /s/ vs. a less marked blocker /k/ in nasal harmony); (2) the implicational relationship between the blocking by /k/ and /p/ (/k/ and /p/ are equally unmarked/marked because both belong to stops). If a listener/learner is introduced to a system showing nasal harmony involving a sound like /s/ as a blocker (i.e., old phonological class) and then is asked whether a sound like /k/ (i.e., new phonological class) will be nasalized, the answer should be "no", since /s/, a fricative, is more marked as a blocker than /k/, a stop. On the other hand, if the listener/learner is exposed to a sound like /k/ as a blocker (i.e., old phonological class) and is asked whether a sound like /s/ (i.e., new phonological class) will be nasalized, there is no prediction, since /s/ is more marked as a blocker than /k/ with respect to the hierarchy. Thus exposure to /s/ as a blocker predicts the blocking by /k/ (abbreviated as s→k) but exposure to /k/ as a blocker does not predict the blocking by /s/ (abbreviated as k→s). Note, however, that it could be the case that both /s/ and /k/ are treated as obstruents instead of fricative and stop. If so, there would be no significant learning difference of direction of generalization (i.e., s→k vs. k→s).

While there is a prediction that between classes, exposure to a more sonorant sound as a blocker implies that a less sonorant sound will also be a blocker, the prediction is different within a class. For instance, whether exposure is to /k/ as a blocker with testing on /p/ (k→p) or vice versa (p→k), similar patterning is expected. However, if k→p and p→k do show significantly different learning, then it would imply that perhaps participants do not use their knowledge of the sonority hierarchy to classify both /k/ and /p/ into the same natural class. Instead, participants might use other knowledge such as place of articulation.

The predictions examined in Experiment 1 are summarized in Table 8 and Table 9, and I explain how to read the table immediately following Table 8. Note that I follow Chuang's (1996) transcription, using the traditional symbol /y/ rather than the IPA /j/. Glide /w/ was always presented in the exposure phase as a target, and glide /y/ was always presented in the test phase.

Table 8. Predictions about the interaction between segments and nasal harmony

	exposure: blocker	test: potential blocker	phonological class	prediction	learnability (> or =)	type A harmony[23]
Pattern A (Group A)	blocker /**s**/	new segment: **/k/** less sonorant than/same class as /s/	(1) /s/ = fricative /k/ = stop	(1) new segment is blocker	(1) Patterns A > B	Type A: A2, A3, A2-A3
			(2) /s, k/ = obstruent	(2) new segment is blocker	(2) Patterns A = B	
Pattern B (Group B)	blocker /**k**/	new segment: **/s/** more sonorant than/same class as /k/	(1) /s/ = fricative /k/ = stop	(1) no prediction	(1) Patterns A > B	Type A: A2, A3, A4, A2-A3
			(2) /s, k/ = obstruent	(2) new segment is blocker	(2) Patterns A = B	

Patterns A and B examine the implicational universals of the blocking effect based on the sonority hierarchy: blocking by the more marked phonological class implies blocking by the less marked phonological class, but not vice versa. Each pattern represents one artificial grammar language. In these two artificial languages, vowels and glides are targets. Stops are blockers. The only difference between A and B is the 'direction' of generalization. That is, Pattern A participants (Group A) are exposed to a blocker /s/, and then tested with /k/, while Pattern B participants (Group B) are exposed to a blocker /k/, and then tested with /s/. This is shown in the columns inside **"exposure: blocker"** and **"test: potential blocker"**.

The hypotheses and predictions of implicational universals are formulated by "patterning" and "learnability". The columns **"phonological class"** and **"prediction"** indicate the possible ways in which a new sound could pattern, assuming sonority classes. For instance, if participants in Pattern A treat a new sound /k/ as a stop and treat /k/ as a blocker, and participants in Pattern B treat a new sound /s/ as a fricative and thus have no basis to determine what to pattern /s/ with, then the learning of Pattern A is expected to be better than the learning of Pattern B (cf. Patterns

[23] Note that this column refers back to Chapter 2. As illustrated in Table 1, Section 2.2, there are six types of Type A nasal harmony (A1, A2, A2-A3, A3, A4), depending on the patterning of blockers, targets and morphological constraints.

A > B, **“learnability”**). Another possibility is that if participants in Pattern A treat a new sound /k/ as an obstruent and treat /k/ as blocker, and participants in Pattern B pattern a new sound /s/ as an obstruent and treat /s/ as blocker, then there is expected to be no significant learning difference (cf. Patterns A = B, “learnability”).

Patterns C and D examine the implicational universals of the blocking effect based on the sonority hierarchy: blocking by the same phonological class (i.e., stops: the same sonority) implies blocking by the same phonological class, and vice versa (see Table 9).

Table 9. Predictions about the interaction between segments and nasal harmony

	exposure: blocker	test: potential blocker	phonological class	prediction	learnability	type A harmony
Pattern C (Group C)	blocker **/p/**	new segment: **/k/** same class as /p/	/k, p/ = stop or obstruent	new segment is blocker	Patterns C = D	Type A: A2, A3, A4, A2-A3
Pattern D (Group D)	blocker **/k/**	new segment: **/p/** same class as /k/	/k, p/ = stop or obstruent	new segment is blocker	Patterns C = D	

The only difference between C and D is the direction of generalization. That is, Pattern C participants (Group C) are exposed to a blocker /p/, and then tested with /k/, while Pattern D participants (Group D) are exposed to a blocker /k/, and then tested with /p/. This is shown in the columns inside **“exposure: blocker”** and **“test: potential blocker”**. Participants in both patterns are predicted to be able to generalize from an old blocker to a new consonant (cf. **“prediction”**) because /k, p/ are both stops/obstruents (cf. **“phonological class”**) and do not involve any change in the sonority hierarchy. Therefore, no significant learning difference is expected (cf. **“learnability”**).

4.1 Predictions

Instead of testing the entire universal implicational nasalized segment hierarchy (Vowels > Glides > Liquids > Fricatives > Stops), I started with the nasalized segment hierarchy with a shorter sonority distance, that between obstruents (stops and fricatives), before testing the hierarchy with a larger sonority distance, that between obstruents and sonorants. In order to determine if there is any directional learning involved, I tested fricatives and stops (cf. Patterns A and B in Table 8), with a sonority distance of 1, and stops and stops (Patterns C and D in Table 9), with a sonority distance of 0. If participants are sensitive to the full sonority hierarchy, then I would expect that they would distinguish fricatives from stops, but not stops from stops. The previous artificial grammar studies all assume that natural classes exist. In studying obstruents /s, k, p/, I ask if participants have a concept of natural classes when learning an artificial grammar, or if they simply treat each sound as a segment. If the latter is true, then the implicational universal based on the nasalized segment hierarchy is not testable. With this concern, the first goal of the current study is to test if participants are sensitive to natural classes. If participants are sensitive to natural classes, then I would expect that they would make natural class generalizations, patterning /s/ and /k/ as fricative and stop respectively and patterning both /k/ and /p/ as stops. Alternatively, they would treat /s, k, p/ all as obstruents.

This chapter presents Experiment 1, looking at Patterns A-D.

4.2 Methods

4.2.1 Language choice

In order to test the implicational universal between sonority and nasal harmony with opaque segments, speakers of Min were chosen as experimental participants. Note that Min is also called Taiwanese and Taiwan Southern Min. I use the term Min.

4.2.1.1 *Inventory, phonotactics and syllable shapes*

Before giving the reasons why I selected this population, I introduce the inventory of Min, given in Table 10 and Table 11.

Table 10. Phonemic consonants in Min (Chung 1996: 3)

			labial	coronal	velar	glottal
stop	voiced		b	l	g	
	voiceless	unaspirated	p	t	k	ʔ
		aspirated	p^h	t^h	k^h	
fricative	voiced			z		
	voiceless			s		h
affricate	voiceless	unaspirated		ts		
		aspirated		ts^h		
nasal	voiced		m	n	ŋ	
glide	voiced		w	y		

Table 11. Phonemic vowels in Min (Chung 1996: 2)

Oral vowels		Nasal vowels	
i	u	ĩ	
e	o	ẽ	
	ɔ		ɔ̃
a		ã	

In Min, there are 20 phonemic consonants, 6 phonemic oral vowels and 4 phonemic nasal vowels. The syllable structure is basically (C)(V)V(C) with /p, t, k, m, n, ŋ/ allowed in coda position. /l/ is classified as a voiced stop, taking the place of /d/.

/p, t, k/ are treated as obstruents since they can appear in a closed syllable with a checked tone, unlike an open syllable or syllable closed by a sonorant (nasal) with unchecked tone (see Hsieh 2003, Hsieh 2006, Tu & Davis 2009, Tu 2013). However, /h, s/ and other fricatives and affricates do not have this complementary distribution with sonorants. Note also that although /ʔ/ can occur in coda position, its pattern of tone sandhi is different from that of /p, t, k/ (see discussion in Chen 2000, Hsieh 2006, Tu 2013).

4.2.1.2 *Limited nasal spreading*

In Min, there is some C-to-V nasal assimilation, but it is limited, and the process is lexicalized (Wang 1993). Among those lexicalized words, the spreading of nasality is bi-directional. Specifically, nasalization can spread leftward within a syllable or within a word (see (9 a-c)), or it can spread rightward to a suffix /a53/ (see (9 d-f)) (numbers represent tones). Consider the forms in (9), from Chung 1996: 171, 173, 175.

(9) Nasal spreading in Min

a. /bi**n**33 + a53/ /tsay31/	→	**mỹã**33 tsay31	'tomorrow' (leftward spreading)
b. /tsa53 + **n**i31/ /ho/	→	**tsãỹ**53 ho53	'so nice' (leftward spreading)
c. /tsay55 + **ỹ**ã53/	→	**tsãỹ**33 ỹã53	'know' (leftward spreading)
d. /bw**ã**13/ + /a53/	→	mw̃ã13 **ã**53	'sesame' (rightward spreading)
e. /**ĩ**13 + a53/	→	ĩ13 **ã**53	'dumpling' (rightward spreading)
f. /**ẽ**55 + a53/	→	ẽ55 **ã**53	'baby' (rightward spreading)

Note that in the examples in (9 a-e), a nasal consonant or vowel spreads [nasal] leftward or rightward. The onsets [b, l, g] and [m, n, ŋ] are in complementary distribution: an oral vowel always follows an oral consonant, as shown in (10) (Cheng 1968, 1973, Ting 1985, Zhang 1989, Pan 2007). This allophonic pattern suggests that in Min voiced stops are targets in onset position (see 9 a, d), but voiceless stops are not.

(10) Nasalization assimilation in Min

/b, l, g/ → [m, n, ŋ]/_ [+ nasal]
V

/m, n, ŋ/ → [b, l, g]/_ [-nasal]
V

Unlike the nasal harmony discussed in Walker (2000), the spreading of nasality across syllables is highly restricted in Min (Chung 1996, Chou 2002).

4.2.1.3 *Reasons for choosing Min speakers as participants*

There are the three reasons for choosing Min speakers as participants in this study. First, I am testing nasalization, and nasal vowels are phonemic in Min. Thus participants should find it straightforward to notice nasality and treat it as phonemic during artificial grammar learning. Lin (2010, 2012) found that the contrast between nasalized and oral segments was easy for Min speakers to hear. Following Lin (2010), for Experiment 1, a post-test was conducted to make sure that participants could distinguish between the two classes.

The second and third reasons are based on a principle of the AG paradigm, namely that the learning should not be inferable from the participants' own language background (nor from feedback in the experiment) – the native language should not give speakers a bias against or in favor of a pattern to be tested in artificial grammar learning. Given this principle, the second and third reasons for having Min speakers as experimental participants are as follows. Second, there are no apparent differences in phonological patterning between sonority classes in Min consonants[24]. This is partly because consonant clusters are prohibited in Min. With a lack of clear evidence for sonority, Min speakers do not bring a bias to an experiment that tests the learning of a hierarchy that is based on sonority.

Third, while Min speakers have phonemic nasalization, they do not have evidence for productive nasal harmony. Thus experience with Min should not influence participants in learning cross-syllabic nasal spreading patterns.

4.2.2 Design

In this section, I introduce the design of the experiment using Patterns A and B in Table 8 for illustration. In Pattern A participants were exposed to /s/ as a blocker and tested on /k/ as a potential blocker (again abbreviated as s→k) and in Pattern B, they were exposed to /k/ and tested on /s/ (k→s).

[24] The literature with respect to nasalization only discusses whether oral vowels and nasal vowels would tend to be grouped into the same category based on differences in vowel height (e.g., Wang 2001), investigates prosodic boundaries in nasality (e.g., Pan 2007), examines historical changes in nasalization (e.g., Lien 2000), and looks at how nasalization influences native Min speakers to learn English coda nasals (e.g., Hsu 2009), etc. But nothing is found having to do with the sonority hierarchy for consonants.

4.2.2.1 *Shapes of words and syllable forms*

The overall design is as follows. Participants were exposed to pairs of morphologically related items (artificial words) with a singular form and a plural form[25]. This morphophonemic design has been used extensively in phonological artificial grammar learning (Kapatsinski 2009, Albright & Do 2013, Albright & Do 2015, Kimper 2015, among others). See Sections 4.2.2.4 and 4.2.2.5 for details. The singular form was composed of oral segments. In the plural, the first vowel was nasalized and, depending on the intervening consonants, subsequent vowels could be nasalized as well.

Words were of the structure $V_1.C_2V_2.C_3V_3$, where V_1 was oral in the singular forms (e.g., /asawa/) and nasal in the plural forms (e.g., /ãsawa/). C_2 and C_3 varied between targets and blockers. [nasal] on V_1 would spread rightward to V_2 and V_3 if there was no blocker interfering (e.g., /ãw̃ãw̃ã/).

Consonants were selected from voiceless stops and fricatives, and glides. In this study, only the oral vowels /a, i, e, ɔ/ and their nasal counterparts were used in stimuli. This is because /u/ vs. /ũ/ is a less robust contrast in terms of frequency in Min. Specifically, /u/ vs. /ũ/ is underrepresented, showing more accidental gaps than other vowel pairs do. Syllable structures conformed to the phonotactics of Min.

Examples of lexical monosyllables in Min are presented in Table 12. Onset consonants are the ones used in this thesis. The gray shading indicates lexical gaps. This clearly shows that monosyllables with /ũ/ have more lexical gaps than the other three nasal vowels.

[25] The participants in my study were familiar with the concept of plurality. Also, I orally explained what singular and plural meant by using Mandarin and English words as examples before they began the experiment.

Table 12. The monosyllables (MS) corresponding to Min lexically[26]

onset	/a/	/ã/	/e/	/ẽ/	/i/	/ĩ/	/ɔ/	/ɔ̃/	/u/	/ũ/
null	a55 'prefix'	ã13 'protect'	e55 'swing'	ẽ55 'baby'	i22 'put'	ĩ13 'round'	ɔ55 'black'	ɔ̃55 'baby's sleep'	u33 'have'	
/k/	ka55 'to cut'	kã55 'basket' (classifier)	ke55 'chicken'	kẽ55 'thick soup'	ki55 'a' (classifier for something long like a stick)	kĩ55 'thick soup'	kɔ55 'high'	kɔ̃31 'to snore'	ku55 'turtle'	
/t/	ta 55 'dry'	tã55 'burden'	te33 'land'	tẽ33 'squeez e'	ti55 'pig'	tĩ55 'sweet'	tɔ33 'road'	tɔ̃33 'a mimic sound to express the situation where the line is busy'	tu55 'pile of'	
/s/	sa55 'to grab'	sã55 'three'	se55 'comb'	sẽ55 'give birth to'	si11 'be'	sĩ 33 'simulative'	sɔ13 'creep'		su 55 'to lose'	
/h/	ha33 'put on'	hã53 'threaten'	he55 'that'	/hẽ/	hi55 'weak'	hĩ33 'ear'	hɔ13 'river'	hɔ̃55 'question marker'	hu11 'attach'	
/w/	wa53 'I'	wã53 'bowl'	we53 'dig'		wi33 'stomach'	wĩ55 'yellow'				
/y/	ya33 'sprinkle'	yã13 'win'					yɔ55 'waist'		yu11 'tender'	yũ13 'sheep'

[26] Note that no liquid exists in Min, rather /l/ is treated as a voiced stop (see Table 10). I did not include voiced stops in the stimuli. Therefore, the possible nasalized segment hierarchy in this study is *NASPLO >> *NASFRIC >> *NASGLI >> *NASVOW.

4.2.2.2 *Stimuli*

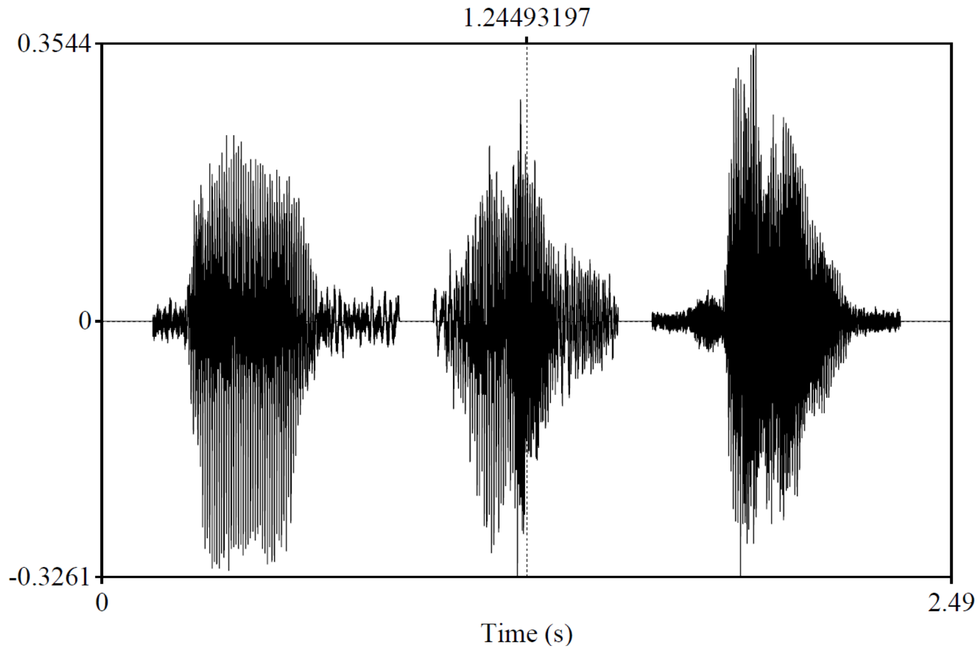
The tri-syllabic nonce-words were created by concatenating syllables spoken in isolation by a male native Min speaker who was naive to the goals of the experiment. The tonal value for each syllable was high level (55), used without any further manipulation[27]. Each 'word' was synthesized individually using Praat (Boersma & Weenink 2012). Following Lin (2012), each minimal pair (eg., /w̃ã/ vs. /wa/) was controlled to match in duration (less than 30 ms difference), and the durations for the whole 'word' were also controlled (less than 30 ms difference)[28]. Concatenation misses appropriate vowel transitions (F1-F3), which might potentially sound odd. To rule out this concern, the minimal pairs of monosyllables were presented to a trained linguist[29] to make sure each pair was contrastive perceptually. The silence before the first syllable was set to 150 ms, the silence between the first and the second syllable was set to 100 ms, and the silence between the second and third syllable was also set to 100 ms[30]. No acoustic checkup was involved (e.g., F1-F3 for transition). A waveform and spectrogram of /ãw̃ãsa/ are shown in Figure 6 and Figure 7 as follows. Following Moreton (2008), no amplitude normalization was applied in order not to disturb the natural intensity difference between high and low vowels.

[27] In Lin (2010), monosyllables were recorded by a native female Min speaker naive to the goal of experiments using Praat recording through a desk-mounted microphone. The tonal value of the monosyllables was set to 207 Hz using Praat (Boersma & Weenink 2008), and then syllables were concatenated to form tri-syllabic artificial words. In addition to the main experiments, Lin (2010) ran a pilot of six participants to confirm that 207 Hz produced by a native Min female speaker was recognized better than 407, 307, 107 Hz in a perceptual judgment task: oral and nasal monosyllables with one of the four tonal values were randomized, and participants heard one monosyllable at a time and were asked to judge whether it was nasalized. The correct rate was the highest for 207 Hz than the others. However, since participants recruited for Lin (2010) claimed that the manipulation of tonal values made oral sounds sound like nasalized sounds and unnatural, in the current experiment I did not manipulate tonal values. The silence intervals 100 ms and 150 ms for the current study were set by praat scripting.

[28] The boundaries were marked by hand, so there is a possibility that some intersyllable boundaries were slightly longer.

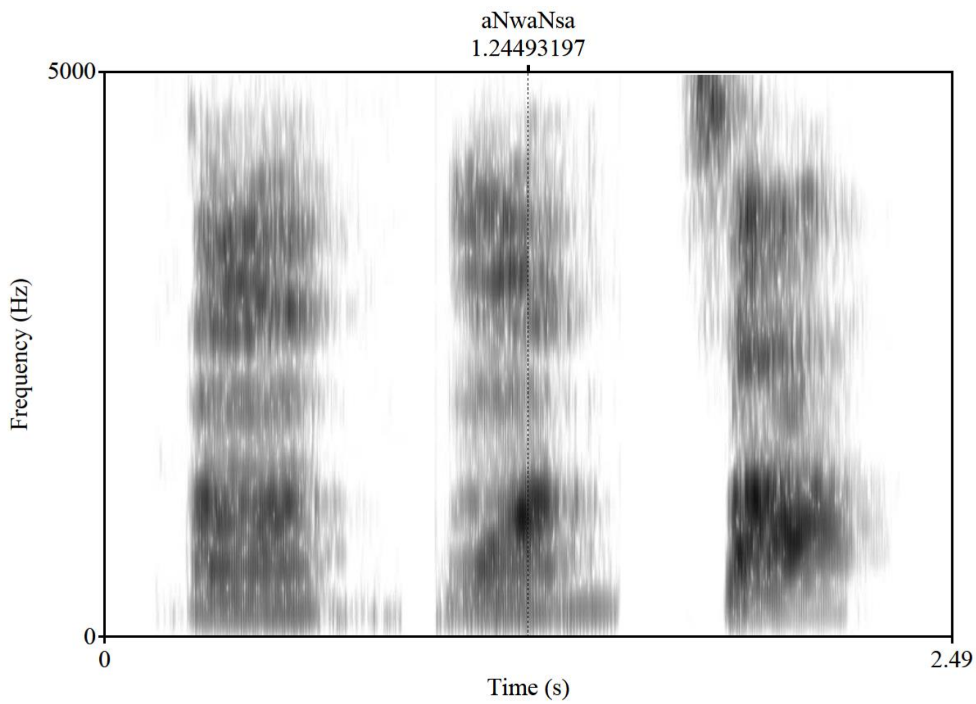
[29] My supervisor Keren Rice.

[30] Lin (2010) points out two concerns. There is a concern that the silence between syllables might sound like plosives to Min speakers, so every item would have blockers everywhere. If so, whatever people are learning, they would not be the learning the patterns that Walker (2000) addresses. However the silence of the stimuli did not sound like plosives to Lin or to James Myers. Another concern is that with silence, the stimulus might sound like three words, not one. However, the stimulus sounded more like one three-syllable word pronounced very slowly. This is good for my purposes, since the listeners can be sure that nasal harmony is not mere coarticulation. Hence, if participants learn any patterns, they might be like the patterns Walker looks at. Following Lin (2010) and Lin (2012), I used a strategy to ensure that participants would treat three-syllable words as one word rather than two or three: Participants were informed that they would hear one trisyllabic word at a time.



ã wã sa

Figure 6. Waveform of /ãw̃ãsa/



ã w̃ã sa

Figure 7. Spectrogram of /ãw̃ãsa/

4.2.2.3 *Task*

Lin (2010) cites Wilson's (2003) claim that a non-metalinguistic memory task (less metalinguistic than a grammaticality judgment task) is closer to a natural language setting than an overt grammaticality task, while Moreton (2008) adopts an overt grammaticality judgment task and suggests that this task can simulate the natural language situation. In this thesis I follow Moreton (2008), using an overt grammaticality judgment task. The reason is that when being told to search for rules, participants would be highly sensitive to old (appeared in the exposure phase) and new items (Lin 2012), not just memorizing old segments. The memory task makes it more difficult to find whether participants generalize from an old segment to a new segment, increasing the possibility that participants simply reject a new item without looking for a pattern.

This study adopts a morphologically driven design (singular vs. plural) that is different from the previous artificial grammar studies on nasal harmony (Lin 2010) in that the design allowed for the testing of the mapping between underlying representations of singulars and plurals instead of the surface representations only. Specifically, in Lin (2010), participants were exposed to one tri-syllabic word at a time (e.g., /a**p**wasa/, /a**m**w̃ãsa/, /a**t**wasa/, /a**n**w̃ãsa/; the nasal coda /m/ or /n/ of the first syllable serves as a trigger), and were asked to figure out what rules conditioned a change in this artificial language. However, there was no clear way for them to detect which segment was a trigger. The current design treats [nasal] as a plural morpheme (trigger), which allows participants to compare a set of singular (without trigger) and plural (with trigger) forms, and it raises the chances for participants to figure out why nasal spreading stops when encountering certain segments (blockers).

4.2.2.4 *Exposure phase*

Based on the above description, two groups of artificial grammars were generated. For both groups, V_1 was /a/ (singular) or /ã/ (plural) (e.g., singular: /awasa/ vs. plural: /ãw̃ãsa/). The direction of nasal harmony for plural forms was always from left to right (triggered by a plural morpheme [nasal] on V_1). Note that although in real languages, the trigger is not necessarily in the first syllable, in the experiments reported in this thesis, the trigger was always in the first syllable. The reason for this is that if the trigger varied in position, the span of nasal harmony would become more narrowed, which would make it more difficult to see a crucial nasality interaction among vowels, glides, fricatives and oral stops.

For the first group of artificial grammars (Pattern A), in the exposure phase, vowels /a, e/ were targets, and glide /w/ was also a target, while fricative /s/ was a blocker.

For the second group of artificial grammars (i.e., Pattern B), in the exposure phase, targets were /a, e/ (vowels) and /w/ (glide), while stop /k/ was a blocker.

The logical combinations used in the exposure phase are four, as in (11). A blocker is underlined.

(11) Four combinations of items in the exposure phase

a. trigger + <u>obstruent</u> + vowel + glide + vowel (e.g., /ãsawa/)

b. trigger + glide + vowel + <u>obstruent</u> + vowel (e.g., /ãw̃ãsa/)

c. trigger + <u>obstruent</u> + vowel + <u>obstruent</u> + vowel (e.g, /ãsasa/)

d. trigger + glide + vowel + glide + vowel. (/ãw̃ãw̃ã/)

The vowel quality in V_2 and V_3 varied between /a/ and /e/. The consonant in C_2 and C_3 also varied between obstruent (/s/ in Pattern A, /k/ in Pattern B) and glide /w/. A total of 16 singular-plural pairs were created (16= 2 (V_2 vowels)*2 (V_3 vowels) * 2 (C_2 consonant) * 2 (C_3 consonant)). 16 exposure (paired) items for Patterns A and B were generated separately, as in Table 13 and Table 14. The exposure items in the exposure phase for (11a) and (11c) were repeated four times in four different blocks, while (11b) and (11d) were repeated eight times in four different blocks in an attempt to increase participants' exposure to combinations with changes in nasalization (i.e., trigger + target + blocker; trigger + target + target). Note that this section only shows four logical combinations for Patterns A and B. The complete stimuli items are given in Appendix I.

Table 13. Positive (i.e., grammatical) stimuli in the exposure phase (Pattern A: /s/ as a blocker (exposure)→/k/ as a blocker (test); participants are expected to be able to generalize from /s/ as a blocker to /k/ as a blocker.)

Set Number	Singular form	Plural form
1st set	(1) /asawa/ 'one apple'	(1') /ãsawa/ 'two apples'
2nd set	(2) /awasa/ 'one bird'	(2') /ãw̃ãsa/ 'two birds'
3rd set	(3) /asasa/ 'one car'	(3') /ãsasa/ 'two cars'
4th set	(4) /awawa/ 'one cat'	(4') /ãw̃ãw̃ã/ 'two cats'

Table 14. Positive (grammatical) stimuli in the exposure phase (Pattern B: /k/ as a blocker (exposure)→/s/ as a blocker (test) – there is no prediction about the generalizability from stop /k/ as a blocker to fricative /s/ as a blocker; participants are expected to be able to generalize from obstruent /s/ to obstruent /k/.)

Set Number	Singular form	Plural form
1st set	(1) /akawa/ 'one apple'	(1') /ãkawa/ 'two apples'
2nd set	(2) /awaka/ 'one bird'	(2') /ãw̃ãka/ 'two birds'
3rd set	(3) /akaka/ 'one car'	(3') /ãkaka/ 'two cars'
4th set	(4) /awawa/ 'one cat'	(4') /ãw̃ãw̃ã/ 'two cats'

4.2.2.5 *Test phase*

The test phase consists of "old" and "new" stimuli. Old stimuli are the same items as the ones presented in the exposure phase while new items have new vowels and consonants in V_2 and V_3 and C_2 and C_3. V_1 was always /a/ for singular (e.g., /ayiki/) and /ã/ for plural items (e.g., /ãỹĩki/). In $C_2V_2.C_3V_3$, C_2 and C_3 were selected from /w, y, s, h, k, t/. V_2 in V_2C_2 and V_3 in V_3C_3 were oral vowels /a, i, e, ɔ/ or nasalized vowels /ã, ĩ, ẽ, ɔ̃/.

For Pattern A (s→k), in the test phase, targets were /i, ɔ/ (vowels) and /y/ (glide), while potential blockers were /k/ (stop) and /h/. /h/ was chosen because both /s/ and /h/ are considered to be fricatives in Min. I assume that participants would treat a blocker /s/ that appeared in the exposure phase and a new segment /h/ the same because both are fricatives. The same logic holds for new vowels /i, ɔ/ and a new glide /y/. Participants would treat old vowels /a, e/ and new vowels /i, ɔ/ the same because they are all vowels. Participants would also group an old glide /w/ and a new glide /y/ into the same class since /w/ and /y/ are both glides. Pattern A is expected to be generalizable with participants generalizing from an old blocker /s/ to a new segment /k/, treating /k/ as a blocker (s→k).

The logical combinations used for new items in the test phase of Pattern A were seven, as in (12).

(12) Seven combinations of items in the test phase

a. trigger + <u>blocker(k)</u> + target (e.g., /ãkiyi/)

b. trigger + <u>blocker(h)</u> + target (e.g., /ãhiyi/)

c. trigger + target + <u>blocker(k)</u> (e.g., /ãỹĩki/)

d. trigger + target + <u>blocker(h)</u> (e.g., /ãỹĩhi/)

e. trigger + <u>blocker(k)</u> + <u>blocker(k)</u> (e.g., /ãkiki/)

f. trigger + <u>blocker(h)</u> + <u>blocker(h)</u> (e.g., /ãhihi/)

g. trigger + target + target (e.g., /ãỹĩỹĩ/)

For Pattern B (k→s), in the test phase, targets were /i, ɔ/ (new vowels) and /y/ (new glide), while the blockers were /s/ (test segment: fricative) and /t/ (new stop). Assuming participants have a concept of natural classes, an old segment /k/ would pattern with a new segment /t/ because both of them belong to stops.

There were 88 test items in the test phase. 44 test items were grammatical, and the other 44 test items were ungrammatical. Among the 44 grammatical items, 16 items were identical to the plural items in the exposure phase, and 28 plural items were new. The 28 new grammatical test items for each group are presented in Table 15 (Pattern A) and Table 16 (Pattern B). The items in grey are crucial items that test whether there is an implicational universal relationship involving fricatives and stops in nasal harmony. Recall that /s̃/ and other nasalized obstruents like /h̃, k̃/ refer to consonants transparent to nasal harmony, not phonetically nasalized fricatives or stops (i.e., transparent non-target). In all cases discussed below, glides are targets.

Table 15. Positive new stimuli in the test phase (For Pattern A, /s/ as a blocker (exposure)→/k/ as a blocker (test), participants are expected to be able to generalize from /s/ to /k/.)

Trigger+/blocker=k/+target (4)	**Trigger+/blocker=h/+target (4)**
/ãkiyi/	/ãhiyi/
/ãkɔyɔ/	/ãhɔyɔ/
/ãkiyɔ/	/ãhiyɔ/
/ãkɔyi/	/ãhɔyi/
Trigger+target+/blocker=k/ (4)	**Trigger+target+/blocker=h/ (4)**
/ãỹĩki/	/ãỹĩhi/
/ãỹɔ̃kɔ/	/ãỹɔ̃hɔ/
/ãỹĩkɔ/	/ãỹĩhɔ/
/ãỹɔ̃ki/	/ãỹĩhi/
Trigger+/blocker=k/+/blocker=k/ (4)	**Trigger+/blocker=h/+/blocker=h/ (4)**
/ãkiki/	/ãhihi/
/ãkɔkɔ/	/ãhɔhɔ/
/ãkikɔ/	/ãhihɔ/
/ãkɔki/	/ãhɔhi/
Trigger+target+target (4)	
/ãỹĩỹĩ/	
/ãỹɔ̃ỹɔ̃/	
/ãỹĩỹɔ̃/	
/ãỹɔ̃ỹĩ/	

Table 16. Positive new stimuli in the test phase (For Pattern B, /k/ as a blocker (exposure)→/s/ as a blocker (test): there is no prediction about the generalizability from stop /k/ to fricative /s/; participants are expected to be able to generalize from obstruent /s/ to obstruent /k/.)

Trigger+/blocker=t/+target (4)	**Trigger+/blocker=s/+target (4)**
/ãtiyi/	/ãsiyi/
/ãtɔyɔ/	/ãsɔyɔ/
/ãtiyɔ/	/ãsiyɔ/
/ãtɔyi/	/ãsɔyi/
Trigger+target+/blocker=t/ (4)	**Trigger+target+/blocker=s/ (4)**
/ãỹĩti/	/ãỹĩsi/
/ãỹɔ̃tɔ/	/ãỹɔ̃sɔ/
/ãỹĩtɔ/	/ãỹĩsɔ/
/ãỹɔ̃ti/	/ãỹĩsi/
Trigger+/blocker=t/+/blocker=t/ (4)	**Trigger+/blocker=s/+/blocker=s/ (4)**
/ãtiti/	/ãsisi/
/ãtɔtɔ/	/ãsɔsɔ/
/ãtitɔ/	/ãsisɔ/
/ãtɔti/	/ãsɔsi/
Trigger+target+target (4)	
/ãỹĩỹĩ/	
/ãỹɔ̃ỹɔ̃/	
/ãỹĩỹɔ̃/	
/ãỹɔ̃ỹĩ/	

The 44 ungrammatical items involve exactly the reverse versions of the 44 grammatical ones (i.e., glides were treated as blockers, and obstruents were treated as targets, cf. Table 17 and Table 18). The motivation for this design is as follows: some scholars say that the reason that participants judge ungrammatical items by chance is because ungrammatical items vary randomly, while grammatical items vary in a systematic way (see Gómez & Schvaneveldt 1994, Gómez & Gerken 1999, Pothos 2007)[31].

Table 17. Pattern A: negative old stimuli in the test phase (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Plural form
(1') /ãs̃ãwa/
(2') /ãwapa/
(3') /ãs̃ãs̃ã/
(4') /ãwawa/
(5') /ãs̃ẽwe/
(6') /ãwepe/
(7') /ãs̃ẽs̃ẽ/
(8') /ãwewe/
(9') /ãs̃ãwe/
(10') /ãwape/
(11') /ãs̃ãs̃ẽ/
(12') /ãwawe/
(13') /ãs̃ẽwa/
(14') /ãwepa/
(15') /ãs̃ẽs̃ã/
(16') /ãwewa/

[31] A caveat should be kept in mind: perhaps participants use fragmentary knowledge to make a correct judgment (i.e., ungrammatical items differ from grammatical items in too consistent a way). With this concern, in this study, if both grammatical and ungrammatical items were designed in a systematic way, then participants would not be able to make any judgments simply based on systematicity. To investigate this more systematically, further research could put one random version for ungrammatical items as a control group.

Table 18. Pattern (A): negative new stimuli in the test phase (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Trigger+/target=k/+blocker (4) /ãk̃ĩyi/ /ãk̃ɔ̃yɔ/ /ãk̃ĩyɔ/ /ãk̃ɔ̃yi/	**Trigger+/target=h/+blocker (4)** /ãh̃ĩyi/ /ãh̃ɔ̃yɔ/ /ãh̃ĩyɔ/ /ãh̃ɔ̃yi/
Trigger+blocker+/target=k/ (4) /ãyiki/ /ãyɔkɔ/ /ãyikɔ/ /ãyɔki/	**Trigger+blocker+/target=h/ (4)** /ãyihi/ /ãyɔhɔ/ /ãyihi/ /ãyihɔ/
Trigger+/target=k/+/target=k/ (4) /ãk̃ĩk̃ĩ/ /ãk̃ɔ̃k̃ɔ̃/ /ãk̃ĩk̃ɔ̃/ /ãk̃ɔ̃k̃ĩ/	**Trigger+/target=h/+/target=h/ (4)** /ãh̃ĩh̃ĩ/ /ãh̃ɔ̃h̃ɔ̃/ /ãh̃ĩh̃ɔ̃/ /ãh̃ɔ̃h̃ĩ/
Trigger+blocker+blocker (4) /ãyiyi/ /ãyɔyɔ/ /ãyiyɔ/ /ãyɔyi/	

In order to see the design more clearly, I put Patterns A and B together by factors, as shown in Table 19. Specifically, I split items into grammatical and ungrammatical. Items that appeared in the exposure phase are old items. Among new items, there are two categories, "same class" and "generalizable". "n" represents the number of items.

For Pattern A, /s/ and /h/ belong to the same phonological class, that is, fricatives. /w/ and /y/ belong to the same class, glides. /s/ and /w/ appeared in the exposure phase, and are regarded as **"old"** items. /h/ and /y/ belong to **"same class"** under **"new"** (/h/ with /s/, /w/ with /y/). Recall that the goal of the study is to see whether participants would be able to generalize from an old class (e.g., /s/ fricative) to a new class (e.g., /k/ stop). Items containing /k/ belong to **"generalizable"** abbreviated as "gen" under "new".

By the same logic, for Pattern B, /k/ and /p/ are grouped in the same class, stops. /w/ and /y/ are grouped in the same class, glides. /k/ and /w/ belong to "old", and /p/ and /y/ belong to "same class" under "new". The goal of the study is to see whether participants would be able to generalize from an old class (i.e., /k/ stop) to a new class (i.e., /s/ fricative). Items containing /s/ belongs to "gen" under "new".

Table 19. Stimuli design for Patterns A and B: test phase

	Pattern A: **s→k** (sk)			Pattern B: **k→s** (ks)		
	new		old	new		old
	gen (n=24)	same class (n=32)	same class (n=32)	gen (n=24)	same class (n=32)	same class (n=32)
grammatical (n=44)	/ã**k**iyi/ /ãỹĩ**k**i/ /ã**k**i**k**i/	/ãhiyi/ /ãỹĩhi/ /ãhihi/ /ãỹĩỹĩ/	/ã**s**awa/ /ãw̃ã**s**a/ /ã**s**a**s**a/ /ãw̃ãw̃ã/	/ã**s**iyi/ /ãỹĩ**s**i/ /ã**s**i**s**i/	/ãpiyi/ /ãỹĩpi/ /ãpipi/ /ãỹĩỹĩ/	/ã**k**awa/ /ãw̃ã**k**a/ /ã**k**a**k**a/ /ãw̃ãw̃ã/
ungrammatical (n=44)	/ã**k̃**ĩyi/ /ãyi**k**i/ /ã**k̃**ĩ**k̃**ĩ/	/ãh̃ĩyi/ /ãyihi/ /ãh̃ĩh̃ĩ/ /ãyiyi/	/ã**s̃**ãwa/ /ãwa**s**a/ /ã**s̃**ã**s̃**ã/ /ãwawa/	/ã**s̃**ĩyi/ /ãyi**s**i/ /ã**s̃**ĩ**s̃**ĩ/	/ãp̃ĩyi/ /ãyipi/ /ãp̃ĩp̃ĩ/ /ãyiyi/	/ã**k̃**ãwa/ /ãwa**k**a/ /ã**k̃**ã**k̃**ã/ /ãwawa/

4.2.2.6 *Post-test*

A post-test was given to ensure that the participants could hear the distinction between nasalized and oral counterparts. In this post-test, each participant heard one pair of monosyllables at a time. They heard a total of 37 pairs. These were identical to the monosyllables used in the main experiment. The participants had to judge which monosyllable was nasalized[32]. This post-

[32] "Nasalized" is a common concept for Taiwanese people (/pʰĩ33-im55/ 'nasalized sounds'), so there is no difficulty in addressing this term directly.

test was done after the main experiment, so that the participants would not receive a clue about the role of nasality in the main experiment.

Following Lin (2010), the correct rate for passing the post-test was set to 75%. The participants who failed to distinguish between /a/ and /ã/ were eliminated, since if participants were not able to distinguish trigger from non-trigger, then it is not appropriate to use their data to test for the implicational universal hierarchy. No one in Experiment 1 failed to pass the post-test.

4.2.3 Participants

The participants were bilingual in Min and Mandarin[33]. Four groups of participants were recruited (Patterns A-D) from National Chung Cheng University and National Sun Yat-sen University in Taiwan. Each participant recruited from National Chung Cheng University received 105 NTD for participating in the experiment. Each participant recruited from National Sun Yat-sen University received 150 NTD for participating in the experiment. The duration of the experiments was 45-50 minutes. All participants reported normal hearing, and all distinguished nasalized monosyllables from their oral counterparts at a rate of 100%. All participants had early childhood Min exposure.

Pattern A: s→k

Eleven college students were recruited. One participant was eliminated because she consistently chose "yes" throughout the test phase. All had studied a foreign language (English 10, Japanese 3) (numbers mean number of speakers). The average age of the participants was 20.3 (SD = 2.4). 3 participants were males and 7 participants were females.

Pattern B: k→s

Ten college students were recruited. All had studied a foreign language (English 10, Japanese 7, German 2, Hakka 2). The average age of the participants was 19.5 (SD = 1.4). 5 participants were males and 5 participants were females.

[33] In Taiwan, Mandarin is an official language, while Min is a home language used in Min populations.

Pattern C: p→k

Ten college students were recruited. All had studied a foreign language (English 10, French 1, Hakka 1). The average age of the participants was 21.1 (SD = 1.9). 8 participants were males and 2 participants were females.

Pattern D: k→p

Ten college students were recruited. All had studied a foreign language (English 10, Japanese 1, German 1, Hakka 1). The average age of the participants was 21.1 (SD = 2.0). 7 participants were males and 3 participants were females.

4.2.4 Procedure

The experiment was run with E-Prime (Schneider et al. 2002) in a lab. The stimuli were played over headphones (SuperLux HD-681F)[34]. In the exposure phase, participants received oral instructions from the experimenter, along with detailed written instructions (given in Appendix III) on the computer screen. The participants first were presented with one of the artificial grammars. The participants were told that they would learn a made-up language that contained tri-syllabic words and they had to search for a singular/plural rule hidden in the exposure phase. In the exposure phase, they heard singular and plural forms in pairs. That is, they heard a word accompanied by a picture of one item (e.g., a picture of one apple), pressed the space bar, and then they heard a word accompanied by a picture of two of the same item (e.g., a picture of two apples) (see Figure 8). When participants heard a word, they were told to pronounce it once as close to the original sounds as possible. They were also informed that only plural forms would change, and singular forms would be treated as base forms. In the test phase, they were told that they would be tested on how well they could learn the singular/plural rule based on a grammaticality judgment test, and that they needed to make a judgment even though certain sounds did not occur in the exposure phase.

[34] Participants were allowed to adjust the volume, but nobody did.

In order to encourage participants to try to generalize to new sounds, participants were given a hint that they should try to group old sounds with new sounds. If an old sound A was grouped with a new sound B, then A and B would share a similar function in this made-up language[35].

During the test phase, one auditory word was presented at a time. The participants were asked to judge whether the words were possible plural forms. Participants would choose '1' if they considered that the word was a possible plural form and '0' if they considered that the word was not. 16 of the grammar-conforming stimuli were identical to those in the exposure phase, and 28 of them were new. 44 items were ungrammatical. E-prime randomly ordered these words. Participants who consistently chose the same answer (i.e., '1' throughout or '0' throughout) during the test phase were eliminated because that means that they were not trying to make a judgment.

After completing the main experiment, a post-test was given to ensure that the participants could hear the distinction between nasalized and oral segments. In this post-test, each participant heard one pair of monosyllables at a time. These were identical to the monosyllables used in the main experiment. The participants had to judge which syllable was nasalized. They pressed button '1' if the first one was nasalized and '2' if the second one was nasalized. Participants who consistently chose the same answer ('1' throughout or '2' throughout) were eliminated, since they were not trying to make a judgment. A post-interview was conducted to examine how they made decisions.

Figure 8. Singular vs. plural picture

[35] This emphasis was not given in pilot experiments that are not reported here. The emphasis to encourage participants to try to make a judgment even though certain sounds did not appear in the exposure is crucial, since the pilot studies (13 pilots in total, failure of passing the post-test: 3 participants) indicated that participants would refuse to make a judgment when faced with any new sounds.

4.2.5 Summary

The consonant segments presented during exposure and test and their predictions are reviewed in Table 20. The bolded segments are the segments under investigation of the implicational universal. During the test I included “old-same class” that were presented in the exposure phase, and “new-same class” and “new-gen”[36].

[36] For “new-same class”, I assumed that participants would have no problem learning it since new-same consonants have the same phonological class as the old-same consonants (e.g., Pattern A: “new-same class” /y/ and “old-same class” /w/; “new-same class” /h/ and “old-same class” /s/). Whether participants could treat both old segment and new-same class the same way is an empirical question. See Section 4.6 for more discussion.

Table 20. Predictions about the interaction between segments and nasal harmony

	exposure	test	prediction	patterning	learnability (> or =)
Pattern A	more sonorant: target /w/ less sonorant: blocker /**s**/	new gen new segment: /**k**/ less sonorant than/same class as /s/ new-same class more sonorant: target /y/ blocker /h/ old-same class blocker /s/	(1) new segment is blocker	(1) /s/ = fricative /k/ = stop	(1) A > B
			(2) new segment is blocker	(2) /s, k/ = obstruent	(2) A = B
Pattern B	more sonorant: target /w/ less sonorant: blocker /**k**/	new-gen new segment: /**s**/ more sonorant than/same class as /k/ new-same class more sonorant: target /y/ blocker /t/ old-same class blocker /k/	(1) no prediction	(1) /s/ = fricative /k/ = stop	(1) A > B
			(2) new segment is blocker	(2) /s, k/ = obstruent	(2) A = B

Pattern C	more sonorant: target /w/ less sonorant: blocker /**p**/	new-gen new segment: /**k**/ same class as blocker new-same class more sonorant: target /y/ blocker /t/ old-same class blocker /p/	new segment is blocker	/k, p/ = stop or obstruent	C = D
Pattern D	more sonorant: target /w/ less sonorant: blocker /**k**/	new-gen new segment: /**p**/ same class as blocker new-same class more sonorant: target /y/ blocker /t/ old-same class blocker /k/	new segment is blocker	/k, p/ = stop or obstruent	C = D

4.3 Results: Patterns A and B

In order to test for the implicational universal on sonority and nasal harmony with opaque segments between fricatives and stops, Experiment 1 used different phonological classes in the exposure phase and the test phase: trained on /s/ as a blocker and tested on /k/ (s→k, Pattern A) and the other way around (k→s, Pattern B) or trained on /p/ as a blocker and tested on /k/ (p→k, Pattern C) and the other way around (k→p, Pattern D). This section focuses on Patterns A and B, with Patterns C and D examined in Section 4.5.

Pattern A is where participants were exposed to fricatives (blocker) (underlined means consonants of interest for testing implicational universals), glides (target), and vowels (target) during the exposure phase, and then they were tested on whether they could generalize the blocking effect to stops (blocker). If this study supports the implicational universal between fricatives and stops, then it is expected that participants would be able to generalize the blocking effect to stops (Pattern A: /s/ as a blocker→/k/).

Pattern B is where participants were exposed to stops (blocker), glides (target), and vowels (target) during the exposure phase, and then they were tested on whether they could generalize the blocking effect to fricatives (blocker). If this study supports the implicational universal between fricatives and stops, then it is expected that participants would not generalize the blocking effect to stops (Pattern B: /k/ as a blocker→/s/), assuming that /k/ and /s/ are in different sonority classes, so would be random in this.

4.3.1 Descriptive statistics: Patterns A and B

In this section I present the descriptive statistics for Patterns A (s→k) and B (k→s). Participants' accuracy was coded based on grammaticality. The test phase had an equal number of correct (grammatical) and incorrect (ungrammatical) plural forms, so if the participants finished the test phase simply by guessing instead of by learning, then the percentage of correct responses would be around 50%. Note that grammaticality is based on the implicational nasalized segment hierarchy.

As discussed in Section 4.2.2, the exposure phase included items of four types, shown in (13). The blocker is underlined.

(13) Four types of items in the exposure phase

a. trigger + obstruent + vowel + glide + vowel

b. trigger + glide + vowel + obstruent + vowel

c. trigger + obstruent + vowel + obstruent + vowel

d. trigger + glide + vowel + glide + vowel

The test items included the exposure items plus new items, with 7 combinations, shown in (14).

(14) Seven types of items in the test phase

a. trigger + <u>blocker(k)</u> + target (e.g., /ãkiyi/)

b. trigger + <u>blocker(h)</u> + target (e.g., /ãhiyi/)

c. trigger + target + <u>blocker(k)</u> (e.g., /ãỹĩki/)

d. trigger + target + <u>blocker(h)</u> (e.g., /ãỹĩhi/)

e. trigger + <u>blocker(k)</u> + <u>blocker(k)</u> (e.g., /ãkiki/)

f. trigger + <u>blocker(h)</u> + <u>blocker(h)</u> (e.g., /ãhihi/)

g. trigger + target + target (e.g., /ãỹĩỹĩ/)

I divided the test items into three classes. The first class is **old-same class** – these are the exposure items. The second is **new-same class**. These are items where the glide was new (/y/ instead of the /w/ seen in the exposure phase) or where the fricative was new (/h/ instead of the /s/ of the exposure phase). Finally, a third class was distinguished, **new-generalizable**. In this case, the stop /k/ is new, and it should be possible to generalize to it.

In order to get a better sense of how participants responded to the above grammatical and ungrammatical items individually, the raw percentages of correct responses for each set are given in Table 21. Note that a response was treated as 'correct' if participants pressed "yes" (i.e., this word could be a possible plural form) to grammatical/old items (e.g., said "yes" to /ãsawa/). A response was treated as 'correct' if participants pressed "no" (i.e., this word could not be a possible plural form) to ungrammatical/old items (e.g., said "no" to */ãs̃ãwa/).

Table 21. The response rates for test items for Pattern A (s→k)

s→k	grammatical items (n=44)			Average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**k**iyi/ /ãỹĩ**k**i/ /ã**k**i**k**i/	/ãhiyi/ /ãỹĩhi/ /ãhihi/ /ãỹĩỹĩ/	/ã**s**awa/ /ãw̃ã**s**a/ /ã**s**a**s**a/ /ãw̃ãw̃ã/		/ã**k̃**ĩyi/ /ãyi**k**i/ /ã**k̃**ĩ**k̃**ĩ/	/ãh̃ĩyi/ /ãyihi/ /ãh̃ĩh̃ĩ/ /ãyiyi/	/ã**s̃**ãwa/ /ãwa**s**a/ /ã**s̃**ã**s̃**ã/ /ãwawa/	
mean	47%	53%	**91%**	65%	29%	41%	**62%**	48%

Specifically, items are divided into two major groups, Memory: Old vs. New, and Grammaticality: Grammatical vs. Ungrammatical. The memory levels "Old" and "New" were used to distinguish exposure-phase items from test-phase new items. If participants learned the artificial grammar, they should be able to generalize from old sounds to new sounds[37]. Specifically, it is expected that the overall accuracy for "New" should be above chance. In addition, the exposure phase might have a memory effect, so a main effect of "Memory" is expected.

Grammaticality, with two levels, "Grammatical" and "Ungrammatical", was used to distinguish grammatical items from ungrammatical ones. If participants learned the artificial grammar, it is expected that the overall accuracy for "Grammatical" should be above random chance. In addition, a main effect of "Grammaticality" is expected, since the previous literature (Moreton 2008, Lin 2010) suggests that participants are often confused about ungrammatical items, but show a bias toward grammatical items. If this is the case, then a significant main effect of "Grammatical" should be found[38]. Recall that the 44 ungrammatical items involve exactly the reverse version of the 44 grammatical ones (i.e., a blocker in the grammatical condition would be treated as a target in the ungrammatical condition, and the same logic applies to targets). Thus I

[37] See Section 4.6 for limitations of the design.
[38] This claim is based on the assumption that if there is an experimental treatment effect, during the test phase grammatical items are expected to be learned better than ungrammatical items since only grammatical items were given during the exposure phase.

coded a grammatical item and its ungrammatical counterpart with the same item number, so that they could be considered to be related to each other.

If a participant responded simply based on whether they heard the segments before in the exposure phase, then they would always chose “yes” for both old-grammatical and old-ungrammatical items, causing the former response rate (old grammatical) to be close to 100%, and the latter one (old ungrammatical) to be close to 0%. However, if participants learned the harmony pattern for the segments presented in the exposure phase (i.e., “old-same”), then it is expected that the response percentage for “old-same” for both grammatical and ungrammatical items would be more than 50%.

As shown in Table 21, for Pattern A (s→k), participants learned old items, since the raw rates for both grammatical and ungrammatical items for the old items are above 50%, suggesting that they did not simply guess.

Similarly, for Pattern B (k→s), according to Table 22, the raw rates for both grammatical and ungrammatical items for the old items are above 50%, suggesting that participants learned old items.

Table 22. The response rates for test items for Pattern B (k→s)

k→s	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**s**iyi/ /ãỹĩ**s**i/ /ã**s**i**s**i/	/ãpiyi/ /ãỹĩpi/ /ãpipi/ /ãỹĩỹĩ/	/ã**k**awa/ /ãw̃ã**k**a/ /ã**k**a**k**a/ /ãw̃ãw̃ã/		/ã**s̃**ĩyi/ /ãyi**s**i/ /ã**s̃**ĩ**s̃**ĩ/	/ãp̃ĩyi/ /ãyipi/ /ãp̃ĩp̃ĩ/ /ãyiyi/	/ã**k̃**ãwa/ /ãwa**k**a/ /ã**k̃**ã**k̃**ã/ /ãwawa/	
mean	46%	64%	**88%**	68%	37%	46%	**57%**	47%

If participants always rejected items with new sounds and accepted items with old sounds, then we would expect that participants would always accept old items no matter whether they were grammatical or not. If so, we could expect that the accuracy of ungrammatical “old-same class” would be close to 0%. However, this is not a case. The results for Patterns A and B suggest that the training did have an effect on participants, because overall, participants did better on old items than on new ones. However, it looks like participants had difficulty generalizing from old to new items. In addition, they generally tended to accept rather than reject items with old segments.

In order to examine the patterns more closely, I compared Patterns A and B by grammaticality. For grammatical items, as shown in Figure 9, the response percentage “new-same” is relatively higher for Pattern B (k→s: 64%), than for Pattern A (s→k: 53%). For the other two categories (“old-same”, “new-gen”), the response percentage of Pattern A (s→k) (91%, 47%) was slightly higher than that of Pattern B (k→s) (88%, 46%).

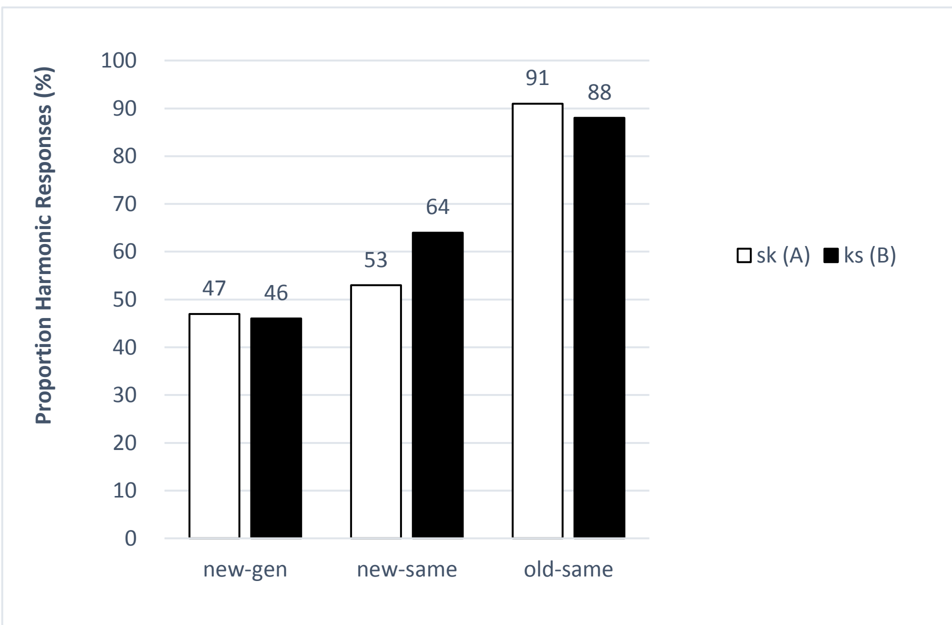


Figure 9. Response percentages for grammatical items for Pattern A: s→k and Pattern B: k→s

Figure 10 shows that for ungrammatical items, the response percentage of old items was slightly better for Pattern A (s→k: 62%) than for Pattern B (k→s: 57%). However, for the other two categories, the response percentage of Pattern B was higher than that of Pattern A (new-same: 46% vs. 41%; new-gen: 37% vs. 29%).

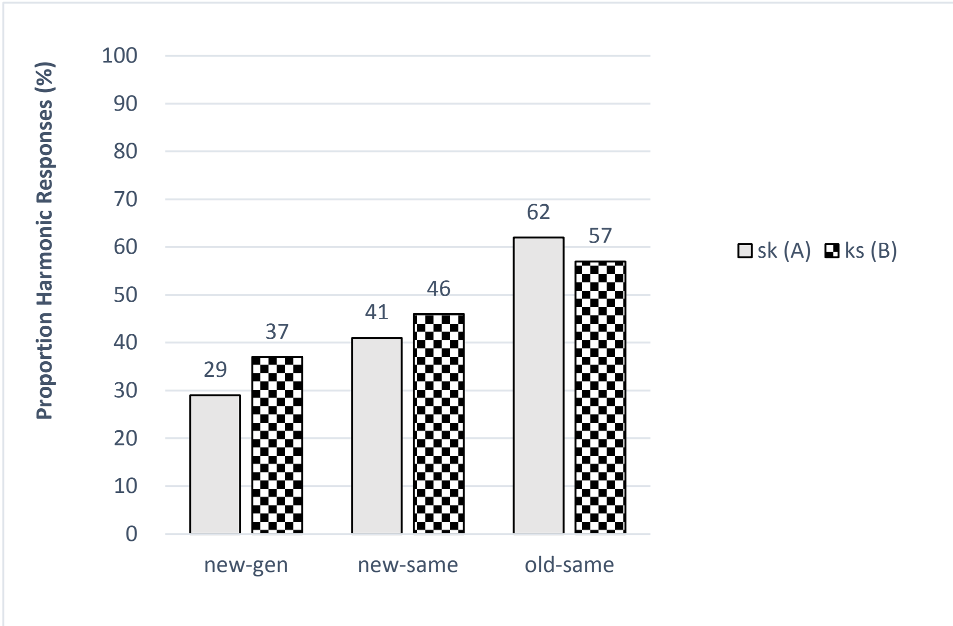


Figure 10. Response percentage for ungrammatical items for Pattern A: s→k and Pattern B: k→s

According to Figure 11, combining grammatical and ungrammatical items together (U stands for ungrammatical, G stands for grammatical), Pattern A and Pattern B participants only learned the old items (the accuracy for G ang U is both over 50%), but failed to generalize from old to new items.

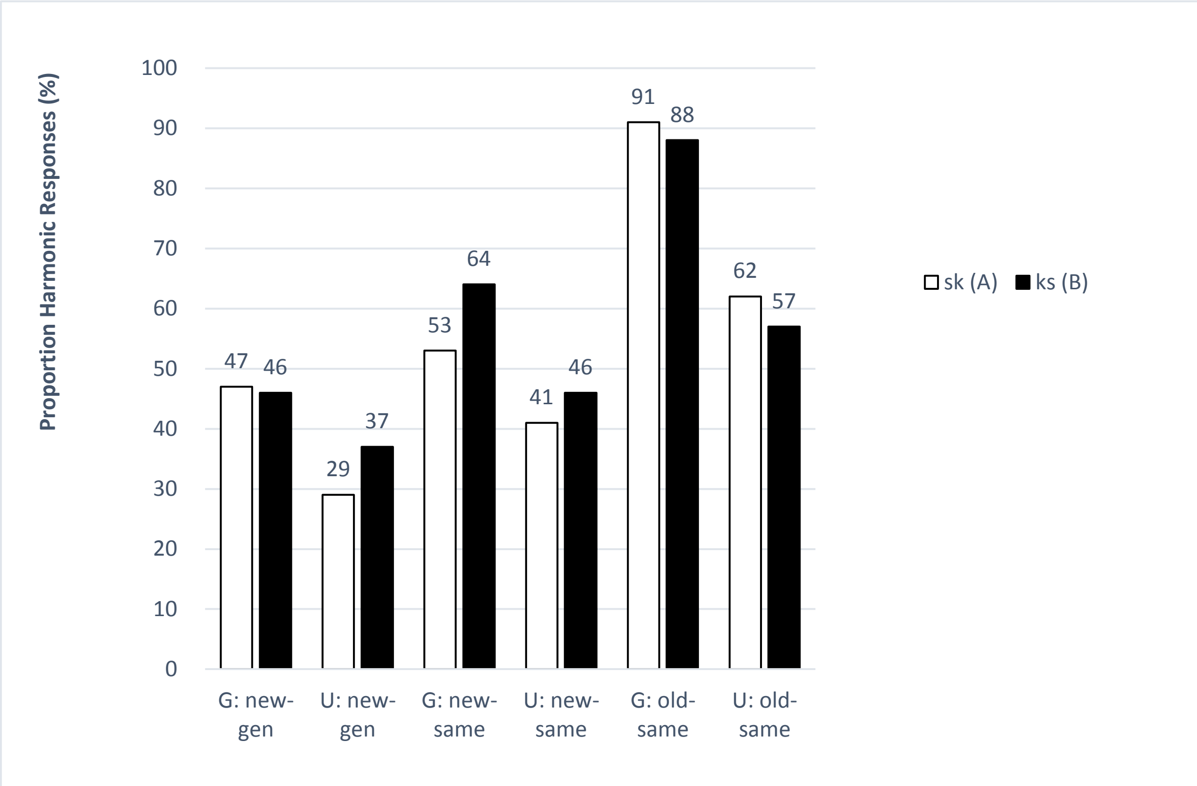


Figure 11. Response percentage for s→k and k→s

In brief, the descriptive statistics for both Patterns A and B failed to show that participants learned the new items. However, note that the overall accuracy is higher for grammatical items than ungrammatical items, suggesting that there was an effect. The next section presents inferential statistics, comparing Pattern A (s→k) with Pattern B (k→s) to see if there is any significant learning difference between the two.

4.3.2 Inferential statistics: comparison between (s→k) and (k→s)

In this section, I discuss the inferential statistics, since the raw response (i.e. descriptive statistics) might not be sensitive enough to test the hypothesis. Specifically, in order to test the hypothesis that there is a learning directional asymmetry between /s/ as a blocker and /k/ as a blocker, this study compared Pattern A with Pattern B. If the results support the hypothesis, then it is expected that participants in Pattern A should learn better than those in Pattern B.

For the statistical analysis, grammatical and ungrammatical items were analyzed separately by using mixed-effect logistic regression modeling (GLMM) (Pinheiro & Bates 2000) implemented in the R package lme4 (Bates & Maechler 2009). In order to test the hypotheses, I compared Pattern A with Pattern B to see if any directional learning was involved (i.e., whether the learnability for s→k would be significantly higher than that for k→s). The dependent variable was accuracy, and the independent variables were "Direction", "Grammar", and "Memory". **"Direction"** was 1 for Pattern A (s→k), and -1 for Pattern B (k→s)[39]. The logic behind this factor is that if the results support the hypothesis of implicational universals investigated in this study, then it is expected that participants in Pattern A should learn better than those in Pattern B, because the sonority for /s/ is higher than that of /k/. In terms of the implicational universal under consideration, the relationship between nasalization and sonority, if a more marked manner class blocks nasal harmony (e.g., fricatives), so will the less marked classes (e.g., stops). By this logic, the baseline for "Direction" should be Pattern B (k→s) instead of Pattern A (s→k). This factor is crucial, since it can directly test whether the experimental results support the hypotheses. A positive effect of "Direction" is thus expected. **"Grammar"** was -1 for test trials that were ungrammatical, 1 for test trials that were grammatical. Recall that the test items had three word types: old-same class, new-same class and new-generizable (new-gen). I coded them as O, NS and NGEN. These three word types were grouped into **"TypeName"**. Note that since I used dummy coding for "TypeName", the reference level of "TypeName" is "NGEN". When I report the main effect of "Direction", it actually refers to the effect of "Direction" when TypeName equals "NGEN", rather than all three word types. For example, a positive main effect of "Direction" means that the learning of new-gen items in Pattern A is learned better than the learning of new-gen items in Pattern B. A similar logic holds for the main effect of "Grammar". This treatment of dummy coding for "TypeName" or for "Grammar" is kept the same for the rest of this thesis.

Wilson (2006)'s logic, if participants learned the pattern, then the accuracy difference between old and new-same class should be not that obvious (i.e., not pure memorization). In addition, if participants generalized to a new consonant, then the accuracy difference between old and new-generalizable class should be not that obvious. This learning trend between old items and

[39] The coding is -1/1 (effect coding) rather than 0/1 (dummy coding) based on James Myers's suggestion, November, 2009.

new-same class items, and between old items and new-generalizable items can be examined by the **interaction of “Direction” and “TypeName”**.

The **interaction of “Grammar” and “TypeName”** is also examined. If there is memory effect, then it is expected to find that old grammatical items are easier to learn than new ungrammatical items.

The formula for both by-participant and by-participant-item analyses was used to carry out the analysis as in (15) and (16). ANOVA was done to compare which model, by-participant or by-participant-and-item analysis, better accounts for the results.

(15) Accuracy ~ Direction + Grammar + Direction*TypeName + Grammar*TypeName + (1|Participant)

(16) Accuracy ~ Direction + Grammar + Direction*TypeName + Grammar*TypeName (1|Participant) + (1|Item)

The ANOVA showed that the by-participant-and-item analysis was significant, suggesting that this model could better account for the results. The analysis is summarized in Table 23. As seen in Table 23, no main effect of “Direction” and no other interactions with “Direction” were found, implying that there were no directional learnability differences (of new-gen) between Pattern A and Pattern B. The main effect of “TypeNameO” was highly and positively significant, indicating that the learning of old test items was significantly higher than that of new-generizable items. The main effect of “TypeNameNS” was positively significant, indicating that the learning of new-same class items was significantly higher than that of new-generizable items.

The main effect of “Grammar” was marginally significant, indicating that the learning of grammatical test items was marginally better than that of ungrammatical ones. The interaction of “Grammar” and “TypeNameO” was positively significant, indicating that the learning of old & grammatical test items was significantly better than that of ungrammatical & new-generalizable items.

Table 23. Effects and interaction for Pattern A (s→k) and Pattern B (k→s) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	**-0.33494**	**0.12905**	**-2.596**	**0.00944 ****
Direction	0.03647	0.12896	0.283	0.77735
Grammar	**0.17830**	**0.09465**	**1.884**	**0.05958.**
TypeNameO	**1.62172**	**0.17888**	**9.066**	**< 2e-16 *****
TypeNameNS	**0.37535**	**0.17353**	**2.163**	**0.03054 ***
Direction: TypeNameO	0.09157	0.16600	0.552	0.58121
Direction: TypeNameNS	-0.19706	0.14645	-1.346	0.17843
Grammar:TypeNameO	**0.74307**	**0.14488**	**5.129**	**2.91e-07 *****
Grammar:TypeNameNS	0.13949	0.12523	1.114	0.26534

'***': p<.001, '**': p<.01, '*': p<.05, '.': p<.1

In short, the null results of "Direction" appear to provide no support for the implicational universal that /k/ is a more likely blocker than /s/ in nasal harmony.

4.4 Discussion: Patterns A and B

Recall that one of the current goals is to investigate whether any directional learning occurred between Pattern A and Pattern B. Specifically, if the overall accuracy for the scenario where participants learned /s/ as a blocker first and then were tested on /k/ as a blocker was significantly higher than the overall accuracy for the one where participants learned /k/ as a blocker first, and then were tested on /s/ as a blocker, then this could be used as evidence to support the implicational universal relationship between sonority and nasal harmony with opaque segments present in the ranking of nasalized segment hierarchy: *NASPLO >> *NASFRIC. Recall that Pattern A is expected to be generalizable, whereas there is no such prediction for Pattern B. If this is the case, I would expect to find a positive effect of "Direction". Nevertheless, the current statistical results showed no main effect of "Direction" or its interactions, implying that no directional relationship of implicational universals was found between Patterns A and B. This does not mean that participants did not learn patterns. The null results of "Direction" just means that there is no evidence to support my hypothesis of directional implicational universals. This seems to contradict the partial nasalized segment hierarchy proposed by Walker (2000): NASPLO >> *NASFRIC. However, the results can be reconciled with the hierarchy (see Table 8). Clements (1990) among others argues that languages can differ in whether obstruents are treated as a class or are more finely differentiated. If obstruents are a single class, then it follows that there would be no significant difference in results between Pattern A and Pattern B with respect to learning. Lin (2012) shows that training on /h/ as a blocker and then testing on /k/ was learned better than the other way around. But no such directional effect was found for s→k vs. k→s. In order to explain why this occurs, Lin (2012) hypothesized that /h/ might pattern as sonorant. If so, perhaps Min speakers only have a two-way opposition – obstruent/sonorant. Combined with the patterning of /s/ and /k/ together as obstruent, the results raise the possibility that the proposed universal implicational nasalized segment hierarchy might simply involve sonorants > obstruents, with finer gradations learned based on exposure to a language rather than innate. This is an important issue in phonological theory – just how much is built in, and what emerges through language acquisition?

Another possibility is that participants did not learn patterns due to potential design limitations that might have led to the null results of "Direction". This possibility is explored in Chapter 5.

4.4.1 Descriptive statistics: Patterns C and D

In this section I present the descriptive statistics for Patterns C and D. Recall that in Pattern C, during the exposure phase, participants were exposed to vowels (targets), glides (targets) and <u>stops</u> (/p/ as a blocker). During the test phase, <u>stops</u> (/k/ as a blocker) were introduced. Participants were expected to be able to generalize /p/ as a blocker to /k/ as a blocker. During the test phase, one auditory word was presented at a time. The participants were asked to judge whether each word was a possible plural form.

As with Patterns C and D, I divided the test items into three word types. The first word type is **old-same class** – these are the exposure items. The second word type is **new-same class**. These are items where the glide was new (/**y/** instead of /w/) or where the stop was new (/**t/** instead of the /p/). Finally, the third word type was distinguished, **new-generalizable**. In this case, the stop /k/ in Pattern C and the stop /p/ in Pattern D are new, and participants should be able to generalize from an old consonant to a new consonant. Note that in Patterns C and D, /p, t, k/ do not involve any change in the sonority hierarchy, and are purely design internal. I arbitrarily treated /k/ as new-generalizable in Pattern C because /k/ is the new segment that only appeared in the new-gen condition, although technically, /k/ has the same phonological class as /t/. /t/ is taken as the neutral one because it is the one that appeared in the new-same condition of the test phase in both Patterns C and D while /k/ and /p/ switched places in Patterns C and D.

As with Patterns C and D, items are divided into two major groups, Memory: Old vs. New, and Grammaticality: Grammatical vs. Ungrammatical. The memory levels "Old" and "New" were used to distinguish exposure-phase items from test-phase new items.

Note that a response was treated as 'correct' if participants pressed "yes" (i.e., this word could be a possible plural form) to grammatical/old items (e.g., said "yes" to /ãpawa/). A response was treated as 'correct' if participants pressed "no" (i.e., this word could not be a possible plural form) to ungrammatical/old items (e.g., said "no" to */ãp̃ãwa/). If a participant responded simply based on whether they heard the segments before in the exposure phase, then they would always chose "yes" for both old-grammatical and old-ungrammatical items, causing the former response rate (old grammatical) to be close to 100%, and the latter one (old ungrammatical) to be close to 0%. However, if participants learned a word type (e.g., "old-same"), then it is expected that the response percentage for "old-same" for both grammatical and ungrammatical items would be over 50%.

In order to get a better sense of how participants responded to the grammatical and ungrammatical items individually, the raw percentages of correct responses for each class are given in Table 24.

Table 24. The response rates for test items for Pattern C (p→k)

p→k	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**k**iyi/ /ãỹĩ**k**i/ /ã**k**i**k**i/	/ãtiyi/ /ãỹĩti/ /ãtiti/ /ãỹĩỹĩ/	/ã**p**awa/ /ãw̃ã**p**a/ /ã**p**a**p**a/ /ãw̃ãw̃ã/		/ã**k̃**ĩyi/ /ãyi**k**i/ /ã**k̃**ĩ**k̃**ĩ/	/ãt̃ĩyi/ /ãyiti/ /ãt̃ĩt̃ĩ/ /ãyiyi/	/ã**p̃**ãwa/ /ãwa**p**a/ /ã**p̃**ã**p̃**ã/ /ãwawa/	
mean	43%	55%	95%	66%	58%	48%	41%	48%

As shown in Table 24, for Pattern C (p→k), participants did not even learn old items, since the raw rates for grammatical and ungrammatical items for the old items are not both above 50%, suggesting that they simply guessed. Specifically, they accepted old grammatical items (95%), and randomly guessed when encountering ungrammatical items (41%).

Similarly, for Pattern D (k→p), according to Table 25, the raw rates for grammatical and ungrammatical items for the old items are not both above 50%, suggesting that participants did not learn old items. They just accepted old grammatical items (94%), and randomly guessed when make a judgment on ungrammatical items (50%).

Table 25. The response rates for test items for Pattern D (k→p)

k→p	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**p**iyi/ /ãỹĩ**p**i/ /ã**p**i**p**i/	/ãtiyi/ /ãỹĩti/ /ãtiti/ /ãỹĩỹĩ/	/ã**k**awa/ /ãw̃ã**k**a/ /ã**k**a**k**a/ /ãw̃ãw̃ã/		/ã**p̃**ĩyi/ /ãyi**p**i/ /ã**p̃**ĩ**p̃**ĩ/	/ãt̃ĩyi/ /ãyiti/ /ãt̃ĩt̃ĩ/ /ãyiyi/	/ã**k̃**ãwa/ /ãwa**k**a/ /ã**k̃**ã**k̃**ã/ /ãwawa/	
mean	53%	55%	94%	69%	27%	45%	50%	44%

In brief, it seems that participants in both Patterns C and D did not learn any word types (i.e., "old-same class", "new-same class", "new-generalizable", even for the old items.

4.4.2 Inferential statistics: comparison between (p→k) and (k→p)

The coding of factors is similar to Patterns A and B. **"Direction"**[40] was -1 for Pattern C (p→k), and 1 for Pattern D (k→p) arbitrarily since both of /p/ and /k/ are stops (same sonority). Participants are expected to be able to generalize from an old segment /p/ as a blocker to /k/ and vice versa. Crucially, no effect of "Direction" is expected to be found.

Other factors such as "Grammar" and "TypeName" are kept the same (cf. Section 4.3.2). "Direction". **"Grammar"** was -1 for test trials that were ungrammatical, 1 for test trials that were grammatical. **"TypeName"** included three word types of test items: old-same class, new-same class and new-generizable (new-gen). I coded them as O, NS and NGEN.

The formula for both by-participant and by-participant-item analyses was used to carry out the analysis as in (17) and (18). ANOVA was done to compare which model, by-participant or by-participant-and-item analysis, better accounts for the results.

[40] For clarity, I bolded factors when I define how I coded the factors. I follow the same convention for the rest of the thesis.

(17) Accuracy ~ Direction + Grammar + Direction*TypeName + Grammar*TypeName + (1|Participant)

(18) Accuracy ~ Direction + Grammar + Direction*TypeName + Grammar*TypeName + (1|Participant) + (1|Item)

The ANOVA showed that the by-participant-and-item analysis was significant, suggesting that this model could better account for the results. The analysis is summarized in Table 26. As shown in Table 26, no main effect of "Direction" was found, implying that there were no directional learnability differences (of new-gen) between Pattern C and Pattern D.

The main effect of "TypeNameO" suggests that old items were learned better than new-generalizable items; however, this learning trend differs between Pattern D (1) and Pattern C (-1) (albeit the interaction is only positively marginally significant).

I suspect that a positive main effect of "TypeNameO" is because the accuracy of grammatical old items (Pattern C: 95%, Pattern D: 94%) was much higher than the accuracy of grammatical new-generalizable items (Pattern C: 43%, Pattern D: 53%), not because participants 'learned' the old items better than new-generalizable items.

The interaction of "Grammar" and "TypeNameO" was highly and positively significant, indicating that the learning of old & grammatical test items was significantly better than that of ungrammatical & new-generalizable items.

Table 26. Effects and interaction for Pattern C: p→k and Pattern D: k→p (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.10381	0.12905	-0.804	0.4212
Direction	-0.12154	0.12905	-0.942	0.3463
Grammar	0.03483	0.09332	0.373	0.7089
TypeNameO	**1.45452**	**0.16424**	**8.856**	**<2e-16 *** **
TypeNameNS	0.13765	0.12349	1.115	0.2650
Direction: TypeNameO	**0.26300**	**0.13971**	**1.883**	**0.0598 .**
Direction: TypeNameNS	0.08989	0.12349	0.728	0.4667
Grammar:TypeNameO	**1.50979**	**0.16426**	**9.191**	**<2e-16 *** **
Grammar:TypeNameNS	0.14095	0.12348	1.141	0.2537

'***': p<.001, '**': p<.01, '*': p<.05, '.': p<.1

4.5 Discussion: Patterns C and D

Recall that one of the current goals is to investigate whether any directional learning of implicational universals occurred between Pattern C and Pattern D. Since no sonority distance is involved for /p/ and /k/, no directional differences are expected to be found.

The inferential statistics did not show a main effect of "Direction", which seems to support my hypothesis. However, the descriptive statistics showed that participants for both Patterns C and D did not learn any word types (i.e., old-same class, new-same class, new-generalizable).

Combining the descriptive and inferential statistics, the current results suggest that participants did not learn patterns. As discussed in Section 4.4, this finding further suggests the possibility that participants did not learn patterns due to potential design limitations. The next section discusses design limitations.

4.6 Design limitations: Patterns A, B, C, D

The results of the experiments presented in this chapter suggest that participants did not learn the patterns under study. In this section I address whether the current results can be used as evidence to argue for or against the hypothesis of the implicational universal that if a more marked blocker class blocks harmony, so do the less marked blocker classes. As discussed in Section 4.2.5, I assumed that new-same class consonants would be treated the same way as old consonants because new-same consonants have the same phonological class as old-same class consonants (/s/ with /h/, /p/ with /t/, /w/ with /y/). This assumption seems to be confounded with one of my research questions, namely whether participants have a concept of natural classes to make a generalization to new-gen consonants, making the results hard to interpret as to whether the results support the hypothesis that, for instance, /s/ as a blocker→/k/ (Pattern A) was expected to be generalizable (i.e., generalizable = whether participants would generalize from /s/ as a blocker to /k/ and thus treat /k/ as a blocker), but /k/ as a blocker→/s/ (Pattern B) was not. For Pattern A, participants were exposed to a singular form /awa**s**a/ and its plural counterpart /ãw̃ã**s**a/; then tested with /k/ to see if they generalized (/ãỹãka/, not */ãyaka/). Similarly, for Pattern B, participants were exposed to /awa**k**a/ and /ãw̃ãka/, and were tested with /s/ replacing /k/ (/ãỹã**s**a/, not */ãya**s**a/).

This confound shows that there might be too many new consonants for participants to draw from for Experiment 1. The following recaps what Pattern A participants were exposed to and tested on, and after that I discuss why I think there are too many new consonants.

In order to encourage participants to generalize to new sounds, participants were given a hint that they could try to group old sounds with new sounds. If an old sound was grouped with a new sound, then they would share a similar function in this made-up language.

Recall that the participants were told that they would learn a made-up language that contained tri-syllabic words and they had to search for a singular/plural rule. In the exposure phase, as shown in Table 13 (Pattern A) in Section 4.2.2.4, they heard singular and plural forms in pairs. For instance, they heard a word like /akawa/ accompanied by a picture of one item, and then they heard a word like /ãkawa/ accompanied by a picture of two of the same item. They were also informed that only plural forms would change, and singular forms would be treated as base forms. During to the test phase, they were told that they would be tested on how well they could learn the singular/plural rule based on a grammaticality judgment test, and that they needed to make a judgment even though certain sounds did not occur in the exposure phase.

During the test phase, participants were asked to judge whether each item they heard was a possible plural form generated by the rules of plurality for the language they were trained on. In order to see the design more clearly, I put two patterns together by factors, as shown in Table 27.

Test items were composed of grammatical and ungrammatical[41] items. Items that appeared in the exposure are "old-same class" items (16 grammatical and 16 ungrammatical items). I do not present old-same class items here since they are the same as exposure items.

[41] Note that I use "ungrammatical" to refer to other variations compared to grammatical items (e.g., for Pattern A, new-same segment-/**y/** is a blocker, and new-generalizable segment-/**k/** is a target; new-same segment-/**y**/ is a blocker, and new-same segment-/**h/** is a target. Those ungrammatical items should not be predicted because of the hierarchy that a segment lower on the hierarchy like stops should be a blocker, while a segment higher on the hierarchy like glides should be a target.

Table 27. Pattern (A): positive (grammatical) and negative (ungrammatical) test stimuli design (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Trigger+/blocker=k/+target (n=4)		**Trigger+/blocker=h/+target** (n=4)	
New-generalizable		**New-same class**	
Grammatical	Ungrammatical	Grammatical	Ungrammatical
(1) /ãkiyi/	(29) /ãk̃ĩyi/	(13) /ãhiyi/	(41) /ãh̃ĩyi/
(2) /ãkɔyɔ/	(30) /ãk̃ɔ̃yɔ/	(14) /ãhɔyɔ/	(42) /ãh̃ɔ̃yɔ/
(3) /ãkiyɔ/	(31) /ãk̃ĩyɔ/	(15) /ãhiyɔ/	(43) /ãh̃ĩyɔ/
(4) /ãkɔyi/	(32) /ãk̃ɔ̃yi/	(16) /ãhɔyi/	(44) /ãh̃ɔ̃yi/
Trigger+target+/blocker=k/ (n=4)		**Trigger+target+/blocker=h/** (n=4)	
New-generalizable		**New-same class**	
Grammatical	Ungrammatical	Grammatical	Ungrammatical
(5) /ãỹĩki/	(33) /ãyiki/	(17) /ãỹĩhi/	(45) /ãyihi/
(6) /ãỹɔ̃kɔ/	(34) /ãyɔkɔ/	(18) /ãỹɔ̃hɔ/	(46) /ãyɔhɔ/
(7) /ãỹĩkɔ/	(35) /ãyikɔ/	(19) /ãỹĩhɔ/	(47) /ãyihi/
(8) /ãỹɔ̃ki/	(36) /ãyɔki/	(20) /ãỹɔ̃hi/	(48) /ãyihɔ/
Trigger+/blocker=k/+/blocker=k/ (n=4)		**Trigger+/blocker=h/+/blocker=h/** (n=4)	
New-generalizable		**New-same class**	
Grammatical	Ungrammatical	Grammatical	Ungrammatical
(9) /ãkiki/	(37) /ãk̃ĩk̃ĩ/	(21) /ãhihi/	(49) /ãh̃ĩh̃ĩ/
(10) /ãkɔkɔ/	(38) /ãk̃ɔ̃k̃ɔ̃/	(22) /ãhɔhɔ/	(50) /ãh̃ɔ̃h̃ɔ̃/
(11) /ãkikɔ/	(39) /ãk̃ĩk̃ɔ̃/	(23) /ãhihɔ/	(51) /ãh̃ĩh̃ɔ̃/
(12) /ãkɔki/	(40) /ãk̃ɔ̃k̃ĩ/	(24) /ãhɔhi/	(52) /ãh̃ɔ̃h̃ĩ/
		Trigger+target+target (n=4)	
		New-same class	
		Grammatical	Ungrammatical
		(25) /ãỹĩỹĩ/	(53) /ãyiyi/
		(26) /ãỹɔ̃ỹɔ̃/	(54) /ãyɔyɔ/
		(27) /ãỹĩỹɔ̃/	(55) /ãyiyɔ/
		(28) /ãỹɔ̃ỹĩ/	(56) /ãyɔyi/

Among “new” items, there are two categories, “same class” and “generalizable”. “n” represents the number of items. There are 28 positive/grammatical items (item number: 1–28), meaning that the items conform to the rules of plurality in Pattern A. However, 28 negative/ungrammatical items (item number 29–56) violate the rules of plurality in Pattern A. Take items (1) vs. (29), and items (9) vs. (37) for instance. In both (1) and (9), vowels and /y/ are targets, and /k/ is a blocker, whereas in (29) and (37), the relationship between /y/ and /k/ is opposite: /y/ was treated as a blocker, and /k/ is a target. /k/ is a new “test” segment. Vowels /i, ɔ/, /y/, and /h/ are counted as new-same class (/y/ is the same phonological class as an old target /w/ and /h/ is the same phonological class as an old blocker /s/).

Even though I assumed that /k/ is the only “test” segment (s→k), there are too many “new” segments (three new consonants /k, h, p/, and two new vowels /i, ɔ/), and it is hard to interpret (1) whether participants failed to generalize from “one” old segment like /s/ to “one” new segment like /k/, as expected by the hypothesis of my study, (2) whether they failed to do so due to the task being too difficult to learn with too many new segments (i.e., floor effect), or (3) whether they did generalize from one old segment to one new segment in some way, but just had difficulty with certain new segments. Specifically, I assumed that participants would know that /s/ and /h/ belong to the same phonological class (fricatives), and that /w/ and /y/ belong to the same phonological class (glides). What if participants did not group /s/ and /h/ together or /w/ and /y/ together, so that they failed to make a generalization to the new phonological class based on the relationship between /s/ and /k/)?

The next chapter presents Experiment 2 with a modified design to address this design difficulty.

Chapter 5 Experiment 2: nasalized segment hierarchy vs. natural classes

In this chapter, I discuss changes in design in response to the limitation of Experiment 1 that there are too many new consonants, and present the results and discussion of Experiment 2 with the modified design.

5.1 Motivation for the follow-up experiments: Experiment 2

In this section, I first discuss six changes that I made in conducting Experiment 2, where I tested the relationship between /w/ and /s/. This is discussed in the methods section (see Section 5.2) in detail, but it is helpful to point out the changes in design here in order to rationalize why Experiment 2 is needed. As discussed in Chapter 4, the complicated design of Experiment 1 made it difficult to determine whether one grammar is learned better than the other. In Experiment 2, I simplified the stimuli. In addition, I analyzed the data taking individual learner type into account.

5.1.1 Changes in the design

First, in the experiments reported in Chapter 4, it was difficult to sort out if directionality mattered with /s/ and /k/. I thus decided to use two segments that clearly differ in sonority: /s/ is a target in the exposure phase and /w/ is a potential target in the test phase (referred to as Pattern 1: S(k)→W(k)[42]), and /w/ is a target in the exposure phase and /s/ is a potential target in the test phase (referred to as Pattern 2: W(k)→S(k)). Note that I use upper case letters to indicate that those segments serve as targets, differentiating them from lower case letters to indicate blockers. The segments outside the parentheses are the targets (upper case) or blockers (lower case) under investigation. The segments inside the parentheses are blockers (lower case) or targets (upper case) that were presented in both exposure and test phases. → means direction: exposure items→test items (e.g., Pattern A: exposure /s/ as target, /k/ as blocker; test /w/ as potential target, /k/ as blocker, S(k)→W(k)). S(k)→W(k) is abbreviated as SW(k). I chose /w/ and /s/ because the distance between them on the sonority hierarchy is relatively large compared to that between /s/ and /k/, and the predictions are clear. That is, /w/ and /s/ are far apart on the hierarchy, with one sonorant and one obstruent, so if there is a hierarchy, this pair might be quite likely to reveal it. /k/ was a blocker in both exposure and test phases in Pattern 1 and Pattern 2.

[42] In order to differentiate Experiment 2 from Experiment 1, I now use numbers for the patterns instead of letters.

Second, the follow up experiments only tested one new consonant per pattern. If I find any significant learning difference between Patterns 1 and 2, then I could be sure that the difference is due to generalizability.

Third, as mentioned in Section 3.2, it is easier to generalize to new vowels than to new consonants (Finley 2011b), so I would expect that applying rules to new vowels would be easier than applying them to new consonants. Based on this assumption, I modified three word-type conditions, namely old-same class, new-same class, and new-generalizable (abbreviated as new-gen). Specifically, one further difference was present between old-same class and new-same class items, namely old-same class items had the vowel /a/ (e.g., S(k)→W(k): /ãs̃ãka/), while for new-same class items, a set of them had a different vowel. Thus, there was a different vowel condition (new-same class) (e.g., S(k)→W(k): /ãs̃ẽka/) as well as a different consonant condition (new-gen) (S(k)→W(k): /ãw̃ãka/). If the learning of the new-same class items (new vowels) is better than that of the new-gen items (new consonants), then this would support the argument of Finley (2011b) that generalizing to new vowels is easier than to new consonants.

Fourth, in the previous design during the test phase participants were asked to judge a single test item per time. The current design uses paired items. During the test participants heard a singular form first (e.g., S(k)→W(k): /asaka/), and then were asked to choose the plural form which obeys the plural rule from two items (e.g., S(k)→W(k): /ãs̃ãka/ vs. */ãsaka/). A singular form serves as a base (non-nasalized form), which might aid participants in picking a plural form, because the participants of my pilot reported that they were confused about whether a sound was nasalized or not because of the lack of a singular form during the test phase. This design modification forces participants to make a judgment even if the sound is new.

Fifth, in the previous design, each test item was presented once. In the current design, 14 test pairs were repeated 4 times, yielding 56 pairs in total. The order was counterbalanced. If participants really made a judgment on a certain pair instead of guessing, I expected that they should get at least 3 out of 4 right (75%)[43].

Sixth, the repetition of exposure items was reduced from four times to three times based on the report of participants in the pilots that four repetitions made them lose their focus and feel too tired to find a rule conditioning the plural change.

Table 28 summarizes the predictions for Patterns 1 and 2. Participants in Pattern 1 are expected to generalize from an old target to a new consonant, treating the new consonant as a target, while there is no such prediction for Pattern 2.

Table 28. Predictions of Patterns 1 and 2

	exposure	test	prediction
Pattern 1 S(k)→W(k)	more sonorant: target /**s**/ less sonorant: blocker /k/	new segment /**w**/: more sonorant than target	new segment is a target
Pattern 2 W(k)→S(k)	more sonorant: target /**w**/ less sonorant: blocker /k/	new segment /**s**/: less sonorant than target	no prediction

5.1.2 Issue of determining which grammar is learned better: learner types

Another issue of artificial grammar learning is as follows. Presumably, we have two artificial grammars, A and B. How do we determine whether learners of Grammar A generalized better than learners of Grammar B? For example, for individual data, in Grammar A, suppose we find that almost every participant got over 60% correct for old items, but the percentage of correct responses for the new phonological class was under 50% correct. In Grammar B, the number of participants who got over 60% correct for the old items was lower than Grammar A, but a few participants got over 60% correct for the new phonological class. What does this mean? Is it valid to say that the

[43] Four completely random coin flips (of a fair coin) will on average come up with 3 or more heads ("successes") 31.25% of the time. A caveat is that in many instances seeing 3-out-of-4 correct results for a particular pair could very easily reflect pure chance (random guessing) on the participant's part. The design of Experiment 1 only used the same item once, which means a 1-out-of-1 correct results for a particular item will on average come up with 1 head ("success") 50% of the time. The reduction of pure guessing from 50% to 31.25% is still an improvement. Ideally, it would be better to use 10 repetitions instead of 4. However, given that Experiment 2 only included 14 test pairs, 4 repetitions had already reached the maximal number of repetitions my participants could accept. My participants reported that they felt like the same items kept appearing, and started to wonder why the experiment had not come to an end.

learning of Grammar B was better than that of Grammar A because a few people for Grammar B could generalize from an old phonological class to a new phonological class? What is a better way to determine which grammar is learned better? In Experiment 2, I introduce individual learner types to determine learnability in Section 5.4.

5.1.3 Summary

In brief, Experiment 2 concerns glides and fricatives: /w/ and /s/. As discussed in Section 4.4, the findings of Experiment 1 lead to two possible interpretations. One interpretation is that there is no evidence to support my hypothesis of directional implicational universals. The other interpretation is that participants did not learn the patterns and failed to extend them to new segment types. Looking at individual data could effectively gauge whether each individual learned the pattern and what learner type each individual belongs to. Especially when we get null results (no significant difference between two groups), it is hard to say whether (1) the results support the hypothesis of implicational universals that Patterns A-D should be equally learnable since their new test segments are obstruents or (2) the design was too hard to learn.

One primary goal of this chapter is to examine whether both a nasalized segment hierarchy and natural classes (manner) play a role in artificial grammar learning. In Section 5.3, the grouped results of Patterns 1 and 2 are presented. The definition of individual learner types is presented in Section 5.4, and the findings of individual leaner types in Patterns 1 and 2 are presented in Section 5.5. General discussion in Section 5.6 identifies four possible ways to interpret the findings of Patterns 1 and 2, namely (1) reference to a nasalized segment hierarchy, (2) reference to natural classes, (3) floor effect, and (4) game strategy. The findings of Patterns 1 and 2 alone are not sufficient to draw conclusions. Therefore, further follow-up experiments, Patterns 3 and 4, are described in Section 5.7. In Sections 5.9 and 5.10, the results of Patterns 3 and 4 are presented. In Sections 5.11 I compare Patterns 1-4, and in Section 5.12 I discuss the implications.

5.2 Methods

5.2.1 Design

Experiment 2 has three tasks: exposure phase, test phase, and post-test. See Section 4.2.2 for the overall design – forced choice method.

In general, participants were exposed to /asaka/ ~ /ãs̃ãka/ (singular ~ plural); then tested with /w/ to see if they generalized (Pattern 1: S(k)→W(k)). Similarly, another group of participants was exposed to /awaka/ ~ /ãw̃ãka/ and tested with /s/ replacing /w/ (Pattern 2: W(k)→S(k)). In both patterns, /k/ was always a blocker throughout the exposure and test. V_1 was always /a/ and V_2 and V_3 varied between /a/ and /i/.

I present the stimuli that participants heard during the exposure and test phases of Experiment 2 in Table 29. During the exposure phase, participants heard a singular form (e.g. S(k)→W(k): /aka**s**a/), and then a corresponding plural form (e.g., S(k)→W(k): /ãka**s**a/). Nasality on the initial /a/ served as a plural marker, triggering nasal spreading rightward.

Table 29. Current stimuli design (exposure)

	Pattern 1: **S(k)→W(k)** (SW(k))		Pattern 2: **W(k)→S(k)** (WS(k))	
	singular	plural	singular	plural
grammatical	/aka**s**a/	/ãka**s**a/	/aka**w**a/	/ãka**w**a/
	/a**s**aka/	/ã**s̃**ãka/	/a**w**aka/	/ã**w̃**ãka/
	/akaka/	/ãkaka/	/akaka/	/ãkaka/
	/a**s**a**s**a/	/ã**s̃**ã**s̃**ã/	/a**w**a**w**a/	/ã**w̃**ã**w̃**ã/
	/aka**s**i/	/ãka**s**i/	/aka**w**i/	/ãka**w**i/
	/a**s**aki/	/ã**s̃**ãki/	/a**w**aki/	/ã**w̃**ãki/
	/akaki/	/ãkaki/	/akaki/	/ãkaki/
	/a**s**a**s**i/	/ã**s̃**ã**s̃**ĩ/	/a**w**a**w**i/	/ã**w̃**ã**w̃**ĩ/
	/aki**s**a/	/ãki**s**a/	/aki**w**a/	/ãki**w**a/
	/a**s**ika/	/ã**s̃**ĩka/	/a**w**ika/	/ã**w̃**ĩka/
	/akika/	/ãkika/	/akika/	/ãkika/
	/a**s**i**s**a/	/ã**s̃**ĩ**s̃**ã/	/a**w**i**w**a/	/ã**w̃**ĩ**w̃**ã/
	/aki**s**i/	/ãki**s**i/	/aki**w**i/	/ãki**w**i/
	/a**s**iki/	/ã**s̃**ĩki/	/a**w**iki/	/ã**w̃**ĩki/
	/akiki/	/ãkiki/	/akiki/	/ãkiki/
	/a**s**i**s**i/	/ã**s̃**ĩ**s̃**ĩ/	/a**w**i**w**i/	/ã**w̃**ĩ**w̃**ĩ/

The logical combinations used in the exposure are four (V.CV.CV) sequences in (19):

(19) Four combinations of items in the exposure phase (singular ~ plural)

(a) trigger + blocker + vowel + <u>target</u> + vowel /akasa/ ~ /ãkasa/

(b) trigger + <u>target</u> + vowel + blocker + vowel /asaka/ ~ /ãs̃ãka/

(c) trigger + blocker + vowel + blocker + vowel /akaka/ ~ /ãkaka/

(d) trigger + <u>target</u> + vowel + <u>target</u> + vowel /asasa/ ~ /ãs̃ãs̃ã/

16 exposure (paired) items for Patterns 1 and 2 were generated separately. The exposure items in the exposure phase for (19a) and (19c) were repeated three times in three different blocks, while (19b) and (19d) were repeated six times in three different blocks in an attempt to increase participants' exposure to combinations with changes in nasalization.

The test items are presented in Table 30. Each trial was aurally presented to participants. During the test phase, they were shown a singular form first (e.g., S(k)→W(k): /asaka/), and then were asked to choose the plural form which obeys the plural rule from two items (e.g., S(k)→W(k): (1) /ãs̃ãka/ vs. *(1') /ãsaka/). If they thought the first test item was correct, they pressed "1", if they thought the second one was correct, then they pressed "2". New-same class items were the same as the old-same class except that one or both of the target vowels (V_2 and V_3) were /e/. As for the new-gen items, a new consonant was introduced. The vowel was always /a/, either or both of C_2 and C_3 was a new segment (/w/ for Pattern 1 and /s/ for Pattern 2) while the other C was /k/ or /s/ (for Pattern 1) and /k/ or /w/ (for Pattern 2).

Table 30. Stimuli design (test items)

	Pattern 1: **S**(k)→**W**(k) (SW(k))			Pattern 2: **W**(k)→**S**(k) (WS(k))		
	new		old	new		old
	gen (n=10, pair=5)	same class (n=10, pair=5)	same class (n=8, pair=4)	gen (n=10, pair=5)	same class (n=10, pair=5)	same class (n=8, pair=4)
grammatical (n=44)	(10) /ã**w̃**ãka/ (11) /ãka**w**a/ (12) /ã**w̃**ãs̃ã/ (13) /ã**w̃**ã**w̃**ã/ (14) /ãs̃ã**w̃**ã/	(5) /ã**s̃**ẽka/ (6) /ãka**s**e/ (7) /ã**s̃**ãke/ (8) /ã**s̃**ẽ**s̃**ẽ/ (9) /ãkesa/	(1) /ã**s̃**ãka/ (2) /ãka**s**a/ (3) /ã**s̃**ã**s̃**ã/ (4) /ãkaka/	(10) /ã**s̃**ãka/ (11) /ãka**s**a/ (12) /ã**s̃**ã**w̃**ã/ (13) /ã**s̃**ã**s̃**ã/ (14) /ã**w̃**ã**s̃**ã/	(5) /ã**w̃**ẽka/ (6) /ãka**w**e/ (7) /ã**w̃**ãke/ (8) /ã**w̃**ẽ**w̃**ẽ/ (9) /ãkewa/	(1) /ã**w̃**ãka/ (2) /ãka**w**a/ (3) /ã**w̃**ã**w̃**ã/ (4) /ãkaka/
ungrammatical (n=44)	(10’) /ã**w**aka/ (11’) /ãk̃ã**w̃**ã/ (12’) /ã**w**asa/ (13’) /ã**w**a**w**a/ (14’) /ã**s̃ã**wa/	(5’) /ã**s**eka/ (6’) /ãk̃ã**s**e/ (7’) /ã**s**ake/ (8’) /ã**s**e**s**e/ (9’) /ãk̃ẽsa/	(1’) /ã**s**aka/ (2’) /ãk̃ã**s**a/ (3’) /ã**s**a**s**a/ (4’) /ãk̃ãk̃ã/	(10’) /ã**s**aka/ (11’) /ãk̃ã**s̃**ã/ (12’) /ã**s**awa/ (13’) /ã**s**a**s**a/ (14’) /ã**w̃**ã**s**a/	(5’) /ã**w**eka/ (6’) /ãk̃ã**w**e/ (7’) /ã**w**ake/ (8’) /ã**w**e**w**e/ (9’) /ãk̃ẽ**w**a/	(1’) /ã**w**aka/ (2’) /ãk̃ã**w**a/ (3’) /ã**w**a**w**a/ (4’) /ãk̃ãk̃ã/

(gen=generalizable, n=number)

Each pair consisted of grammatical (conforming to the pattern) and ungrammatical items. Grammaticality is based on the implicational nasalized segment hierarchy hypothesis. For example, for Pattern 1: S(k)→W(k), a participant would hear /awaka/, a singular form, first, and then a pair of choices for plural forms (e.g. (10) /ã**w̃**ãka/ (grammatical) vs. (10’) /ã**w**aka/ (ungrammatical)). The correct answer was “1”. The number and order between the grammatical and the ungrammatical choices were balanced. In half of the items, the grammatical choice was presented first while for the other half, the ungrammatical choice was presented first.

Consider the S(k)→W(k) pattern (see Table 30). Participants are trained on /s/ as a target and tested on whether they treat /w/ as a target. If participants treat /w/ as a target, then they will favor (12) /ãw̃ãs̃ã/ over (12’) */ãwasa/ and favor (14) /ãs̃ãw̃ã/ over (14’) */ãs̃ãwa/. However, if they treat /w/ as a blocker, then their preference will be the other way around (favoring (12’) */ãwasa/ over (12) /ãw̃ãs̃ã/ and favoring (14’) */ãs̃ãwa/ over (14) /ãs̃ãw̃ã/). Because the test involves a pairwise forced choice, participants have the freedom to determine whether they treat a new segment /w/ as a target or as a blocker. If there is a bias toward /w/ as a target, then the hypothesis that participants are able to generalize from /s/ as a target to /w/ as a target would receive support.

I am also interested to see if participants use any strategies to "guess". In items (11) and (11') in Table 30, if participants think that segments found in the exposure phase should always be nasalized, then I would expect that they would nasalize an old segment /k/, even though it was not nasalized in the exposure phase (favoring (11') */ãk̃ã**w̃**ã/ over (11) /ãka**w**a/ in Pattern 1).

I use "ungrammatical" to refer to other variations compared to grammatical items, assuming the nasalized segment hierarchy hypothesis. For old and new-same class items, the ungrammatical counterparts are just the reverse relationship between targets and blockers. That is, /w/ in Pattern 1 and /s/ in Pattern 2 were blockers, and /k/ was a target. But different from Experiment 1, new-gen was designed in a way that could test different strategies (e.g., in Pattern 2 /k/ is nasalized, and /s/ is nasalized like (11') */ãk̃ãs̃ã/; /w/ is a target, and /s/ is a blocker like (14') */ãw̃ãsa/); /s/ is a blocker and /k/ is a blocker like (10') */ãsaka/. Since I did not have a specific prediction about whether in Pattern 2 the new segment /s/ would be a target or a blocker, technically I could not refer (14') */ãw̃ãsa/ and (10') */ãsaka/ as "ungrammatical", since both of them are "possible" patterning in nasal harmony. However, I still treat (14') */ãw̃ãsa/ and (10') */ãsaka/ as "ungrammatical" based on the direction of learning: W(k)→S(k).

5.2.2 Stimuli

The monosyllabic stimuli for Experiment 2 are the stimuli recorded for Experiment 1 (see Section 4.2.2.2) by the same male native speaker of Min. The concatenation of tri-syllabic words was done to fit the current design.

5.2.3 Procedure

Different from Experiment 1, during the test phase in Experiment 2, participants were presented with a singular form followed by two potential plural forms. They were told that they needed to select one of the two plural forms where the singular form serves as its base. Participants would choose '1' if they considered that the first plural was a possible plural form and '2' if they considered that the second plural was correct. Four of the grammar-conforming stimuli were identical to those in the exposure phase, and 10 of them were new, all of which were paired with

14 ungrammatical items. E-prime randomly ordered these pairs. Participants who consistently chose the same answer (i.e., '1' throughout or '2' throughout) during the test phase were eliminated because that means that they were not trying to make a judgment.

5.2.4 Participants

The participants were bilingual in Min and Mandarin. Two groups of participants were recruited (Patterns 1 and 2). All reported normal hearing, and all distinguished nasalized monosyllables from their oral counterparts at a rate of 100%. All had early childhood Min exposure. Each participant recruited from National Sun Yat-sen University received 150 NTD for participating in the experiment, which lasted about 45-50 minutes.

Pattern 1: S(k)→W(k)

Ten college students were recruited from National Sun Yat-sen University in Taiwan. All had studied a foreign language (English 10, Japanese 1). The average age of the participants was 21.9 (SD = 2.8). 8 participants were males and 2 participants were females.

Pattern 2: W(k)→S(k)

Ten college students were recruited from National Sun Yat-sen University in Taiwan. All had studied a foreign language (English 10, Japanese 1, French 1). The average age of the participants was 21.1 (SD = 2.2). 8 participants were males and 2 participants were females.

5.3 Grouped statistics: Patterns 1 and 2

If participants make a judgment based on the nasalized segment hierarchy, Pattern 1 is expected to be learned better than Pattern 2. The response percentages of the test items for Pattern 1: S(k)→W(k) and Pattern 2: W(k)→S(k) can be found in Figure 12. Note that the accuracy for each item is based on the implicational nasalized segment hierarchy hypothesis[44]. Patterns 1 and 2 participants did well on old test items and new test items containing a new vowel[45]. Pattern 2

[44] The implicational nasalized segment hierarchy hypothesis is formed the initial motivation for this work, but natural classes turned out to be important as well.
[45] As mentioned in Section 1.2 and Section 5.1.1, it is easier to generalize to new vowels than to new consonants, so I would expect that applying rules to new vowels would be easier than to new consonants.

participants performed poorly on the new-gen test items, while Pattern 1 participants were able to generalize to new-gen test items.

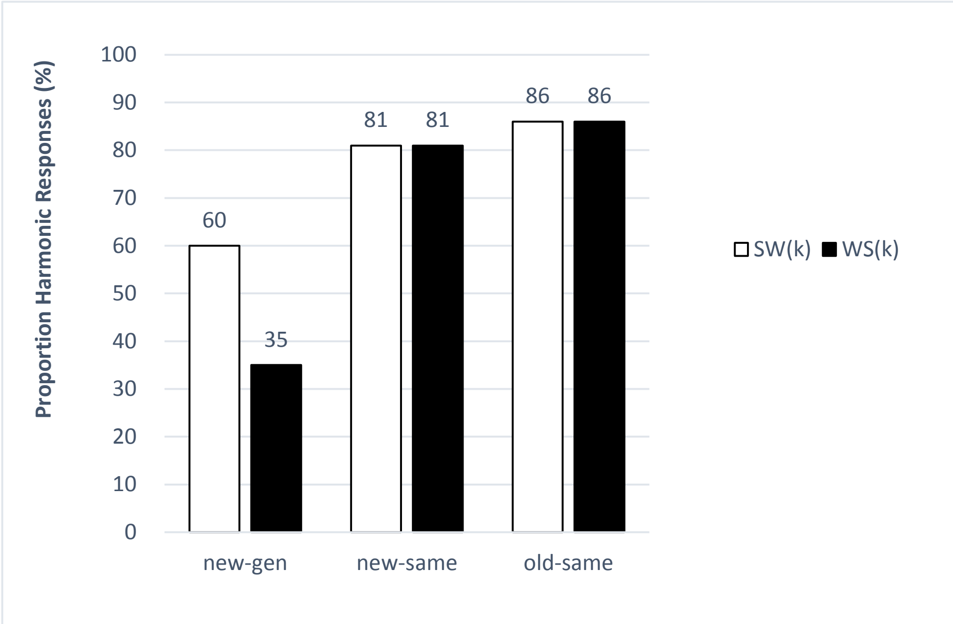


Figure 12. Response percentage of the test items for Pattern 1: SW(k) and Pattern 2: WS(k)

Test items were analyzed using mixed-effect logistic regression modeling (GLMM) (Pinheiro & Bates 2000) implemented in the R package lme4 (Bates & Maechler 2009). **"Direction"** was -1 for W(k)→S(k), and 1 for Patterns S(k)→W(k). The logic behind this factor is that if the results support the hypotheses about implicational universals investigated in this study, then it is expected that participants in the S(k)→W(k) group should learn better than those in the W(k)→S(k) group, because the sonority of /s/ is lower than that of /w/. Thus blocking by /w/ implies blocking by /s/, but not vice versa. In addition, if /s/ is a target, this implies that /w/ is also a target. By this logic, the baseline (i.e., coding: -1) for "Direction" should be W(k)→S(k). If we find a significant "positive effect", then it means that S(k)→W(k) is learned better than W(k)→S(k). This factor is crucial, since it could directly test whether the experimental results support the hypotheses.

Similar to Experiment 1, the test items had three types: old-same class, new-same class and new-generizable (new-gen). I coded them as O, NS and NGEN. These three categories were grouped into **"TypeName"**. The interaction of "Direction" and "TypeName" also needs to be considered to examine whether the accuracy of O, NS, and NGEN differs by "Direction". If a positive effect of "TypeNameO" and a negative effect of the interaction of "Direction" and "TypeNameO" are found, then this suggests that the old items were learned better than the new-same class items, especially in Pattern 2: W(k)→S(k), and this learning trend is not that obvious for Pattern 1: S(k)→W(k). This interaction effect would strengthen the hypothesis of directionality of implicational universals because if participants learned the pattern, the accuracy for both old and new-same class items should be at least equal or over 60%, and the accuracy difference between old and new-same class should be not that obvious (i.e., not pure memorization).

The formula for both by-participant and by participant & item analyses was used to carry out the analysis, as in (20) and (21).

(20) Accuracy ~ Direction * TypeName + (1|Participant)

(21) Accuracy ~ Direction * TypeName + (1|Participant) + (1|Item)

The ANOVA showed that the by-participant-and-item analysis was significant, suggesting that this model could better account for the results. According to Table 31, the main effect of "Direction" is positively significant, indicating that S(k)→W(k) (coding: 1) was learned better than W(k)→S(k) (coding: -1), which supports my hypothesis that a more marked segment (/s/ as a target) implies a less marked segment (/w/ as a target), but not vice versa. "TypeName" has three levels: O (old), NS (new-same class), and NGEN (new-gen), and those three are coded as O, NS, and NGEN. A significantly positive effect of "TypeNameO" was found, indicating that old items were learned better than new-gen items. The effect of "TypeNameNS" was significantly positive, indicating that new-same class items were learned better than new-gen items. Note that, as mentioned in Section 4.4.2, since I used dummy coding for "TypeName", the reference level of "TypeName" is "NGEN". When I report the main effect of "Direction", it actually refers to the effect of "Direction" when TypeName equals "NGEN" rather than all three word types.

Table 31. Effects and interaction for S(k)→W(k) and W(k)→S(k) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.03393	0.41058	0.083	0.9341
Direction	**0.80322**	**0.39608**	**2.028**	**0.0426 ***
TypeNameO	**2.34313**	**0.27853**	**8.412**	**< 2e-16 *****
TypeNameNS	**1.90302**	**0.42261**	**4.503**	**6.7e-06 *****
Direction: TypeNameO	-0.56201	0.43598	-1.289	0.1974
Direction:TypeNameNS	-0.50297	0.40740	-1.235	0.2170

'***': *p*<.001, '**': *p*<.01, '*': *p*<.05, '.': *p*<.1

In brief, the inferential statistics of Patterns 1 (S(k)→W(k)) and 2 (W(k)→S(k)) supports the hypothesis that a nasalized segment hierarchy plays a role.

5.4 Individual learner types

Several claim that typological universals do not equate with absolute grammatical universals, with typological universals reflecting statistical tendencies, not absolutes (see, for instance, Berent 2013: 147 for discussion). For this reason, I undertook an analysis of individual results. Specifically, though my hypothesis concerns implicational universals, it is also the case that individuals might vary in their learning strategies. Finley (2011a) reported that some participants generalized to a pattern not predicted by her hypothesis. This suggests that individuals might develop different strategies when making their judgments even though they were exposed to the same stimuli. As discussed in Section 3.2.2, both Finley and Berent's studies suggest that individuals do not pattern identically.

Individual differences in artificial grammar learning is an underresearched area (Visser et al. 2009, Zimmerer et al. 2011: 492). In the literature on artificial grammar learning area and second language acquisition, researchers have ascribed individual differences to declarative memory (knowledge about facts and events) and procedural memory (e.g., motor skills and habit learning) (see Reber 1967, Ullman, 2004, 2005, Conway et al. 2010, Ettlinger et al. 2014, 2015, Morgan-Short et al. 2014). Procedural memory is argued to reflect mental grammar such as phonology, syntax, morphology, long-distance dependencies, etc. (see Morgan-Short et al. 2014). However, the primary interest of this thesis is to examine phonological generalizability instead of memory. I will not discuss the assessment of declarative memory and procedural memory. Instead, I examine the results of the experiment involving /s/ and /w/ and group learners into different categories based on these results, focusing on how they classify new items. I propose two major types of learners, what I call categorization learners and statistical learners.

A **categorization learner** tried to group a new segment with certain old segments. The grouped segments share the same role in nasal harmony (i.e., either target or blocker). **Statistical learner** means that participants used fragmentary knowledge such as syllable combinations (e.g., phonotactics) to make a judgment (see discussion of statistical learning in Saffran et al. 1996, Gómez & Gerken 1999, Saffran et al. 1999, Fiser & Aslin 2002, Yang 2004, among others). Categorization learners are further broken down into three subcategories, *pattern learner*, *generalizer*, and *generalizer (opposite)*. Statistical learners are further broken down into *positional statistician* and *unbound nasalizer*. If a participant randomly guessed, I refer to them as **random**. The predictions of learner types about old-same class for Pattern 1 are shown in Table 32 and for Pattern 2 in Table 33. The discussion of subtypes follows Table 32 and Table 33. "Grammatical" items in Pattern 1 conform to the pattern of nasal harmony where /s/ and vowels are targets and /k/ is a blocker, whereas "ungrammatical" items are the pattern of harmony where /s/ is a blocker, and /k/ and other vowels are targets. For instance, the categorization learners learned the pattern, and judge old grammatical forms (1, 2, 3, 4) as grammatical and ungrammatical forms as ungrammatical (1', 2', 3', 4'). That is, when categorization learners make a judgment on pairs, they would correctly choose grammatical old-same class items (check mark ✓ in the cell) over ungrammatical items (blank in the cell)[46]. The statistical learners, on the other hand, are sensitive to the position or number of nasalized syllables.

[46] The subtypes of categorization learners are not predicted to show different patterning for the old-same class items, but they would be distinguished once new-same class and new-gen items are taken into account. For the purpose of clarity, I illustrate the old-same class first, followed by the others.

<u>Old-same class:</u>

Table 32. Pattern 1 predictions of raw responses for test items by learner type: old-same class

	Pattern 1: **S**(k)→**W**(k) (SW(k))					
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	Old-same class					
grammatical	(1) /ã**s̃**ãka/	✓	✓	✓	✓	✓
	(2) /ãka**s**a/	✓	✓	✓		
	(3) /ã**s̃**ã**s̃**ã/	✓	✓	✓	✓	✓
	(4) /ãkaka/	✓	✓	✓		
ungrammatical	(1’) /ã**s**aka/					
	(2’) /ãk̃ã**s**a/					✓
	(3’) /ã**s**a**s**a/					
	(4’) /ãk̃ãk̃ã/					✓

Table 33. Pattern 2 predictions of raw responses for test items by learner type: old-same class

		Pattern 2: **W**(k)→**S**(k) (WS(k))				
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	Old-same class					
grammatical	(1) /ã**w̃**ãka/	✓	✓	✓	✓	✓
	(2) /ãka**w**a/	✓	✓	✓		
	(3) /ã**w̃**ã**w̃**ã/	✓	✓	✓	✓	✓
	(4) /ãkaka/	✓	✓	✓		
ungrammatical	(1’) /ã**w**aka/					
	(2’) /ãk̃ã**w**a/					✓
	(3’) /ã**w**a**w**a/					
	(4’) /ãk̃ãk̃ã/					✓

Consider now the subtypes of categorization learners. *Pattern learner* means that participants learn the pattern, but fail to generalize from an old target to a new test consonant (/w/ in Pattern 1, /s/ in Pattern 2). In this case, I would expect that when choosing a word out of a pair, participants would tend to favor grammatical old-same class (1, 2, 3, 4) over ungrammatical old-same class items (1’, 2’, 3’, 4’) (see Table 32 and Table 33).

Generalizer means that participants not only learn the pattern, but also succeed in generalizing from an old target to a new test consonant. Like the *pattern learners*, they would also favor grammatical old-same class items (1, 2, 3, 4) over ungrammatical items (1’, 2’, 3’, 4’) (see Table 32 and Table 33).

Generalizer (opposite) means that participants learn the pattern and are able to generalize from an old target to a new consonant. However, they make a generalization in the opposite way, treating a new test segment as a blocker instead of as a target. Again, they would tend to choose grammatical old-same class (1, 2, 3, 4) than ungrammatical old-same class items (1’, 2’, 3’, 4’) (see Table 32 and Table 33).

Consider now the subtypes of statistical learners. *Positional statistician* means that participants use statistical knowledge such as phonotactics to make their decision. For instance, they might make the assumption that whenever the second syllable with a target is nasalized, it should be counted as “correct”. Such participants would consider a word with the second syllable nasalized as grammatical. Consider Pattern 1 (S(k)→W(k)) for instance (see Table 32). A positional statistician judges /s̃ã/ on the second syllable (1, 3) as grammatical, but not /sa/ (1’, 3’). But they would have no idea about how to choose between the pair (2) /ãkasa/ and (2’) /ãk̃ãsa/ since target /s/ is not in the second syllable. Other positionally-based patterns are possible as well.

Unbound nasalizer involves the interpretation that nasalization equals plurality (i.e., morphological learner, nasalization = plurality). The more syllables that are nasalized within a word, the better the word is. This learner calculates the number of nasalized syllables and favors a word with a larger number of nasalized syllables over a word with a smaller number of nasalized syllables. That is, a word with three nasalized syllables is chosen over one with one or two nasalized syllable (1, 2’, 3, 4’). Consider Pattern 2 (W(k)→S(k)) for instance (see Table 33). Participants favor grammatical (1) /ã**w̃**ãka/ over ungrammatical (1’) /ã**w**aka/ because (1) has two nasalized syllables, while (1’) has only one nasalized. However, they also favor ungrammatical (2’) /ãk̃ã**w**a/ over grammatical (2) /ãka**w**a/ since the number of nasalized syllables in the former is larger than in the latter.

<u>New-same class (different vowel):</u>

Now I turn to the predictions of new-same class items for Pattern 1 in Table 34 and for Pattern 2 in Table 35. Being able to learn the new-same class items suggests that participants learned the pattern since they could apply rules to a new vowel, not simply memorizing old items.

Table 34. Pattern 1 predictions of raw responses for test items by learner type: new-same class

		Pattern 1: **S**(k)→**W**(k) (SW(k))				
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	New-same class					
grammatical	(5) /ã**s̃**ẽka/	✓	✓	✓	✓	✓
	(6) /ãka**se**/	✓	✓	✓		
	(7) /ã**s̃**ãke/	✓	✓	✓	✓	✓
	(8) /ã**s̃**ẽ**s̃**ẽ/	✓	✓	✓	✓	✓
	(9) /ãkesa/	✓	✓	✓		
ungrammatical	(5’) /ã**s**eka/					
	(6’) /ãk̃ã**se**/					✓
	(7’) /ã**s**ake/					
	(8’) /ã**sese**/					
	(9’) /ãk̃ẽsa/					✓

Table 35. Pattern 2 predictions of raw responses for test items by learner type: new-same class

	Pattern 2: **W**(k)→**S**(k) (WS(k))					
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	New-same class					
grammatical	(5) /ã**w̃**ẽka/	✓	✓	✓	✓	✓
	(6) /ãka**we**/	✓	✓	✓		
	(7) /ã**w̃**ãke/	✓	✓	✓	✓	✓
	(8) /ã**w̃**ẽ**w̃**ẽ/	✓	✓	✓	✓	✓
	(9) /ãkewa/	✓	✓	✓		
ungrammatical	(5') /ã**w**eka/					
	(6') /ãk̃ã**we**/					✓
	(7') /ã**w**ake/					
	(8') /ã**w**e**w**e/					
	(9') /ãk̃ẽ**w**a/					✓

Recall that *pattern learner*, *generalizer*, *and generalizer (opposite)* all learned the pattern. I expect that participants for these three learner types would tend to favor grammatical new-same class items over ungrammatical ones (cf. Table 34 and Table 35). This suggests that participants would be able to apply the rules conditioning plural change on new vowels.

The *positional statistician* might make the assumption that whenever the second syllable with a target is nasalized, it should be counted as "correct".

The *unbound nasalizer* would favor (6') /ãk̃ã**we**/ over (6) /ãka**we**/ in Pattern 2 (see Table 35) because the former word has two nasalized syllables, whereas the latter has only one nasalized syllable.

<u>New-gen class (different consonant):</u>

Now I turn to the predictions of new-gen class items for Pattern 1 in Table 36 and for Pattern 2 in Table 37. This one is important since it differentiates the sub-types of categorization learners. Recall that a check mark ✓ in a cell means the item was chosen by participants, and a blank in a cell means the item was not chosen by participants. A check mark ✓ inside parentheses means that either choice (e.g., (14) or (14’)) is possible.

Table 36. Pattern 1 predictions of raw responses for test items by learner type: new-gen

	Pattern 1: **S**→ **W** (SW)					
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	New-gen					
grammatical	(10) /ã**w̃**ãka/		✓		✓	✓
	(11) /ãka**w**a/		✓	✓		
	(12) /ã**w̃**ãs̃ã/		✓		✓	✓
	(13) /ã**w̃**ã**w̃**ã/		✓		✓	✓
	(14) /ãs̃ã**w̃**ã/		✓		(✓)	✓
ungrammatical	(10’) /ã**w**aka/			✓		
	(11’) /ãk̃ã**w̃**ã/					✓
	(12’) /ã**w**asa/			✓		
	(13’) /ã**w**a**w**a/			✓		
	(14’) /ã**s̃ã**wa/			✓	(✓)	

Table 37. Pattern 2 predictions of raw responses for test items by learner type: new-gen

	Pattern 2: **W**(k)→**S**(k) (WS(k))					
		categorization learner			statistical learner	
		pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
	New-gen					
grammatical	(10) /ã**s̃**ãka/		✓		✓	✓
	(11) /ãka**s**a/		✓	✓		
	(12) /ã**s̃**ã**w̃**ã/		✓		✓	✓
	(13) /ã**s̃**ã**s̃**ã/		✓		✓	✓
	(14) /ã**w̃**ã**s̃**ã/		✓		(✓)	✓
ungrammatical	(10') /ã**s**aka/			✓		
	(11') /ãk̃ã**s̃**ã/					✓
	(12') /ã**s**awa/			✓		
	(13') /ã**s**a**s**a/			✓		
	(14') /ã**w̃ã**sa/			✓	(✓)	

The *pattern learner* learns the pattern, but fails to generalize from an old target to a new test consonant (/w/ in Pattern 1, /s/ in Pattern 2). I would predict that pattern learners would have no preference for new-gen items.

The *generalizer* not only learns the pattern, but also succeeds in generalizing from an old target to a new test segment. Specifically, generalizers pattern an old target with a new test segment and treat a new test consonant as a target, meaning that they would tend to favor grammatical new-gen over ungrammatical ones (see Table 36 and Table 37).

The *generalizer (opposite)* learns the pattern, but makes a generalization in an opposite way: treating a new test consonant as a blocker instead of a target. Since they treat a new segment as a blocker, it is expected that they would be inclined to favor ungrammatical new-gen items over grammatical ones (see Table 36 and Table 37) except for (11) where the old segment /k/ should be a blocker.

The *positional statistician* uses statistical knowledge such as phonotactics to make their decision. For new-gen items, they would have no difficulty in choosing a word with a nasalized second syllable over a word with a non-nasalized second syllable when the second syllable contained a new segment /w/ in Pattern 1 and /s/ in Pattern 2 as in (10-10',12-12',and, 13-13'). However, they would have difficulty in choosing one from a particular pair: (14) /ãs̃ãw̃ã/ vs. (14') /ãs̃ãwa/ (Pattern 1), and (14) /ãw̃ãs̃ã/ vs. (14') /ãw̃ãsa/ (Pattern 2) because in both words, a second syllable is nasalized. Specifically, when the second syllable contains an old segment /s/ in Pattern 1 and /w/ in Pattern 2, as in (14-14'), the learner would prefer nasalization on the second syllable, but would not be able to choose between nasalization (14) or no nasalization (14') on the third syllable. This is why I put the check mark in brackets (✓).

The *unbound nasalizer* favors a word with a larger number of nasalized syllables over a word with a smaller number of nasalized syllables. For instance, participants would favor grammatical (12) /ã**s̃**ã**w̃**ã/ over ungrammatical (12') /ã**s**awa/ in Pattern 2 because the former word has three nasalized syllables, whereas the latter has only one nasalized syllable. However, participants would favor ungrammatical (11') /ãk̃ã**s̃**ã/ over (11) /ãka**s**a/, depending on which word has more nasalized syllables.

The raw percentages of 'ideal' learner types are summarized by word type in Table 38.

Table 38. Pattern 2 predictions of raw responses for test items by learner type: new-gen

Pattern					
	categorization learner			statistical learner	
	pattern learner	generalizer	generalizer (opposite)	positional statistician	unbound nasalizer
Old-same	100%	100%	100%	varies (50%)[47]	50%
New-same	100%	100%	100%	varies (60%)	60%
New-gen	50%	100%	20%	varies (60%)	80%

The three sub-types of categorization learners can be distinguished by new-gen items. The *positional statistician* and the *unbound nasalizer* can be differentiated in new-gen items and whether /k/ is nasalized. Specifically, /k/ could be nasalized for the *unbound nasalizer*. /k/ cannot

[47] The percentages are based on one kind of positional statistician (i.e., nasalizing the second syllable with a target), but participants could be using other kinds of statistical strategies in their learning.

be nasalized for the *positional statistician* for whom the second syllable with a target always gets nasalized, but a second syllable with a blocker is not nasalized. Note that when the distinction of correct rate for the positional statistician and the unbound nasalizer was somtimes close for new-gen, the post-interview helped me make a decision as to which category the person belonged in. For the unbound nasalizer, the participant would say something to indicate that they preferred a word with more nasalized syllables. For positional statistician, they would say that they noticed that the second syllable was nasalized most of the time. They were certain that a syllable with a target should be nasalized in the second syllable, but did not understand why sometimes the third syllable was nasalized and sometimes it was not.

5.5 Individual data: Patterns 1 (S(k)→W(k)) and 2 (W(k)→S(k))

As discussed in Section 5.4, I define two major types of participants, what I call categorization learners and statistical learners. Categorization learners are further broken down into three categories, pattern learner, generalizer, and generalizer (opposite). Statistical learners too are further broken down into positional statistician and unbound nasalizer. Recall that if participants randomly guess, I refer to them as random.

Table 39 summarizes the distribution of participant types by subcategories (i.e., pattern learner, generalizer, generalizer (opposite), positional statistician, unbound nasalizer, and random for Patterns 1 (S(k)→W(k)) and 2 (W(k)→S(k)). For instance, in Pattern 1 (S(k)→W(k)), 1 participant belongs to pattern learner, 3 belong to generalizer, 2 belong to generalizer (opposite), 2 belong to positional statistician, 1 belongs to unbound nasalizer, and 1 belongs to random.

Specifically, I look at whether categorization learners show an asymmetry between Pattern 1 and Pattern 2. Note that for categorization learners, if more than 5 participants belong to generalizer or opposite generalizer, I conclude that they had a tendency to pattern a new segment with old segments in some way.

By looking at the learner type, for each pattern, basically, the major difference between the two patterns is that Pattern 2: W(k)→S(k) seems to have more opposite generalizers than Pattern 1: S(k)→W(k). What does that mean? I will further discuss each subtype based on each participant's correct rates (recall 'correct' refers to the predictions of the implicational nasalized segment hierarchy) as well as actual responses, and discuss this question following Table 39 and Table 40.

Table 39. Pattern 1: SW(k) vs. Pattern 2: WS(k) (Individual difference)

	Pattern 1: SW(k)			Pattern 2: WS(k)		
		exposure	**test**		exposure	**test**
	k as blocker	s as target	**w as target**	k as blocker	w as target	**s as target**
pattern learner		n=1			n=0	
generalizer		n=3			n=2	
generalizer (opposite)		n=2			n=6	
positional statistician		n=2			n=2	
unbound nasalizer		n=1			n=0	
random		n=1			n=0	

Table 40 shows the descriptive statistics for individual data in detail. I will discuss each learner type by percentages. I bold any percentages equal to or higher than 60%.

Table 40. Pattern 1: SW(k) vs. Pattern 2: WS(k) (Individual difference)

	SW(k)	NGEN	NS	NO	sum		
Subj	1	**95%**	**100%**	**100%**	**98%**	categorization learner	generalizer
Subj	2	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	3	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	4	20%	**85%**	**100%**	**66%**	categorization learner	generalizer (opposite)
Subj	5	20%	**95%**	**100%**	**70%**	categorization learner	generalizer (opposite)
Subj	6	50%	**100%**	**100%**	**82%**	categorization learner	pattern learner
Subj	7	**60%**	55%	**63%**	59%	statistical learner	positional statistician
Subj	8	50%	**70%**	**94%**	**70%**	statistical learner	positional statistician
Subj	9	**65%**	55%	50%	57%	statistical learner	unbound nasalizer
Subj	10	40%	45%	50%	45%	random	

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

	WS(k)	NGEN	NS	NO	sum		
Subj	11	**65%**	**70%**	**75%**	**70%**	categorization learner	generalizer
Subj	12	**60%**	**100%**	**100%**	**86%**	categorization learner	generalizer
Subj	13	35%	**60%**	**75%**	55%	categorization learner	generalizer (opposite)
Subj	14	20%	**80%**	**100%**	**64%**	categorization learner	generalizer (opposite)
Subj	15	20%	**95%**	**94%**	**68%**	categorization learner	generalizer (opposite)
Subj	16	20%	**100%**	**100%**	**71%**	categorization learner	generalizer (opposite)
Subj	17	20%	**60%**	**100%**	57%	categorization learner	generalizer (opposite)
Subj	18	20%	**95%**	**100%**	**70%**	categorization learner	generalizer (opposite)
Subj	19	55%	**65%**	**63%**	**61%**	statistical learner	positional statistician
Subj	20	35%	**80%**	56%	57%	statistical learner	positional statistician

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

5.5.1 Categorization learner-generalizer

For generalizers, the percentage of correct responses for old-same class, new-same class, and new-generalizable are all high. They appeared to know that /k/ blocked spreading (i.e., they were "learners"). Several factors can be considered in thinking about how a participant determined whether a consonant was a target or a blocker.

Here I include reports that participants gave in post-interviews about how they made decisions. These self-reported strategies provide interesting insight into what participants thought they did but may not reflect what they actually did.

(a) Place of articulation:

These learners tried to articulate a new segment, and compared how close it was to an old segment that they were exposed to. For example, they thought that the place of articulation of the new segment /w/ was similar to that of the old segment /s/ because both are produced around the lip area[48]. During the post-interview, subject 1 reported that /w/ and /s/ were both articulated by opening the mouth and moving somewhere around the lip areas, but /k/ was articulated by closing the mouth and not moving somewhere around the lip areas. Subject 3 reported that "deeper" sounds (produced in a more back position) like /k/ did not need to be nasalized.

(b) Ease of articulation:

Some learners treated a new segment /s/ or /w/ as a target, saying that nasalizing /s/ or /w/ was easy for them to articulate. This is my speculation, since not all generalizers could explicitly state why they chose the pattern they did. They just said they "felt" this new sound should act like

[48] In Min, /w/ seems to pattern with labials phonologically. There is a labial co-occurrence restriction that two labials are not allowed within a constituent (two labials are under nucleus node: */uo/, */uau/) or are not allowed to "be dominated by a common node and one labial is immediately dominated by one and only one node which does not dominate the other labial" (Bao 2000: 110) (nucleus node for */tup/, */uap/); rhyme node for */pam/). Other than that, two labials are allowed (/pu/, /pau/).

a target when producing it. (They did not call it a target, but referred to an old consonant that functioned as a target)[49].

5.5.2 Categorization learner-generalized (opposite)

Like other categorization learners, these learners knew that /k/ was a blocker and they generalized from an old vowel to a new vowel. They knew old items. If participants classified a new segment (S(k)→W(k): /w/, W(k)→S(k): /s/) as a target, I would call them a generalizer, but if they classified it as a blocker, then I would call them an opposite generalizer. Consider subjects 15 and 16 for instance.

Subject 15 reported that /s/ was harder to nasalize and thus they decided not to nasalize it.

Subject 16 and subject 18 reported that they thought /s/ and /k/ sounded similar. Since /k/ in Pattern 2 was not nasalized, they thought they should not nasalize /s/.

5.5.3 Categorization learner-pattern learner

For all categorization learners, the percentages of correct responses for old-same class and new-same class are high (over 60%), suggesting that they learned the pattern they were exposed to. Pattern learners knew that /k/ blocked spreading and could generalize to new vowels, but they failed to generalize to new consonants (new-gen: 50% (chance level)). For example, they knew that /s/ was nasalized in /ã**s̃**ãka/, but not in /ãka**s**a/, since /k/ blocks spreading**.** They performed well with old items. They were also able to generalize to new vowels (i.e., /e/). Consider subject 6 for example. Subject 6 got 100% right for old items (e.g., /a**s**aka/ ~ /ãs̃ãka/ ~ */ã**s**aka/), and was able to generalize nasalization to a new vowel (e.g., /aseka/ ~ /ãs̃**ẽ**ka/ ~ */ãseka/), but treated /w/ as a

[49] Two additional strategies that learners might use but are not relevant to this thesis are presented in this footnote. First, any new segments should be targets, and /k/ was just an exception. These learners had the impression that the main difference between singulars and plurals was nasalization, making them more inclined to treat new segments as targets. Since targets had changes (nasalization), but blockers did not (no nasalization), nasalization was the preferred strategy. For example, during a post-interview, subject 2 reported that /k/ was the only consonant that would block nasal spreading. Other consonants including new consonants should pattern together in allowing harmony in the same way: Stop nasalizing when encountering /k/, otherwise nasalize. Subject 11 treated /k/ as an exception, and other new consonants as targets (since /k/ was the only special consonant that could block spreading). Subject 12 thought that /s/ was a "new" segment and therefore a target (i.e., any "new" consonant should be treated as a target). This has nothing to do with the implicational relationship between an old target /w/ and a new segment /s/. Second, I suspect that learners could simply use a strategy that /w/ or /s/ should be a target, since the majority of new segments obey this trend (i.e., the majority of items had both a target and a blocker, so participants assumed that a new consonant needed to be a target when combined with a blocker).

target half the time (e.g., /awaka/ ~ /ãw̃ãka/ ~ *ãwaka). The generalization to new vowels suggests that this participant was not simply memorizing old items, but treated vowels as a class. However, choosing /w/ as a target half the time suggests that this participant (pattern learner) failed to generalize from the old consonant /s/ to /w/.

5.5.4 Statistical learner-positional statistician

If participants did not realize that /k/ was a blocker, they used fragmentary knowledge such as certain syllable combinations to make a judgment on a new consonant. Since they did not learn the pattern based on sonority, consistent performance is not expected in terms of sonority. The percentage of correct responses for new-same class or new-generalizable may be relatively high, or the percentage of correct responses for both new categories may be high. Sometimes the percentages of correct responses for old-same class and new-same class are both high. What differentiates the positional statistician from the pattern learner is that they did not know the patterning of old consonants in the new-gen condition. This suggests that they did not learn the pattern under investigation.

Such learners were puzzled about why /s/ or /w/ was sometimes nasalized, and sometimes not. They used strategies such (a) second syllable tending to be nasalized, (b) same-consonant tending to be nasalized, (c) new consonant tending not to be nasalized, or (d) new consonant tending to be nasalized all the time (unbound nasalizer), combined with random choices, or with fragmentary knowledge (e.g., /k̃ã/ at the second syllable tending to be thought as "correct" because of participant's sensitivity to the fact that the crucial change of nasalization lies on the second syllable).

5.5.5 Statistical learner-unbound nasalizer

Subject 9 compared which word had more nasalized syllables out of a pair (binary choice). Containing more nasalized syllables in a word was thought to create a better plural form (nasality = plurality).

In brief, learner types reflect different strategies used by the participants. This thesis focuses on the results of categorization learners.

5.6 Discussion: learner types

In this section, I present four possible ways of interpreting the individual data for Patterns 1 and 2.

5.6.1 Possibility 1: reference to the sonority hierarchy type

WS(k) and SW(k) patterns had roughly the same number of learners (categorization learning: pattern learner, generalizer, generalizer (opposite), but the difference between these two patterns is that WS(k) had more opposite generalizers than SW(k) did, which means that participants for WS treated a new consonant /s/ as a blocker rather than a target.

What does that suggest? The first possibility (implicational universals) is that the reason that the learning of Pattern 2: WS(k) was significantly worse than that of Pattern 1: SW(k) is because generalizing from /w/ as a target to /s/ as a potential target is a harder pattern to learn compared to /s/ as a target→/w/ as a potential target. Therefore, in Pattern 2 participants had difficulty in generalizing from /w/ as a target to /s/ as a potential target. If so, it would support my hypothesis of the nasalized segment hierarchy. Since the nasalized segment hierarchy is highly correlated with the sonority hierarchy, I treat Possibility 1 **reference to the sonority hierarchy**. For the rest of the thesis, I call the implicational nasalized segment hierarchy hypothesis the sonority hierarchy (type) hypothesis.

5.6.2 Possibility 2: reference to sonority natural classes

The second possibility (patterning) is that for the WS(k) pattern, learners patterned /k/ as a blocker with /s/ since both /k/ and /s/ belong to the natural classes of obstruents. That is why WS(k) had more opposite generalizers (i.e., treating /s/ as a blocker instead of target). Since this shows that participants are sensitive to the sonority distinction between sonorants and obstruents and pattern a new segment by comparing with sonority natural classes of the old segments, I treat Possibility 2 as **reference to sonority natural classes** (obstruent or sonorant). The sonority natural classes I discuss here involve the feature [sonorant], which distinguishes sonorants and obstruents.

5.6.3 Possibility 3: floor effect

The third possibility is that the artificial grammar is just too hard to learn (i.e., floor effect). That is why I got few generalizers in both Patterns 1 and 2.

5.6.4 Possibility 4: game strategy

The more extreme view (game strategy) is that perhaps no phonological processing was involved for these two patterns. What participants were trying to do for "NGEN" is just to determine whether a "new" segment needs to be "nasalized" or "non-nasalized (i.e., new = nasalized or new = non-nasalized). That is, participants treat the experiment as a game instead of as language learning.

5.6.5 Possibility 1 vs. Possibility 2

Assuming Possibilities 3 and 4 being ruled out, it seems that the sonority natural class hypothesis is more plausible. For the SW(k) pattern, few participants generalized from /s/ as a target to /w/ as a target, which poses a challenge to the first possibility that WS(k) is harder to learn/generalize than SW(k) (see Table 40). Maybe the reason that few participants made this generalization is because they thought that /s/ and /w/ do not belong to the same phonological class, so they tended not to treat the new consonant /w/ as a target. They also failed to pattern /k/ as a blocker with /w/, because /k/ and /w/ do not belong to the same phonological class. In brief, the current descriptive statistics seem to suggest that instead of the sonority hierarchy being involved, participants tried to compare whether a new segment's phonological class is close to any old segment's phonological class. If yes, then participants patterned the two segments together, assigning the same status of nasality function (i.e, target or blocker) (i.e., among old and new segments, there is reference to natural classes).

5.7 Follow-up experiments: Patterns 3 (k(S)→t(S)) and 4 (k(W)→t(W))

The finding of Patterns 1 and 2 cannot rule out Possibilities 3 and 4. In order to disentangle the four possibilities, sonority hierarchy type, sonority natural classes, floor effect, and game strategy, I ran two follow-up experiments involving the consonants /k/ and /t/. These experiments are similar to those discussed in the previous chapter using /k/ and /p/ (see Table 20 in Section

4.2.5), but the experiments reported here use the design where in the test phase the singular was heard, and participants chose between two possible plurals (exposure: /k/ as a blocker→test: whether /t/ could be a blocker).

Possibilities 3 (floor effect) and 4 (game strategy) are considered to be confounds. If these confounds are ruled out, then phonological processing is the only possibility that is involved. Possibilities 1 (sonority hierarchy type) and 2 (sonority natural classes) could be possible only if this artificial grammar learning involves phonological processing. In Section 5.7.1 and 5.7.3 I present the hypotheses to test possibilities 3 and 4.

The directionality of learning (sonority hierarchy type) for Patterns 3 and 4 (k→t, t→k) focuses on the generalizability of blocking effects. The difference between the two patterns is that Pattern 3 had /s/ as a target for both exposure and test, while Pattern 4 had /w/ as a target for both exposure and test (i.e., k(S)→t(S) vs. k(w)→t(W)).

In general, participants were exposed to /asak/ ~ /ãs̃ã**k**a/; then they were tested with /t/ to see if they generalized (singular /asa**t**a/ ~ plural /ãs̃ã**t**a/ ~ plural */ãs̃ã**t̃**ã/) (**Pattern 3: k(S)→t(S)**) (abbreviated as kt(S), see Table 41). Similarly, the other group of participants were exposed to /awaka/ ~ /ãw̃ã**k**a/ and tested with /t/ (/awa**t**a/ ~ /ãw̃ã**t**a/ ~ */ãw̃ã**t̃**ã/) (**Pattern 4: k(W)→t(W)**) (abbreviated as kt(W), see Table 41). For Pattern 3, /s/ was always a target throughout the exposure and test, while for Pattern 4, /w/ was always a target.

Table 41. Predictions of the sonority hierarchy type for Patterns 3 and 4

	exposure	test	prediction
Pattern 3 (kt(S))	more sonorant: target /s/ less sonorant: blocker /**k**/	new segment: /**t**/ same class as blocker	new segment is blocker
Pattern 4 (kt(W))	more sonorant: target /w/ less sonorant: blocker /**k**/	new segment: /**t**/ same class as blocker	new segment is blocker

Table 42 presents the exposure phase, and the test phase is shown in Table 43.

Table 42. Stimuli design (exposure)

	Pattern 3: **k**(S)→**t**(S) kt(S)		Pattern 4: **k**(W)→**t**(W) kt(W)	
	singular	plural	singular	plural
grammatical	/a**k**asa/	/ã**k**asa/	/a**k**awa/	/ã**k**awa/
	/asa**k**a/	/ãs̃ã**k**a/	/awa**k**a/	/ãw̃ã**k**a/
	/a**k**a**k**a/	/ã**k**a**k**a/	/a**k**a**k**a/	/ã**k**a**k**a/
	/asasa/	/ãs̃ãs̃ã/	/awawa/	/ãw̃ãw̃ã/
	/a**k**asi/	/ã**k**asi/	/a**k**awi/	/ã**k**awi/
	/asa**k**i/	/ãs̃ã**k**i/	/awa**k**i/	/ãw̃ã**k**i/
	/a**k**a**k**i/	/ã**k**a**k**i/	/a**k**a**k**i/	/ã**k**a**k**i/
	/asasi/	/ãs̃ãs̃ĩ/	/awawi/	/ãw̃ãw̃ĩ/
	/a**k**isa/	/ã**k**isa/	/a**k**iwa/	/ã**k**iwa/
	/asi**k**a/	/ãs̃ĩ**k**a/	/awi**k**a/	/ãw̃ĩ**k**a/
	/a**k**i**k**a/	/ã**k**i**k**a/	/a**k**i**k**a/	/ã**k**i**k**a/
	/asisa/	/ãs̃ĩs̃ã/	/awiwa/	/ãw̃ĩw̃ã/
	/a**k**isi/	/ã**k**isi/	/a**k**iwi/	/ã**k**iwi/
	/asi**k**i/	/ãs̃ĩ**k**i/	/awi**k**i/	/ãw̃ĩ**k**i/
	/a**k**i**k**i/	/ã**k**i**k**i/	/a**k**i**k**i/	/ã**k**i**k**i/
	/asisi/	/ãs̃ĩs̃ĩ/	/awiwi/	/ãw̃ĩw̃ĩ/

(k,t=blockers; S,W=targets)

Table 43. Stimuli design (test items)

	Pattern 3: **k**(S)→**t**(S) (kt(S))			Pattern 4: **k**(W)→**t**(W) (kt(W))		
	new		old	new		old
	gen (n=10, pair=5)	same class (n=10, pair= 5)	same class (n=8, pair=4)	gen (n=10, pair= 5)	same class (n=10, pair= 5)	same class (n=8, pair=4)
grammatical (n=44)	/ã**t**asa/	/ãs̃ẽ**k**a/	/ãs̃ã**k**a/	/ã**t**awa/	/ãw̃ẽ**k**a/	/ãw̃ã**k**a/
	/ãs̃ã**t**a/	/ã**k**ase/	/ã**k**asa/	/ãw̃ã**t**a/	/ã**k**awe/	/ã**k**awa/
	/ã**t**aka/	/ãs̃ã**k**e/	/ãs̃ãs̃ã/	/ã**t**aka/	/ãw̃ã**k**e/	/ãw̃ãw̃ã/
	/ã**t**a**t**a/	/ãs̃ẽs̃ẽ/	/ã**k**a**k**a/	/ã**t**a**t**a/	/ãw̃ẽw̃ẽ/	/ã**k**a**k**a/
	/ãka**t**a/	/ã**k**esa/		/ãka**t**a/	/ã**k**ewa/	
ungrammatical (n=44)	/ã**t̃**ãs̃ã/	/ãse**k**a/	/ãsa**k**a/	/ã**t̃**ãw̃ã/	/ãwe**k**a/	/ãwa**k**a/
	/ãs̃ã**t̃**ã/	/ã**k̃**ãse/	/ã**k̃**ãsa/	/ãw̃ã**t̃**ã/	/ã**k̃**ãwe/	/ã**k̃**ãwa/
	/ã**t̃**ãka/	/ãsa**k**e/	/ãsasa/	/ã**t̃**ãka/	/ãwa**k**e/	/ãwawa/
	/ã**t̃**ã**t̃**ã/	/ãsese/	/ã**k̃**ã**k̃**ã/	/ã**t̃**ã**t̃**ã/	/ãwewe/	/ã**k̃**ã**k̃**ã/
	/ãk̃ã**t̃**ã/	/ã**k̃**ẽsa/		/ãk̃ã**t̃**ã/	/ã**k̃**ẽwa/	

(gen=generalizable, n=number)

In the following three subsections I present three hypotheses.

5.7.1 Hypothesis 1: game strategy

If artificial grammar learning does not involve phonology at all but is simply a game strategy, then I would expect that participants would tend to (1) nasalize a new consonant all the time (unbound nasalizer) or (2) never nasalize a new consonant[50].

5.7.2 Hypothesis 2: phonological processes

If artificial grammar learning does involve phonological processes, then Possibility 2 (sonority natural classes) is also possible in addition to Possibility 1 (sonority hierarchy type). My

[50] These represent two possibilities of game strategies. It could be that participants use other types of game strategies as well.

hypothesis is that if participants pattern a new segment depending on phonological classes of old segments (that is, taking both old target and blocker into account – reference to sonority natural classes), then it is expected that kt(S) might be harder to learn than kt(W), since the sonority distance between /s/ and the new segment /t/ is shorter than that between /w/ and /t/. Specifically, for the kt(S) group, although /k/ and /t/ belong to the same sonority class, stop, the fact that /s/ is a fricative might create difficulty for participants to determine whether /t/ should pattern with /k/ as a blocker or with /s/ as a target (i.e., /t, k, s/ belong to obstruents). On the contrary, for the kt(W) group, the distance between /w/ and /t/ on the sonority hierarchy is further, which might make it easier for participants to pattern a new segment /t/ with /k/ as a blocker.

5.7.3 Hypothesis 3: floor effect

If the design of artificial grammars is too hard to learn (floor effect), then I would expect that there will be few generalizers for both Pattern 3: k(S)→t(S) and Pattern 4: k(W)→t(W), just like Patterns 1: S(k)→W(k) and 2: W(k)→S(k).

5.8 Methods

See Section 5.2 for the overall design and forced choice method.

5.8.1 Participants

The participants were bilingual in Min and Mandarin. Two groups of participants were recruited (Patterns 3 and 4). All reported normal hearing, and all distinguished nasalized monosyllables from their oral counterparts at a rate of 100%. All had early childhood Min exposure. Each participant recruited from National Sun Yat-sen University received 150 NTD for participating in the experiment, which lasted about 45-50 minutes.

Pattern 3: k(W)→t(W)

Ten college students were recruited from National Sun Yat-sen University in Taiwan. All had studied a foreign language (English 10, Japanese 1, French 1, German 1). The average age of the participants was 21.3 (SD = 2.5). 7 participants were males and 3 participants were females.

Pattern 4: k(S)→t(S)

Ten college students were recruited from National Sun Yat-sen University in Taiwan. All had studied a foreign language (English 10, Japanese 3, German 1). The average age of the participants was 21.6 (SD = 2.5). 8 participants were males and 2 participants were females.

5.9 Grouped statistics: Patterns 3 kt(S) and 4 kt(W)

Now I turn to the comparison of Patterns 3 k(S)→t(S) and 4 k(W)→t(W).

As discussed above, to tease apart the four possibilities: (1) reference to sonority hierarchy type, (2) reference to sonority natural classes, (3) floor effect, and (4) game strategy, I ran two follow-up experiments (kt(S) vs. kt(W)). The direction of learning (sonority hierarchy type) is the same for both experiments (k→t), with the difference between the two experiments being that Pattern 3 had /s/ as a target for both exposure and test, while Pattern 4 had /w/ as a target for both exposure and test (i.e., k(S)→t(S) vs. k(W)→t(W)).

The sonority natural class hypothesis predicts that Pattern 4: kt(W) would be learned better than Pattern 3: kt(S) since the distance of sonority of /w/ is larger than that of /s/, making it easier for participants to pattern /k/ with /t/ in the kt(W) group. Therefore, I coded kt(W) as 1, and coded kt(S) as -1. Note that the hypothesis that there is reference to the sonority hierarchy type does not make a direct prediction for kt(S) and kt(W). The hypothesis for reference to the sonority hierarchy type predicts that **k**(S)→**t**(S) vs. **t**(S)→**k**(S) will be equally learnable/generalizable, as will **k**(W)→**t**(W) vs. **t**(W)→**k**(W). I discuss this in Chapter 6.

The percentages of correct responses of the test items for Pattern 3: k(S)→t(S) and Pattern 4: k(W)→t(W) can be found in Figure 13. Patterns 3 and 4 participants did well on all three word conditions, namely old-same, new-same, and new-gen test items.

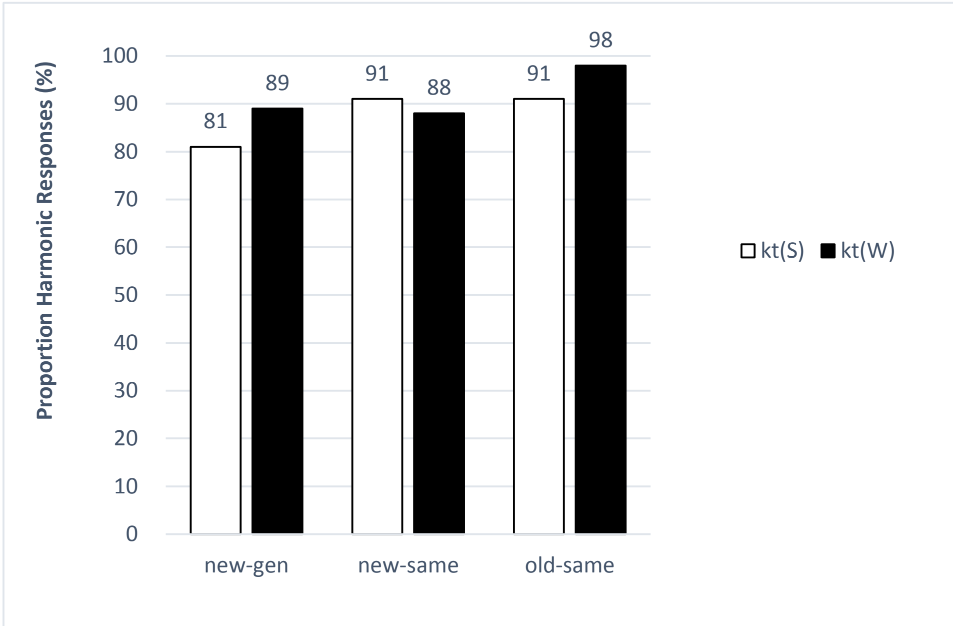


Figure 13. Response percentage of the test items for Pattern 3: kt(S) and Pattern 4: kt(W)

According to Table 44, there is no main effect of "Direction", suggesting that there was no significant learning difference between kt(S) and kt(W) patterns.

Recall that "TypeName" has three levels: O (old), NS (new-same class), and NGEN (new-gen), coded as O, NS, and NGEN. A significantly positive main effect of "TypeNameO" was found, suggesting that old items were learned better than new-gen items.

Table 44. Effects and interaction for k(W)→t(W) and k(S)→t(S) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.5122	0.4462	5.630	1.8e-08 ***
Direction	0.3180	0.4305	0.739	0.460047
TypeNameO	**1.4733**	**0.4411**	**3.340**	**0.000839 *** **
TypeNameNS	0.3709	0.3367	1.102	0.270667
Direction:TypeNameO	0.5580	0.4412	1.264	0.206053
Direction:TypeNameNS	-0.5386	0.3367	-1.600	0.109651

'***': p<.001, '**': p<.01, '*': p<.05, '.': p<.1

In brief, the inferential statistics appear to support the sonority hierarchy type hypothesis that Pattern 3 kt(S) and Pattern 4 kt(W) are equally generalizable.

5.10 Individual data: Patterns 3 kt(S) and 4 kt(W)

Table 45 summarizes the distribution of participant types by subcategories except for random. For instance, in Pattern 3, 2 participants belong to pattern learner, 4 belong to generalizer, and so forth. In general, the major difference between the two patterns is that **Pattern 4**: **k (W)→t (W)** has more generalizers than **Pattern 3: k (S)→t (S**). I show each participant's correct rates in Table 46.

Table 45. Pattern 3: kt(S) vs. Pattern 4: kt(W) (Individual difference)

	Pattern 3: kt(S)			Pattern 4: kt(W)		
	exposure		test	exposure		test
	s as target	k as blocker	t as blocker	**w as target**	k as blocker	t as blocker
pattern learner		n=2			n=0	
generalizer		n=4			n=9	
generalizer (opposite)		n=1			n=1	
positional statistician		n=3			n=0	
unbound nasalizer		n=0			n=0	
random		n=0			n=0	

Table 46 presents the individual differences[51]. I bold any percentages equal to or higher than 60%.

Table 46. Pattern 3: kt(S) vs. Pattern 4: kt(W) (Individual difference)

	kt(S)	NGEN	NS	O	sum		
Subj	21	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	22	**95%**	**100%**	**100%**	**98%**	categorization learner	generalizer
Subj	23	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	24	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	25	45%	**100%**	**100%**	**80%**	categorization learner	generalizer (opposite)
Subj	26	**65%**	**85%**	**94%**	**80%**	categorization learner	pattern learner
Subj	27	**65%**	**100%**	**100%**	**88%**	categorization learner	pattern learner
Subj	28	**75%**	**80%**	**69%**	**75%**	statistical learner	positional statistician
Subj	29	**95%**	**75%**	**63%**	**79%**	statistical learner	positional statistician
Subj	30	**70%**	**65%**	**81%**	**71%**	statistical learner	positional statistician

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

	kt(W)	NGEN	NS	O	sum		
Subj	31	**100%**	**65%**	**100%**	**88%**	categorization learner	generalizer
Subj	32	**95%**	**100%**	**100%**	**98%**	categorization learner	generalizer
Subj	33	**90%**	**95%**	**94%**	**93%**	categorization learner	generalizer
Subj	34	**90%**	**95%**	**94%**	**93%**	categorization learner	generalizer
Subj	35	**100%**	**95%**	**100%**	**98%**	categorization learner	generalizer
Subj	36	**90%**	**60%**	**94%**	**80%**	categorization learner	generalizer
Subj	37	**100%**	**70%**	**100%**	**89%**	categorization learner	generalizer
Subj	38	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	39	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	40	25%	**95%**	**100%**	**71%**	categorization learner	generalizer (opposite)

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

Combined with data from the previous two patterns (SW(k) vs. WS(k)), the findings suggest that the design of my artificial grammars was learnable, since this time there were more generalizers for both Pattern 3: kt(S) and Pattern 4: kt(W) than there were for Pattern 1: SW(k) and Pattern 2:

[51] Note that subject 25 was a generalizer for the first block of learning, but was an opposite generalizer for the other three blocks.

WS(k). This suggests that the manipulation of different consonant combinations for new and old consonants did affect the learning outcome.

It seems that there is phonological processing involved. Most of the participants tended to treat a new consonant /t/ as a blocker consistently (i.e., participants knew that consonants would not be nasalized after /t/), suggesting that participants patterned /t/ as a blocker with /k/.

Now that we know that phonological processing is involved, we can consider whether the findings support Possibility 1 (sonority hierarchy type) or 2 (sonority natural class). The results show that the learning of kt(S) and kt(W) are much better than the learning of WS(k) and SW(k). What does that imply? Does it imply that /k/ as blocker→/t/ as potential blocker is easier for participants to generalize than /s/ as target→/w/ as potential target or /w/ as target→/s/ potential target?

kt(W) had more generalizers than kt(S). Does that mean that the findings support the hypothesis concerning reference to sonority natural classes, with the learning of kt(W) expected to be better than that of kt(S)?

The current individual data for Patterns 3 and 4 appears to support the hypothesis that sonority natural classes play a role. In the next section, I will combine individual data with inferential statistics to further determine which possibility, possibility 1 or 2, is more plausible or whether both possibilities could hold.

5.11 General discussion: Patterns 1-4

In Section 5.6.5, I argued that the sonority natural class hypothesis seems more plausible if only Patterns 1 and 2 are taken into account. In this section I argue that both sonority hierarchy type and sonority natural classes play a role in artificial grammar learning once Patterns 1-4 are taken into account.

5.11.1 Inferential statistics

If we take a closer look at the inferential statistics, they do not falsify the hypothesis concerning the direction of implicational universals (i.e., there is reference to the sonority hierarchy type: Vowels > Glides > Liquids > Fricatives > Stops). There is a positive main effect of "Direction" between SW(k) and WS(k) ($p<.05$) (see Table 31 in Section 5.3), implying that participants in the SW(k) group learned better than participants in the WS(k) group. This corresponds to the sonority hierarchy type prediction that SW(k) is expected to be learned better than WS(k).

However, this could be an epiphenomenon of having /k/ as a blocker, with obstruent /k/ priming obstruent /s/ as a blocker but not sonorant /w/. If this is the case, the inferential statistics would support the sonority natural class hypothesis that in Pattern 2 an old blocker /k/ would prime a potential target /s/ to be a blocker because /k/ and /s/ are obstruents, but no such priming effect would occur in Pattern 1. Under this account, the learning of Pattern 1: SW(k) was better than that of Pattern 2: WS(k), as can be seen from the percentages of correct responses for the new-gen items for 6 opposite generalizers in Pattern 2 are close to 20%. That is, it is not because Pattern 1: SW(k) has more generalizers than Pattern 2: WS(k), as predicted by the sonority hierarchy type hypothesis. Rather, it is because Pattern 2: WS(k) has more opposite generalizers than Pattern 1: SW(k), as predicted by the sonority natural class hypothesis.

The hypothesis of sonority hierarchy type does not make a direct prediction about the relationship between Pattern 3: kt(S) and Pattern 4: kt(W). It only predicts that there will be no generalization privilege in kt(S) vs. tk(S), nor in kt(W) vs. tk(W).

5.11.2 Learner types

The descriptive statistics support the hypothesis of sonority natural classes that participants were trying to compare whether a new segment's phonological class is closer to any old segment's phonological class. Recall the expectation – participants would pattern the two segments together, assigning the same status, either target or blocker. Consider Pattern 1: S(k)→W(k) and Pattern 2: W(k)→S(k) for example (see Table 39 and Table 40 in Section 5.5). Both patterns have few generalizers (S(k)→W(k): 3 generalizers; W(k)→S(k): 2 generalizers). But W(k)→S(k) has more opposite generalizers than S(k)→W(k): 2 opposite generalizers; W(k)→S(k): 6 opposite generalizers). From the perspective of reference to sonority natural classes, this suggests that for the WS(k) pattern, participants tried to pattern /k/ as a blocker with /s/, since both /k/ and /s/ are obstruents. Thus WS(k) had more opposite generalizers (i.e., treating /s/ as a blocker instead of as a potential target) because speakers patterned /k/ with /s/, disregarding /w/. On the other hand, for the SW(k) pattern, maybe the reason that few participants made a generalization is because participants thought that /s/ and /w/ did not belong to the same phonological class, so they tended not to treat a new consonant /w/ as a target. They failed to pattern /k/ as a blocker with /w/ (opposite generalizer), because /k/ and /w/ did not belong to the same phonological class.

The hypothesis about reference to sonority natural classes is also applicable to the follow-up experiments, Patterns 3 and 4: k(S)→t(S) vs. k(W)→t(W). The descriptive statistics show that kt(S) has 4 generalizers, while kt(W) has 9 generalizers. These results support the prediction that if participants pattern a new segment depending on the phonological class of old segments (that is, taking both old target and blocker into account), then it is expected that kt(S) might be harder to learn than kt(W), since the sonority distance between /s/ and a new segment /t/ is shorter than that between /w/ and /t/[52]. Specifically, for the kt(S) group, though /k/ and /t/ belong to the same sonority class, stop, the fact that /s/ is a fricative might create difficulty for participants to determine whether /t/ should pattern with /k/ as a blocker or with /s/ as a target (i.e., /t, k, s/ belong to obstruents). On the contrary, for the kt(W) group, the distance between /w/ and /t/ on the hierarchy is further, which makes it easier for participants to pattern a new segment /t/ with /k/ as a blocker.

In brief, the results of Patterns 1-4 support the sonority natural class hypothesis. The sonority hierarchy effects are present but not strong, as I did not find that many generalizers in Pattern 1 (although I did find fewer generalizers in Pattern 2, as predicted). Weak sonority hierarchy effects do not mean that the sonority hierarchy type hypothesis is rejected. The sonority hierarchy type hypothesis would be rejected only if I found counterevidence that there were many opposite generalizers in Pattern 1. Specifically, finding many opposite generalizers in Pattern 1 would suggest that Pattern 1 participants did not generalize to a new consonant as a target, as the sonority hierarchy type hypothesis would predict. Rather, they generalized to a new consonant as a blocker. This contradicts the sonority hierarchy hypothesis, which we can reject in light of the counterevidence.

5.12 Discussion: interactive approach vs. pure sonority natural classes

The inferential statistics provide evidence that the sonority hierarchy type plays an important role in learning since a main effect of "Direction" is found between Pattern 1: SW(k) and Pattern 2: WS(k), as predicted. The individual responses (learner types) provide evidence that sonority

[52] /s/ and /t/ share the same place of articulation ([coronal]).

natural classes (new segment is of the same sonority natural class as the blocker) play an important role in learning. In Patterns 3: k(S)→t(S) and 4: k(W)→t(W), sonority differences in blocker do not occur (/k, t/ = obstruents), and thus direction is not involved and evidence for natural classes is clear. Though the sonority hierarchy type influences Patterns 1 (S(k)→W(k)) and 2 (W(k)→S(k)) more, sonority natural classes also affect the learning. This is evident in Pattern 2, which has more opposite generalizers – participants group test segment /s/ with /k/ (referring to Table 39 and Table 40 in Section 5.5) as are both obstruents. These findings suggest that the sonority hierarchy and sonority natural classes both are involved in learning. In the rest of thesis, I call this **the interactive approach**.

However, there is also a possibility that the inferential statistics simply corroborate the prediction of sonority natural classes that participants in Pattern 2: WS(k) tend to be opposite generalizers, but confusion about the patterning of the new segment occurs in Pattern 1: SW(k). If this is the case, then the findings support the hypothesis of sonority natural classes alone without evidence for sonority hierarchy type. I call this **the approach of pure sonority natural classes**.

5.13 Revisiting the results of Experiment 1

In Section 5.4 I mentioned that the group analysis masks some important generalizations such as whether participants learn a pattern and generalize from an old consonant to a new consonant, and I established criteria to define individual learner types. I now use the individual learner types proposed in Section 5.4 to revisit the results of Patterns A and B (s(W)→k(Y) vs. k(W)→s(Y)) and Patterns C and D (p(W)→k(Y) vs. k(W)→p(Y) discussed in Chapter 4.

In Section 4.3, the averaged data for each pattern (Pattern A vs. Pattern B) was presented. However, this fails to capture the fact that different individuals might use different learning strategies to make judgments. Therefore, here I focus on each individual's percentages of correct responses, as in Table 47 (Pattern A) and Table 48 (Pattern B).

I bold any percentage equal to or higher than 60% and underline if both grammatical and corresponding ungrammatical percentages are equal or over 60%. The results are messy. Participants only got old-same class right (Pattern A: participants 2, 4, 6, 9 and 10; Pattern B: participants 2, 4, 7, and 10), not mentioning new items. Though subject 7 in Pattern B got new-same class right, s/he did not get old-same class items right, suggesting that s/he did not learn the pattern.

This further confirms the concern that participants failed to make generalizations, likely because there were too many new segments in the test phase.

Table 47. Pattern A: the individual response rates for test items (s→k)

s→k	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**k**iyi/ /ãỹĩ**k**i/ /ã**k**i**k**i/	/ãhiyi/ /ãỹĩhi/ /ãhihi/ /ãỹĩỹĩ/	/ã**s**awa/ /ãw̃ã**s**a/ /ã**s**a**s**a/ /ãw̃ãw̃ã/		/ã**k̃**ĩyi/ /ãyi**k**i/ /ã**k̃**ĩ**k̃**ĩ/	/ãh̃ĩyi/ /ãyihi/ /ãh̃ĩh̃ĩ/ /ãyiyi/	/ã**s̃**ãwa/ /ãwa**s**a/ /ã**s̃**ã**s̃**ã/ /ãwawa/	
subj1	**67%**	**69%**	94%	77%	50%	31%	**69%**	50%
subj2	0%	38%	<u>**100%**</u>	50%	50%	44%	<u>**94%**</u>	**64%**
subj3	25%	38%	**100%**	57%	**92%**	56%	19%	52%
subj4	50%	38%	<u>**88%**</u>	59%	42%	31%	<u>**81%**</u>	52%
subj5	50%	50%	**100%**	**68%**	25%	**75%**	38%	48%
subj6	50%	**69%**	<u>**69%**</u>	**64%**	25%	31%	<u>**75%**</u>	45%
subj7	**100%**	**100%**	**88%**	**95%**	0%	0%	38%	14%
subj8	33%	50%	**81%**	57%	25%	44%	50%	41%
subj9	**75%**	50%	<u>**94%**</u>	**73%**	33%	44%	<u>**63%**</u>	48%
subj10	17%	31%	<u>**100%**</u>	52%	50%	56%	<u>**94%**</u>	**68%**
mean	47%	53%	**91%**	65%	29%	41%	**62%**	48%

Table 48. Pattern B: the individual response rates for test items (k→s)

k→s	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**s**iyi/ /ãỹĩ**s**i/ /ã**s**i**s**i/	/ãpiyi/ /ãỹĩpi/ /ãpipi/ /ãỹĩỹĩ/	/ã**k**awa/ /ãw̃ã**k**a/ /ã**k**a**k**a/ /ãw̃ãw̃ã/		/ã**s̃**ĩyi/ /ãyi**s**i/ /ã**s̃**ĩ**s̃**ĩ/	/ãp̃ĩyi/ /ãyipi/ /ãp̃ĩp̃ĩ/ /ãyiyi/	/ã**k̃**ãwa/ /ãwa**k**a/ /ã**k̃**ã**k̃**ã/ /ãwawa/	
subj1	**83%**	44%	**100%**	**75%**	33%	44%	88%	57%
subj2	25%	**63%**	**<u>75%</u>**	57%	**67%**	50%	**<u>75%</u>**	**64%**
subj3	58%	**69%**	**94%**	**75%**	8%	19%	44%	25%
subj4	58%	**63%**	**<u>94%</u>**	**73%**	42%	50%	**<u>69%</u>**	55%
subj5	50%	56%	**100%**	**70%**	33%	44%	31%	36%
subj6	50%	**69%**	**100%**	**75%**	33%	56%	44%	45%
subj7	58%	**<u>63%</u>**	**88%**	**70%**	50%	**<u>69%</u>**	38%	52%
subj8	0%	**94%**	**88%**	**66%**	50%	31%	56%	45%
subj9	17%	56%	**44%**	41%	8%	44%	50%	36%
subj10	58%	**63%**	**<u>100%</u>**	**75%**	42%	50%	**<u>75%</u>**	57%
mean	46%	64%	**88%**	68%	37%	46%	**57%**	47%

Table 49 shows that in Pattern C subjects 3, 8, and 9 learned old segments, as shown by accuracy percentages equal to or higher than 60%, but none of them learned new items. That is, Pattern C participants identified old items as correct but did not generalize to new segments (see Table 49).

Table 49. Pattern C: the response rates for test items (p→k)

p→k	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**k**iyi/ /ãỹĩ**k**i/ /ã**k**i**k**i/	/ãtiyi/ /ãỹĩti/ /ãtiti/ /ãỹĩỹĩ/	/ã**p**awa/ /ãw̃ã**p**a/ /ã**p**a**p**a/ /ãw̃ãw̃ã/		/ã**k̃**ĩyi/ /ãyi**k**i/ /ã**k̃**ĩ**k̃**ĩ/	/ãt̃ĩyi/ /ãyiti/ /ãt̃ĩt̃ĩ/ /ãyiyi/	/ã**p̃**ãwa/ /ãwa**p**a/ /ã**p̃**ã**p̃**ã/ /ãwawa/	
subj1	33%	50%	**94%**	**61%**	58%	25%	0%	25%
subj2	17%	44%	**100%**	57%	**92%**	56%	6%	48%
subj3	**83%**	**81%**	**100%**	**89%**	33%	31%	**69%**	45%
subj4	58%	50%	**100%**	**70%**	25%	56%	6%	30%
subj5	25%	50%	**81%**	55%	**67%**	31%	56%	50%
subj6	42%	31%	**94%**	57%	**75%**	**69%**	19%	52%
subj7	33%	44%	**100%**	**61%**	**67%**	50%	44%	52%
subj8	58%	**81%**	**100%**	**82%**	58%	38%	**69%**	55%
subj9	50%	50%	**100%**	**68%**	58%	**63%**	**94%**	**73%**
subj10	33%	**69%**	**81%**	**64%**	42%	**63%**	44%	50%
mean	43%	55%	**95%**	**66%**	58%	48%	41%	48%

Table 50 (Pattern D) shows that subjects 5, 6, 7, and 8 learned old segments, as shown by accuracy percentages equal to or above 60%. Only subject 8 learned the new-same class items.

Table 50. Pattern D: the response rates for test items (k→p)

k→p	grammatical items (n=44)			average	ungrammatical items (n=44)			average
	new		old		new		old	
	gen (n=12)	same class (n=16)	same class (n=16)		gen (n=12)	same class (n=16)	same class (n=16)	
	/ã**p**iyi/ /ãỹĩ**p**i/ /ã**p**i**p**i/	/ãtiyi/ /ãỹĩti/ /ãtiti/ /ãỹĩỹĩ/	/ã**k**awa/ /ãw̃ã**k**a/ /ã**k**a**k**a/ /ãw̃ãw̃ã/		/ã**p̃**ĩyi/ /ãyi**p**i/ /ã**p̃**ĩ**p̃**ĩ/	/ãt̃ĩyi/ /ãyiti/ /ãt̃ĩt̃ĩ/ /ãyiyi/	/ã**k̃**ãwa/ /ãwa**k**a/ /ã**k̃**ã**k̃**ã/ /ãwawa/	
subj1	17%	50%	**94%**	57%	33%	50%	19%	34%
subj2	**75%**	**69%**	**94%**	**80%**	42%	31%	50%	41%
subj3	**67%**	**75%**	**75%**	**73%**	0%	38%	25%	23%
subj4	50%	31%	94%	59%	50%	50%	31%	43%
subj5	**67%**	56%	**94%**	**73%**	33%	**63%**	**69%**	57%
subj6	**67%**	56%	**100%**	**75%**	58%	**75%**	**94%**	**77%**
subj7	8%	31%	**100%**	50%	42%	38%	**88%**	57%
subj8	**83%**	**63%**	**100%**	**82%**	42%	**63%**	**69%**	59%
subj9	50%	**69%**	**100%**	**75%**	33%	19%	44%	32%
subj10	50%	50%	**88%**	**64%**	25%	25%	13%	20%
mean	53%	55%	**94%**	**69%**	27%	45%	50%	44%

The above four tables clearly show that most participants for Patterns C and D did not generalize to new segments since the accuracy for the pairs of grammatical and ungrammatical counterparts for new-same class and new-gen items is not equal to or higher than 60% (except for subject 8).

5.14 General discussion of Experiment 1

Recall from Section 4.4.2 that no main effect of "Direction" was found in Experiment 1. This could be interpreted in two ways. The first interpretation is that participants in both Patterns C and D learned the pattern equally well, so there is no main effect of "Direction". The second interpretation is that participants in these patterns did not learn the pattern and performed equally poorly, so no main effect of "Direction" was found. The failure of learning new items for the individual data for Patterns C and D further confirms a concern that participants did not make generalizations. I attribute this to design problems since there is learning with the design for Experiment 2.

A more precise way to see which interpretation is more plausible is to look at individual learner types. If we simply look at the logistic regression results shown in Section 4.4.2 or at the grouped descriptive statistics in Section 4.4.1, then it is hard to tell whether individuals learned the pattern even if taking random variables, participants and items into account and incorporate them in logistic regression modeling. But when we examine individual data, it is easy to tell whether individuals learned the pattern or not in terms of the predictions of learner types. Specifically, the individual data show that at least some participants performed well on the old items, a point missed by the grouped descriptive statistics.

5.15 Summary

In this chapter I considered two hypotheses to account for the results of Experiment 2, sonority hierarchy and sonority natural classes. In brief, the hypothesis of sonority natural classes fully accounts for the findings of Patterns 1-4. The fewer generalizers in Pattern 1: SW(k) could also be ascribed to a weak effect of sonority hierarchy, however, suggesting that the sonority hierarchy also plays a role. In this case, the interactive approach is supported.

The next chapter presents follow-up experiments to further test the interaction of the sonority hierarchy type and sonority natural classes-sonority/manner. It also explores the possibility that participants use other natural classes – [continuant] and/or place of articulation in patterning.

Chapter 6 Experiment 3: sonority effects

The primary goal of this chapter is to better test the interactive approach – sonority hierarchy type & sonority natural classes and the approach of pure sonority natural classes proposed in Chapter 5 (see Section 5.12). Specifically, there are two major types of effects that might occur, hierarchy type and natural class type. The sonority hierarchy type hypothesis predicts that if a more sonorant consonant is a blocker, a less sonorant consonant will also be. The sonority natural class hypothesis predicts two classes, sonorants and obstruents. As discussed in Chapter 5, the findings of Experiment 2 show that the sonority natural class hypothesis accounts for the data completely. On the contrary, the sonority hierarchy type does not seem to account for the findings of Patterns 1 and 2 well, given the fact that Pattern 1 does not have substantially more generalizers than Pattern 2.

It is useful to consider the consonants that I am testing in terms of features, shown in Table 51. For convenience, I use binary features.

Table 51. Natural classes & sonority hierarchy type

a. natural classes /t, k, s, w/

	t	k	s	w
sonorant	-	-	-	+
continuant	-	-	+	+
labial	-	-	-	+
coronal	+	-	+	-
dorsal	-	+	-	+

b. sonority hierarchy type

	w	> s	> t, k
sonorant	+	-	-
continuant	+	+	-

In terms of manner natural classes, /t/ and /k/ are expected to pattern together as both are [-sonorant, -continuant]. Crucially, /s/ might pattern with /t/ and /k/ (both [-sonorant]), or with /w/ (both [+continuant]). In this section I test some additional patterns to try to distinguish the hypotheses. A full set of combinations, 8 patterns, is presented in Table 52. In Chapter 5 I reported the results for Patterns 1: S(k)→W(k), Pattern 2: W(k)→S(k), Pattern 3: k(S)→t(S), and Pattern 4: k(W)→t(W). I test Patterns 5, 6, 7, and 8 in this chapter.

Table 52. Full sets of patterns: features

	exposure	test
Pattern 1 SW(k)	target /s/ [-sonorant, +continuant, +coronal] blocker /k/ [-sonorant, -continuant, +dorsal]	new segment /w/ [+sonorant, +continuant, +dorsal]
Pattern 2 WS(k)	target /w/ [+sonorant, +continuant, +dorsal] blocker /k/ [-sonorant, -continuant, +dorsal]	new segment /s/ [-sonorant, +continuant, +coronal]
Pattern 3 kt(S)	target /s/ [-sonorant, +continuant, +coronal] blocker /k/ [-sonorant, -continuant, +dorsal]	new segment /t/ [-sonorant, -continuant, +coronal]
Pattern 4 kt(W)	target /w/ [+sonorant, +continuant, +dorsal] blocker /k/ [-sonorant, -continuant, +dorsal]	new segment /t/ [-sonorant, -continuant, +coronal]
Pattern 5 tk(S)	target /s/ [-sonorant, +continuant, +coronal] blocker /t/ [-sonorant, -continuant, +coronal]	new segment /k/ [-sonorant, -continuant, +dorsal]
Pattern 6 tk(W)	target /w/ [+sonorant, +continuant, +dorsal] blocker /t/ [-sonorant, -continuant, +coronal]	new segment /k/ [-sonorant, -continuant, +dorsal]
Pattern 7 SW(t)	target /s/ [-sonorant, +continuant, +coronal] blocker /t/ [-sonorant, -continuant, +coronal]	new segment /w/ [+sonorant, +continuant, +dorsal]
Pattern 8 WS(t)	target /w/ [+sonorant, +continuant, +dorsal] blocker /t/ [-sonorant, -continuant, +coronal]	new segment /s/ [-sonorant, + continuant, +coronal]

Overall these eight patterns involve testing between sonority classes (exposure to [-sonorant] as a target→testing if [+sonorant] would be treated as a target or vice versa, see Patterns 1, 2, 7, 8) and within a sonority class (exposure to [-sonorant] as a blocker→testing if [-sonorant] would be treated as a blocker, see Patterns 3, 4, 5, 6).

While I have focused on the feature [sonorant] in examining natural classes, note that the feature [continuant] also plays a role in the sonority hierarchy. I examine whether [continuant] is important in Section 6.10.

Patterns 1-8 are minimally different. In Patterns 3 versus 5 and Patterns 4 versus 6 the place of articulation of the stops in the exposure vs. test phases is switched. Patterns 7-8 and are minimally different from Patterns 1-2 in that the blocker is a coronal /t/ rather than a dorsal /k/. This switch of place of articulation allows us to examine whether place of articulation influences the learning. This topic will be addressed in Section 6.12.

In brief, the experiments reported in Chapter 5 led me to distinguish (1) an interactive approach where sonority hierarchy type being equal, sonority natural classes are seen, and (2) a pure sonority natural class approach. Experiment 3, reported in this chapter, is designed to further test predictions of the sonority hierarchy type and sonority natural class hypotheses. Continuancy and place of articulation will be discussed in this chapter.

6.1 Predictions: directionality (sonority hierarchy type)

In this section I present the predictions for patterns 1-8 under the hierarchy type hypothesis. Recall that the sonority hierarchy type can be directly tested by directionality ("Direction") by comparing two patterns. The predictions about sonority hierarchy type for the eight patterns are presented in Table 53.

For instance, recall from Section 5.11.1 that the sonority hierarchy type hypothesis does not make a direct prediction about the relationship between Pattern 3: kt(S) and Pattern 4: kt(W). It only predicts that there will be no generalization privilege in Pattern 3: kt(S) vs. Pattern 5: tk(S), nor in Pattern 4: kt(W) vs. Pattern 6: tk(W) because the directionality of sonority distance is the same for /k/ and /t/ (obstruents): no main effect of "Direction" is expected between Patterns 3 and 5, or between Patterns 4 and 6.

On the other hand, the sonority directionality between Patterns 1: SW(k) and 2: WS(k) is the same as Patterns 7: SW(t) and 8: WS(t). Therefore, it is predicted that a new consonant /w/ (sonorant) is generalizable in Pattern 7, but there is no such prediction for /s/ (obstruent) for Pattern 8. In this case, a main effect of "Direction" is expected to be found: the learning of Pattern 7 would be significantly better than that of Pattern 8.

Table 53. Full sets of patterns: sonority hierarchy type

	exposure	test	prediction: sonority hierarchy type
Pattern 1 SW(k)	more sonorant: target /**s**/ less sonorant: blocker /k/	new target /**w**/: more sonorant than target	new segment is a target
Pattern 2 WS(k)	more sonorant: target /**w**/ less sonorant: blocker /k/	new target /**s**/: less sonorant than target	no prediction
Pattern 3 kt(S)	more sonorant: target /s/ less sonorant: blocker /**k**/	new blocker /**t**/: same class as blocker	new segment is blocker
Pattern 4 kt(W)	more sonorant: target /w/ less sonorant: blocker /**k**/	new blocker /**t**/: same class as blocker	new segment is blocker
Pattern 5 tk(S)	more sonorant: target /s/ less sonorant: blocker /**t**/	new blocker /**k**/: same class as blocker	new segment is blocker
Pattern 6 tk(W)	more sonorant: target /w/ less sonorant: blocker /**t**/	new blocker /**k**/: same class as blocker	new segment is blocker
Pattern 7 SW(t)	more sonorant: target /**s**/ less sonorant: blocker /t/	new target /**w**/: more sonorant than target	new segment is a target
Pattern 8 WS(t)	more sonorant: target /**w**/ less sonorant: blocker /t/	new target /**s**/: less sonorant than target	no prediction

Patterns 5-8 allow us to examine if there is a learning asymmetry when two patterns have the opposite directionality of sonority classes (Pattern 7: obstruent as a target→sonorant and Pattern 8: sonorant as a target→obstruent) and if there is no significant learning difference when two patterns involve directionality within a sonority class (obstruent→obstruent, Patterns 5 and 6).

6.2 Sonority natural classes:

In this section I summarize the predictions for Patterns 1-8 given the sonority natural class hypothesis in Table 54. For clarity, bolded segments within each pattern means that those segments are under study. In general, participants would try to pattern a new segment with old segments in terms of [sonorant]. If they would find a shared value for the feature [sonorant] between a new segment and an old segment, then they would pattern these two together. However, if there is not a one-to-one matching relation, they would be confused about what to pattern a new segment with.

Table 54. Predictions about sonority natural classes for Patterns 1-8

	exposure	test	prediction: sonority natural classes
Pattern 1 SW(k)	**[-sonorant]**: target /**s/** [-sonorant]: blocker /k/	**[+sonorant]**: new target /**w/**	confused about what to pattern test segment /w/ with (/w/ = sonorant)
Pattern 2 WS(k)	**[+sonorant]**: target /**w/** [-sonorant]: blocker /k/	**[-sonorant]**: new target /**s/**	opposite generalizers (pattern test segment /s/ with /k/) (/s, k/ = obstruents)
Pattern 3 kt(S)	[-sonorant]: target /s/ **[-sonorant]**: blocker /**k**/	**[-sonorant]**: new blocker /**t**/	confused what to pattern test segment /t/ with) (/t, s, k/ = obstruents)
Pattern 4 kt(W)	[+sonorant]: target /w/ **[-sonorant]**: blocker /**k**/	**[-sonorant]**: new blocker /**t**/	generalizers (pattern test segment /t/ with /k/) (/t, k/ = obstruents)
Pattern 5 tk(S)	[-sonorant]: target /s/ **[-sonorant]**: blocker /**t**/	**[-sonorant]**: new blocker /**k**/	**confused about what to pattern test segment /k/ with** (/k, t, s/ = obstruents
Pattern 6 tk(W)	[+sonorant]: target /w/ **[-sonorant]**: blocker /**t**/	**[-sonorant]**: new blocker /**k**/	**generalizers** (pattern test segment /k/ with /t/) (/k, t/ = obstruents)
Pattern 7 SW(t)	**[-sonorant]**: target /**s**/ [-sonorant]: blocker /t/	**[+sonorant]**: new target /**w**/	**confused about what to pattern test segment /w/ with** (/w/ = sonorant; /s, t/ = obstruents)
Pattern 8 WS(t)	**[-sonorant]**: target /**w**/ [-sonorant]: blocker /t/	**[-sonorant]**: new target /**s**/	**opposite generalizers** (pattern test segment /s/ with /t/) (/s, t/ = obstruents)

I expect that there will be more generalizers for Pattern 6: **t(W)→k(W)** than for Pattern 5: **t(S)→k(S)**, similar to Pattern 4: kt(W) vs. Pattern 3: kt(S). Specifically, for Pattern 6: tk(W), the sonority of the old segment /w/, a sonorant, is distinct from that of the new segment /k/ and of the old segment /t/, both obstruents, making it easier for participants to pattern a new segment /k/ with an old segment /t/ that acts as a blocker, as shown in Table 54 under "prediction: sonority natural classes". However, for Pattern 5: tk(S), /t, s, k/ are all obstruents, which will confuse participants about what the new segment /k/ should pattern with, as shown in Table 54 under "prediction: sonority natural classes".

As for Pattern 7, participants would have no idea about what to pattern a new segment /w/ with because /s/ and /t/ are both obstruents. With respect to Pattern 8, the hypothesis about sonority natural classes predicts that since /s, t/ are obstruents, while /w/ is a sonorant, participants would be inclined to pattern a new segment /s/ with an old blocker /t/ rather than with /w/, yielding opposite generalizers. The predictions are presented in Table 54 under "prediction: sonority natural classes".

Recall that sonority natural classes are tested by individual learner types. I look at whether the number of categorization learners is asymmetrical between the two patterns. For categorization learners, if more than 5 participants belong to generalizer or opposite generalizer, I conclude that they had a tendency to pattern a new segment with an old segment in some way. The subcategories of generalizers and opposite generalizers are crucial, suggesting that participants are able to generalize from an old vowel to a new vowel (new-same class items) as well as from an old consonant to a new consonant (new-gen items) (see Section 5.1.1).

In Table 54, bolding of a segment within each pattern indicates that that segment is under study. For instance, in Pattern 5: t(S)→k(S), bolded /t/ in exposure means that /t/ as a blocker appeared in the exposure phase, while non-bolded /s/ as a target appeared in both exposure and test phases. Under the sonority natural class hypothesis, participants are expected to pattern a new potential blocker /k/ with an old blocker /t/ (generalizers) rather than an old target /s/ (opposite generalizers).

6.3 Continuancy

In Section 4.4 I noted that Min speakers have a two-way opposition – obstruent/sonorant. However, the full sonority hierarchy type requires the features [sonorant] and [continuant] (see Table 52). The feature [continuant] is not involved in defining natural classes in this experiment. If this is the case, I could further modify the argument proposed in Section 4.4 – Min participants do not show the three-way opposition between [+sonorant, +continuant], [-sonorant, +continuant], and [-sonorant, -continuant], but rather, they simply distinguish [+sonorant] and [-sonorant] in this experiment. I consider the continuancy natural class hypothesis below.

Patterns 2: WS(k) and 8: WS(t) are particularly interesting because the involvement of the feature [continuant] would have different predictions from the sonority hierarchy type hypothesis. Specifically, /s/ is [-sonorant, +continuant] and /w/ is [+sonorant, +continuant]. Crucially, as mentioned at the beginning of this chapter, if [continuant] is considered, /s/ might pattern with /t/ and /k/ (both [-sonorant]), or with /w/ (both [+continuant]). Specifically, in both Patterns 2: WS(k) and 8: WS(t), /s/ and /w/ could pattern together as [+continuant]. In this case, it is expected to find no main effect of "Direction" between Pattern 1: SW(k) and Pattern 2: WS(k), and between Pattern 7: SW(t) and Pattern 8: WS(t). This is different from the sonority hierarchy type hypothesis, predicting that it is expected to find a main effect of "Direction" between Patterns 1 and 2, and between Patterns 7 and 8. In general, participants are expected could be able to generalize to a new target (Patterns 1, 2, 7, 8), and to a new blocker (Patterns 3, 4, 5, 6). Therefore, no main effect of "Direction" is expected between Patterns 1 and 2, between Patterns 7 and 8, between Patterns 3 and 5, and Patterns 5 and 6.

Similarly, the sonority natural class findings match the predictions of natural class based on [sonorant], but not those of natural class based on [continuant] for the crucial patterns, Patterns 2, 3, 5, and 8. The continuancy natural classes hypothesis predicts that participants in these patterns would be able to generalize to a test target (generalizers, Patterns 1, 2, 7, 8), and to a test blocker (generalizers, Patterns 3, 4, 5, 6), grouping together /s and /w/, as shown in Table 55.

Table 55. Continuancy natural classes

	exposure	test	prediction: continuancy natural class
Pattern 1 SW(k)	**[+continuant]**: target /**s**/ [-continuant]: blocker /k/	new target /**w**/: (**[+continuant]**)	many generalizers (pattern test target /w/ with /s/ = [+continuant])
Pattern 2 WS(k)	**[+continuant]**: target /**w**/ [-continuant]: blocker /k/	new target /**s**/: (**[+continuant]**)	many generalizers (pattern test target /s/ with /w/) (/s, w/ = [+continuant])
Pattern 3 kt(S)	[+continuant]: target /s/ **[-continuant]**: blocker /**k**/	new blocker /**t**/: (**[-continuant]**)	many generalizers (pattern test blocker /t/ with /k/) (/t, k/ = [-continuant])
Pattern 4 kt(W)	[+continuant]: target /w/ **[-continuant]**: blocker /**k**/	new blocker /**t**/: (**[-continuant]**)	many generalizers (pattern test blocker /t/ with /k/) (/t, k/ = [-continuant])
Pattern 5 tk(S)	[+continuant]: target /s/ **[-continuant]**: blocker /**t**/	new blocker /**k**/: (**[-continuant]**)	**many generalizers** (pattern test blocker /k/ with /t/) (/k, t/ = [-continuant])
Pattern 6 tk(W)	[+continuant]: target /w/ **[-continuant]**: blocker /**t**/	new blocker /**k**/: (**[-continuant]**)	**many generalizers** (pattern test blocker /k/ with /t/) (/k, t/ = [-continuant])
Pattern 7 SW(t)	**[+continuant]**: target /**s**/ [-continuant]: blocker /t/	new target /**w**/: (**[+continuant]**)	**many generalizers** (pattern test target /w/ with /s/) (/w, s/ = [+continuant]
Pattern 8 WS(t)	**[+continuant]**: target /**w**/ [-continuant]: blocker /t/	new target /**s**/: (**[+continuant]**)	**many generalizers** (pattern test segment /s/ with /w/) (/s, w/ = [+continuant])

6.4 Summary

In summary, Experiment 3 aims to test the predictions of the sonority hierarchy type hypothesis and the sonority natural class hypothesis to see if the interactive approach or the pure sonority natural class approach is upheld, ruling out continuancy and place of articulation as confounds.

In Section 6.5, methods are presented. For the purpose of clarity, I discuss directionality (sonority hierarchy type) for Patterns 3, 4, 5, and 6 in Section 6.6, and Patterns 7 and 8 in Section 6.7. I then discuss sonority natural classes in Sections 6.8 and 6.9. In Section 6.10, continuancy is examined. In Section 6.12, place of articulation is examined.

6.5 Methods

See Section 5.2 for the overall design and the forced choice method. The monosyllabic stimuli for Experiment 3 are the stimuli recorded for Experiment 1 (see Section 4.2.2.2) by the same male native speaker of Min. The concatenation of tri-syllabic words was done to fit the current design.

6.5.1 Materials

In this section, I present the materials for Patterns 5, 6, 7, and 8.

In general, Patterns 5 tk(S) and 6 tk(W) are similar to Patterns 3 kt(S) and 4 kt(W). The only difference between these concerns whether participants were exposed to /t/ or to /k/, and then were tested on /k/ or on /t/. Specifically, participants were exposed to /asata/ ~ /ãs̃ã**t**a/; then tested with /k/ to see if they generalized (/asa**k**a/ ~ /ãs̃ã**k**a/ ~ *ãs̃ãk̃ã) (Pattern 5: t(S)→k(S)). Similarly, the other group of participants were exposed to /awata/ ~ /ãw̃ãta/ and tested with /k/ (/awaka/ ~ /ãw̃ãka/ ~ */ãw̃ãk̃ã/) (Pattern 6: t(W→k(W)). For Pattern 5, /s/ was always a target throughout the exposure and test, while for Pattern 6, /w/ was always a target.

Table 56 and Table 57 and present the exposure items and the test items.

Table 56. Stimuli design (exposure)

	Pattern 5: **t**(S)→ **k**(S)		Pattern 6: **t**(W)→ **k**(W)	
	singular	plural	singular	plural
grammatical	/a**t**asa/	/ã**t**asa/	/a**t**awa/	/ã**t**awa/
	/asa**t**a/	/ãs̃ã**t**a/	/awa**t**a/	/ãw̃ã**t**a/
	/a**t**a**t**a/	/ã**t**a**t**a/	/a**t**a**t**a/	/ã**t**a**t**a/
	/asasa/	/ãs̃ãs̃ã/	/awawa/	/ãw̃ãw̃ã/
	/a**t**asi/	/ã**t**asi/	/a**t**awi/	/ã**t**awi/
	/asa**t**i/	/ãs̃ã**t**i/	/awa**t**i/	/ãw̃ã**t**i/
	/a**t**a**t**i/	/ã**t**a**t**i/	/a**t**a**t**i/	/ã**t**a**t**i/
	/asasi/	/ãs̃ãs̃ĩ/	/awawi/	/ãw̃ãw̃ĩ/
	/a**t**isa/	/ã**t**isa/	/a**t**iwa/	/ã**t**iwa/
	/asi**t**a/	/ãs̃ĩ**t**a/	/awi**t**a/	/ãw̃ĩ**t**a/
	/a**t**i**t**a/	/ã**t**i**t**a/	/a**t**i**t**a/	/ã**t**i**t**a/
	/asisa/	/ãs̃ĩs̃ã/	/awiwa/	/ãw̃ĩw̃ã/
	/a**t**isi/	/ã**t**isi/	/a**t**iwi/	/ã**t**iwi/
	/asi**t**i/	/ãs̃ĩ**t**i/	/awi**t**i/	/ãw̃ĩ**t**i/
	/a**t**i**t**i/	/ã**t**i**t**i/	/a**t**i**t**i/	/ã**t**i**t**i/
	/asisi/	/ãs̃ĩs̃ĩ/	/awiwi/	/ãw̃ĩw̃ĩ/

(t, k=blockers; S,W=targets)

Table 57. Stimuli design (test items)

	Pattern 5: **t**(S)→**k**(S)			Pattern 6: **t**(W)→**k**(W)		
	new		old	new		old
	gen (n=10, pair=5)	same class (n=10, pair= 5)	same class (n=8, pair=4)	gen (n=10, pair= 5)	same class (n=10, pair= 5)	same class (n=8, pair=4)
grammatical (n=44)	/ã**k**asa/	/ãs̃ẽ**t**a/	/ãs̃ã**t**a/	/ã**k**awa/	/ãw̃ẽ**t**a/	/ãw̃ã**t**a/
	/ãs̃ã**k**a/	/ã**t**ase/	/ã**t**asa/	/ãw̃ã**k**a/	/ã**t**awe/	/ã**t**awa/
	/ã**k**aka/	/ãs̃ã**t**e/	/ãs̃ãs̃ã/	/ã**k**aka/	/ãw̃ã**t**e/	/ãw̃ãw̃ã/
	/ã**k**a**k**a/	/ãs̃ẽs̃ẽ/	/ã**t**a**t**a/	/ã**k**a**k**a/	/ãw̃ẽw̃ẽ/	/ã**t**a**t**a/
	/ãka**k**a/	/ã**t**esa/		/ãka**k**a/	/ã**t**ewa/	
ungrammatical (n=44)	/ã**k̃**ãs̃ã/	/ãse**t**a/	/ãsa**t**a/	/ã**k̃**ãw̃ã/	/ãwe**t**a/	/ãwa**t**a/
	/ãs̃ã**k̃**ã/	/ã**t̃**ãse/	/ã**t̃**ãsa/	/ãw̃ã**k̃**ã/	/ã**t̃**ãwe/	/ã**t̃**ãwa/
	/ã**k̃**ãka/	/ãsa**t**e/	/ãsasa/	/ã**k̃**ãka/	/ãwa**t**e/	/ãwawa/
	/ã**k̃**ã**k̃**ã/	/ãsese/	/ã**t̃**ã**t̃**ã/	/ã**k̃**ã**k̃**ã/	/ãwewe/	/ã**t̃**ã**t̃**ã/
	/ãk̃ã**k̃**ã/	/ã**t̃**ẽsa/		/ãk̃ã**k̃**ã/	/ã**t̃**ẽwa/	

(gen=generalizable, n=number)

Pattern 7 SW(t) and Pattern 8 WS(t) are similar to Pattern 1 SW(k) and Pattern 2 WS(k). The only difference between these concerns whether an old blocker is /t/ or /k/.

Table 58 and Table 59 present the exposure items and the test items.

Table 58. Stimuli design (exposure)

	Pattern 7: **S**(t)→ **W**(t)		Pattern 8: **W**(t)→ **S**(t)	
	singular	plural	singular	plural
grammatical	/ata**s**a/	/ãta**s**a/	/ata**w**a/	/ãta**w**a/
	/a**s**ata/	/ã**s̃**ãta/	/a**w**ata/	/ã**w̃**ãta/
	/atata/	/ãtata/	/atata/	/ãtata/
	/a**s**a**s**a/	/ã**s̃**ã**s̃**ã/	/a**w**a**w**a/	/ã**w̃**ã**w̃**ã/
	/ata**s**i/	/ãta**s**i/	/ata**w**i/	/ãta**w**i/
	/a**s**ati/	/ã**s̃**ãti/	/a**w**ati/	/ã**w̃**ãti/
	/atati/	/ãtati/	/atati/	/ãtati/
	/a**s**a**s**i/	/ã**s̃**ã**s̃**ĩ/	/a**w**a**w**i/	/ã**w̃**ã**w̃**ĩ/
	/ati**s**a/	/ãti**s**a/	/ati**w**a/	/ãti**w**a/
	/a**s**ita/	/ã**s̃**ĩta/	/a**w**ita/	/ã**w̃**ĩta/
	/atita/	/ãtita/	/atita/	/ãtita/
	/a**s**i**s**a/	/ã**s̃**ĩ**s̃**ã/	/a**w**i**w**a/	/ã**w̃**ĩ**w̃**ã/
	/ati**s**i/	/ãti**s**i/	/ati**w**i/	/ãti**w**i/
	/a**s**iti/	/ã**s̃**ĩti/	/a**w**iti/	/ã**w̃**ĩti/
	/atiti/	/ãtiti/	/atiti/	/ãtiti/
	/a**s**i**s**i/	/ã**s̃**ĩ**s̃**ĩ/	/a**w**i**w**i/	/ã**w̃**ĩ**w̃**ĩ/

(S,W=targets; t=blocker)

Table 59. Stimuli design (test items)

	Pattern 7: **S**(t)→**W**(t)			Pattern 8: **W**(t)→**S**(t)		
	new		old	new		old
	gen (n=10, pair=5)	same class (n=10, pair= 5)	same class (n=8, pair=4)	gen (n=10, pair= 5)	same class (n=10, pair= 5)	same class (n=8, pair=4)
grammatical (n=44)	(10) /ã**w̃**ãta/ (11) /ãta**w**a/ (12) /ã**w̃**ãs̃ã/ (13) /ã**w̃**ã**w̃**ã/ (14) /ãs̃ã**w̃**ã/	(5) /ã**s̃**ẽta/ (6) /ãta**s**e/ (7) /ã**s̃**ãte/ (8) /ã**s̃**ẽ**s̃**ẽ/ (9) /ãtesa/	(1) /ã**s̃**ãta/ (2) /ãta**s**a/ (3) /ã**s̃**ã**s̃**ã/ (4) /ãtata/	(10) /ã**s̃**ãta/ (11) /ãta**s**a/ (12) /ã**s̃**ã**w̃**ã/ (13) /ã**s̃**ã**s̃**ã/ (14) /ã**w̃**ã**s̃**ã/	(5) /ã**w̃**ẽta/ (6) /ãta**w**e/ (7) /ã**w̃**ãte/ (8) /ã**w̃**ẽ**w̃**ẽ/ (9) /ãtewa/	(1) /ã**w̃**ãta/ (2) /ãta**w**a/ (3) /ã**w̃**ã**w̃**ã/ (4) /ãtata/
ungrammatical (n=44)	(10') /ã**w**ata/ (11') /ãt̃ã**w̃**ã/ (12') /ã**w**asa/ (13') /ã**w**a**w**a/ (14') /ã**s̃ã****w**a/	(5') /ã**s**eta/ (6') /ãt̃ã**s**e/ (7') /ã**s**ate/ (8') /ã**s**e**s**e/ (9') /ãt̃ẽsa/	(1') /ã**s**ata/ (2') /ãt̃ã**s**a/ (3') /ã**s**a**s**a/ (4') /ãt̃ãt̃ã/	(10') /ã**s**ata/ (11') /ãt̃ã**s̃**ã/ (12') /ã**s**awa/ (13') /ã**s**a**s**a/ (14') /ã**w̃ã****s**a/	(5') /ã**w**eta/ (6') /ãt̃ã**w**e/ (7') /ã**w**ate/ (8') /ã**w**e**w**e/ (9') /ãt̃ẽ**w**a/	(1') /ã**w**ata/ (2') /ãt̃ã**w**a/ (3') /ã**w**a**w**a/ (4') /ãt̃ãt̃ã/

(gen=generalizable, n=number)

6.5.2 Participants

The participants were bilingual in Min and Mandarin. Four groups of participants were recruited (Patterns 5, 6, 7, 8). All reported normal hearing. The rate of distinguishing nasalized from oral counterparts (post-test) was 100%. All participants had early childhood Min dialect exposure. 20 participants for Patterns 5 and 6 (10 for each) recruited from National Sun Yat-sen University received 300 NTD for participating in the experiment. 20 participants for Patterns 7 and 8 (10 for each) were recruited from McGill University and the University of Toronto, receiving 15 CAD. The duration of the experiments was 45-50 minutes.

Pattern 5: t(S)→k(S)

All participants had studied a foreign language (English 10, Japanese 2). The average age of the participants was 23.2 (SD = 5.3). 6 participants were males and 4 participants were females.

Pattern 6: t(W)→k(W)

All had studied a foreign language (English 10, Japanese 1). The average age of the participants was 20.9 (SD = 4.5). 5 participants were males and 5 participants were females.

Pattern 7: S(t)→W(t)

All had studied a foreign language (English 10, Japanese 4). The average age of the participants was 29.3 (SD = 2.7). 5 participants were males and 5 participants were females.

Pattern 8: W(t)→S(t)

All had studied a foreign language (English 10, Japanese 4, Spanish 1). The average age of the participants was 28.1 (SD = 4.2). 1 participant was male and 9 participants were females.

6.6 Testing directionality (sonority hierarchy type): Patterns 3-6

In this section I examine the inferential statistics of Patterns 3-6 to see if the findings support the directionality predictions of the sonority hierarchy type hypothesis discussed in Section 6.1. The sonority hierarchy type hypothesis predicts that there will be no generalization privileges in Pattern 3: kt(S) vs. Pattern 5: tk(S), nor in Pattern 4: kt(W) vs. Pattern 6: tk(W) since the direction from /t/ to /k/ and /k/ to /t/ does not involve a change in the sonority hierarchy (/t, k/ = obstruents). I expect to find no main effect of "Direction" between Pattern 3: kt(S) and Pattern 5: tk(S), and no main effect of "Direction" between Pattern 4: kt(W) and Pattern 6: tk(W).

Since the sonority hierarchy type hypothesis does not predict any learning asymmetry between the two patterns, I arbitrarily coded tk(S) as 1, and kt(S) as -1[53] for **"Direction"**. This factor is crucial, since it could directly test whether the experimental results support the sonority hierarchy type hypothesis and sonority natural class hypothesis. The test items were of three types: old-same class, new-same class and new-generalizable. I coded them as O, NS and NGEN. These three categories were grouped into **"TypeName"**. The formula for both by-participant and by participant & item analyses was used to carry out the analysis as in (22) and (23).

[53] I coded t(S)⟶k(S) as 1, simply because /t/ and /s/ share the coronal place of articulation.

(22) Accuracy ~ Direction * TypeName + (1|Participant)

(23) Accuracy ~ Direction * TypeName + (1|Participant) + (1|Item)

No main effect of "Direction" between Pattern 3 and Pattern 5 was found (see Table 60), suggesting that there was no significant learning difference between Pattern 3: tk(S) and Pattern 5: kt(S). This supports the prediction of reference to sonority hierarchy that both patterns would be learned equally well or poorly.

The main effect of "TypeNameO" is positively significant, indicating that old-same items were learned significantly better than new-generalizable items.

The main effect of "TypeNameNS" is positively significant, indicating that new-same-class items were learned significantly better than new-generalizable items.

Table 60. Effects and interaction for Pattern 5: t(S)→k(S) and Pattern 3: k(S)→t(S) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.66803	0.46288	3.604	0.000314 ***
Direction	0.40184	0.38213	1.052	0.292981
TypeNameO	**1.30597**	**0.45368**	**2.879**	**0.003994 ****
TypeNameNS	**1.15539**	**0.41887**	**2.758**	**0.005809 ****
Direction:TypeNameO	-0.13850	0.23291	-0.595	0.552097
Direction:TypeNameNS	-0.04944	0.21268	-0.232	0.816190

'***': *p*<.001, '**': *p*<.01, '*': *p*<.05, '.': *p*<.1

As for Pattern 4 and Pattern 6, since the sonority hierarchy type hypothesis does not predict any learning asymmetry between the two patterns, for "Direction", I arbitrarily coded kt(W) as 1, and tk(W) as -1[54]. No main effect of "Direction" was found, suggesting that there was no significant learning difference between tk(S) and kt(S), as expected by the hypothesis.

[54] I coded k(W)⟶t(W) as 1, simply because /k/ and /w/ share the dorsal place of articulation.

The main effect of “TypeNameO” is positively significant, indicating that old-same items were learned significantly better than new-generalizable items (see Table 61).

Table 61. Effects and interaction for Pattern 6: t(W)→k(W) and Pattern 4: k(W)→t(W) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.91370	0.59112	4.929	8.26e-07 ***
Direction	0.07267	0.39629	0.183	0.85450
TypeNameO	**2.00700**	**0.75843**	**2.646**	**0.00814 ** **
TypeNameNS	0.11126	0.62259	0.179	0.85816
Direction:TypeNameO	0.03260	0.43473	0.075	0.94023
Direction:TypeNameNS	-0.34458	0.25002	-1.378	0.16813

‘***’: p<.001, ‘**’: p<.01, ‘*’: p<.05, ‘.’: p<.1

As predicted, there is no effect of direction. In the next subsection I compare Pattern 7: SW(t) and Pattern 8: WS(t) to see if the results support the sonority hierarchy type hypothesis.

6.7 Testing directionality (sonority hierarchy): Patterns 7 and 8

“Direction” is coded 1 for Pattern 7: SW(t), and -1 for Pattern 8: WS(t) because Pattern 7 is predicted to be generalizable, whereas there is no prediction for Pattern 8. The inferential statistics for both patterns are presented in Table 62.

Table 62. Effects and interaction for Pattern 7: S(t)→W(t) and Pattern 8: W(t)→S(t) (by-participant & item analysis)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.12936	0.44992	-0.288	0.7737
Direction	**0.91627**	**0.44281**	**2.069**	**0.0385 ***
TypeNameO	**4.04935**	**0.39616**	**10.221**	**< 2e-16 *****
TypeNameNS	**2.22938**	**0.48204**	**4.625**	**3.75e-06 *****
Direction:TypeNameO	-0.09115	0.56235	-0.162	0.8712
Direction:TypeNameNS	-0.08329	0.47147	-0.177	0.8598

'***': p<.001, '**': p<.01, '*': p<.05, '.': p<.1

The main effect of "Direction" is positively significant, suggesting that the learning of Pattern 7 was significantly better than that of Pattern 8. This supports the sonority hierarchy type hypothesis.

The main effect of "TypeNameO" is positively significant, suggesting that old items were learned better than new items. This means that memory plays a role in learning.

The main effect of "TypeNameNS" is positively significant, suggesting that the learning of new-same class items was better than that of new-gen items, as expected based on Experiment 2 and on findings in the literature (Finley 2011b).

6.8 Testing sonority natural classes: learner type (Patterns 5 & 6)

Recall from Section 6.2 that individual learner type can be used to test the hypothesis about sonority natural classes. This hypothesis predicts that participants would compare whether a new segment's sonority natural class is close to any old segment's sonority natural class (obstruents/sonorants). If yes, then participants would pattern the two segments together, assigning the same status of nasality function (e.g., blocker). Pattern 5: tk(S) participants are expected to be confused about what to pattern a new segment /k/ with because /t, k, s/ are all obstruents. In this case, I would expect to find the number of generalizers or of opposite generalizers not more than 5. Pattern 6: tk(W) participants are expected to be generalizers because a new segment /k/ can be patterned with an old blocker /t/. That is, the number of generalizers for Pattern 6 is expected to be over 5.

Table 63 shows that Pattern 5 has 5 generalizers, 2 opposite generalizers, and 3 positional statisticians. Pattern 6 has 9 generalizers, and 1 opposite generalizer.

Table 63. Pattern 5: t(S)→k(S) vs. Pattern 6: t(W)→k(W) (individual learner type)

	Pattern 5: tk(S)			Pattern 6: tk(W)		
	exposure		test	exposure		test
	s as target	t as blocker	k as blocker	**w as target**	t as blocker	k as blocker
pattern learner		n=0			n=0	
generalizer		n=5			n=9	
generalizer (opposite)		n=2			n=1	
positional statistician		n=3			n=0	
unbound nasalizer		n=0			n=0	
random		n=0			n=0	

Table 64 presents the individual differences for Patterns 5 and 6.

Table 64. Pattern 5: tk(S) vs. Pattern 6: tk(W) (individual learner type)

	tk(S)	**NGEN**	**NS**	**O**	sum		
Subj	41	**95%**	**100%**	**100%**	**98%**	categorization learner	generalizer
Subj	42	**95%**	**95%**	**100%**	**96%**	categorization learner	generalizer
Subj	43	**95%**	**80%**	**88%**	**88%**	categorization learner	generalizer
Subj	44	**100%**	**85%**	**69%**	**86%**	categorization learner	generalizer
Subj	45	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	46	40%	**100%**	**100%**	**79%**	categorization learner	generalizer (first; the rest: opposite[55])
Subj	47	20%	**100%**	**100%**	**71%**	categorization learner	generalizer (opposite)
Subj	48	**95%**	**60%**	44%	**68%**	statistical learner	positional statistician
Subj	49	45%	**65%**	**81%**	**63%**	statistical learner	positional statistician
Subj	50	40%	**70%**	**94%**	**66%**	statistical learner	positional statistician

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

	tk(W)	**NGEN**	**NS**	**O**	sum		
Subj	51	**95%**	**75%**	**100%**	**89%**	categorization learner	generalizer
Subj	52	**80%**	**85%**	**88%**	**84%**	categorization learner	generalizer
Subj	53	**100%**	**65%**	**88%**	**84%**	categorization learner	generalizer
Subj	54	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	55	**95%**	**100%**	**100%**	**98%**	categorization learner	generalizer
Subj	56	**85%**	**90%**	**100%**	**91%**	categorization learner	generalizer
Subj	57	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	58	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	59	**95%**	**95%**	**100%**	**96%**	categorization learner	generalizer
Subj	60	20%	**100%**	**100%**	**71%**	categorization learner	generalizer (opposite)

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

According to Table 64, Pattern 5: tk(W) has more generalizers than Pattern 6: tk(S) (9 vs. 5), similar to Pattern 3: kt(W) vs. Pattern 4: kt(S) (9 vs. 4). Combined with the data from the previous

[55] Subject 46 acted as a generalizer for the first block, but acted as an opposite generalizer for the rest of three repetitions, so I categorized this participant into "opposite generalizer".

four patterns (SW(k) vs. WS(k); kt(S) vs. kt(W)), this result further strengthens the argument that the design of my artificial grammars was learnable and phonological learning occurred.

The manipulation of different consonant combinations for new and old consonants did affect the learning outcome. There were more generalizers for Pattern 3: kt(S) versus Pattern 4: kt(W) than there were for Pattern 1: SW(k) versus Pattern 2: WS(k) correspondingly. Similarly, there were more generalizers for both Patterns 5: tk(S) and Pattern 6: tk(W) than were Pattern 1: SW(k) vs. Pattern 2: WS(k). In brief, participants consistently tended to treat a new segment /k/ or /t/ as a blocker. The patterning of new segments /s/ and /w/ in Patterns 1 and 2 is relatively unstable (compared with Patterns 3-6), but it is still more common for Pattern 2: WS(k) participants to pattern a new segment /s/ with an old blocker /k/ than for Pattern 1: SW(k) participants to pattern a new segment /w/ with an old blocker /k/, suggesting the role of the sonority natural classes. This also shows evidence that [continuant] was not referred to, otherwise /s/ and /w/ could easily pattern together as [+continuant].

The results of Patterns 5 and 6 further confirm the prediction I made in Section 6.2: Pattern 6 tk(W) would be expected to have more generalizers than Pattern 5: tk(S) (9 vs. 5), similar to Pattern 4: kt(W) vs. Pattern 3: kt(S) (9 vs. 4) because sonorant /w/ in Pattern 6: tk(W) would make it easier for participants to group a new obstruent /k/ with an old blocker, obstruent /t/ (generalizers), rather than with /w/. However, obstruent /s/ in Pattern 5: tk(S) would make it harder for participants to decide what to group a new segment /k/ with because a new segment /k/, an old blocker /t/, and an old target /s/ are all obstruents. This supports the hypothesis of sonority natural classes (old blocker /t/, either target /s/ in Pattern 5 or target /w/ in Pattern 6). Specifically, this suggests that participants also tried to compare a new segment (/k/) with an old segment and to group the new segment with one of the old segments in a sonority natural class.

In sum, the results support the sonority natural class hypothesis that participants show a clear patterning if they could group a new consonant with an old consonant in the same sonority natural class ([sonorant]), setting aside another old consonant in a different sonority natural class. If three consonants are in the same sonority natural class or in natural classes that are close in sonority as in fricatives and stops, then participants would show confusion. As predicted, participants in Pattern 5: tk(S) were confused about whether they should pattern a new segment /k/ with old blocker /t/ or old target /s/: the number of generalizers and opposite generalizers is both below 6. Participants in Pattern 6: tk(W) were inclined to be generalizers with 9 generalizers, suggesting that participants were able to pattern a new obstruent /k/ with an old blocker /t/, setting an old target /w/ aside. The next section examines Patterns 7 and 8.

6.9 Testing sonority natural classes: learner type (Patterns 7 & 8)

The hypothesis about sonority natural classes predicts that Pattern 7: SW(t) participants would be confused about the patterning of a new segment /w/ because /s, t/ are both obstruents. Pattern 8: WS(t) would tend to have opposite generalizers because a new segment /s/ and an old blocker /t/ are obstruents.

In this section Patterns 7: S(t)→W(t) and 8: W(t)→S(t) are examined by looking at individual data. Table 65 summarizes the distribution of learner type for Patterns 7 and 8. In Pattern 7 there are 5 generalizers, 4 opposite generalizers, and 1 positional statistician. In Pattern 8 there are 1 learner, 6 opposite generalizers, 2 positional statisticians, and 1 random.

Table 65. Pattern 7: SW(t) vs. Pattern 8: WS(t) (individual learner type)

	Pattern 7: SW(t)			Pattern 8: WS(t)		
		exposure	**test**		exposure	**test**
	t as blocker	s as target	**w as target**	t as blocker	w as target	**s as target**
pattern learner		n=0			n=1	
generalizer		n=5			n=0	
generalizer (opposite)		n=4			n=6	
positional statistician		n=1			n=2	
unbound nasalizer		n=0			n=0	
random		n=0			n=1	

The following presents the individual accuracy by word conditions (NGEN, NS, O) as seen in Table 66.

Table 66. Pattern 7: SW(t) vs. Pattern 8: WS(t) (individual learner type)

	SW(t)	**NGEN**	**NS**	**O**	sum		
Subj	61	**85%**	55%	**94%**	**77%**	categorization learner	generalizer
Subj	62	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	63	**80%**	**100%**	**94%**	**91%**	categorization learner	generalizer
Subj	64	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	65	**100%**	**100%**	**100%**	**100%**	categorization learner	generalizer
Subj	66	25%	**95%**	**100%**	**71%**	categorization learner	generalizer (opposite)
Subj	67	20%	**60%**	**100%**	57%	categorization learner	generalizer (opposite)
Subj	68	30%	**95%**	**94%**	**71%**	categorization learner	generalizer (opposite)
Subj	69	20%	**100%**	**94%**	**70%**	categorization learner	generalizer (opposite)
Subj	70	45%	**65%**	**75%**	**61%**	statistical learner	positional statistician

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

	WS(t)	**NGEN**	**NS**	**O**	sum		
Subj	71	30%	**75%**	**100%**	**66%**	categorization learner	generalizer (opposite)
Subj	72	35%	**75%**	**88%**	**64%**	categorization learner	generalizer (opposite)
Subj	73	20%	**100%**	**100%**	**71%**	categorization learner	generalizer (opposite)
Subj	74	20%	**100%**	**100%**	**71%**	categorization learner	generalizer (opposite)
Subj	75	20%	**60%**	**100%**	57%	categorization learner	generalizer (opposite)
Subj	76	20%	**60%**	**100%**	57%	categorization learner	generalizer (opposite)
Subj	77	45%	**90%**	**100%**	**77%**	categorization learner	pattern learner
Subj	78	40%	**65%**	**81%**	**61%**	statistical learner	positional statistician
Subj	79	**65%**	**75%**	**81%**	**73%**	statistical learner	positional statistician
Subj	80	55%	45%	**69%**	55%	random	

(NGEN=new-generalizable, NS=new-same class; O=old-same class)

According to Table 66, participants for Pattern 7 performed better than those for Pattern 8 in generalizability based on the fact that in Pattern 7 there are 5 generalizers, but in Pattern 8 there are none. The number of opposite generalizers is slightly more than for Pattern 8 (number = 6) than for Pattern 7 (number = 4), suggesting that there is a higher possibility for Pattern 8 participants to wrongly generalize an old blocker /t/ (instead of an old target) to a test segment.

Whether this learning asymmetry is significant is discussed in Section 6.7 on inferential statistics. Following the criterion of learnability (over 6 participants being the same learner type), I could argue that Pattern 8 participants were inclined to favor the opposite patterning of a new consonant, but I could not argue that Pattern 7 participants showed a clear tendency since neither the number of generalizers or of opposite generalizers is over 5. It is hard to say whether Pattern 7 participants learned better than Pattern 8 participants in terms of learner types without inferential statistics since Pattern 7 participants showed mixed results: 5 generalizers; 4 opposite generalizers. However, with a positive effect of "Direction" being found between Patterns 7 and 8, it suggests that the learning of Pattern 7 is better than that of Pattern 8.

In brief, in Pattern 7, the number of both generalizers and opposite generalizers is below 6: 5 generalizers and 4 opposite generalizers. This means that Pattern 7 participants were confused about what to pattern a new segment /w/ with. As for Pattern 8, the number of opposite generalizers is 6, showing a tendency toward opposite generalizers. The results of Pattern 7 and Pattern 8 support the hypothesis of sonority natural classes. The sonority hierarchy type hypothesis is not rejected since I do not find counterevidence that there were over 5 opposite generalizers in Pattern 7. Combined with the findings of Patterns 1 and Pattern 2 (see a similar argument in Section 5.11.2 for Pattern 1), the results of Patterns 1 and 7 can be interpreted as a weak effect of the sonority hierarchy. Specifically, the reason that there were not many generalizers is because Min has a weak effect of the sonority hierarchy, not because the sonority hierarchy type is contradicted and rejected. I would predict that a language with a stronger effect of the sonority hierarchy would result in many generalizers in Patterns 1 and 7.

In conclusion, the interactive approach is supported.

6.10 Testing the continuancy natural class hypothesis

The continuancy natural class hypothesis cannot account for the patterning of /s/ bases on the feature [+continuant]. Specifically, participants in each pattern would be expected to pattern a new consonant with an old consonant with the same direction of continuancy ([+continuant]→[+continuant] for Patterns 1, 2, 7, 8, [-continuant]→[-continuant] for Patterns 3, 4, 5, 6).

The results of Patterns 2, 8 do not support the prediction of natural class based on [continuant]. Participants did not pattern a new consonant with an old consonant based on [continuant] (/s/ = [+continuant]). Rather, they patterned consonants based on [sonorant] (/s/ = [-sonorant]).

In addition, the continuancy natural class hypothesis does not account for the learning asymmetry between Patterns 3 and 5 and Patterns 4 and 6, with were more generalizers in Patterns 4 and 6 than in Patterns 3 and 5. This learning asymmetry is accounted for by the sonority natural class hypothesis.

6.11 Summary

The findings of learner type for Patterns 1-8 match the predictions of the sonority natural class hypothesis (see Table 67). The findings cannot be ascribed to and explained fully by the continuancy natural class hypothesis. However, the sonority hierarchy hypothesis is not rejected because I failed to find counterevidence that there were many opposite generalizers in Pattern 7: SW(t) or Pattern 1: SW(k).

I modify the two-way opposition between obstruents and sonorants discussed in Section 4.4. Min participants do not show a three-way opposition between [+sonorant, +continuant], [-sonorant, +continuant], and [-sonorant, -continuant]. They simply distinguish [+sonorant] and [-sonorant] in this experiment.

The checkmark "✓" in Table 67 means the results of Experiments 2 and 3 match with the prediction of the hypothesis. The cross mark X means the results contradict the prediction of the hypothesis (counterevidence). The triangle mark ∆ means that I failed to find counterevidence.

Table 67. Full sets of patterns: predictions supported for the three hypotheses

	sonority hierarchy hypothesis		sonority natural classes hypothesis	continuancy natural class hypothesis	
	learner type	"Direction"	learner type	learner type	"Direction"
Pattern 1 SW(k)	△	✓	✓ confused about what to pattern a new target /w/ with (/w/ = [+sonorant], /s, k/ = [-sonorant]).	△	x generalizers (/s, w/ = [+continuant], /k/ = [-continuant])
Pattern 2 WS(k)	✓	✓	✓ 6 opposite generalizers (a new target /s/, an old blocker /k/ = [-sonorant])	x generalizers	x generalizers
Pattern 3 kt(S)	△	✓	✓ confused about what to pattern a new blocker /t/ with (/k, t, s/ = [-sonorant]).	△	✓
Pattern 4 kt(W)	✓	✓	✓ 9 generalizers (a new blocker /t/, an old blocker /k/ = [-sonorant])	✓ (/t, k/ = [-continuant])	✓
Pattern 5 tk(S)	△	✓	✓ confused about what to pattern a new blocker /k/ with (/k, t, s/ = [-sonorant]	△	✓
Pattern 6 tk(W)	✓	✓	✓ 9 generalizers (a new blocker /k/, an old blocker /t/ = [-sonorant])	✓ (/k, t/ = [-continuant])	✓
Pattern 7 SW(t)	△	✓	✓ confused about what to pattern a new target /w/ with (/w/ = [+sonorant], /s, t/ = [-sonorant]).	△	x generalizers
Pattern 8 WS(t)	✓	✓	✓ 6 opposite generalizers (a new target /s/, an old blocker /t/ = [-sonorant])	✓ (/s, t/ = [-continuant])	x generalizers

6.12 Confound: place of articulation

Patterns 3 versus 5, and Patterns 4 versus 6 are minimally different in that the place of articulation of the stops in the exposure vs. test phases are switched. Patterns 7-8 are minimally different from Patterns 1-2 in that the blocker is a coronal /t/ rather than a dorsal /k/. This switch of place of articulation allows us to examine whether place natural classes influence the learning[56]. /s, t/ are of special interest because there is no controversy that the place feature [coronal] is shared by /s/ and /t/. The features of /w/ are less clear – it could be [labial] only (see note 48 in Section 5.5.1 on a labial co-occurrence restriction by Bao (2000)) or both [labial] and [dorsal]. I thus focus here on whether there is a coronal effect.

[56] Place of articulation is seldom used as a hierarchy related to sonority. An exception is van der Hulst (1994a,b), who argues that coronal is a true consonant (structure in oral cavity), dorsal is vowel-like because dorsal consonants and vowels share [dorsal] and labial lies in between (structure outside of oral cavity). This thesis is concerned about place natural classes, so I will not discuss a place hierarchy here.

6.12.1 Testing place natural classes

One as yet unexamined possibility to account for the patterning observed is that participants match a new segment with old segments in terms of place classes (labials, coronals, and dorsals). The hypothesis and predictions are presented in Table 68.

Table 68. Eight patterns: place natural classes

	exposure	test	prediction	predictions supported
Pattern 1 SW(k)	**coronal**: target /**s**/ dorsal: blocker /k/	new segment /**w**/ (**labial, dorsal**): same dorsal place as /k/	opposite generalizer /w/ = blocker /k/ (/w, k/ = dorsal)	x
Pattern 2 WS(k)	**labial, dorsal**: target /**w**/ dorsal: blocker /k/	new segment /**s**/ (**coronal**): different place from dorsal /w, k/	confused (neither of /w/ or /k/ are coronal)	x
Pattern 3 kt(S)	coronal: target /s/ **dorsal**: blocker /**k**/	new segment /**t**/ (**coronal**): same coronal place as /s/	opposite generalizer /t/ = target /s/ (/t, s/ = coronal)	x
Pattern 4 kt(W)	labial, dorsal: target /w/ **dorsal**: blocker /**k**/	new segment /**t**/ (**coronal**): different place from /w/ or /k/	confused (/w, k/ = dorsal)	x
Pattern 5 tk(S)	coronal: target /s/ **coronal**: blocker /**t**/	new segment /**k**/ (**dorsal**): different place from coronal /s, t/	confused (/s, t/ = coronal)	✓
Pattern 6 tk(W)	labial, dorsal: target /w/ **coronal**: blocker /**t**/	new segment /**k**/ (**dorsal**): same dorsal place as /w/	opposite generalizer /k/ = target (/k, w/ = dorsal)	x
Pattern 7 SW(t)	**coronal**: target /**s**/ coronal: blocker /t/	new segment /**w**/ (**labial, dorsal**): different place from coronal /s, t/	confused (/s, t/ = coronal)	✓
Pattern 8 WS(t)	**labial, dorsal**: target /**w**/ coronal: blocker /t/	new segment /**s**/ (**coronal**): same coronal place as /t/	opposite generalizer /s/ = blocker (/s, t/ = coronal)	✓

The predictions about place natural classes for Pattern 5: tk(S) and Pattern 7: SW(t) are the same as those of the sonority natural class hypothesis (see Table 68). Pattern 5: tk(S) and Pattern 7: SW(t) participants would be confused about how to pattern a new segment. Specifically, place natural classes predict that since the place of articulation of old consonants /s, t/ is coronal in Patterns 5 and 7, participants would have no idea how to pattern a new dorsal /k/ (Pattern 5) or a new labial dorsal /w/ (Pattern 7) because /k/ and /w/ do not match the old segments in place of articulation.

As for Pattern 8: WS(t), a coronal /s/ is expected to pattern with an old blocker, coronal /t/ (opposite generalizers). This is supported by the results of learner type: 6 opposite generalizers. The results also support the sonority natural classes because the test segment, obstruent /s/, is expected to pattern with an old blocker, obstruent /t/.

The predictions about place natural classes on other patterns (Patterns 1, 2, 3, 4, 6) are different from the predictions of the sonority natural class hypothesis (see "predictions supported" in Table 68). As for Pattern 1: SW(k), the place natural class hypothesis predicts that participants would tend to pattern a test segment /w/ with /k/ because both share [dorsal] (opposite generalizers), whereas the hypothesis of sonority natural classes predicts that participants would be confused about the patterning of /w/ because /s, k/ are obstruents. The results of learner type contradicts the place natural class hypothesis but supports sonority natural class hypothesis. 2 generalizers and 3 opposite generalizers in Pattern 1 suggest that participants were confused about what to pattern a test segment /w/ with. In brief, the sonority natural class effect we see in Patterns 5, 7, 8 cannot be attributed to a place effect in disguise because in Pattern 1 for instance, where the two natural class types (sonority and place) make different predictions, the sonority based natural class hypothesis makes the correct prediction.

As to Pattern 2: WS(k), the place natural class hypothesis predicts that participants would be confused about what to pattern the test segment coronal /s/ with, because neither of the old segments (target /w/, blocker /k/) is coronal, whereas the sonority natural class hypothesis predicts an inclination to opposite generalizers (/s, k/ = obstruents). The results of learner type do not support the place natural class hypothesis of: 6 opposite generalizers.

As for Pattern 3: kt(S), a new segment coronal /t/ would be matched with coronal /s/ as a target, so opposite generalizers are expected. The results of learner type do not support the place natural class hypothesis: 4 generalizers; 2 opposite generalizers. Instead, the results support the sonority natural class hypothesis (/s, k, t/ = obstruents).

As for Pattern 4: kt(W), participants would be confused about what to pattern a test segment coronal /t/ with because neither of /k/ or /w/ is coronal. This is not supported by the results of learner type: 9 generalizers. Instead, the results support the sonority natural class hypothesis (/t, k/ = obstruents).

As for Pattern 6: tk(W), the place natural class hypothesis predicts that the test segment dorsal /k/ would pattern with dorsal /w/ instead of coronal /t/. This is not supported by the results of learner type: 9 generalizers. The results also support the sonority natural class hypothesis (/k, t/ = obstruents).

In brief, the results of learner type for do not match with all the predictions of place natural classes. That is, the results of Patterns 1, 2, 3, 4, 6 are not supported by the predictions[57].

6.13 Summary: Patterns 1-8

The primary purpose of Experiment 3 was to investigate how sonority hierarchy type and sonority natural classes play a role in artificial grammar learning.

This chapter further strengthens the view that participants actively use sonority natural classes when making a judgment on a new segment. Specifically, Min participants actively use [sonorant] in patterning new segments. This is important evidence for natural classes since Min does not have evidence for a consonant sonority hierarchy or for sonority natural classes from phonotactics or phonological processes (Chapter 2).

On the other hand, the sonority hierarchy type hypothesis is not rejected. Not finding positive evidence of many generalizers for Patterns 1 and 7 does not reject this hypothesis (i.e. null results). Only finding counterevidence that there were many opposite generalizers in Patterns 1 and 7 could successfully reject the sonority hierarchy hypothesis. Specifically, the findings support the interactive approach. Patterns 2 and 8 are still supported by the sonority hierarchy type hypothesis because there were fewer generalizers in both patterns as predicted (2 generalizers in Pattern 2, none in Pattern 8).

[57] I also compared Patterns 5 and 6, Patterns 1 and 7, and Patterns 2 and 8, no main effect of "Direction" is found. This further shows that place of articulation of /k/ and /t/ did not affect the learning (see the full tables of these comparisons in Tables 86-88 in Appendix II).

Sonority hierarchy type being equal (Patterns 3: kt(S) vs. 5: tk(S), Patterns 4: kt(W) vs. 6: tk(W)), evidence for sonority natural classes is seen. Though the sonority hierarchy type influences Patterns 1 SW(k) versus 2 WS(k), and Patterns 7 SW(t) versus 8 WS(t) more, sonority natural classes also affect the learning. This is evident in Patterns 2 and 8 with more opposite generalizers. These findings suggest that the sonority hierarchy interacts with sonority natural classes.

The continuancy natural class hypothesis is rejected because /s/ and /w/ do not pattern together as a class of [+continuant]. The results of Patterns 2 and 8 contradict the continuancy natural class prediction. Participants did not pattern a new consonant /s/ to an old consonant /w/ based on [continuant] (/s, w/ = [+continuant]). Rather, they patterned consonants based on [sonorant] (/s, k/ = [-sonorant]).

The place of articulation hypothesis is ruled out because the learner type findings of Patterns 1, 2, 3, 4, and 6 contradict the predictions of the place natural class hypothesis.

Chapter 7 Conclusions

7.1 Summary of the thesis

In recent phonological research, an artificial grammar (AG) paradigm (e.g., Wilson 2006, Moreton 2008, Nevins 2010, Finley 2011a Moreton & Pater 2012a,b) has been used to test language universals. This paradigm allows the study of aspects of proposed universals that can be difficult to test with real language. My research examines one proposed universal, the implicational nasalized segment hierarchy, testing whether this hierarchy is found with speakers of a language with no clear evidence for a nasalized segment hierarchy, or even for a role for sonority differences between consonants.

This thesis assumes that the nasalized segment hierarchy mirrors the sonority hierarchy, given that the nasalization of segments is highly correlated with the sonority hierarchy. The current study adopts a morphologically driven design, with [nasal] as a floating plural morpheme attached to the leftmost vowel of the plural form.

The thesis focuses on the nature of consonant-vowel nasal harmony with opaque segments by using an artificial grammar learning paradigm. I conducted a full set of experiments using the segments /s, w, k, t/ to test the roles of sonority hierarchy type and sonority natural classes in artificial grammar learning. A continuancy natural class hypothesis and a place natural class hypothesis were also tested.

The sonority hierarchy predictions are based on expectations if the sonority hierarchy is universal: it should be easier to learn a grammar if in the test phase the new segment is more sonorant than the target or equivalent in sonority to the blocker in the exposure phase. If the test segment is less sonorant than the target, then there is essentially no prediction.

The sonority natural class hypothesis predicts that participants compare whether a new segment's phonological class is close to any old segment's phonological class in term of the feature [sonorant]. If yes, then participants would pattern the two segments together, assigning the same status of nasality function (i.e, target or blocker).

The continuancy natural class hypothesis predicts that /s/ and /w/ would pattern together as [+continuant]. The place natural class hypothesis predicts that /s, t/ could pattern together as [coronal].

The current findings show that Min participants actively used [sonorant] instead of [continuant] when patterning a new consonant with old segments. It also shows that the sonority hierarchy, as seen by directionality effects, also plays a role. This thesis concludes that the interactive approach, involving both the feature [sonorant] and the sonority hierarchy, can account for the results.

The continuancy natural class hypothesis[58] and the place natural class hypothesis are ruled out since the predictions contradict the results of the eight patterns (see Section 6.10 and Table 68 in Sections 6.12.1). In addition to testing these hypotheses, the current study aims to offer a way to quantitatively and qualitatively evaluate the results of artificial grammar learning. Quantitatively, this study adopts logistic regression to test directional learning. Qualitatively, this study classifies individual participants into learner types to account for individual differences and different learning strategies adopted by participants, an issue which is seldom addressed in the literature on phonological artificial grammar learning.

It is important to keep in mind that Min does not have obvious evidence for the sonority hierarchy, and there does not appear to be phonological evidence for the feature [sonorant]. Nevertheless, Min speakers group segments together on the basis of sonority, providing the type of evidence required in artificial grammar.

7.2 Testing a language with sonority hierarchy effects

The experiments with Min participants show overall a featural effect ([sonorant]) in the language, with weaker evidence for the sonority hierarchy. Does this suggest that the dichotomy of obstruents and sonorants is universal cross-linguistically? What would happen if participants spoke a language with sonority hierarchy effects? Would the results from speakers of languages with and without evidence for sonority show the same tendencies in patterning or would they exhibit language-specific patterning?

In order to investigate this puzzle, I carried out a pilot experiment with Quebec French speakers. The reason that I chose Quebec French is because nasality is phonemic in Quebec French, and sonority effects are found, making it a good comparison to Min, which has no obvious sonority hierarchy effects. My question is: would Quebec French participants use the sonority hierarchy more than Min participants, since their phonology has clear sonority hierarchy effects? If Quebec

[58] One interpretation for the failure of [continuant] to play a role in these experiments can be given within the Contrastive Hierarchy hypothesis (Dresher 2009), namely that the feature [sonorant] has scope over the feature [continuant].

French participants rely on reference to the sonority hierarchy more than Min participants, I expect that there will be more generalizers in Pattern 1: S(k)→W(k) than in Pattern 2: W(k)→S(k), and this difference in number of generalizers between the two patterns for French speakers should be more obvious than for Min speakers.

Concerning sonority natural classes, Quebec French has richer consonant inventory than Min, namely /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ, m, n, ɲ, ŋ, l, ʁ/ (Walker 1984: 10). It also has clusters that show sonority sequencing effects. Therefore, I would expect that French participants would have more fine-grained distinctions of sonority natural classes, especially for stops and fricatives. For instance, the sonority natural class hypothesis, focusing on the sonority natural class shared between a new consonant and old consonants, would make different predictions about Pattern 1: SW(k) and Pattern 2: WS(k). Specifically, as for Pattern 1 (see Table 69), given that the sonority distance between a new glide /w/ and an old fricative /s/ is closer than the distance between the new glide /w/ and an old stop /k/, the sonority natural class predicts that participants would pattern the new glide /w/ with the old fricative /s/. That is, generalizers are expected.

As for Pattern 2 (see Table 69), given that the sonority distance between a new fricative /s/ and an old blocker /k/ is closer to that between a new fricative /s/ and an old target /w/, the sonority natural class hypothesis predicts that participants would pattern the new fricative /s/ with the old blocker /k/. That is, opposite generalizers are expected.

Table 69. Sonority distance between new and old consonants

a sonority distance

	new /**w**/ (Pattern 1)	new /**s**/ (Pattern 2)
old /k/	Δs = 2	Δs = 1
old /s/	Δs = 1	n/a
old /w/	n/a	Δs = 1

b Sonority hierarchy of /w, s, k/

glide > fricative > stop

In the following section, I report the results of my preliminary work, Pattern 1: SW(k) and Pattern 2: WS(k).

7.3 Is there more than the classes that is universal?

I recruited 14 Quebec French speakers for Pattern 1: SW(k) and for Pattern 2: WS(k). The results show that the majority of participants had difficulty perceiving a contrast between nasalized and oral segments. Even though I switched the post-test for distinguishing the nasalized monosyllables from their nonnasalized counterparts to a pre-test in order to enhance the possibility of participants noticing the contrast, it did not help participants to learn the patterns.

I then examined the phonetic inventories of Quebec French and Min with respect to vowels. The inventory of nasal and oral vowels differs between Quebec French and Min. In Min, ease of perception for nasal vowel and its counterpart is: a/ã, e/ẽ > i/ĩ, ɔ/ɔ̃ (Chang 2008). In Quebec French, there is no nasal high vowel (Walker 1984: 81). Nasal vowels are all diphthongized (Walker, Desmeules-Trudel 2015), suggesting that nasal vowels are in general longer than their oral counterparts.

With this perceptual difference between vowels in French and Min, I speculate that the Quebec French participants' failure to perceive the contrast between nasal and oral vowels is because French does not have the high nasal vowels that are found in Min. Thus, French participants need to build up phonological realizations in order to learn nasal harmony in my experiments. The relevant phonemes do exist in the Min inventory, so Min speakers do not need to build up phonological realizations; instead, they could concentrate on the patterns. Specifically, in the exposure phase, I used the Min vowels /a/ and /i/. In the test, I used Min vowels /a/ and /e/. If the Quebec French participants had difficulty hearing the difference between /i/ and /ĩ/, it would be harder for them to identify a harmony pattern, which would further impact their patterning of a new consonant/vowel. Only one participant got 100% on the pre-test, and this participant generalized to a new consonant (as an opposite generalizer). This suggests that participants' perception of the contrast between nasal and oral counterparts is related to the ability to learn a pattern.

The above findings on French suggest that before testing sonority hierarchy type and sonority natural classes, I need to carefully examine the relationship among nasal/oral vowels and degrees of diphthongization when conducting a French artificial grammar experiment.

In sum, the results of the pilot experiment show that some language-specific factors such as lacking high nasal vowels/cross-linguistic perceptual nuances of sounds come into play in artificial grammar learning. Future work is needed to figure out what language-specific factors would affect the current experiment for Quebec French speakers. After those factors are controlled, the language universality of sonority effects in nasal harmony could be tested. In addition to sonority effects, a language with a rich place inventory might allow better control of place of articulation. A language like French with overt sonority effects might reveal more about the effect of place of articulation and help to understand the balance between language particular and universal effects.

7.4 Hierarchy and natural classes

As discussed in Section 3.3.1, most of the work in artificial grammar tests natural classes alone (e.g., Wilson 2006, Finley 2011a). The current study is designed to examine the relationship between natural classes as well.

This thesis shows that the sonority hierarchy and sonority natural classes interact to some extent. The influence of sonority natural classes is easily seen when no directionality effects are present: testing with a different consonant within the same sonority natural class shows that learners group a new consonant with an old consonant into classes according to sonority natural classes. The evidence for the sonority hierarchy is more subtle. Future work is needed to test a language with a fuller sonority hierarchy represented such as Quebec French to see if [continuant] plays a role in nasal harmony.

How can we be sure that my current analysis for the sonority hierarchy and sonority natural class hypotheses is appropriate and does not miss important generalizations? Moreton & Pater (2012b: 709) point out that calculations of substantive bias are strongly influenced by the features that the experimenter adopts. I propose a way to test the sonority hierarchy type hypothesis and the sonority natural class hypothesis separately, but sonority natural classes and the sonority hierarchy are not separable. Sonority natural classes form the foundation of the sonority hierarchy (obstruent < sonorant in the case of Min). Future work is needed to tease apart the contribution of the sonority hierarchy and sonority natural classes in artificial grammar learning. The current design has a confound in that the patterning of a new consonant and an old consonant only involves [-sonorant] (/s, k, t/). That is, participants patterned all [-sonorant] consonants together, setting [+sonorant] /w/ aside (i.e., Patterns 1, 2, 7, 8). I could design patterns that would allow one to test the interaction between [sonorant] and [continuant]. For instance, would [continuant] be involved in artificial grammar when participants' language shows phonological evidence for [continuant] distinctions in obstruents (fricative as [+continuant], stop as [-continuant])? Or would there be any learning asymmetry in directionality experiments if participants' language shows phonological evidence for [continuant] distinctions in sonorants (glide as [+continuant], lateral as [-continuant])?

7.5 Is artificial phonology phonology?

Moreton & Pater (2012b) remark that formal complexity bias (the number of features) is domain-general. On the other hand, substantive bias is domain-specific (specific to languages). If artificial grammar learning could capture substantive bias, it would show strong evidence that participants learn the patterns using cognitive processes exclusive to natural-language phonology. However, Moreton & Pater also indicate that substantive bias is weak (Moreton & Pater 2012b: 710). This is based on the fact that among the substantive bias studies they reviewed, there were no successes and 5 failures, while among the formal complexity studies they reviewed, there were 8 successes and 1 failure.

This thesis aims to test whether learnability matches with an implicational universal involving substantive bias. Specifically, will a pattern that is predicted by an implicational hierarchy be easier to learn than one that is not predicted or that is indeterminate? If the implicational hierarchy is supported, this can be used as evidence to argue that phonological artificial grammar learning is exclusive to natural language (domain-specific).

Though evidence for the sonority hierarchy type is weak for Min participants, it nevertheless provides evidence about substantive bias. All the patterns I compare involve the same number of features, only the directionality of segments under consideration is switched. If there is any learning difference between the two patterns I compare, it must be phonetic naturalness that drives this learning bias. Whether phonetic naturalness help participants to learn the pattern that is predicted or to impede the pattern that is not predicted or that is indeterminate is an empirical question.

7.6 Could participants learn any patterns?

One working assumption of artificial learning is that participants could learn patterns in artificial grammar learning without inferring from their own language-specific knowledge. Motivated by this working assumption, another research question could be proposed: could participants learn a pattern in an artificial grammar paradigm without inferring from their own language-specific knowledge or are they restricted by phonetically universal constraints during artificial grammar experiments?

I assumed that participants would perceive and learn a distinction between /s, k, t/ vs. /s̃, t̃, k̃/ (recall that these are not actually nasalized, but transparent), /w, y/ vs. /w̃, ỹ/, /a, i, e, ɔ/ vs. /ã, ĩ, ẽ, ɔ̃/ in an artificial grammar learning without inferring from their own language-specific knowledge. However, it is also possible that participants just used their own language-specific distinction between consonants and vowels. This concern about language-specific issues is validated by the fact that Quebec French speakers failed to learn the patterns because they could not distinguish nasalized and oral vowel counterparts well.

A possible artificial grammar experiment could be conducted to test whether participants could learn anything even if it is not phonetically motivated or predicted by Universal Grammar. It is possible to reverse the nasalized segment hierarchy of the current experimental design to test whether participants could learn the opposite version. That is, now the ungrammatical items would be grammatical in the exposure phase. This design violates Walker (2000)'s universal nasalized constraint hierarchy and thus the opposite patterns of blockers/targets for this experiment are not found in real languages[59].

If the results show that participants do not learn this pattern, this suggests that humans are able to learn the pattern only when it conforms to phonetically universal constraints. If the results show that participants treat blockers and targets in a reverse way, then on the one hand, the result could be used as evidence to support the assumption that participants could learn a distinction between oral and nasal counterparts in an exposure phase even though this design does not obey the universal nasalized constraint hierarchy (i.e., participants learn any pattern during the experiment). The results could also be used as evidence against the hypothesis that artificial grammar paradigms provide a way to test natural markedness (i.e., more natural = easier to learn), since humans could learn 'unnatural' patterns that violate phonetically universal constraints[60].

[59] Previous studies only test logically possible but unattested patterns (e.g., Moreton 2008, Moreton 2010, Wilson 2003). However, this follow-up experiment examines one pattern that is not logically possible and is unattested.

[60] Seidl & Buckley (2005) tested 8.5-9.5-month-olds infants, showing that infants can learn an arbitrary (non-phonologically-grounded) pattern. This thesis focuses on adult findings and does not present infant studies in detail in this section (see discussion about infants and adults in artificial grammar learning in Section 3.2.3).

7.7 Design

The current design presents a consonant in one sonority class, and tests on a consonant in a different or the same sonority class. An interesting design for follow up would involve a pattern where the segments were the same, with a hierarchy conforming to the nasalized segment hierarchy (e.g., (tW→kY)), and with a hierarchy violating the nasalized segment hierarchy (Wt→Yk)[61].

7.8 Exposure and robustness

Moreton & Pater (2012b) note that most of the artificial grammar studies use under 30 minutes of exposure. This is true of my thesis as well. They conclude that "artificial grammar phonology appears easier to lose than first-language phonology and easier to acquire than second-language phonology" (Moreton & Pater 2012b: 711). Ettilinger et al. (2015) shows that artificial grammar learning is equivalent to second-language learning. Artificial grammar learning is in this way more like short-term memory. However, the trained knowledge could become part of long-term memory if it involves sleep (Gómez et al. 2006, St. Clair & Monaghan 2008, Wagner et al. 2004). Sharon & Martin (2016) shows that sleep helps phonological learning in phonologically natural pattern (vowel harmony).

A future study can explore the issue of whether increasing duration of exposure phase and scheduling the experiment on different days would influence the effects of the sonority hierarchy. Would Min participants be able to generalize to Pattern 1: SW(k) and to Pattern 7: SW(t) more if I increased the duration of the exposure phase or scheduled the experiments on different days?

In Section 5.4, I noted that procedural memory and declarative memory belong to long-term memory. However, only procedural memory could reflect mental grammar, including phonology, syntax, morphology, long-distance dependencies (see Morgan-Short et al. 2014). If this is the case, would procedural memory influence my learning results? Specifically, would participants with better procedural memory be able to generalize to Pattern 1: SW(k) and Pattern 7: SW(t)?

---

[61] Thanks my committee Yoonjung Kang for this suggestion.

7.9 Conclusion

In conclusion, this thesis shows that the feature [sonorant] is psychologically real in phonological artificial grammar learning. Participants were able to actively compare a new segment with an old segment in the same sonority natural classes.

There are some major questions left to answer for future work. Assuming that the sonority hierarchy is phonological, how much is universal, and how much is language-specific, emerging from language acquisition? If a more fine-grained sonority hierarchy has to emerge from exposure, why are sonority effects so robust cross linguistically? Why do languages that have reverse sonority hierarchy effects not occur typologically? Assuming the nasalized segment hierarchy mirrors the sonority hierarchy, where does nasal harmony come from? Future work with Quebec French participants would show a better understanding of language-specific properties about nasal harmony. Future work on illogical and unattested patterns (ungrammatical pattern – the reverse nasalized segment hierarchy) could probe into a question about whether the sonority hierarchy is universal and irreversible.

References

Ajíbóyè, Oladiipo. (2002). Mòbà nasal harmony. Manuscript, UBC.

Albright, Adam & Young Ah Do. (2013). Biased learning of phonological alternations. The 21th Manchester Phonology Meeting, May 23–25, Manchester, UK.

Albright, Adam & Young Ah Do. (2015). The source of the paradigm uniformity preference. Annual Meeting on Phonology (AMP), October 9–11, Vancouver, Canada.

Anderson, Stephen R. (1985). *Phonology in the twentieth Century: Theories of rules and theories of representations.* Chicago: University of Chicago Press.

Archangeli, Diana & Douglas Pulleyblank. (1994). *Grounded phonology*. Cambridge, Mass.: MIT Press.

Baković, Eric & Colin Wilson. (2000). Transparency, strict locality, and targeted constraints. *Proceedings of WCCFL* 19: 43–56.

Bao, Zhiming. (2000). Sub-syllabic processes and the Southern Min syllable. In Debao Xu (ed.), *Chinese phonology in generative grammar*. NewYork: Academic Press. 95–116.

Bateman, Nicoleta. (2007). A Crosslinguistic Investigation of Palatalization. Doctoral dissertation, University of California, San Diego.

Bates, Douglas & Martin Maechler. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-32. http://CRAN.R-project.org/package=lme4

Bennett, Ryan. (2012). *Foot-conditioned phonotactics and prosodic constituency*. Doctoral dissertation, University of California, Santa Cruz.

Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104, 591–630.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel Moreno A. & Paul Smolensky. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences* 105. 5321–5325.

Berent, Iris. (2013). *The phonological mind*. Cambridge Press: New York.

Blevins, Juliette. (1995). The syllable in phonological theory. In John Goldsmith (ed.). *Handbook of phonological theory*. Blackwell. 206–244.

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

Blevins, Juliette. (2006). A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics* 32. 117–165.

Boersma, Paul & David Weenink. (2008). *Praat: A system for doing phonetics by computer* /Computer program/. http://www.praat.org/

Boersma, Paul & David Weenink. (2012). *Praat: A system for doing phonetics by computer* /Computer program/. http://www.praat.org/

Boersma, Paul, Kateřina Chladkova & Titia Benders. (2013). Learning phonological structures from sound-meaning pairs. The 21th Manchester Phonology Meeting, May 23–25, Manchester, UK.

Boersma, Paul. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Doctoral dissertation, University of Amsterdam.

Boersma, Paul. (2003). Nasal harmony in functional phonology. In Jeroen van de Weijer, Vincent J. van Heuven & Harry van der Hulst (eds.) *The phonological spectrum*, Vol. 1: *Segmental structure*. Amsterdam: John Benjamins. 3–35

Botma, Bert. (2011). Sonorants. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.). The Blackwell Companion to Phonology. Oxford: Wiley-Blackwell. 171–194.

Butcher, Andrew & Marija Tabain. (2004). On the back of the tongue: Dorsal sounds in Australian languages. *Phonetica* 61. 22–52.

Bybee, Joan. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Chambers, Kyle, Onishi, Kristine H. & Fisher, Cynthia. (2010). A vowel is a vowel: Generalization to newly learned phonotactic constraints to new contexts. *Journal of Experimental Psychology: Learning, Memory and Cognition* 36(3). 821–828.

Chang, Yung-Hsiang Shawn. (2008). An acoustic and perceptual study of vowel nasality in Taiwanese. *USTWPL* 4, 17–26. Urbana-Champaign: Department of Linguistics University of Illinois.

Chen, Matthew Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.

Cheng, Robert. (1968). Tone sandhi in Taiwanese. *Linguistics* 41. 19–42.

Cheng, Robert. (1973). Some notes on tone sandhi in Taiwanese. *Linguistics* 100. 5–25.

Chomsky, Noam & Morris Halle. (1968). *The sound patterns of English*. New York: Harper & Row.

Chomsky, Noam. (1972). *Language and mind* (English edition). Harcourt Brace Jovanovich: New York.

Chomsky, Noam. (1980). On cognitive structures and their development: A reply to Piaget. In Massimo Piatelli-Palmarini (ed.) *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Harvard University Press: Cambridge, MA. 35–52.

Chomsky, Noam. (2007). Review of Margaret Boden's mind as machine: A History of Cognitive Science. *Artificial Intelligence* 171, 1094–1103.

Chou, Tai-yun. (2002). An optimality-theoretic analysis on Southern Min nasality. Unpublished master's thesis. National Kaohsiung Normal University.

Chung, Raung-fu. (1996). *The Segmental phonology of Southern Min in Taiwan*. Taipei: Crane.

Clements, George N. & Elizabeth Hume V. (1995). The internal organization of speech sounds. In Goldsmith (ed), *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 245–306.

Clements, George N. (1990). The role of the sonority cycle in core syllabification. In John Kingston & Mary Beckman (eds.) *Papers in laboratory phonology: Between the grammar and physics of speech*. Cambridge: Cambridge University Press. 1838–1865.

Cohn, Abigail (1990). *Phonetic and phonological rules of nasalization*. Doctoral dissertation, UCLA. [UCLA Working Papers in Phonetics 76]

Cohn, Abigail (1993). The status of nasalized continuants. In Marie K. Huffman & Rena Krakow (eds.), *Nasals, nasalization, and the velum. Phonetics and phonology*. Vol. 5. San Diego: Academic Press. 329–367.

Cole, Jennifer & Loren Trigo. (1988). Parasitic harmony. Features, segmental structure, and harmony processes. In Harry van der Hulst & Norval Smith (eds.) *Features, segmental structure and harmony processes*. Dordrecht: Foris. 19–39.

Conway, Christopher M., Althea Bauernschmidt, Sean S. Huang & David B. Pisoni. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition* 114(3). 356–371.

Cristiá, Alejandrina & Amanda Seidl. (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development* 4(3). 203–227.

Cser, András. (2003). *The typology and modelling of obstruent lenition and fortition processes*. Budapest: Akadémiai Kiadó.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrmann. (2011). Explaining sonority projection effect. *Phonology* 28. 197–234.

Davidson, Lisa, Peter Jusczyk & Paul Smolensky. (2004). The initial and final states: Theoretical implications and experimental explorations of Richness of the Base. In Rene Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press. 321–368.

de Lacy, Paul. (2000). Markedness in prominent positions. In Ora Matushansky, Albert Costa, Javier Martin-Gonzalez, Lance Nathan & Adam Szczegielniak (eds.) HUMIT 2000, *MIT Working Papers in Linguistics* 40. Cambridge, MA: MITWPL. 53–66.

de Lacy, Paul. (2006). *Markedness: Reduction and preservation in phonology*. Cambridge Studies in Linguistics 112. Cambridge University Press.

Do, Young Ah & Adam Albright. (2015). Paradigm uniformity in the lab: prior bias, learned preference, or L1 transfer? Annual Meeting on Phonology, October 9–11, Vancouver, Canada.

Dresher, Elan B. (2009). *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Ettlinger, Marc, Ann R. Bradlow & Patrick C. M. Wong. (2014). Variability in the learning of complex morphophonology. *Applied psycholinguistics* 35(4). 807–831.

Ettlinger, Marc, Kara Morgan-Short, Mandy Faretta-Stutenberg & Patrick C. M. Wong. (2015). The relationship between artificial and second language learning. *Cognitive Science*. 1–26.

Evans, Nicolas & Stephen C. Levinson. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Science* 32. 429–492.

Feldman, Jacob. (2000). Minimization of Boolean complexity in human concept learning. *Nature* 407. 630–633.

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

Finley, Sara & William Badecker. (2010). Linguistic and non-linguistic influences on learning biases for vowel harmony. In Ohlsson Stellan & Richard Catrambone (eds.) *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society. 706–711.

Finley, Sara. (2008). *Formal and cognitive restrictions on vowel harmony*. Doctoral dissertation, John Hopkins University.

Finley, Sara. (2011a). The privileged status of locality in consonant harmony. *Cognitive Science*. 3268–3273.

Finley, Sara. (2011b). Generalizing to novel consonants in artificial grammar learning. *Journal of Memory and Language* 65. 74–83.

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

Finley, Sara. (2013). Generalization to unfamiliar talkers in artificial grammar learning. *Psychonomic Bulletin & Review*.

Finley, Sara. (2015). Learning non-adjacent dependencies in phonology: Transparent vowels in vowel harmony. *Language 91,* 48–72.

Fiser, József & Aslin, Richard N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences* 99, 15822–15826.

Fodor Jerry A. & Merrill Garrett F. (1966). Some reflections on competence and performance. In John Lyons & Roger Wales (eds.) *Psycholinguistics papers*. Edinburgh University Press.

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

Gafos, Adamantios I. (1996). *The articulatory basis of locality in phonology*. Doctoral dissertation, Johns Hopkins University. /Published 1999, New York: Garland.

Gerken, LouAnn & Alex Bollt. (2008). Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Leanring and Development* 4(3). 228–248.

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

Gómez, Rebecca L. & Roger W. Schvaneveldt. (1994). What is learned from artificial grammars? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2). 396–410.

Gómez, Rebecca L., Richard R. Bootzin & Lynn Nadel. (2006). Naps promote abstraction in language-learning infants. *Psychological Science* 17(8). 670–674.

Gordon, Mathew. (1999). *Syllable weight: Phonetics, phonology and typology*. Doctoral dissertation, UCLA.

Greenberg, Joseph H. (1978). Some generalizations concerning initial and final consonant clusters. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.) *Universals of human language*. Stanford University Press. 243–279

Guion, Susan G. (1996). *Velar palatalization: Coarticulation, perception, and sound change*. Doctoral dissertation, University of Texas at Austin.

Guion, Susan G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, 55. 18–52.

Hansson, Gunnar Ólafur. (2001). *Theoretical and typological issues in consonant harmony*. Doctoral dissertation, University of California, Berkeley, Berkeley.

Hansson, Gunnar Ólafur. (2004). Long-distance voicing agreement: An evolutionary perspective. Berkeley Linguistics Society 30. In Marc Ettlinger, Nicholas Fleischer & Mischa Park-Doob (eds.) *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 130–141.

Harms, Phillip L. (1985). Epena Pedee (Saija): Nasalization. In Ruth M. Brend (ed.) *From phonology to discourse: Studies in six Colombian languages*. Dallas: Summer Institute of Linguistics. 13–18.

Harms, Phillip L. (1994). *Epena Pedee syntax*. Arlington: Summer Institute of Linguistics and University of Texas at Arlington.

Hayes Bruce, Kie Zuraw, Péter Siptár & Zsuzsa Londe. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4). 822–863.

Hayes, Bruce & Donca Steriade. (2004). The phonetic bases of phonological markedness. In Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.) *Phonologically based phonology*. Cambridge: Cambridge University Press. 1–33.

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

Henke, Eric, Ellen M. Kaisse & Richard Wright. (2012). Is the Sonority Sequencing Principle an epiphenomenon? In Steven Parker (ed.), *The sonority controversy*. (Phonology and Phonetics 18). Berlin/Boston: De Gruyter Mouton. 65–100.

Honeybone, Patrick. (2008). Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In Brandão de Carvalho, J., Scheer, T. & Ségéral, P. (eds). *Lenition and Fortition*. Berlin: Mouton de Gruyter. 9–93.

Hooper, Joan B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In William M. Christie (ed.) *Current progress in historical linguistics*. Amsterdam: North Holland. 96–105.

Hsieh, Feng-Fan. (2003). Loanword tonology in Taiwanese. Paper presented at the 15th North American Conference on Chinese Linguistics, Michigan State University.

Hsieh, Feng-Fan. (2006). High infidelity: The non-mapping of Japanese accent onto Taiwanese tone. In Feng-fan Hsieh & Michael Kenstowicz (eds.), *MIT Working Papers in Linguistics* 52: Studies in Loanword Phonology. Cambridge: Department of Linguistics and Philosophy, Massachusetts Institute of Technology. 1–27.

Hsu, Hsiu-Ling. (2009). A study of English coda nasals: Taiwanese learning English as a foreign language. Academy Colloquium *Language Acquisition and Optim*ality 7.

Hume, Elizabeth & Keith Johnson. (2001). A model of the interplay of speech perception and phonology. In Elizabeth Hume & Keith Johnson (eds). *The role of speech perception in phonology*. San Diego: Academic Press. 3–26.

Jensen, John T. (1974). A constraint on variables in phonology. *Language* 50(4, Part 1). 675–86.

Kapatsinski, Vsevolod. (2009). *The architecture of grammar in artificial grammar learning: Formal biases in the acquisition of morphophonology and the nature of the learning task*. Doctoral dissertation, Indiana University.

Kaun, Abigail R. (1995). *The typology of rounding harmony: An optimality theoretic approach*. Doctoral dissertation, Los Angeles, CA: UCLA.

Kaun, Abigail R. (2004). The typology of rounding harmony. In B. Hayes & Donca Steriade (eds.) *Phonetics in phonology*. Cambridge: Cambridge University Press. 87–116.

Keating, Patricia & Aditi Lahiri. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica* 50. 73–101.

Kimper, Wendell. (2015). Asymmetrical generalisation of harmony triggers. Annual Meeting on Phonology (AMP), October 9–11, Vancouver, Canada.

Kochetov, Alexei. (2011). Palatalisation. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.). The Blackwell Companion to Phonology. Oxford: Wiley-Blackwell. 1666–1690.

Koo, Hahn & Lydia Callahan. (2011). Tier adjacency is not a necessary condition for learning non-adjacent dependencies. *Language and Cognitive Processes*. 1425–1432.

Kuhl, Patricia K. (2001). Language, mind and brain: Experience alters perception. In Charles A. Nelson & Monica Luciana (eds.) *Handbook of developmental cognitive neuroscience*. Cambridge: MIT Press. 99–115.

Kuhl, Patricia K., Karen Williams A., Francisco Lacerda, Kenneth Steven F. & Björn Lindblom. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255. 606–698.

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

Ladefoged, Peter. (2001). *A Course in Phonetics*. San Diego: Harcourt Brace.

Lien, Chin-Fa. (2000). Denasalization, vocalic nasalization and related issues in Southern Min: A dialectal and comparative perspective. In Ting, Pang-Hsin & Anne O. Yue (eds.) *Essays of linguistic change and the Chinese dialects*. Academic Sinica, Taipei.

Lin, Yu-Leng. (2009). Tests of Analytic Bias in Native Mandarin and Native Southern Min Speakers. In Yun Xiao (ed.), *Proceedings of the 21st North American Conference on Chinese Linguistics* (*NACCL-21*), 81–92. 2 Volumes. Bryant University, Smithfield, Rhode Island. Distributed by *NACCL Proceedings Online*, The Ohio State University, Columbus, Ohio.

Lin, Yu-Leng. (2010). *Testing Universal Grammar in phonological artificial grammar learning*. Master thesis. National Chung Cheng University, Taiwan.

Lin, Yu-Leng. (2011). Tests of Analytic Bias in native Mandarin and native Southern Min Speakers. In Yun Xiao, Liang Tao & Hooi ling Soh (eds.) *Current issues in Chinese Linguistics*. Cambridge Scholars Publishing.

Lin, Yu-Leng. (2012). *Nasal harmony with opaque segments: Evidence from wug tests*, Generals paper, University of Toronto.

Lin, Yu-Leng. (2013). Nasal harmony with opaque segments and learnability: Evidence from wug tests. Presented at *21st Manchester Phonology Meeting*. May 23–25, 2013, University of Manchester, UK.

Lin, Yu-Leng. (2014). The ambiguous status of laryngeals in nasal vowel-consonant harmony. Poster presented at Workshop on Learning Biases in Natural and Artificial Language at 2014 Annual Meeting of the Linguistics Association of Great Britain (2014 LAGB). September 1–5, 2014. University of Oxford, UK.

Lin, Yu-Leng. (2016). What matters in artificial grammar learning, sonority hierarchy or natural classes? *Proceedings of 2015 Annual Meeting on Phonology (AMP 2015)*. University of British Columbia, Vancouver, Canada.

Lonbina, David. (2011). Recursion and competence/performance distinction in AGL tasks. *Langauge and Cognitive Processes* 26(10). 1563–1586.

Love, Bradley C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin and Review* 9(4). 829–835.

Majerus, Steve, Martial Van der Linden, Ludivine Mulder, Thierry Meulmans & Frédéric Peters. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language* 51. 297–306.

McCarthy, John. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12. 373–418.

McMullin, Kevin & Gunnar Ólafur Hansson. (2014). Locality in long-distance phonotactics: Evidence for modular learning. In Leland Kusmer & Jyoti Iyer (eds.) *Proceedings of NELS 44*, Vol. 2. Amherst, MA: GLSA Publications, University of Massachusetts. 1–14.

Mielke, Jeff. (2011). Distinctive features. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.). *The Blackwell companion to phonology*. Wiley-Blackwell. 391–415.

Moreton, Elliott & Joe Pater. (2012a). Structure and substance in artificial-phonology learning, Part I: Structure. *Language and Linguistics Compact*. 686–701.

Moreton, Elliott & Joe Pater. (2012b). Structure and substance in artificial-phonology learning, Part I: Substance. *Language and Linguistics Compact*. 702–718.

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

Moreton, Elliott. (2010). Underphonologization and modularity bias. In Steve Parker (ed.) *Phonological argumentation: Essays on evidence and motivation*. London: Equinox. 79–101.

Moreton, Elliott. (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* 67(1). 165–83.

Morgan-Short, Kara, Mandy Faretta-Stutenberg, Katherine Brill-Schuetz, Helen Carpenter & Wong, Patrick C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition* 17(1), 56–72.

Neisser, Ulric & Paul Weene. (1962). Hierarchies in concept attainment. *Journal of Experimental Psychology* 64(6). 640–645.

Nevins, Andrew. (2009). On formal universals in phonology. *Behavioral and Brain Sciences* 32. 461–462.

Nevins, Andrew. (2010). Two case studies in phonological universals: A view from artificial grammars. *Biolinguistics* 4. 217–232.

Nevins, Andrew. (2013). Restrictive theories of harmony. The 21th Manchester Phonology Meeting, May 23–25, Manchester, UK.

Newton, Ashley M. & Jill de Villiers G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science* 18(7). 574–579.

Nosofsky, Robert M., Mark Gluck A., Thomas Palmeri J., Stephen McKinley C. & Paul Gauthier. (1994a). Comparing models of rulebased classification learning: A replication and extension of Shepard, Hovland, and Jenkins 1961. *Memory and Cognition* 22(3). 352–369.

Nosofsky, Robert M., Thomas Palmeri J. & Stephen McKinley C. (1994b). Rule-plus-exception model of classification learning. *Psychological Review* 101(1). 53–79.

Odden, David. (1995). Adjacency parameters in phonology. *Language* 70(2). 289–330.

Ohala, John J. (1992). What's cognitive, what's not, in sound change. In Günter Kellermann & Michael D. Morrissey (eds.) *Diachrony within synchrony: Language history and cognition. Duisberger Arbeiten zur Sprach- und Kulturwissenschaft 14*. Frankfurt: Peter Lang. 309–355.

Onn, Farid M. (1976). *Aspects of Malay phonology and morphology*. Doctoral dissertation, University of Illinois.

Pan, Ho-Hsien. (2007). The effects of prosodic boundaries on nasality in Taiwan Min. *Journal of the Acoustical Society of America* 121(6). 3755–3769.

Parker, Steve. (2002). *Quantifying the sonority hierarchy*. Doctoral dissertation, University of Massachusetts, Amherst.

Parker, Steve. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics 36*. 55–90.

Parker, Steve. (2011). Sonority. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.). *The Blackwell companion to phonology*. Oxford: Wiley-Blackwell. 1160–1184.

Peperkamp, Sharon & Dupoux, Emmanuel. (2007). Learning the mapping from surface to underlying representations in artificial language. In Jennifer Cole & José Ignacio Hualde (eds.), Laboratory Phonology 9. 315–338.

Peperkamp, Sharon, Le Calvez, Rozenn, Nadal, Jean-Pierre & Dupoux, Emmanuael. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*. 31–41.

Pierrehumbert, Janet B. (1993). Dissimilarity in the Arabic verbal roots. In A. Schafer (ed). *Proceedings of the Northeast Linguistics Society*, Volume 23. Amherst, Mass: Graduate Linguistics Students Association. 367–81.

Pierrehumbert, Janet B. (2002). An unnatural process. *The 8th Meeting on Laboratory Phonology*.

Piggott, Glyne L. & Harry van der Hulst. (1997). Locality and the nature of nasal harmony. *Lingua* 103. 85–112.

Piggott, Glyne L. (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* 10. 33–77.

Piggott, Glyne L. (2003). Theoretical implications of segment neutrality in nasal harmony. *Phonology 20.* Cambridge University Press. UK. 375–424.

Pinheiro, José & Bates, Douglas. (2000). Mixed-effects Models in S and S-Plus. New York: Springer.

Pothos, Emmanuel M. (2007). Theories of artificial grammar learning. *Psychological Bulletin* 133. 227–244.

Prince, Alan & Paul Smolensky. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Rutgers Technical Report TR-2. New Brunswick, NJ, Rutgers Center for Cognitive Science [also Rutgers Optimality Archive 537]. Published in 2004 by Blackwell.

Prince, Alan & Paul Smolensky. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Blackwell.

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

Reber, Arthur S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning*. 855–863.

Reber, Arthur S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory* 2. 88–94.

Reber, Arthur S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118. 219–235.

Ren, Jie, Liqun Gao & James L. Morgan. (2010). Mandarin speakers' knowledge of the sonority sequencing principle. Paper presented at the 20th Colloquium of Generative Grammar, University of Pompeu Fabra, Barcelona.

Rice, Karen. (2007). Markedness in phonology. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 79–97.

Robins, Robert H. (1957). Vowel nasality in Sundanese: A phonological and grammatical study. *Studies in Linguistic Analysis*. Oxford: Blackwell. 87–103.

Rose, Sharon & Rachel Walker. (2002). A typology of consonant agreement as correspondence. Submitted. /ROA-458-0801/

Rose, Sharon & Rachel Walker. (2004). A typology of consonant agreement as correspondence. *Language* 80(3). 475–531.

Saffran, Jenny R. & Erik Thiessen D. (2003). Pattern induction by infant language learners. *Developmental Psychology* 39(3). 484–94.

Saffran, Jenny R., Elizabeth K. Johnson, Richard N. Aslin & Elissa L. Newport. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*. 27–52.

Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport. (1996). Statistical learning by8-month-old infants. *Scienc*e 274. 1926–1928.

Schneider, Walter, Amy Eschman & Anthony Zuccolotto. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Seidl, Amanda & Eugene Buckley. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development* 1(3 & 4). 289–316.

Sharon Peperkamp & Alexander Martin. (2016). Sleep-dependent consolidation in the learning of natural vs. unnatural phonological rules. The 15th Conference on Laboratory Phonology (LabPhon15). July 13-1, 2016, Cornell University.

Shepard, Roger N., Carl Hovland L. & Herbert Jenkins M. (1961). Learning and memorization of classifications. *Psychological Monographs* 75 (13, Whole No. 517). 1–42.

Skoruppa, Katrin & Peperkamp, Sharon. (2011). Adaptation to novel accents: Feature-based learning in context-sensitive phonological regularities. *Cognitive Science* 35(2), 348–366.

Smith, David J., Paul Minda J. & David Washburn A. (2004). Category learning in rhesus monkeys: A study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology: General* 133(3). 398–404.

Smith, Jennifer. (2002). *Phonological augmentation in prominent positions*. Doctoral dissertation, University of Massachusetts, Amherst.

St. Clair, Michelle C. & Padraic Monaghan. (2008). Language abstraction: Consolidation of language structure during sleep. *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. In Bradley C. Love, Kateri McRae & Vladimir M. Sloutsky (eds.). 727–732. Austin Texas: Cognitive Science Society.

Steriade, Donca. (2001). Directional asymmetries in place assimilation: A perceptual account. In Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. San Diego: Academic Press. 219–250.

Terbeek, Dale. (1977). A cross-language multidimensional scaling study of vowel perception. *UCLA Working Papers in Linguistics*, 37. 1271.

Ternes, Elmar. (1973). The phonemic analysis of Scottish Gaelic. Hamburg: Buske.

Ternes, Elmar. (2006). *The phonemic analysis of Scottish Gaelic*. 3rd edition. Dublin: Dublin Institute for Advanced Studies.

Ting, Pang-Hsin. (1985). *Taiwan Yuyan Yuanliou (Sources of Languages in Taiwan)*. Student: Taipei.

Toro, Juan M., Marina Nespor, Jacques Mehler & Luca Bonatti L. (2008a). Finding words and rules in a speech stream. *Psychological Science* 19(2). 137–44.

Toro, Juan. M., Mohinish Shukla, Marina Nespor & Ansgar Endress D. (2008b). The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception and Psychophysics* 70(8). 1515–25.

Tu, Jung-Yueh & Stuart Davis. (2009). Japanese loanwords into Taiwanese Southern Min. In *Proceedings of the Second International Conference on East Asian Linguistics.* Simon Fraser University.

Tu, Jung-Yueh. (2013). *Word prosody in loanword phonology: Focus on Japanese borrowings into Taiwanese Southern Min*. Doctoral dissertation. Indiana University.

Ullman, Michael T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92(1–2). 231–270.

Ullman, Michael T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In Christina Sanz (ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

van der Hulst, Harry. (1994a). Radical CV phonology: The locational gesture. *UCL Working Papers in Linguistics* 6. 439–477.

van der Hulst, Harry. (1994b). An Introduction to Radical CV Phonology. In Susana Shore & Marisa Vilkuna (eds.) Sky 1994: Yearbook of the Linguistics Association of Finland. Helsinki: The Linguistics Association of Finland. 23–56.

Visser, Ingmar, Maartje E. J. Raijmakers & Emmanuel M. Pothos. (2009). Individual strategies in artificial grammar learning. *American Journal of Psychology* 122(3). 293–307.

Wagner, Ullrich, Stephen Gais, Hlide Haider, Rolf Verleger & Jan Born. (2004). Sleep inspires insight. *Nature* 427. 352–355.

Walker, Douglas C. (1984). *The pronunciation of Canadian French*. Ottawa: University of Ottawa Press.

Walker, Rachel. (1998). *Nasalization, neutral segments, and opacity effects*. Doctoral dissertation, University of California, Santa Cruz.

Walker, Rachel. (2000). *Nasalization, neutral segments and opacity effects.* New York: Garland.

Walker, Rachel. (2011). Nasal harmony. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell. 1838–1865.

Wang, Samuel. (1993). Nasality as an autosegment in Taiwanese. Paper presented at the First International Conference on Languages in Taiwan, March, National Normal University, Taipei (Also in Wang 1995, 66–85).

Wang, Samuel. (1995). Experimental studies in Taiwanese phonology. Taipei: The Crane Publishing Co. Ltd.

Wang, Samuel. (2001). A concept formation experiment on the nasality of vowels in Taiwan Min. *Language and Linguistics*, 2.1, 183–195.

Warker, Jill A. & Gary Dell S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32, 387–398.

Warker, Jill A., Christine A. Whalen & Samantha Gereg. (2008). Limits on learning phonotactic constraints from recent production experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(5). 1209–1295.

White, James. (2013). *Bias in phonological learning: Evidence from saltation*. Doctoral dissertation. UCLA.

Williamson, Kay. (1965). *A grammar of the Kolokuma dialect of Ijo*. Cambridge: Cambridge University Press.

Williamson, Kay. (1969). Ijo. In Elizabeth Dunstan (ed.) *Twelve Nigerian languages*. New York: Africana Publishing. 97–114.

Williamson, Kay. (1987). Nasality in Ijo. In David Odden (ed.) *Current approaches to African linguistics*, vol. 4. Dordrecht: Foris. 397–415.

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

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

Yang, Charles. (2004). Universal grammar, statistics, or both. *Trends in cognitive sciences* 8, 451–456.

Zec, Draga. (2007). The syllable. In Paul de Lacy (ed,), *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 161–193.

Zhang, Zheng-Xiang. (1989). *Notes on Southern Min dialects of Taiwan (Language of publication: Chinese)*. The Liberal Arts Press: Taipei.

Zhao, Xu & Iris Berent. (2011). Are markedness constraints universal? Evidence from Mandarin Chinese speakers. Paper presented at the Boston University Conference on Language Development. Boston, MA.

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

Appendix I

In this Appendix I, I include the complete stimuli for Experiment 1.

Table 70. Pattern (A): **exposure phase**

Positive stimuli in the exposure phase (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Set Number	Singular form	Plural form
1st set	(1) /asawa/ ‘one apple’	(1’) /ãsawa/ ‘two apples’
2nd set	(2) /awasa/ ‘one bird’	(2’) /ãw̃ãsa/ ‘two birds’
3rd set	(3) /asasa/ ‘one car’	(3’) /ãsasa/ ‘two cars’
4th set	(4) /awawa/ ‘one cat’	(4’) /ãw̃ãw̃ã/ ‘two cats’
5th set	(5) /asewe/ ‘one flower’	(5’) /ãsewe/ ‘two flowers’
6th set	(6) /awese/ ‘one ball’	(6’) /ãw̃ẽse/ ‘two balls’
7th set	(7) /asese/ ‘one dog’	(7’) /ãsese/ ‘two dogs’
8th set	(8) /awewe/ ‘one bowl’	(8’) /ãwewe/ ‘two bowls’
9th set	(9) /asawe/ ‘one pencil’	(9’) /ãsawe/ ‘two pencils’
10th set	(10) /awase/ ‘one jacket’	(10’) /ãw̃ãse/ ‘two jackets’
11th set	(11) /asase/ ‘one ring’	(11’) /ãsase/ ‘two rings’
12th set	(12) /awawe/ ‘one book’	(12’) /ãw̃ãw̃ẽ/ ‘two books’
13th set	(13) /asewa/ ‘one pig’	(13’) /ãsewa/ ‘two pigs’
14th set	(14) /awesa/ ‘one box’	(14’) /ãw̃ẽsa/ ‘two boxes’
15th set	(15) /asesa/ ‘one watch’	(15’) /ãsesa/ ‘two watches’
16th set	(16) /awewa/ ‘one bag’	(16’) /ãw̃ẽw̃ã/ ‘two bags’

Table 71. Pattern (A): **negative old** stimuli in the **test phase** (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Plural form
(1') /ãs̃ãwa/
(2') /ãwapa/
(3') /ãs̃ãs̃ã/
(4') /ãwawa/
(5') /ãs̃ẽwe/
(6') /ãwepe/
(7') /ãs̃ẽs̃ẽ/
(8') /ãwewe/
(9') /ãs̃ãwe/
(10') /ãwape/
(11') /ãs̃ãs̃ẽ/
(12') /ãwawe/
(13') /ãs̃ẽwa/
(14') /ãwepa/
(15') /ãs̃ẽs̃ã/
(16') /ãwewa/

Table 72. Pattern (A): **positive new** stimuli in the **test phase** (/s/ as a blocker (exposure)→/k/ as a blocker (test))

The items in grey are crucial items that test whether there is an implicational universal relationship (in this case, a new consonant /k/ with a target).

Trigger+/blocker=k/+target (4)	**Trigger+/blocker=h/+target (4)**
/ãkiyi/	/ãhiyi/
/ãkɔyɔ/	/ãhɔyɔ/
/ãkiyɔ/	/ãhiyɔ/
/ãkɔyi/	/ãhɔyi/
Trigger+target+/blocker=k/ (4)	**Trigger+target+/blocker=h/ (4)**
/ãỹĩki/	/ãỹĩhi/
/ãỹɔ̃kɔ/	/ãỹɔ̃hɔ/
/ãỹĩkɔ/	/ãỹĩhɔ/
/ãỹɔ̃ki/	/ãỹĩhi/
Trigger+/blocker=k/+/blocker=k/ (4)	**Trigger+/blocker=h/+/blocker=h/ (4)**
/ãkiki/	/ãhihi/
/ãkɔkɔ/	/ãhɔhɔ/
/ãkikɔ/	/ãhihɔ/
/ãkɔki/	/ãhɔhi/
Trigger+target+target (4)	
/ãỹĩỹĩ/	
/ãỹɔ̃ỹɔ̃/	
/ãỹĩỹɔ̃/	
/ãỹɔ̃ỹĩ/	

Table 73. Pattern (A): **negative new** stimuli in the **test phase** (/s/ as a blocker (exposure)→/k/ as a blocker (test))

Trigger+/target=k/+blocker (4) /ãk̃ĩyi/ /ãk̃ɔ̃yɔ/ /ãk̃ĩyɔ/ /ãk̃ɔ̃yi/	**Trigger+/target=h/+blocker (4)** /ãh̃ĩyi/ /ãh̃ɔ̃yɔ/ /ãh̃ĩyɔ/ /ãh̃ɔ̃yi/
Trigger+blocker+/target=k/ (4) /ãyiki/ /ãyɔkɔ/ /ãyikɔ/ /ãyɔki/	**Trigger+blocker+/target=h/ (4)** /ãyihi/ /ãyɔhɔ/ /ãyihi/ /ãyihɔ/
Trigger+/target=k/+/target=k/ (4) /ãk̃ĩk̃ĩ/ /ãk̃ɔ̃k̃ɔ̃/ /ãk̃ĩk̃ɔ̃/ /ãk̃ɔ̃k̃ĩ/	**Trigger+/target=h/+/target=h/ (4)** /ãh̃ĩh̃ĩ/ /ãh̃ɔ̃h̃ɔ̃/ /ãh̃ĩh̃ɔ̃/ /ãh̃ɔ̃h̃ĩ/
Trigger+blocker+blocker (4) /ãyiyi/ /ãyɔyɔ/ /ãyiyɔ/ /ãyɔyi/	

Table 74. Pattern (B): **exposure phase**

Positive stimuli in the exposure phase (/k/ as a blocker (exposure)→/s/ as a blocker (test))

Set Number	Singular form	Plural form
1st set	(1) /akawa/ ‘one apple’	(1’) /ãkawa/ ‘two apples’
2nd set	(2) /awaka/ ‘one bird’	(2’) /ãw̃ãka/ ‘two birds’
3rd set	(3) /akaka/ ‘one car’	(3’) /ãkaka/ ‘two cars’
4th set	(4) /awawa/ ‘one cat’	(4’) /ãw̃ãw̃ã/ ‘two cats’
5th set	(5) /akewe/ ‘one flower’	(5’) /ãkewe/ ‘two flowers’
6th set	(6) /aweke/ ‘one ball’	(6’) /ãw̃ẽke/ ‘two balls’
7th set	(7) /akeke/ ‘one dog’	(7’) /ãkeke/ ‘two dogs’
8th set	(8) /awewe/ ‘one bowl’	(8’) /ãwewe/ ‘two bowls’
9th set	(9) /akawe/ ‘one pencil’	(9’) /ãkawe/ ‘two pencils’
10th set	(10) /awake/ ‘one jacket’	(10’) /ãw̃ãke/ ‘two jackets’
11th set	(11) /akake/ ‘one ring’	(11’) /ãkake/ ‘two rings’
12th set	(12) /awawe/ ‘one book’	(12’) /ãw̃ãw̃ẽ/ ‘two books’
13th set	(13) /akewa/ ‘one pig’	(13’) /ãkewa/ ‘two pigs’
14th set	(14) /aweka/ ‘one box’	(14’) /ãw̃ẽka/ ‘two boxes’
15th set	(15) /akeka/ ‘one watck’	(15’) /ãkeka/ ‘two watckes’
16th set	(16) /awewa/ ‘one bag’	(16’) /ãw̃ẽw̃ã/ ‘two bags’

Table 75. Pattern (B): **negative old** stimuli in the **test phase** (/k/ as a blocker (exposure)→/s/ as a blocker (test))

Plural form
(1’) /ãk̃ãwa/
(2’) /ãwapa/
(3’) /ãk̃ãk̃ã/
(4’) /ãwawa/
(5’) /ãk̃ẽwe/
(6’) /ãwepe/
(7’) /ãk̃ẽk̃ẽ/
(8’) /ãwewe/
(9’) /ãk̃ãwe/
(10’) /ãwape/
(11’) /ãk̃ãk̃ẽ/
(12’) /ãwawe/
(13’) /ãk̃ẽwa/
(14’) /ãwepa/
(15’) /ãk̃ẽk̃ã/
(16’) /ãwewa/

Table 76. Pattern (B): **positive new** stimuli in the **test phase** (/k/ as a blocker (exposure)→/s/ as a blocker (test))

The items in grey are crucial items that test whether there is an implicational universal relationship (in this case, a new consonant /s/ with a target).

Trigger+/blocker=t/+target (4)	**Trigger+/blocker=s/+target (4)**
/ãtiyi/	/ãsiyi/
/ãtɔyɔ/	/ãsɔyɔ/
/ãtiyɔ/	/ãsiyɔ/
/ãtɔyi/	/ãsɔyi/
Trigger+target+/blocker=t/ (4)	**Trigger+target+/blocker=s/ (4)**
/ãỹĩti/	/ãỹĩsi/
/ãỹɔ̃tɔ/	/ãỹɔ̃sɔ/
/ãỹĩtɔ/	/ãỹĩsɔ/
/ãỹɔ̃ti/	/ãỹĩsi/
Trigger+/blocker=t/+/blocker=t/ (4)	**Trigger+/blocker=s/+/blocker=s/ (4)**
/ãtiti/	/ãsisi/
/ãtɔtɔ/	/ãsɔsɔ/
/ãtitɔ/	/ãsisɔ/
/ãtɔti/	/ãsɔsi/
Trigger+target+target (4)	
/ãỹĩỹĩ/	
/ãỹɔ̃ỹɔ̃/	
/ãỹĩỹɔ̃/	
/ãỹɔ̃ỹĩ/	

Table 77. Pattern (B): **negative new** stimuli in the **test phase** (/k/ as a blocker (exposure)→/s/ as a blocker (test))

Trigger+/target=t/+blocker (4)	**Trigger+/target=s/+blocker (4)**
/ãt̃ĩyi/	/ãs̃ĩyi/
/ãt̃ɔ̃yɔ/	/ãs̃ɔ̃yɔ/
/ãt̃ĩyɔ/	/ãs̃ĩyɔ/
/ãt̃ɔ̃yi/	/ãs̃ɔ̃yi/
Trigger+blocker+/target=t/ (4)	**Trigger+blocker+/target=s/ (4)**
/ãyiti/	/ãyisi/
/ãyɔtɔ/	/ãyɔsɔ/
/ãyitɔ/	/ãyisɔ/
/ãyɔti/	/ãyɔsi/
Trigger+/target=s/+/target=s/ (4)	**Trigger+/target=t/+/target=t/ (4)**
/ãs̃ĩs̃ĩ/	/ãt̃ĩt̃ĩ/
/ãs̃ɔ̃s̃ɔ̃/	/ãt̃ɔ̃t̃ɔ̃/
/ãs̃ĩs̃ɔ̃/	/ãt̃ĩt̃ɔ̃/
/ãs̃ɔ̃s̃ĩ/	/ãt̃ɔ̃t̃ĩ/
Trigger+blocker+blocker (4)	
/ãyiyi/	
/ãyɔyɔ/	
/ãyiyɔ/	
/ãyɔyi/	

Table 78. Pattern (C): **exposure phase**

Positive stimuli in the exposure phase (/p/ as a blocker (exposure)→/k/ as a blocker (test))

Set Number	Singular form	Plural form
1st set	(1) /apawa/ 'one apple'	(1') /ãpawa/ 'two apples'
2nd set	(2) /awapa/ 'one bird'	(2') /ãw̃ãpa/ 'two birds'
3rd set	(3) /apapa/ 'one car'	(3') /ãpapa/ 'two cars'
4th set	(4) /awawa/ 'one cat'	(4') /ãw̃ãw̃ã/ 'two cats'
5th set	(5) /apewe/ 'one flower'	(5') /ãpewe/ 'two flowers'
6th set	(6) /awepe/ 'one ball'	(6') /ãw̃ẽpe/ 'two balls'
7th set	(7) /apepe/ 'one dog'	(7') /ãpepe/ 'two dogs'
8th set	(8) /awewe/ 'one bowl'	(8') /ãw̃ẽw̃ẽ/ 'two bowls'
9th set	(9) /apawe/ 'one pencil'	(9') /ãpawe/ 'two pencils'
10th set	(10) /awape/ 'one jacket'	(10') /ãw̃ãpe/ 'two jackets'
11th set	(11) /apape/ 'one ring'	(11') /ãpape/ 'two rings'
12th set	(12) /awawe/ 'one book'	(12') /ãw̃ãw̃ẽ/ 'two books'
13th set	(13) /apewa/ 'one pig'	(13') /ãpewa/ 'two pigs'
14th set	(14) /awepa/ 'one box'	(14') /ãw̃ẽpa/ 'two boxes'
15th set	(15) /apepa/ 'one watch'	(15') /ãpepa/ 'two watches'
16th set	(16) /awewa/ 'one bag'	(16') /ãw̃ẽw̃ã/ 'two bags'

Table 79. Pattern (C): **negative old** stimuli in the **test phase** (/p/ as a blocker (exposure)→/k/ as a blocker (test))

Plural form
(1’) /ãp̃ãwa/
(2’) /ãwapa/
(3’) /ãp̃ãp̃ã/
(4’) /ãwawa/
(5’) /ãp̃ẽwe/
(6’) /ãwepe/
(7’) /ãp̃ẽp̃ẽ/
(8’) /ãwewe/
(9’) /ãp̃ãwe/
(10’) /ãwape/
(11’) /ãp̃ãp̃ẽ/
(12’) /ãwawe/
(13’) /ãp̃ẽwa/
(14’) /ãwepa/
(15’) /ãp̃ẽp̃ã/
(16’) /ãwewa/

Table 80. Pattern (C): **positive new** stimuli in the **test phase** (/p/ as a blocker (exposure)→/k/ as a blocker (test))

The items in grey are crucial items that test whether there is an implicational universal relationship (in this case, a new consonant /k/ with a target).

Trigger+/blocker=k/+target (4) /ãkiyi/ /ãkɔyɔ/ /ãkiyɔ/ /ãkɔyi/	**Trigger+/blocker=t/+/blocker=t/ (4)** /ãtiti/ /ãtɔtɔ/ /ãtitɔ/ /ãtɔti/
Trigger+target+/blocker=k/ (4) /ãỹĩki/ /ãỹɔ̃kɔ/ /ãỹĩkɔ/ /ãỹɔ̃ki/	**Trigger+target+/blocker=t/ (4)** /ãỹĩti/ /ãỹɔ̃tɔ/ /ãỹĩtɔ/ /ãỹĩti/
Trigger+/blocker=k/+/blocker=k/ (4) /ãkiki/ /ãkɔkɔ/ /ãkikɔ/ /ãkɔki/	**Trigger+/blocker=t/+target (4)** /ãtiyi/ /ãtɔyɔ/ /ãtiyɔ/ /ãtɔyi/
Trigger+target+target (4) /ãỹĩỹĩ/ /ãỹɔ̃ỹɔ̃/ /ãỹĩỹɔ̃/ /ãỹɔ̃ỹĩ/	

Table 81. Pattern (C): **negative new** stimuli in the **test phase** (/p/ as a blocker (exposure)→/k/ as a blocker (test))

Trigger+/target=k/+blocker (4) /ãk͠ĩyi/ /ãk͠ɔ̃yɔ/ /ãk͠ĩyɔ/ /ãk͠ɔ̃yi/	**Trigger+/target=t/+blocker (4)** /ãt͠ĩyi/ /ãt͠ɔ̃yɔ/ /ãt͠ĩyɔ/ /ãt͠ɔ̃yi/
Trigger+blocker+/target=k/ (4) /ãyiki/ /ãyɔkɔ/ /ãyikɔ/ /ãyɔki/	**Trigger+blocker+/target=t/ (4)** /ãyiti/ /ãyɔtɔ/ /ãyitɔ/ /ãyiti/
Trigger+/target=k/+/target=k/ (4) /ãk͠ĩk͠ĩ/ /ãk͠ɔ̃k͠ɔ̃/ /ãk͠ĩk͠ɔ̃/ /ãk͠ɔ̃k͠ĩ/	**Trigger+/target=t/+/target=t/ (4)** /ãt͠ĩt͠ĩ/ /ãt͠ɔ̃t͠ɔ̃/ /ãt͠ĩt͠ɔ̃/ /ãt͠ɔ̃t͠ĩ/
Trigger+blocker+blocker (4) /ãyiyi/ /ãyɔyɔ/ /ãyiyɔ/ /ãyɔyi/	

Table 82. Pattern (D): **exposure phase**

Positive stimuli in the exposure phase (/k/ as a blocker (exposure) →/p/ as a blocker (test))

Set Number	Singular form	Plural form
1st set	(1) /akawa/ 'one apple'	(1') /ãkawa/ 'two apples'
2nd set	(2) /awaka/ 'one bird'	(2') /ãw̃ãka/ 'two birds'
3rd set	(3) /akaka/ 'one car'	(3') /ãkaka/ 'two cars'
4th set	(4) /awawa/ 'one cat'	(4') /ãw̃ãw̃ã/ 'two cats'
5th set	(5) /akewe/ 'one flower'	(5') /ãkewe/ 'two flowers'
6th set	(6) /aweke/ 'one ball'	(6') /ãw̃ẽke/ 'two balls'
7th set	(7) /akeke/ 'one dog'	(7') /ãkeke/ 'two dogs'
8th set	(8) /awewe/ 'one bowl'	(8') /ãw̃ẽw̃ẽ/ 'two bowls'
9th set	(9) /akawe/ 'one pencil'	(9') /ãkawe/ 'two pencils'
10th set	(10) /awake/ 'one jacket'	(10') /ãw̃ãke/ 'two jackets'
11th set	(11) /akake/ 'one ring'	(11') /ãkake/ 'two rings'
12th set	(12) /awawe/ 'one book'	(12') /ãw̃ãw̃ẽ/ 'two books'
13th set	(13) /akewa/ 'one pig'	(13') /ãkewa/ 'two pigs'
14th set	(14) /aweka/ 'one box'	(14') /ãw̃ẽka/ 'two boxes'
15th set	(15) /akeka/ 'one watch'	(15') /ãkeka/ 'two watches'
16th set	(16) /awewa/ 'one bag'	(16') /ãw̃ẽw̃ã/ 'two bags'

Table 83. Pattern (D): **negative old** stimuli in the **test phase** (/k/ as a blocker (exposure)→/p/ as a blocker (test))

Plural form
(1') /ãk̃ãwa/
(2') /ãwaka/
(3') /ãk̃ãk̃ã/
(4') /ãwawa/
(5') /ãk̃ẽwe/
(6') /ãweke/
(7') /ãk̃ẽk̃ẽ/
(8') /ãwewe/
(9') /ãk̃ãwe/
(10') /ãwake/
(11') /ãk̃ãk̃ẽ/
(12') /ãwawe/
(13') /ãk̃ẽwa/
(14') /ãweka/
(15') /ãk̃ẽk̃ã/
(16') /ãwewa/

Table 84. Pattern (D): **positive new** stimuli in the **test phase** (/k/ as a blocker (exposure)→/p/ as a blocker (test))

The items in grey are crucial items that test whether there is an implicational universal relationship (in this case, a new consonant /p/ with a target).

Trigger+/blocker=t/+target (4)	**Trigger+/blocker=p/+target (4)**
/ãtiyi/	/ãpiyi/
/ãtɔyɔ/	/ãpɔyɔ/
/ãtiyɔ/	/ãpiyɔ/
/ãtɔyi/	/ãpɔyi/
Trigger+target+/blocker=t/ (4)	**Trigger+target+/blocker=p/ (4)**
/ãỹĩti/	/ãỹĩpi/
/ãỹɔ̃tɔ/	/ãỹɔ̃pɔ/
/ãỹĩtɔ/	/ãỹĩpɔ/
/ãỹɔ̃ti/	/ãỹɔ̃pi/
Trigger+/blocker=t/+/blocker=t/ (4)	**Trigger+/blocker=p/+/blocker=p/ (4)**
/ãtiti/	/ãpipi/
/ãtɔtɔ/	/ãpɔpɔ/
/ãtitɔ/	/ãpipɔ/
/ãtɔti/	/ãpɔpi/
Trigger+target+target (4)	
/ãỹĩỹĩ/	
/ãỹɔ̃ỹɔ̃/	
/ãỹĩỹɔ̃/	
/ãỹɔ̃ỹĩ/	

Table 85. Pattern (D): **negative new** stimuli in the **test phase** (/k/ as a blocker (exposure)→/p/ as a blocker (test))

Trigger+/target=t/+blocker (4) /ãt̃ĩyi/ /ãt̃ɔ̃yɔ/ /ãt̃ĩyɔ/ /ãt̃ɔ̃yi/	**Trigger+/target=p/+blocker (4)** /ãp̃ĩyi/ /ãp̃ɔ̃yɔ/ /ãp̃ĩyɔ/ /ãp̃ɔ̃yi/
Trigger+blocker+/target=t/ (4) /ãyiti/ /ãyɔtɔ/ /ãyitɔ/ /ãyɔti/	**Trigger+blocker+/target=p/ (4)** /ãyipi/ /ãyɔpɔ/ /ãyipɔ/ /ãyɔpi/
Trigger+/target=t/+/target=t/ (4) /ãt̃ĩt̃ĩ/ /ãt̃ɔ̃t̃ɔ̃/ /ãt̃ĩt̃ɔ̃/ /ãt̃ɔ̃t̃ĩ/	**Trigger+/target=p/+/target=p/ (4)** /ãp̃ĩp̃ĩ/ /ãp̃ɔ̃p̃ɔ̃/ /ãp̃ĩp̃ɔ̃/ /ãp̃ɔ̃p̃ĩ/
Trigger+blocker+blocker (4) /ãyiyi/ /ãyɔyɔ/ /ãyiyɔ/ /ãyɔyi/	

Appendix II

In this Appendix II, include the inferential statistics for the comparison between Patterns 5 (Exp. 3) and 6 (Exp. 3), Patterns 1 (Exp. 2) and 7 (Exp. 3), and Patterns 2 (Exp. 2) and 8 (Exp. 3).

Table 86. Effects and interaction for Pattern 6: t(W)→k(W) and Pattern 5: t(S)→k(S) (by-subject & item analysis) ("Direction: coded as 1 for Pattern 6, coded as -1 for Pattern 5)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.0799	0.4384	4.745	2.09e-06 ***
Direction	0.6364	0.4296	1.482	0.138472
TypeNameO	**1.5852**	**0.4585**	**3.458**	**0.000545 *** **
TypeNameNS	**0.7241**	**0.3814**	**1.898**	**0.057654 .**
Direction: TypeNameO	0.3699	0.4573	0.809	0.418550
Direction:TypeNameNS	-0.2179	0.3814	-0.571	0.567786

'***': *p*<.001, '**': *p*<.01, '*': *p*<.05, '.': *p*<.1

Table 87. Effects and interaction for Pattern 7: S(t)→W(t) and Pattern 1: S(k)→W(k) (by-subject & item analysis) ("Direction": coded as 1 for Pattern 7, coded as -1 for Pattern 1)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.28585	0.62150	2.069	0.038551 *
Direction	-0.01456	0.54027	-0.027	0.978502
TypeNameO	**2.37019**	**0.50267**	**4.715**	**2.41e-06 *** **
TypeNameNS	**1.67652**	**0.45364**	**3.696**	**0.000219 *** **
Direction: TypeNameO	**0.91508**	**0.35403**	**2.585**	**0.009745 ** **
Direction:TypeNameNS	0.44270	0.27496	1.610 0.	0.107389

'***': *p*<.001, '**': *p*<.01, '*': *p*<.05, '.': *p*<.1

Table 88. Effects and interaction for Pattern 8: W(t)→S(t) and Pattern 2: W(k)→S(k) (by-subject & item analysis) ("Direction": coded as 1 for Pattern 8, coded as -1 for Pattern 2)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.4380	0.3797	-1.154	0.248661
Direction	0.1859	0.1862	0.998	0.318112
TypeNameO	**2.7872**	**0.5380**	**5.181**	**2.21e-07 *** **
TypeNameNS	**1.9068**	**0.4918**	**3.877**	**0.000106 *** **
Direction: TypeNameO	0.1436	0.3092	0.465	0.642219
Direction:TypeNameNS	-0.3995	0.2537	-1.575	0.115348

'***': *p*<.001, '**': *p*<.01, '*': *p*<.05, '.': *p*<.1

Appendix III

Instructions

Welcome to participate in this experiment.

This experiment is about learning a made-up language.

Specifically, you would learn a rule of "sounds" conditioning singular and plural forms in this made-up language.

At Stage 1, you would be familiarized with vocabulary from this made-up language, and then at Stage 2, you would be tested on whether you could learn the rule.

The rule of "sounds" conditioning singular and plural forms

Example of English:

Cat (singular form)   Cats (plural form)

A sound "s" needs to be added to cat to create a plural form cats.

Adding s is a rule about how to make singular become plural forms.

Note that the volume, speed and tone differences are just because of the defects of the sound files, you can just ignore them while learning.

This experiment is irrelevant to any languages you learned or acquired before.

Treat it as learning a brand-new language.

Stage 1: learning this language (exposure)

Every time one oral tri-syllabic vocabulary and its picture will be presented: the first tri-syllabic vocabulary with its picture is a singular, and then a plural with its picture (i.e., in pairs) will be presented.

You need to repeat every oral vocabulary just after you hear it. You should try to imitate it as closely as possible (Order: singular→plural).

Your repetition will be recorded, but no feedback will be given.

There's no semantic relationship between sounds and pictures. Pictures are just used to tell you which two vocabulary are grouped together as a pair of singular and plural. Focus on the rule of sounds.

When you see cross symbol "+", it means the next oral vocabulary with its picture will be presented.

Stage 2: test

In this stage, you will be tested on whether you learn the rule conditioning singulars and plurals.
When you hear a vocabulary, you have to judge whether it is "a possible plural form".
If you think it could be a possible plural form, press "1".
If not, press "0".
If you are not sure, make the best judgment you could based on the intuition of this language you will learn in Stage 1.
Note that in this stage, no singular/plural pairs with pictures will be presented.

Post-interview & post-test

After finishing the experiment, you will be asked what your rule is. And then you will be given a post-test.

Stage 1: learning this language (exposure)
Press space to begin.

Repeat it once. Imitate it as closely as possible.

Repeat it once. Imitate it as closely as possible.