Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21). 2009.
Volume 1. Edited by Yun Xiao. Smithfield, Rhode Island: Bryant University. Pages 81-92.

Tests of Analytic Bias in Native Mandarin Speakers and Native Southern Min Speakers*

Yu-leng Lin
Institute of Linguistics, National Chung Cheng University

In this study, I did the experiment in two language groups. Groups 1 were the speakers whose first language is Mandarin and Group 2 were speakers whose first language is Southern Min. In the experiment, the subjects had to learn two artificial languages, HH (a vowel agrees with another vowel in height) such as [titi], and HV (a high vowel is followed by a voiced consonant, or a low vowel is followed by a voiceless consonant) such as [tidi]. My goal is twofold. The first goal is to figure out whether HH is learned better than HV due to typological asymmetry. The second goal is to find out whether language-specific phonology such as Mandarin and Southern Min help subjects to learn HH and HV. Mandarin generally has no voicing contrast, but Southern Min has voicing contrast. Both languages have no vowel harmony. The findings showed that Group 1 and 2 learned HH better than HV. However, Group 2 did not learn HV better than Group 1 did. Two implications could be inferred. First, L1 phonology plays no role in learning HH and HV, because Group 2 did not learn HV better than Group 1 did. Second, both groups learned HH better than HV.

1. Introduction

Analytic bias and channel bias have been considered as two factors giving rise to typological differences in phonology (Moreton 2008, in press). The former is systematic predispositions like Universal Grammar, which help people to learn some patterns but restrain people from learning other patterns (Steriade 2001 and Wilson 2003). The latter is phonetically systematic errors, which occur when phonological representations are transmitted between speakers and hearers, caused by phonetic interactions, which act as precursors for phonologization (Ohala 1993).

Moreton (2008) ran the experiment to figure out which bias can affect typology: analytic bias, channel bias or both. The experiment was to test native English speakers to learn two patterns, height-height and voice-voice, both of which were designed in two artificial languages. The height-height patterns mean that a vowel agrees with

* I would like to thank Professor James Myers for suggestions and comments. Thanks also to Professor Elliott Moreton for clarifying his experiment design and his helpful suggestions. None of them should be responsible for defects. I take responsibility for errors.

another vowel in height such as [p**iki**] or [p**iku**]. The height-voice patterns mean that a high vowel is followed by a voiced consonant such as [p**ig**o], or a non-high vowel is followed by a voiceless consonant such as [p**ok**o]. The result showed that native English speakers learned the height-height patterns better than the height-voice patterns and this result was consistent with the fact that the height-height patterns are typologically more frequent than the height-voice patterns. The claim for typology frequency difference for these two patterns was confirmed by Moreton (2008). He tested eighteen language families and the height-height patterns outnumbered the height-voice patterns by fifteen language families to three language families.

The typological asymmetry for the height-height and height-voice patterns can result from channel bias or analytic bias. If the phonetic precursor of height-height is larger than the precursor of height-voice, then channel bias could be the cause. This hypothesis follows Ohala (1994), who claims that the more the precursor is, the more chances occur for phonologization, and therefore the more frequent the phonological pattern is. However, Moreton surveyed 7 studies, and the precursors of the height-height and height-voice patterns were calculated by measuring the vowel F1. The results showed that the vowel F1 for the height-height patterns was not larger than the height-voice patterns. That is, the phonetic precursor for the height-height patterns was not larger than the phonetic precursor for the height-voice patterns. Hence, the typological asymmetry for the height-height and height-voice patterns were not due to channel bias, because channel bias such as phonetic precursor could not assist native English speakers in learning the height-height patterns better than the height-voice patterns.

In this study, I followed the method of the experiment of Moreton (2008) and ran this experiment in two groups. Group 1 is the speakers whose first languages are Mandarin and Group 2 is the speakers whose first languages are Southern Min. The reason why I ran this experiment is that because Moreton (2008) only tested native English speakers, and he claimed that English phonology, which is irrelevant to typology, could not explain his experimental results; however, I doubt his claim. It is also possible that the height-height patterns will not be learned better than the height-voice patterns by different language speakers. That is, if the height-height patterns are not learned better than the height-voice patterns or there is no significant difference for learning the height-height and height-voice patterns, then Moreton's results are only specific to English phonology, rather than language-universal.

In order to figure out whether language-specific phonology affects height-height vs. height-voice learning, I preferred to run this experiment in two language groups, one is Mandarin and the other is Southern Min. Mandarin generally has no voicing contrast except for [ʂ] and [ʐ] while Southern Min has voicing contrast, and both languages have no vowel harmony. If language-specific phonology really plays a role in learning the height-height and height-voice patterns, then native Southern Min speakers are supposed to learn at least the height-voice patterns better than native Mandarin speakers. The reason is that Southern Min has voicing contrast, so it is easier for them to notice the relationship between vowel height and voicing in the height-voice patterns. Furthermore, both languages have no vowel harmony, so it is also impossible for the phonologies of Mandarin and of Southern Min to help both native speakers to learn the height-height patterns better.

In terms of the above assumptions, there are two goals in my study. The first goal is to find out whether both native Southern Min speakers and native Mandarin speakers learn the height-height patterns better than the height-voice patterns. If not, the results can suggest that the results of Moreton (2008) are specific to English phonology instead of language-universal. If yes, then the second goal is to find out whether L1 phonology results play a role. If native Southern Min speakers learn the height-voice patterns better than native Mandarin speakers do. Then the results suggest Southern Min phonology help the subjects to learn the height-voice patterns. However, if native Southern Min speakers do not learn the height-voice patterns better than native Mandarin speakers do, then it implies that L1 phonology has no help for native Southern Min speakers to learn the height-voice patterns. Besides, L1 phonology does not affect both languages to learn the height-height patterns better, because both languages have no vowel harmony. If the experiment rules out L1 phonology as a factor, then I can suggest that this learning asymmetry for the height-height and height-voice patterns is language-universal. In that case, analytic bias can be the only factor to lead to the asymmetry for the height-height and height-voice patterns because the channel bias such as phonetic precursors is ruled out in terms of Moreton (2008).

The paper is organized as follows. Section 2 presents Group 1's (native Mandarin speakers) results and discussion. Section 3 presents Group 2's (native Southern Min speakers) results and discussion. Section 4 concludes this paper.

2. Group 1: native Mandarin speakers

The height-height (HH) and height-voice (HV) patterns were designed in two artificial languages, and this experiment used the Artificial Grammar (AG) paradigm (Reber 1989) to compare learning of HH and HV. Wilson (2003) said a typical AG experiment includes two phases. One is the study phase, and the other is the test phase. In the study phase, subjects are exposed to stimuli which have been generated with a grammar. Then in the test phase, subjects are tested on their ability to distinguish novel stimuli (not occur in the study phase), which conform to the same grammar of the study phase from the stimuli, which does not obey the same grammar of the study phase. Besides, the AG paradigm does not have explicit negative evidence (i.e., feedback) when subjects do not choose the correct stimuli in the test phase. Hence, AG paradigm is like natural first-language acquisition.

In this study, the experiment had two language groups, native Mandarin speakers and native Southern Min speakers. In this section, I introduce Group 1, native Mandarin speakers about the method, results and discussion as follows.

2.1. Method

2.1.1. Design

The 'words' used in two artificial 'languages' had phonological structure C1V1C2V2. C1 and C2 were selected from the set /t d k g/, and V1 and V2 from the set /i u æ ɔ/. Within these limited sets, 256 'words' were possible. A word was HH-conforming if V1 and V2 were both phonologically high (/i u/) or phonologically non-high (/æ ɔ/). A word was HV-conforming if V1 and C2 were high and voiced, or non-high and voiceless. Therefore, there were 64 'words' that were both HH- and HV-conforming, 64 that were HH- but not HV-conforming, 64 that were HV- but not HH-conforming and 64 that were neither HH- nor HV-conforming. Half of the subjects were be tested the HH artificial language, and another half of subjects were be tested the HV artificial language.

In the HH artificial language, for each subject, 32 HH-conforming 'words' was randomly chosen for use in a study phase, which allowed subjects to be familiar with this artificial language. These 32 'words' were subject to the constraint designed by Moreton (2008:99) (See table 1 below): (a) vowels agree in height and {V1≠V2}, (b) vowels agree in height and {V1=V2}, (c) vowels disagree in height and {V1≠V2}, and (d) vowels disagree in height and {V1=V2}. In the HV artificial language, an analogous procedure was followed, 32 HV-conforming 'words' were chosen and conformed to the constraint: (a) V1 high iff C2 voiced and {V1≠V2}, (b) V1 high iff C2 voiced and {V1=V2}, (c) V1 high iff C2 voiceless and {V1≠V2}, and (d) V1 high iff C2 voiceless and {V1=V2} (See table 1). Note that the number (8 or 16) listed in the table 1 means that how many stimuli were put in each cell. The reason why stimuli were designed in this way is that if the study phase were designed as table 1, which had two factors, {V1=V2} and {V1≠V2}, then it is easier to see whether "Same-Vowel" affects the results. If not, then the results show that the subjects really learn the height-height and height-voice patterns, rather than depending on the patterns, which have the same vowels.

In the test phase, stimuli also obey the selection restrictions of the table 1. Another 32 HH-conforming 'words' as positive test items, which did not occur in the study phase were chosen in the HH artificial language, so did the HV artificial language. Finally, 64 'words' which were neither HH- nor HV-conforming were randomly selected for the HH and HV artificial languages as negative test items. That is, the HH artificial language had 32 negative test items in its test phase, and the HV artificial language also had 32 negative test items.

Table 1

		HV Artificial Language		HH Artificial Language	
		HH-non-conformity		HV-non-conformity	
Same-Vowel	Order	(vowels agree in height)	(vowels disagree in height)	(V1 high iff C2 voiced)	(V1 high iff C2 voiceless)
{V1≠V2}	(1st half) (2nd half)	[tidu] (N=8)	[tidæ] (N=16)	[tidu] (N=8)	[titu] (N=8)
{V1=V2)	(1st half) (2nd half)	[tidi] (N=8)	*impossible*	[tidi] (N=8)	[titi] (N=8)

Note that Moreton (2008) tested each participant to learn both artificial languages, HV and the HH. However, in my experiment, I separated the experiment into two small experiments, HH and HV. In that case, participants only learned one artificial language, either HH or HV, because learning two artificial languages were too time-consuming and tiring for a subject.

2.1.2. Subjects

Twenty participants were recruited from the students at the National Chiayi University, and National Cheng Kung University in Taiwan. All reported Mandarin as their first language and normal hearing, and all of them did not major in English or other foreign languages. All participants had early childhood dialect exposure (HV Artificial Language: Southern Min 8, and Hakka 2; HH Artificial Language: Southern Min 8, and Hakka 2). All had studied a foreign language (HV Artificial Language: English 10 and Japanese 1; HH Artificial Language: English 10 and Japanese 1). Ten participants were tested in the HV artificial language, and another ten participants were tested in the HH artificial language. The average age for the subjects of the HV artificial language was 25.4 (SD= 1.8) and the average age for the subjects of the HH artificial language was 24.8 (SD=2.0). Participants were rewarded with chocolate for the experiment, which lasted about twenty minutes.

2.1.3. Stimuli

I adopted Moreton (2008)'s stimuli, which were synthesized using the MBROLA diphone concatenative synthesizer (Dutoit et al. 1996), voice is 'US 3' (a male speaker of American English), and each 'word' was synthesized respectively. The duration of the consonant is 100 ms, the duration for the vowel is 225 ms, the duration for silence is 150 ms, and silence occurred initially and finally. Hence, the total duration for C1V1C2V2 is 950 ms (150 + 100 + 225 +100 + 225 +150). Furthermore, in order not to disturb the natural intensity difference between high and low vowels, no amplitude normalization was applied. In that case, every subject heard each stimulus with the same voice quality and duration, both of which might potentially affect the empirical results.

2.1.4. Procedure

The experiment was run by E-Prime (Schneider et al. 2002). The experiment had two parts. The first part was a study phase and the second part was a test phase. For the study phase, there were totally 32 words in this phase. Native Mandarin speakers heard a word, and pronounced it back once. The second part was a test phase, which was to test how well they could recognize 'words'. The test phase has 32 positive stimuli, which were different from stimuli in the study phase, and 32 negative stimuli. The computer said two words sequentially. One is a word of the artificial language, and the other is not. Subjects would choose '1' if it was the first word, '2' if it was the second word. The words, which belong to the artificial language in the test phase, are not the same as the words in the study phase. Half of the positive stimuli were designed to be the first word, and another half of the positive stimuli were designed to be the second word. E-Prime randomly chose these positive stimuli, so the subjects could not be able to detect the order.

2.2. Results and discussion

The result for the HV artificial language and the HH artificial language had two parts respectively. One was the raw percentage of correct response for total subject responses for HV and HH. The other was the raw percentage of correct response for four types of subject responses (like Table 1). Correct response means that the subject chooses stimuli conforming to the artificial language, rather than stimuli not conforming to the artificial language.

2.2.1 Correct response for total subject responses

The raw percentage of correct response for total subject responses and their averages for the HV artificial language and the HH artificial language are given by table 2. Mandarin speakers really learned HH better than HV (73.3% vs. 53.1%).

Table 2

	HV Artificial Language: Mandarin	HH Artificial Language: Mandarin
Average	49.1	73.3

2.2.2 Correct response for four types of subject responses

The raw percentage of correct response for four types of subject responses and their averages for the HV artificial language and HH artificial language are given by table 3 and table 4. The averages indicated that native Mandarin speakers learned HH better than HV.

Table 3

	HV Artificial Language: Mandarin			
	vowels agree in height and (V1$\neq$V2)	vowels agree in height and (V1$=$V2)	vowels disagree in height and (V1$\neq$V2)	vowels disagree in height and (V1$=$V2)
Average	55.0	40.0	51.9	*impossible*

Table 4

	HH Artificial Language: Mandarin			
	V1 high iff C2 voiced and (V1$\neq$V2)	V1 high iff C2 voiced and (V1$=$V2)	V1 high iff C2 voiceless and (V1$\neq$V2)	V1 high iff C2 voiceless and (V1$=$V2)
Average	68.8	75.0	72.5	73.8

3. Group 2: native southern Min speakers

My first goal is to find whether the results for learning the height-height and height-voice patterns are consistent with Moreton (2008). In Group 1, the results for testing native Mandarin speakers showed that the HH Artificial language was learned better than the HV artificial language.

In section 3, I want to find out whether native Southern Min speakers learn the height-height patterns better than the height-voice patterns. If yes, it implies that the learning asymmetry was not specific to English phonology.

In addition to the first goal, the second goal second goal is to find out whether

Southern Min phonology such as voicing helps native Southern Min speakers to learn the height-voice patterns better than native Mandarin speakers.

3.1. Method

Twenty participants are recruited from the community at South Region Water Resources Office, Pingtung, Taiwan. The experiment followed the same procedure as Group 1 in all respects. All reported Southern Min as their first language and normal hearing, and all of them did not major in English or other foreign languages. All had early childhood language exposure around age seven (HV Artificial Language: Mandarin 10; HH Artificial Language: Mandarin 10 and Hakka 2) and all had studied a foreign languages (HV Artificial Language: English 9 and Japanese 2; HH Artificial Language: English 8 and Japanese 2). Ten participants were tested the HV artificial language, and another ten participants were tested the HH artificial language. The average age for the subjects of the HV artificial language was 42.8 (SD= 5.1) and the average age for the subjects of the HH artificial language was 41.3 (SD= 5.7).

3.2 Results and discussion

The result for the HV artificial language and the HH artificial language also had two parts respectively as Group 1 did. The first part was the raw percentage of correct response for total subject responses for HV and HH, and the second part was the raw percentage of correct response for four types of subject responses in HV and HH.

3.2.1 Correct response for total subject responses

The averages for the HH and HV artificial languages showed that native Southern Min speakers really learned HH better than HV (66.4% vs. 43.8%). Consider table 5 as below.

Table 5

	HV Artificial Language: Southern Min	HH Artificial Language: Southern Min
Average	45.0	66.4

3.2.2 Correct response for four types of subject responses

The averages for the HV artificial language and the HH artificial language respectively demonstrated that native Southern Min speakers learned HH better than HV. Consider table 6 and table 7 as below.

Table 6

	HV Artificial Language: Southern Min			
	vowels agree in height and ($V1 \neq V2$)	vowels agree in height and ($V1 = V2$)	vowels disagree in height and ($V1 \neq V2$)	vowels disagree in height and ($V1 = V2$)
Average	47.5	36.3	48.8	*impossible*

Table 7

	HH Artificial Language: Southern Min			
	V1 high iff C2 voiced and (V1$\neq$V2)	V1 high iff C2 voiced and (V1$=$V2)	V1 high iff C2 voiceless and (V1$\neq$V2)	V1 high iff C2 voiceless and (V1$=$V2)
Average	61.3	52.5	63.75	66.3

4. General discussion

In this section, I give the summary of the main findings of the experiment for two groups and try to rule out two possibilities other than analytic bias, which also lead to the leaning asymmetry for height-height and height-voice patterns in terms of the empirical results.

4.1. Summary of empirical results

In the Group 1, ten native Mandarin speakers and another ten native Mandarin speakers were tested the HV and HH artificial languages respectively. The averages toward the raw percentage of correct response for total subject responses and the one for four types of subject responses corresponded to Moreton's result, which showed that the height-height patterns were learned better than the height-voice patterns.

In Group 2, ten native Southern Min speakers and another ten native Southern Min speakers were tested the HV artificial language and the HH artificial language individually. The average toward the raw percentage of correct response for total subject responses and the average for four types of subject responses were also consistent with Moreton's result.

Above all, native Southern Min speakers did not learn the height-voice patterns better than the native Mandarin speakers did, which suggested that voicing contrast did not affect the results.

4.2. Possibilities other than analytic bias

Two possible reasons other than analytic bias can explain why Group 1 and Group 2 learned height-height patterns better than height-voice patterns.

First, the subjects in the HH artificial language heard only HH-conforming positive test items, but the subjects in the HV artificial language heard HH-conforming and HH-non-conforming positive test items. In that case, the better performance in the HH artificial language might have no relations with learning in the experiment; instead, this result was due to a pre-existing preference for HH-conforming test items. If so, subjects in the HV artificial language would be likely to choose the HH-conforming positive test items. That is, the average for both (vowels agree in height and {V1$\neq$V2}) (55.0 %) and (vowels agree in height and {V1$=$V2}) (40.0) would be higher than that of (vowels disagree in height and {V1$\neq$V2}) (51.9%). However, in the Group 1 and Group 2, the average for HH-non-conforming positive test items (vowels disagree in height and {V1$\neq$V2}) was not the least. Consider table 8 and 9.

Table 8: correct response for four types of subject responses

	HV Artificial Language: Mandarin			
	vowels agree in height and (V1≠V2)	vowels agree in height and (V1=V2)	vowels disagree in height and (V1≠V2)	vowels disagree in height and (V1=V2)
Average	55.0	40.0	51.9	*impossible*

Table 9: correct response for four types of subject responses

	HV Artificial Language: Southern Min			
	vowels agree in height and (V1≠V2)	vowels agree in height and (V1=V2)	vowels disagree in height and (V1≠V2)	vowels disagree in height and (V1=V2)
Average	47.5	36.3	48.8	*impossible*

Second, in half of the study phase and positive test items in the HH artificial language, the stimuli, which had the identical vowels, occurred 50% (e.g. in the [titi] and [tidi] cells in Table 1). In the HV artificial language, only 25% of the stimuli that had the same vowel occurred (e.g. the [tidi] cell). Maybe the subjects in the HH artificial language did not learn to recognize stimuli, which agreed in height, but only learned to recognize stimuli, which had identical vowels. By the same logic, the better performance in the HH artificial language might have nothing to do with learning in the experiment; instead, this result was due to a pre-existing preference for repeated vowels. If so, the subjects in the HH artificial language would be likely to choose positive items whose vowels are the same. In Experiment 1, the average for items whose vowels are identical (V1 high iff C2 voiced and (V1=V2) & V1 high iff C2 voiceless and (V1=V2)) was really higher than that of items whose vowels are different (V1 high iff C2 voiced and (V1≠V2) & V1 high iff C2 voiceless and (V1≠V2)). Nevertheless, in Experiment 2, the average for items whose vowels are identical was not always higher than the average whose vowels are different. Consider table 10 and table 11.

Table 10: correct response for four types of subject responses

	HH Artificial Language: Mandarin			
	V1 high iff C2 voiced and (V1≠V2)	V1 high iff C2 voiced and (V1=V2)	V1 high iff C2 voiceless and (V1≠V2)	V1 high iff C2 voiceless and (V1=V2)
Average	68.8	75.0	72.5	73.8

Table 11: correct response for four types of subject responses

	HH Artificial Language: Southern Min			
	V1 high iff C2 voiced and (V1 ≠ V2)	V1 high iff C2 voiced and (V1 = V2)	V1 high iff C2 voiceless and (V1 ≠ V2)	V1 high iff C2 voiceless and (V1 = V2)
Average	61.3	52.5	63.75	66.3

4.3. Theoretical implications

According to the results of two experiments, two implications are presented. First, in terms of my assumptions mentioned in the introduction, if language-specific factor such as voicing contrast plays a role, then native Southern Min speakers learn at least height-voice patterns better than native Mandarin speakers. However, the results showed that native Southern Min speakers learned height-voice pattern worse than native Mandarin speakers (see table 12-14). In that case, it suggested that the language-specific phonology such as voicing in Southern Min did not help native Southern Min speakers to learn the height-voice patterns better than native Mandarin speakers. Besides, if language-specific phonology such as vowel harmony plays role, then I expected that native Mandarin and native Southern Min speakers native do not learn the height-height patterns better, because there is no vowel harmony in both languages. However, my results demonstrated that both native Southern Min and native Mandarin speakers learned the height-height patterns better than the height-height patterns. In that case, analytic bias such as Universal Grammar can explain the learning asymmetry for the height-height and height-voice patterns. The channel bias such as the phonetic precursor can not explain my results, because according to Moretion (2008) as I mentioned in the introduction, the phonetic precursor for the height-height patterns is not larger than the phonetic precursor for the height-voice patterns.

Table 12: correct response for total subject responses

	HV Artificial Language: Mandarin	HV Artificial Language: Southern Min
Average	49.1	45.0

Table 13: correct response for four types of subject responses

	HV Artificial Language: Mandarin			
	vowels agree in height and (V1 ≠ V2)	vowels agree in height and (V1 = V2)	vowels disagree in height and (V1 ≠ V2)	vowels disagree in height and (V1 = V2)
Average	55.0	40.0	51.9	*impossible*

Table 14: correct response for four types of subject responses

	HV Artificial Language: Southern Min			
	vowels agree in height and ($V1 \neq V2$)	vowels agree in height and ($V1 = V2$)	vowels disagree in height and ($V1 \neq V2$)	vowels disagree in height and ($V1 = V2$)
Average	47.5	36.3	48.8	*impossible*

Second, analytic bias prefers phonological (structural) simplicity, which means that it is easier to learn the patterns, which has one place feature, than the patterns, which has more than one place feature. That is, if the patterns have more than one place feature, then the patterns are complex (Gordon 2004). In the previous literature, linguists observed that many languages consider certain syllable types to be heavier than others (Allen 1973, and Levin 1985). Thus, Gordon (2004) used syllable weight to clarify what phonological simplicity is. He claimed that many languages regard all syllables having long vowels as heavy. Some languages regard CVV and CVC as heavy, because both of them have branching rhymes (nucleus + coda), which are the only legal positions to get moras (Hyman 1985, and Hayes 1989). Some languages regard all syllables having a certain vowel quality like treating low vowels as heavy. Nevertheless, there are no attested languages which regard all syllables having long vowels and are closed by a lateral as heavy, because no single feature can include long vowels and the syllables closed by a lateral. That is, no place feature can have [+syllabic] and [+lateral] at the same time. The phonological simplicity can explain why the height-height patterns are learned height-voice patterns. The reason is that the height-height patterns involve one place feature [height], but the height-voice patterns involve two place features [height] and [voice]. That is, the height-height patterns are phonological simple, but the height-voice patterns are phonological complex.

4.4. Unsolved problems

First, if L1 phonology plays no role, then the results of Group 1 and of Group 2 should be equal. However, native Mandarin speakers learned both the height-height and height-voice patterns better than native Southern Min speakers did. This suggested that maybe the speakers who I chose in the Group 1 and Group 2 lead to these learning difference.

Second, in the Group 1 and Group 2, both subjects speak Mandarin. Although in the Group 2, I tried to choose the native Southern Min speakers who started to learn Mandarin around 7 years old, I could not avoid the possibility for Mandarin learning experiences affect the results. The possible solution is that to broadcast Mandarin talk shows before the subjects in the Group 1 run the experiment, and that to broadcast Southern Min talk shows before the subjects in the Group 2 run the experiment.

REFERENCES

Allen, W. Sidney. 1973. *Accent and rhyme*. Cambridge: Cambridge University Press.

Dutoit, T., V. Pagel, N. Pierret, F. Bataille & O. van der Vreken. 1996. The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*. Vol. 3. 1393-1396.

Gordon, Matthew. 2004. "Syllable weight." In Hayes Bruce, Robert Kirchner and Donca Steriade (eds.), *Phonetically-Based Phonology*, 277-312. Cambridge: Cambridge University Press.

Hyman, Larry M. 1985. *A theory of phonological weight*. Dordrecht: Foris. (1992). Moraic mismatches in Bantu. *Phonology* 9, 255-266.

Hayes, Bruce. 1989. "Compensatory lengthening in moraic phonology." *Linguistic Inquiry* 20, 253-306.

Levin, Juliette. 1985. "A metrical theory of syllabicity." PhD dissertation, MIT.

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

Moreton, Elliott (in press). "Underphonologization and modularity bias." In Steve Parker (ed.) *Phonological argumentation: essays on evidence and motivation*. London, UK: Equinox.

Ohala, John J. 1993. "The phonetics of sound change." In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237-278. London: Longman.

Ohala, John J. 1994. "Hierarchies of environments for sound variation; plus implications for 'neutral' vowels in vowel harmony." *Acta Linguistica Hafniensia* 27, 371-382.

Reber, Arthur S. 1989. "Implicit learning and tacit knowledge." *Journal of Experimental Psychology: General* 118, 219-235.

Schneider, W., Eschman, A., & Zuccolotto, A. 2002. *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Steriade, Donca. 2001. "Directional asymmetries in place assimilation: A perceptual account." In Hume, Elizabeth & Keith Johnson (eds.), *The role of speech perception in phonology*, 219-250. San Diego: Academic Press.

Wilson, Colin. 2003. "Experimental investigation of phonological naturalness." *WCCFL* 22, 533-546.