Perception and Production of Yorùbá Tones by Young and Adult native Yorùbá speakers and native speakers of non-tone languages

by

Saliu Ayowumi Shittu

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics
University of Alberta

Abstract

Young participants (8-16-year-old) and adults (18 years +) from three different language backgrounds (Yorùbá speakers in Nigeria, Yorùbá speakers in Canada, and English speakers in Canada), participated in four perception experiments (one AX Discrimination, two Identification, and one Lexical), and one production (imitation) experiment. These experiments investigate the effects of age and tone language context on participants' tone perception and production accuracy. The experiments further investigate participants' perception and production of the three Yorùbá basic tones and the two complex tones in addition to the effects of attrition on the perception and production accuracy of the young Yorùbá speakers in Canada. The results reveal that globally, adults are more accurate than the young participants in tone perception and production. The results also reveal that while the young Yoruba speakers in Nigeria outperform the other young groups in all experiments, there is no significant difference between the Yorùbá-speaking and the English-speaking adults in tone discrimination and in the identification of the three basic tones. In the identification of the contour tones and in tone production, the Yorùbá speakers in Nigeria are more accurate than the Yorùbá speakers and the English speakers in Canada. The Yorùbá speakers in Canada are more accurate than the English speakers in Canada only in tone production. In the lexical recognition experiment, which involves only the Yorùbá-speaking groups, the Yorùbá speakers in Nigeria are more accurate than the Yorùbá speakers in Canada. The perception of the three basic tones reveals that while there is no significant difference in the identification accuracy between the High and the Low tones, these two tones are identified more accurately than the Mid tone. The results further reveal that the adult English speakers in Canada tend to identify the High tones with significantly more accuracy. Further analysis reveals High-Mid confusion in addition to the effects of the mid-range position of the

Mid tone; participants often misperceived both the High and the Low tones as Mid tone. The production results highlight the importance of the salient acoustic cues of the Low-tone as the tone is produced more accurately than either the High or the Mid tone. While also revealing High-Mid tone confusion, the production results further reveal that both the High and the Low tones are often misrepresented as Mid tones while Low tones rarely attract misrepresentation of the other tones. In the identification of the contour tones, the contour High tones (in LH pattern) are identified more accurately than the contour Low (in HL pattern) with further observation that the English-speaking Canadian adults tend to identify the contour-High tones significantly more accurately than the contour Low. Similarly, in Production, participants imitate the contour-High tones more accurately than the contour-Low tones. The performance of the English-speaking Canadian adults concerning the basic High tone and the contour High tones lends some support to the PAM-S model. The basic High tone is possibly assimilated to the English primary stress, while the contour High tone is comparable to the English iambic metre. Finally, the demographic factors reveal that 'longer stay in Canada' by the young Yorùbá speakers in Canada significantly lowers the group's accurate performance in tone perception and production. Older age at testing contributes to improved tone perception but not production while being born outside Nigeria does not significantly affect participants' tone perception and production skills.

Preface

This thesis is an original work by Saliu Ayowumi Shittu. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board,

Project Name:

"A study of the effects of language attrition on the perception of Yorùbá tones"

Approval No. Pro 00026974,

Approval Date: June 11, 2013

Renewals & Amendment:

REN 3: June 09, 2014

AME 4: September 21, 2015

REN 5: May 14, 2015

REN 6: May 31, 2016

The research conducted in this thesis was completed in collaboration with Dr. Anne-Michelle Tessier, Dr. Juhani Järvikivi, and Dr. Benjamin V. Tucker.

Dr. Anne-Michelle Tessier assisted in the concept formulation for the study. She also suggested the ideas for the experimental type and design, stimuli composition, and the experimental procedure.

Dr. Juhani Järvikivi assisted in suggesting and applying the appropriate statistical models used in analyzing the results for all the five experiments, discussion of the results of the experiments, chapter reviews, and manuscript editing.

Dr. Benjamin V. Tucker assisted in the initial experimental design, and in suggesting and reviewing the statistical models for the experiments, statistical analyses of the results, revising the chapters, and editing the manuscripts.

Dr. Evangelia Daskalaki, a member of my supervisory committee, assisted with the review of the thesis especially the sections related to bilingualism and heritage language.

I am responsible for the experimental design, data collection, data analysis, and manuscript writing.

Dedication

I dedicate this thesis to the memory of my father

Acknowledgement

This medium is not enough to mention the names of every person who has contributed in various ways to the success of the work reported in this thesis. I am forever grateful.

I will begin by thanking Allah for His mercy in granting me the strength and the ability to accomplish this despite my weaknesses. All praise and glory are due to HIM alone.

Many thanks are also due to my current supervisors Dr. Juhani Järvikivi and Dr. Benjamin V. Tucker. I am grateful to Dr. Järvikivi for his efforts in suggesting the best statistical models for the analyses of the results. His constructive suggestions and inputs about interpreting the results of the statistical models are very helpful to me at various stages of the analyses. I am also grateful for his time in going through the thesis multiple times. His reviews and feedbacks and comments help keep the writing in good shape.

I am also grateful to Dr. Benjamin Tucker, my co-supervisor with Dr. Järvikivi for installing the software on my laptop to enable me to conduct the experiments off-campus. He was also of immense help whenever I face problem during the early stages of the experimental design. His co-ordination of the sound booths was of enormous assist in the smooth running of subjects for the experiments in the phonetics/phonology laboratory. Along with Dr. Järvikivi, his contributions to the statistical analyses aspect, especially the R functions, are highly appreciated. I also thank Ben for his insightful comments and feedback during the review of the thesis.

To the other member of my supervisory committee - Dr. Evangelia Daskalaki, I thank you for reviewing the draft and for your advice and suggestions to improve the thesis.

I am also grateful to my former supervisor, Dr. Anne-Michelle Tessier who suggested the idea for the initial study, which evolved into the present thesis. She was very instrumental in recommending the types of experiments for the research as well as the experimental design using the E-prime software. Her insightful views about the preliminary findings are precious in shaping the ideas discussed in the experiments.

Dr. Herb Colston, the Chair of the department also deserves a special THANK YOU for his wisdom and support. He was very supportive during one of the tense moments in my program when all hopes of finding a supervisor seemed lost.

Coming to the Department of Linguistics has been a pleasant experience for me. I cherish the opportunity given me to conduct my research in an environment with the necessary modern facilities and especially among people full of goodness and warmth. The warm embrace offered by the Department compensates for the cold Edmonton winter.

I am using this medium to express my sincere gratitude to Mr. S. Aderounmu, the Principal of Umar Bin Khattab (UBK) Private School in Shaki as well as the students and teachers of the school who took part in the experiments. I am also very thankful to Mr. Jamiu Olabode and Engineer M.B Fasasi, the Director of Academic Affairs, Oke Ogun Polytechnic, Shaki for their efforts in recruiting the adult Nigerian participants for the study.

Recruiting young Canadians for the study would not have been possible without the kindness of Mrs. Mona Nashman-Smith, the Principal, Edmonton Islamic Academy, who facilitated the recruitment exercise at the school. I also acknowledge the efforts of the Vice Principal of the Academy, Mrs. Tayyaba Zahid, as well as the School Librarian, Mrs. Maryam al-Sheikh Ali, in coordinating the smooth running of the experiments in the school library. My gratitude also goes to all the students who participated as well as their parents who granted their consents. I should also acknowledge the efforts of the Department of Linguistics in making the recruitment of the adult Canadian participants possible through SONA.

I am also very thankful to my contact, Mrs. Abbas Adeladan for her efforts in recruiting participants in Calgary, Alberta, as well as to all the Canadian-based Yorùbá parents and children who took part in the experiments. The tone evaluators, Mr. Tokunbo Omiyinka and Mr. Olumide, who helped at different stages of the study also deserve heartfelt votes of gratitude.

Finally, I thank my family especially my mother for her prayers, her patience and most of all, her strength and faith through all those times she did not see me or her grandchildren. My dear wife also deserves special thanks for her love, care, patience, prayer, support, and faith in me. My children, my brothers and sisters, my relatives and friends have all been very supportive throughout the program. Thus, to everyone who has contributed to the success of this thesis in one way or another, I say THANK YOU!

Table of Contents

List of Tables

List of Figures

List of symbols

ACC = Accuracy

SD= Standard Deviation

BTC= Better than the Chance Score

N= Number of participants in the population group

H= High tone

M= Mid tone

L=Low tone

HH= High-High tone pattern

HM= High-Mid tone pattern

HL= High-Low tone pattern

MH= Mid-High tone pattern

MM= Mid-Mid tone pattern

ML= Mid-Low tone pattern

LH= Low-High tone pattern

LM= Low-Mid tone pattern

LL= Low-Low tone pattern

AD / Ad = adult participants

YG / Yg = young participants

YorNgr = Nigerian-based Yorùbá speaking participants

YorCan = Canadian-based Yorùbá-speaking participants

EngCan = Canadian-based English-speaking participants

POB= Place Of Birth

(ˊ)= An acute accent used to signify the Yorùbá High tone

(ˉ)= A macron used to signify the Yorùbá Mid tone

(ˋ)= A grave accent used to signify the Yorùbá Low tone

REF= Reference level

Cont= contour tone

NonCont= non-contour tone

Cont-L = contour Low (i.e. the L-tone in Yorùbá HL pattern)

Cont-H= contour High (i.e. the H- tone in Yorùbá LH pattern)

Chapter One: Introduction

Tone is considered one of the suprasegmental features of a language.[1] When linguists talk about tone in a linguistic context, we are generally referring to how fundamental frequency (F0) is used to indicate lexical and grammatical meanings. According to Hyman (2001:1368), a tone language (e.g., Yorùbá, Igbo, Zulu, Mandarin, Cantonese, Vietnamese), "is a language in which an indication of pitch enters into the lexical realization of at least some morphemes" (also see, e.g., Yip, 2002; Fromkin 1978, for a detailed review of tone). Most tone languages also have intonation; in intonation languages such as English, pitch is used to indicate syntactic (sentential) and semantic as well as pragmatic information (see, e.g., Cole, 2015; Cutler, Dahan & Van Donselaar, 1997 for more on intonation languages). Acoustically, tone in tone languages is signalled mainly through fundamental frequency (F0) (Ladefoged, 1962), in addition to duration and amplitude (Whalen & Xu, 1992).

The purpose of the present thesis is to investigate how listeners from different linguistic backgrounds recognize the three basic lexical tones in Yorùbá, a tone language spoken in West Africa. A number of studies have focussed on phonetic and phonological aspects of Yorùbá (e.g., Pulleyblank, 2004; Laniran and Clements, 2003; Hayward, Watkins, and Oyetade, 2003; Connell, 2002; Akinlabi and Liberman, 2000; Bakare, 1995; Connell and Ladd, 1990; Pulleyblank, 1986; Akinlabi, 1985; and Hombert, 1976, 1977). Other studies involving other African tone languages include Connell (2002: Ibibio, Kunama, Mambila, and Dschang), Connell (2000: Mambila), Hombert (1988: Bulu), and Omozuwa (1991: Edo). While most of these works focus on describing and analyzing the tone inventories as well as the tone features of these tone languages, a few of them (e.g., Connell, 2000, 2002; Hayward et al., 2003; Omozuwa 1991; and Hombert, 1976, 1977, 1988) focus on perception of these tone languages. Thus, one major aim of this thesis is to contribute to the still relatively sparse literature on tone perception in African tone languages, and in Yorùbá, more precisely.

[1]In the present study, suprasegmentals (also called prosodic features), refer to speech features that occur simultaneously with segments (vowels and consonants) such as stress, tone, word juncture, nasalization etc. Segments (segmentals) also called phonemes, refer to the units of sound that can be analyzed such as the sounds of the consonants and the vowels.

Of more precise concern to this thesis, in addition to the tone perception patterns of native tone language speakers versus those of speakers with non-tone language backgrounds, is the effect of age, language context, language-internal as well as demographic factors on participants' tone perception accuracy. To this end, the study includes young and adult participants of three different language backgrounds: native speakers of Yorùbá residing in Southwestern Nigeria, bilingual speakers of English and Yorùbá residing in Western Canada, and monolingual speakers of English. Thus, the thesis deals with three population groups with different tone language contexts and comprises of two age brackets: young and adult speakers. In addition to the factors related to participants' ages and their tone language context, demographic factors such as the participants' ages at the time of testing, their ages at time of arrival in Canada, length of stay in Canada, and place of birth are also considered. These data are obtained through an adapted Alberta Language Environment Questionnaire (ALEQ, (Paradis, Emmerzael, & Sorenson Duncan, 2010) (see Appendix A1 for the adapted ALEQ questionnaire). Finally, the study also considers language-internal factors specific to Yorùbá in terms of types of tone pairs, tone space of the three tones, syllable counts, tone contours and other factors. The present thesis is expected to contribute to the body of literature on tone perception and production from different perspectives. Theoretically, it will enhance our understanding of tone perception and production by exploring the relevance of age and language context. It will show how the individual differences of these population groups modulate their performances in tone perception and production. Methodologically, it will present experimental procedures (experimental types, participants, and materials) for investigating the perception and production of an African tone language by native and non-native tone language speakers. Empirically, it will add to the relatively sparse literature on tone perception and production in African languages.

The structure of the present thesis is as follows: Chapter 1 presents relevant background information about Yorùbá, the language under study. Chapter 2 introduces the current study focussing on the participants, materials, the experiments, and the factors that informed the research questions. Chapters 3 to 7 present the five experiments while chapter 8 presents the demographic factors for the Yorùbá-Canadian young participants. Chapter 9 presents the general discussion and the conclusion.

1.1 Background on the Yorùbá Language

Yorùbá belongs to the Niger-Congo language family spoken by about 30 million people. It is a tone language with about twenty varieties spread across Southwestern Nigeria, Southern Benin Republic and Central Togo (Orie, 2012). According to Bamgbose (1966) the varieties of Yorùbá differ from each other phonologically, lexically, and, to some extent, grammatically. Focussing on phonology, according to Orie (2012), all varieties have three tone distinctions and seven phonemic oral vowels, but phonetically, the varieties may exhibit up to nine vowel systems. The consonantal inventory is relatively stable in many varieties. Although minor differences may exist, all varieties have the same places of articulation for stops, fricatives, nasals, laterals, taps, and glides, (Orie, 2012:38). However, some varieties do not have the syllabic nasals.

As for vocabulary, one notable difference mentioned by Orie (2012) is that some varieties have lost the consonant /r/. As such, all lexical items produced with the /r/ sound in other dialects are pronounced without the consonant. For example, ārā 'body'(SY) is pronounced as āā 'body' in these other dialects. Thus, in these varieties, the minimal CV requirement for a syllable in Yorùbá is relaxed (for more examples of the phonological and lexical differences, see Orie (2012: 31-52). The section on dialect in Orie (2012) also includes a map showing the geographical locations of the Yorùbá dialects.

In this thesis I will be focussing on the Standard variety, otherwise referred to as Standard Yorùbá (SY). SY is a koine language based mainly on the variety spoken in the Oyo-Ibadan axis. [2] It is the variety learnt at school and used in Yorùbá literary works. It is used in regional print media, on radio and TV, and in academic or social conferences where members of all the other dialects meet (Bamgbose, 1967).

Standard Yorùbá has three distinctive tones, High, Mid and Low (Bamgbose, 1966; Akinlabi & Liberman, 2000). Throughout the present thesis, the High tone is marked by an acute accent as in *á*; a grave accent marks the Low tone as in *à*, and a macron marks the Mid tone as in *ā*. These

[2] The label 'Standard' does not indicate any significant difference between the written form and the form spoken by the larger population in these regions.

three tone markings are placed on the vowel or the nasal vowel portion of the syllable, or on the syllabic nasal.

Tone is an important element of the language. As a result, any change in the tone changes the meaning of the word, as shown in the minimal tone pairs presented in Table 1.1. Tone is also prominent in all grammatical aspects of the language such as phonology, morphology, syntax, semantics and pragmatics (Akinlabi & Liberman, 2000). Thus, Yorùbá is a tone language that uses pitch to signal both lexical and grammatical information.

Table 1. 1 Same phonemes with different tones realizing different pronunciations and different meanings

Word	**Gloss**	**Word**	**Gloss**	**Word**	**Gloss**	**Word**	**Gloss**	**Word**	**Gloss**
īg͡bá	calabash	ìg͡bà	time	ìg͡bá	garden egg	īg͡bā	200	īg͡bà	climbing rope
ọ̄wọ́	Hand	ọ̄wọ̀	broom	ọ̀wọ̀	respect	ọ̀wọ́	same		
mú	Take	Mù	enter into	mū	drink				
ìlú	town	Ìlù	drum	ìlū	opener				
āk͡pá	Arm	àk͡pá	mark	àk͡pà	wasteful				

The Standard variety has 18 basic consonants,[3] and seven contrastive oral vowels namely *a, e, ẹ, i, o, ọ*, and *u*. Four of the seven oral vowels have nasalized counterparts in Standard Yorùbá as in [*ã, ũ, ĩ, ọ̃*]. One issue of phonological relevance is Yorùbá vowel harmony. In Yorùbá vowel-initial words, the language's phonotactic rules allow for a sequence of some vowels while not allowing for others as shown in Table 1.2 (see Awobuluyi, 1967; Bamgbose, 1967; Awobuluyi and Bamgbose, 1967). According to the rule, words such as /*ọbẹ, ẹkọ, ẹba* and *ọwọ* / are acceptable while sequences such as /**obẹ, *ekọ, *eba* and **ọwo/* are not.

[3] SY orthography does not use the consonants *c, q, v, x, and z* while English does not have the labial-velar double articulation *kp* and *gb*. Orthographically *kp* is written as *p* in SY, but articulatorily it is a co-articulated labial-velar (e.g., *apá* = *āk͡pá*) and not a bilabial stop as in English. *gb*, represented as *gb* in SY, is a double articulation beginning with /g/ and ending with the labial /b/ as in *igba*= *ig͡ba*.

Table 1. 2 Distributional facts of [ATR] harmony for Yorùbá disyllabic vowel-initial nouns. (V1=the first vowel (the vertical axis); V2=second vowel (the horizontal axis); (✓)=a sequence that occurs; (×)= a sequence that does not occur) (Awobuluyi, 1967)

		V2						
	Vowels	ɪ	e	ẹ	ɑ	ọ	o	u
V1	ɪ	✓	✓	✓	✓	✓	✓	✓
	e	✓	✓	×	×	×	✓	✓
	ẹ	✓	×	✓	✓	✓	×	✓
	ɑ	✓	✓	✓	✓	✓	✓	✓
	ọ	✓	×	✓	✓	✓	×	✓
	o	✓	✓	×	×	×	✓	✓

The examples in Table 1.1 also indicate that there are only three syllable structures in Yorùbá: V, CV, or CṼ. Other notable phonological peculiarities of the language include the fact that Yorùbá does not have words beginning with the [u] vowel. Additionally, Yorùbá does not permit the High tone in a word-initial V-syllable structure as shown in the minimal pairs in (1a), a pattern which is permissible in a CV syllable structure as shown in the near minimal pairs in (1b).

(1) *Non-permissibility of the H tone in Yorùbá word-initial V-syllable*

a. i. *ó.dō H.M ii. *ó.dó H.H iii. *ó.dò H.L

b. i. g͡bó H ‘be ripe’ ii. g͡bó.ná H.H ‘be hot’ iii. g͡bó.yà H.L ‘be brave’

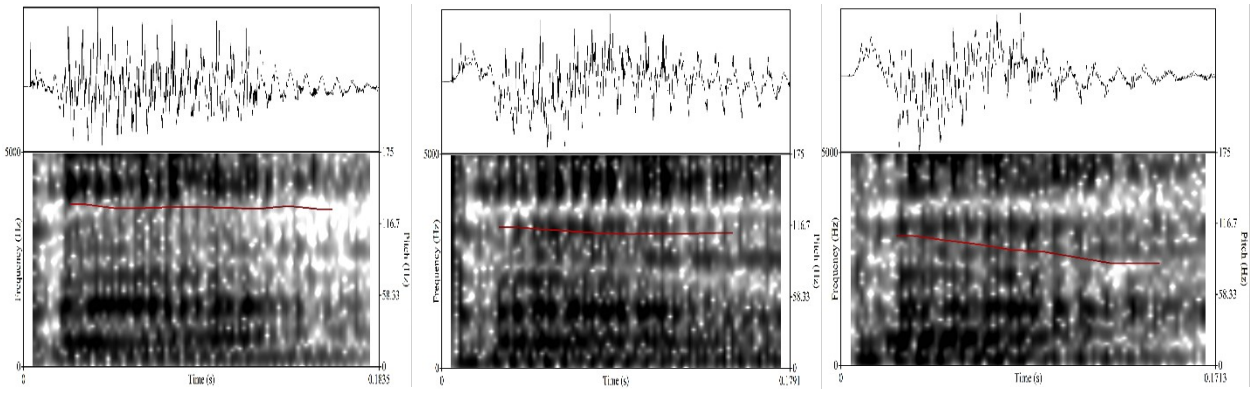
1.1.1 The Three Basic Tones

In their study of the phonetic interpretation of the three Yorùbá tones, Hayward, Watkins, and Oyetade (2003) found that the Yorùbá High tone is marked by a more prominent higher frequency than the other two tones. The Low tone, on the other hand, is marked by a more prominent low frequency and a distinctive voice quality, namely creaky phonation. The Mid tone, which to some extent, shares breathy phonation with the High tone, may be considered the least marked of the three Yorùbá tones. Bakare (1995:52), in a related finding, revealed that the best differentiator of these tones in Yorùbá is the fundamental frequency (F0) followed by

intensity, duration and formant frequencies. Bakare (1995) further revealed that the High tone has the highest intensity, the highest fundamental frequency, the highest third formant frequency and the shortest duration. To further highlight the differences between the three Yorùbá tones, Figures 1.1a to c, show the F0 measurements for the syllable '*ba*' accompanied by each of the three Yorùbá tones, taken at the mid-point of the utterance. A male native Yorùbá speaker produced the words. The spectrograms show the F0 variations and the pitch lines for the three tones. The High tone (Figure 1.1) records an F0 of 130 Hz, the Mid tone (Figure 1.1b) has an F0 of 111 Hz and the Low tone (Figure 1.1c), has an F0 of 88 Hz.

Figure 1.1 Spectrogram of the syllable 'ba' with the three Yorùbá tones

a. bá = (130Hz) b. bā = (111Hz) c. bà = 88Hz



Studies of the tones (e.g., Hombert, 1977; Connell, 2002), have reported the presence of gradient intrinsic fundamental frequency (IF0) across tones. IF0 refers to the claimed universal tendency in which high vowels have higher F0 than low vowels. In relation to Yorùbá tones, Hombert (1977) observed that the IF0 of a given tone varies according to vowel heights. The study revealed that the IF0 difference between high and low vowels is more pronounced with High tone than with Low tone. However, the scope of the present thesis will focus on how accurately the participants perceive the tones regardless of the other segmental features. In other words, the study will not pursue the effects of the preceding consonants or the effects of the vowel heights (cf. Hombert, 1977) on participants' tone perception accuracy.

Other studies of these three basic Yorùbá tones (e.g., Akinlabi, 1985; Pulleyblank, 1986, 1988; Orie, 1997; Akinlabi & Liberman, 2000) also reveal that phonologically, they vary in strength.

The High tone is the strongest and the most stable while the Mid tone is the weakest and most unstable. The High tone is easier to perceive than either the Mid or the Low tone. One instance often cited as an indicator of Mid-tone weakness is the vowel deletion context in Yorùbá. This references the rule where, across word boundaries, a vowel is deleted when it is adjacent to a second vowel. In such a context, the High and the Low tones are usually retained while the Mid tones are usually deleted as shown in (2) below.

(2) *M-tone deletion across word boundaries.*

a. dé īlé	(HMH)	→	délé	(HH) 'reach home'
b. jū ìwé	(MLH)	→	jùwé	(LH) 'throw a book'
c. wá ōkō	(HMM)	→	wókō	(HM) 'come to the farm
d. lū ìlù	(MLL)	→	lùlù	(LL) 'beat the drum'

More evidence of the strength of the High and the Low tones is revealed by Orie (2006b), who observed adult native English speakers learning Yorùbá as a second language (L2) at the initial stage of acquisition. The pattern shown by these L2 learners was to impose the High and the Low tones on all nouns based on the English stress pattern as in (3a, b) (Orie, 2006b:122). As the examples show, the Mid tone is realized as either a High tone (3a, b) or a Low tone (3c, d). In the same study, these learners also identified utterance- or word-initial Mid tones as High tones as in (4a, b). Utterance- or word-final Mid tones were also identified as Low tones as in (4c, d).

(3) *Imposition of H and L on nouns (names) by L2 learners.*

Yorùbá names			Adaptation by L2 learners	
a.Yéwáñdé	(HHMH)	→	Yèwáńdè	(LHHL)[4]
b. Títílāyọ̀	(HHML)	→	Títìláyọ̀	(HLHL)
c. Ọ̄básẫjọ́	(MHMH)	→	Ọ̀bàsẫ́jọ̀	(LLHL)
d. Ādébáyọ̀	(MHHL)	→	Àdèbáyọ̀	(LLHL) or (HLHL)

(4) *L2 learners identifying M initial as H and M-final as L.*

a. ōwó	(MH)	→	ówò	(HL)	'money'
b. ādé	(MH)	→	ádè	(HL)	'crown'
c. tóbī	(HM)	→	tóbì	(HL)	'be big'
d. sẹ́gũ̄	(HM)	→	sẹ́gũ̀	(HL)	'be victorious'

[4] Examples 3a and b are from Orie (2006)

1.1.2 The Complex Tones: Rising LH and Falling HL

In addition to the three basic Yorùbá tones, there are two other pitch modifications obtainable in derived contexts, contour tones. In the first context, when a High tone follows a Low tone, the High tone becomes a rising tone as in (5). In the second context, when a Low tone follows a High tone, it is realized as a falling tone as in (6).

(5) ìwé (LH) → ìwě (L. $_{L}$H)[5] 'book'

(6) mídò (HL) → mídô (H. $_{H}$L) 'nonce word'

One explanation for these complex tones is that the tone spreads rightward. For example, in the production of a LH sequence, the rising tone starts off from the Low tone of the first syllable, then spreads rightwards onto the High tone of the second syllable to create the contour tone. Consequently, the resulting contour does not allow for a full realization of the High tone of the second syllable. Similarly, the falling tone sets off from the High tone of the first syllable, then spreads rightwards onto the Low tone of the second. The spread from the High tone to the Low reduces the full realization of the Low tone of the second syllable (see Yip, 2002: 147 for more on Yorùbá contour tones).[6] The following paragraphs and Table 1.3 will further demonstrate how these contour tones differ from the basic tones.

By way of illustration, Table 1.3 presents pitch measurements for the last syllable '*ba*' of the word '*aba*' produced with different tone patterns ending with either the H or L tone. Thus, we have HL, LL, and ML as well as HH, MH, and LH.[7] The measurements were taken at the onset, middle, and offset of the pitch levels. A male native Yorùbá speaker produced the words. The F0 frequencies were measured in Hertz (Hz).

[5] This representation indicates the rising movement in the second syllable. The smaller font size points to a residual tone from the Low tone of the first syllable to the High tone of the second. It does not in any way indicate an elongation of the vowel. A similar process applies to the falling tone in (6).

[6] According to Yip (2002) "The effect is much greater than can be accounted for by simple coarticulation, and it is clear that what probably starts as simple perseveratory coarticulation has been phonologized" (p.147)

[7] M-final patterns are omitted here and in Figure 1.2 because the aim is to highlight only the F0 differences between Contour High and the other two H-final patterns (HH and MH), and between the Contour Low and the other two L-final patterns (LL and ML).

Table 1. 3 Pitch measurements (F0) for the final syllables ‘ba’ of the Yorùbá word ‘aba’

Word	Gloss	Onset (Hz)	Middle (Hz)	Offset (Hz)
ābáMH	‘a Nigerian city’	156.5	150.8	157.5
àbá**LH**	‘suggestion’	113.8	157.1	212.7
ábáHH	‘nonce word’	152.9	148.47	156.77
ābàML	‘farm settlement’	102.78	93.39	83.537
ábà**HL**	‘nonce word’	169.3	132.97	98.26
àbàLL	‘nonce word’	97.39	90.52	80.65

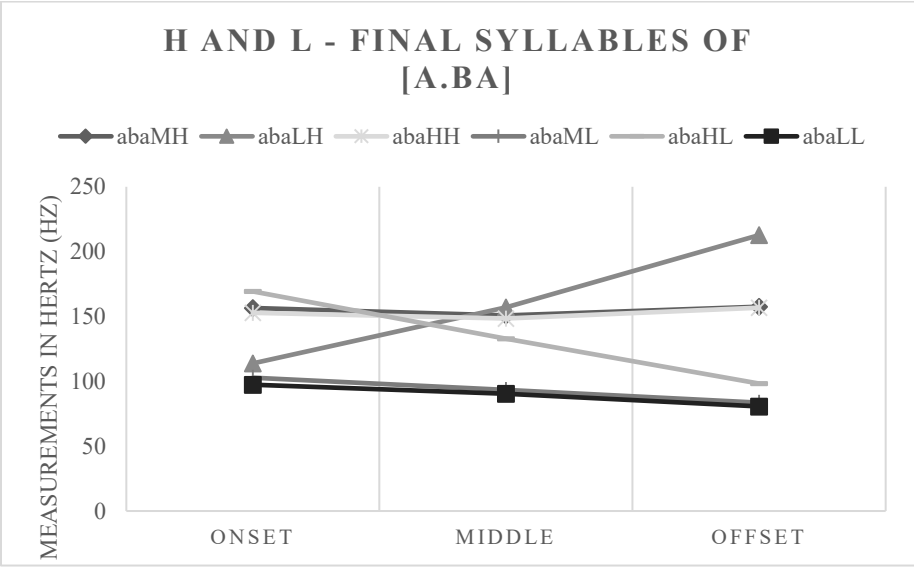
The numbers in the table show that the High tones in both MH and HH recorded similar pitch measurements at the onset (156.5; 152.9), at the middle (150.8; 148.4), and at the offset (157.5; 156.7). In comparison, the High tone in LH begins with a lower onset (113.8) than what is observed for MH and HH, but with higher mid (157.1) and offset (212.7) pitch measurements than those of MH and HH. The relatively lower onset pitch recorded for the High tone in LH is evidence of the spread of the L tone from the preceding syllable. The F0 for the H in LH is also higher but closer to the other High tones.

Similarly, the pitch levels for the word-final Low tones in both ML and LL are also similar at the onset (102.7; 97.3), at the middle (93.3; 90.5), and at the offset (83.5; 80.6). In comparison, the Low tone in HL records a higher onset (169.3), middle (132.9) and offset (98.2) than what is observed for the Low tones in ML and LL. Moreover, the onset pitch for the Low tone in HL is relatively high, showing evidence of the spread of the preceding High tone in the word. It is also observed that the F0 is much higher than for the other Low tones, as is the final pitch level. This is further evidence of the lingering effects of the preceding High tone.

To get a more precise grasp of the difference between the contour tones and the simple tones, the frequency measurements in Table 1.3 are plotted in Figure 1.4 below to give a visual representation. As shown in the graph, the last tones (H tone) in both MH and HH are closer together and are almost parallel at the onset, the middle, and the offset. In comparison, the last tone in LH starts off closer to the Low tones in ML and LL, and much lower than those of MH and HH. At mid-point, it is much closer to MH and HH, but at the offset, it ascends far higher

than both MH and HH. Similarly, the last tones (L tone) in both ML and LL begin at almost the same point of the scale, and then continue a slight descent in an almost parallel way to the middle and offset points. In comparison, the last tone in HL has the highest onset of all, a point even higher than the H tones in MH, HH, and LH. Its mid-point is much farther than those of ML and LL, and closer to those of the H tones. However, its offset fails to reach the level of the other Low tones in ML and LL. Other interesting points portrayed by the graph include the fact that the H in LH shows a clear ascending slope, while the Low in HL shows a clear descending slope. Additionally, the mid-points for the last tones in both LH and HL (H and L respectively) are closer to the H-final in MH-HH than the L-final in ML-LL.

Figure 1.2 Average F0 for H- and L-final syllables for 'ba' in 'aba' at onset, middle and offset



Among the studies that examined the local effects exerted by the three Yorùbá tones on each other are Connell and Ladd (1990), and Laniran and Clements (2003). These studies suggested that of all the tones, neighbouring H and L are found to exert local influences on each other. A look at the measurements obtained for the stimuli in Table 1.3 and Figure 1.2 appears to confirm these findings. In the HL pattern, the Low tone following the High tone triggers a raising effect. The preceding H tone is relatively higher in this utterance and the lingering effect of the pitch height is passed on to the following Low tone, hence an onset frequency higher than those of the other Low tones. The other local effect resulting from H and L sequence is observed in the LH pattern. It appears that the pitch movement from the preceding L tone to the following H tone triggers a substantial raising of the H tone, leading to a frequency higher than the H tones in MH

and HH. The behaviour of these two tones will be relevant in the way they are perceived by listeners. Since the present thesis mostly hinges on tone perception of an African tone language, the next section presents a review of some of the previous studies undertaken in the area of tone perception and production of some African tone languages.

1.2 Tone perception and production in African tone languages.

Studies that examine the perception and production of African tone languages are relatively few. Among the studies that focus on production (e.g., Hayward, Watkins & Oyetade, 2003; Connell, 2002; Hombert, 1977), there are those that attempt to describe the voice quality (F0 and phonation type) used by native speakers to mark a particular tone (e.g., Hayward et al., 2003). There are also studies that investigate the claimed universality of intrinsic fundamental frequency or intrinsic pitch (IF0) as it relates to tone languages, to be discussed in greater detail below. Other studies focus on tone perception by investigating the frequency range at which speakers of African tone languages perceive the native tones (e.g., Hombert, 1976; Connell, 2000; Omozuwa, 1991; and Hombert, 1988).

Hayward, Watkins and Oyetade (2003) investigated voice quality - pitch association in an African tone language. The purpose of the study was to investigate the voice quality and the phonation type used to mark the three Yorùbá tones. The experiment was based on recordings of three male speakers with native competence in Standard Yorùbá. The speakers were asked to read out a 42-item list of nonsense CV syllables containing all seven Yorùbá vowels combined with the three Yorùbá tones. The tokens were preceded by two initial consonants (/t/ and /l/) and were placed in a carrier phrase. The items were presented in written form with all the tone marks included. Digital audio and laryngograph recordings of participants' tokens were made simultaneously and later analyzed. The results revealed that while the vowel quality and the initial consonant exerted some influence on voice quality, the Low tone had a distinctive voice quality for all three speakers. The Yorùbá Low tone is marked by a distinctive voice quality which involves more prominent low frequency and somewhat creakier phonation. The High tone is marked by a more prominent higher frequency than the Mid tone and the two tones appear to have more or less similar breathy phonation.

Two studies, namely Connell (2002) and Hombert (1977) examined the issue of IF0 in tone languages. IFO refers to the correlation between vowel height and fundamental frequency (F0) which is considered a universal phenomenon (Whalen & Levitt, 1995). The principal claim is that high vowels are realized with higher F0 than low vowels. For example, in a language with a vowel system comprising of three vowel height levels, the IF0 of the high vowels will be higher than that of the mid vowels; the IF0 of mid vowels will be higher than that of the low vowels.

One of the studies that focused on IF0 in tone languages is Connell (2002). The study examined the effects of the size of tone inventory on IF0, as well as the extent to which Low tone neutralizes IF0. The study also investigated whether there is an interaction between IF0 and tone. Connell (2002) involved four adult native speakers from each of four African tone languages namely Ibibio (2 tones), Kunama (3 tones), Mambila (4 tones), and Dschang (4 tones: for Dschang, only the high and low vowels were used). The participants were asked to read naturally occurring words representing all possible vowels and tone combinations in their respective languages. F0 measurements for the tonal targets were taken and analyzed. This was done by identifying the F0 levels of the syllables and the presumed tonal targets while taking account of the effects of the perturbations of the preceding or following obstruents. The results showed that three of the four languages (Ibibio, Kunama, and Dschang) show IF0 for most or all of the participants, while Mambila did not show significant IF0. Connell (2002) suggests that the size of the Mambila tone inventory (4 level tones) could not have affected the results because one of the four languages - Dschang - also has 4 level tones and the IF0 observed for this language was higher than that of Ibibio with 2 level tones. Connell (2002:123) observed that the crowded tonal space, in addition to the "degree of F0 modulation used in producing tonal contrasts, appears to be the primary factor". The results further indicated that IF0 is generally reduced or almost neutralized for Low tones. Regarding the interaction between IF0 and tone, the variation observed for the four languages appears to show that there is little or no evidence to support the relationship between IF0 and tone. Finally, Connell (2002) argues that these findings indicate that a different physiological mechanism is applied in the production of Low tones.

Similar research was conducted on Yorùbá by Hombert (1977). The first experiment examined the effect of consonant types on the three Yorùbá tones. The study included two participants who were asked to read a word list of 42 CV tokens (2 consonants /g, k/ x 7 vowels x 3 tones)

inserted in a carrier phrase. Each token was read five times yielding 70 measurements (7 vowels x 5 repetitions x 2 subjects). The results revealed that voiced consonants have greater effects on High tones while voiceless consonants have greater effects on Low tones. In addition, the duration of these effects is shorter in Yorùbá than in English. The second experiment investigated the IF0 of vowels with the three Yorùbá tones. The results revealed that the IF0 difference between high and low vowels are smaller with a Low tone than with a High tone, confirming the presence of vowel gradient effect. The study further revealed that in addition to pitch, secondary cues are likely to be used to identify the tones. Such cues include phonation type (cf. Hayward et al., 2003), loudness, duration, and direction and speed of change of the tones. Finally, the study also revealed that the Yorùbá Low tone is in fact a Low-falling tone. Thus, these characteristic features of the tones may contribute more or less to the accurate perception of the tones.

The next group of studies to be discussed relate to perception of tones in African tone languages. These studies focused on the frequency range at which native listeners of these tone languages perceive the native language tones in bisyllabic nouns. One such study is Hombert's (1976) work on Yorùbá. In this study, seven Yorùbá native speakers were asked to listen to fifteen minimal pairs of Yorùbá VCV nouns previously recorded by a Yorùbá speaker. Each pair was presented five times in a randomized fashion. Participants were then asked to evaluate (on a scale of 0 to 10) the perceptual distance between the two tone patterns. The results revealed that native Yorùbá speakers perceived the bisyllabic nouns based on the direction of change of F0 of the second vowel which could either be falling (the L tone in LL and ML); level (the M and H tones in MM, MH, LM), or rising (the H tone in LH). The results further revealed that native speakers perceived bisyllabic nouns based on a combination of (a) the amount of F0 change of the second vowel (contour vs. level), (b) the average F0 value of the second vowel, and (c) the F0 difference between the onset of the second vowel and the offset of the first vowel. The second part of the study investigated whether by modifying the amplitude or the duration, or the magnitude of F0 contour of the final Low tones in LL and ML, these modified Low tones would be perceived as Mid tones. The results revealed that there was no misperception when the final Low tones were modified with increased amplitude or increased duration. Misperception occurred only when the final Low tones were modified with either steady F0 or with a combination of steady F0, increased duration and increased amplitude. The results suggest that falling F0 is the primary

perceptual cue to identifying Low tones in word-final position. The study further suggests that amplitude information is not an essential cue to distinguish between Low and Mid tones in word-final position. It needs to be mentioned that because the stimuli in Hombert (1976) involved real Yorùbá bisyllabic nouns, they could not accommodate all the phonotactically-licit Yorùbá bisyllabic patterns. Three of the nine possible patterns (all H-initial pattern: HH, HM, and HL) were excluded in the stimuli because they are not available in real Yorùbá VCV noun patterns adopted for the study. Using nonce words as is the case with the present thesis will accommodate all the possible disyllabic patterns.

Another study in this regard is Connell (2000) who examined the perception of lexical tones in Mambila, a tone language with four-level lexical tones (High, Mid-High, Mid-Low, and Low), using synthesised tones of ten level frequencies in steps of 5Hz inserted in a carrier phrase. Six native Mambila speakers and six native English speakers were pre-trained for the experiment which was similar to the one used in Hombert (1988). The participants' task was to listen to synthesized stimuli over the headphones and then group this continuum of stimuli into Mambila's four tonal categories (High, Mid-High, Mid-Low, and Low), according to the perceived height of each stimulus. The results showed that pitch height is the essential cue for the perception of lexical tones in Mambila. The results further revealed that while the identification of both the High and the Low tones were well defined by the native listeners, the identification of the two Mid tones appeared less defined. One explanation given was that the two Mid tones had considerably smaller perceptual space than either the High or the Low tone. Moreover, the study further found that overall, native tone language listeners (Mambila speakers) did not show significantly greater sensitivity to pitch than the non-tone language listeners (English speakers). However, the native Mambila speakers exhibited the effects of linguistic experience in their response patterns. They showed consistency in placing the tones into different perceptual spaces while the English participants showed more variations in placing the tones.

Omozuwa (1991) investigated the perception and production (repetition) of the two tones in Edo, a language spoken in Edo State in Southern Nigeria. The study used Edo bisyllabic nouns and it involved native Edo speakers and non-Edo speakers from five different linguistic groups namely Yorùbá, Ika, Hausa, French, and Chinese speakers. In the experiment, recorded Edo bisyllabic nouns were played back to these non-Edo participants in a listening and repetition task aimed at

simulating Edo tones. Then, the repeated tokens were presented to ten Edo native speakers/hearers to analyze. They were asked to use their knowledge of their native language in assessing the acceptability or non-acceptability of the utterances by the non-Edo speakers. The mistakes made by these non-Edo speakers served as the basis for the identification of the perceptual cues for tone perception in Edo nouns. The analysis revealed that an upward movement of the F0 in the same frequency zone, whether intrasyllabic or intersyllabic is an essential acoustic cue for the perception of the Edo HH pattern. For the perception of the LL pattern, a decrease in F0 value from the beginning to the end of F0 on each of the two syllables is essential. For the LH pattern, the major acoustic cue is an F0 difference of not less than a quarter of a musical tone between the Low tone of the first syllable and the High tone of the second syllable. Finally, in the HL pattern, the Low tone of the second syllable is realized as a High-falling tone. In addition to highlighting the importance of the form and the direction of change of F0 in Edo tone perception, the results further indicated that participants' L1 influenced their perception of Edo tones. Participants whose L1 shared some affinity with Edo (e.g., Yorùbá and Ika) performed better than those whose L1 did not (e.g., Hausa, French and Chinese). Furthermore, among these participants whose L1 were not closely related to Edo, those with experience in tone manipulation (e.g., phonetician, musicologist, speech therapist), performed well in pitch perception and repetition, regardless of their L1.

Finally, Hombert (1988) investigated the perception of monosyllabic and disyllabic nouns in Bulu, a Bantu language spoken in Southern Cameroon. Bulu has three register tones, two of which are in the same low frequency zone. The third tone occupies the high frequency zone. The study investigated the frequency range at which Bulu native speakers perceive the three tones in synthetic stimuli involving monosyllabic and bisyllabic words. One native Bulu speaker was trained for the study and thirteen monosyllabic stimuli (5 stable tones, 4 rising tones and 4 descending tones were used. In the disyllabic tone patterns, the 13 monosyllabic stimuli were preceded by either a stable Low tone (110Hz), a medium Low tone (130Hz), or a High tone (150Hz). The participant was given a response sheet on which to select the drawings associated with the item that had the tonal schemes of the stimulus. One extra space (marked X) was also included in case there was no match. The results showed that the monosyllabic stimuli were perceived as the first Low tone when the production of the Low tone was accompanied by a slightly downward frequency of about 10Hz. They were perceived as the second Low tone when

the Low frequency was stable. The stimuli were perceived accurately as High tone when the production of the tone was accompanied by a High frequency which can be either stable or slightly rising. In the bisyllabic patterns, the native speakers' perception was affected by the F0 of the preceding tone. The first pattern (A) was hardly perceived due to the nature of its frequency (F0) which required that the F0 of the second syllable be much lower than that of the first. This frequency area was not covered in the range of the synthetic stimuli. The second pattern (B) was identified more accurately when the two Low tones in the bisyllabic pattern were of the same frequency range. Pattern C was identified more accurately when the second tone (High) begins with a higher F0 (of at least 10Hz) than that of the first syllable.

Further on the perception of Yorùbá tones, one of the studies that shares some similarities in purpose to the present thesis is Bakare (1995). The study involved forty participants comprised of ten Yorùbá speakers, ten Chinese speakers, ten music students (without tone language background), and ten English monolinguals. The four groups, each with different tonal experience, took part in both discrimination and identification of Yorùbá tones using real Yorùbá monosyllabic morphemes. The morphemes consisted of Yorùbá CV syllables, using ten consonants (*kp, t, k, gb, d g, m, n, l, s*), three vowels (*i, a, u*), and the three Yorùbá tones (*High, Mid, Low*). The discrimination experiment adopted the S-ABC psychophysical procedure. Listeners were asked to choose which of the three stimuli (ABC) corresponded to the sample stimulus (S). For the identification experiment, listeners were asked to choose the diacritic marking of High (´), Mid (¯), or Low (`) that corresponded to the tone they heard. The results showed that while all groups were able to perceive the Yorùbá tones, there was a perceptual hierarchy among the groups: Yorùbá > Chinese > Music > English monolingual. Experience with a tone language was a significant factor in lexical tone perception. The results also revealed that the order of discrimination accuracy of the three tones by the Yorùbá participants was High > Low > Mid, while for the Chinese group, the music students, and the monolingual students, the order of perception was High > Mid > Low. Thus, the High tone attracted the best perception accuracy by all participants. Mid and the Low tones were the hardest to perceive for the Yorùbá participants and the other groups, respectively. The study also suggests that there was no significant difference in the identification of the three tones. The results further observed that overall, discrimination of Yorùbá tones was easier for participants than identification. In addition, tone discrimination did not show the type of toneme difficulty observed in tone

identification. Thus, similar to Bakare (1995), this thesis will adopt the discrimination and the identification method to investigate tone perception by native Yorùbá and non-tone language speakers. It will also use monosyllabic stimuli in the discrimination and lexical experiments while in other experiments, it will include disyllabic as well as trisyllabic stimuli to fit with the study's purpose. However, unlike Bakare (1995), the present thesis will not adopt real Yorùbá words, in order not to give undue advantage to one group over the other.

The studies summarized above have contributed to the research on African tone languages in different ways. These contributions range from the description of the voice quality to mark the tones of an African tone language (e.g., Hayward et al., 2003) up to the interaction between intrinsic fundamental frequency (IF0) and tones in African tone languages (e.g., Hombert, 1977; Connell, 2002). In other contributions, studies such as Hombert (1976, 1988) and Connell (2000) have laid out the methodology for investigating the acoustic cues used by native speakers of African tone languages in the perception of their native tones. These studies have revealed the presence of IF0 with respect to tones in Yorùbá as well as in other African tone languages. They have also shown that while pitch height is essential for the perception of the level tones, other secondary cues such as voice quality and phonation type, duration and intensity are also important acoustic features necessary for the perception and production of African tones. The perception studies have revealed the frequency range at which native listeners of these languages perceive the tones and the tone patterns in their languages. Bakare (1995) has also shown that the Yorùbá High tone attracted higher rate of perceptual accuracy than the other two tones. It is yet to be seen to what extent the knowledge of these features as revealed in these previous studies will assist listeners with diverse tone language background (both native and non-native tone language speakers) in the perception and production of these tones. It is also yet to be seen how participants' age difference (young versus old) will affect their perception and production of these tones.

The purpose of these perception studies was to examine the frequency range at which native speakers of these tone languages perceive the tones using mostly synthesized (and in some case, e.g., Hombert, 1976), real stimuli. In contrast, the purpose of the present thesis is to investigate the perception of Yorùbá tones by young and adult native Yorùbá speakers as well as non-native Yorùbá speakers. To this end, the design of the present thesis uses non-synthesized stimuli

produced by a native speaker for the perception and the production experiments. Moreover, nonce words are used to allow a wide range of possibilities in stimuli choice. For example, the use of real Yorùbá nouns in Hombert (1976) led to the exclusion of three Yorùbá bisyllabic patterns, not found in the language's noun formation. As these patterns are important to the present thesis, using nonce words as stimuli allows for their inclusion.

The present thesis also includes more participants to add to the reliability of the results. In addition, as one of the objectives is to investigate the extent to which differing degrees of exposure to tone language affects its perception, the study includes participants with diverse exposure to Yorùbá. Furthermore, the present thesis uses nonce words in experiments that involve both native tone language speakers and non-tone language speakers. Using real words would have resulted in an undue advantage for the native speakers. Finally, despite having level tones as in Yorùbá, the tone languages investigated in some of the studies either have larger tone inventories (e.g., Mambila, 4 level tones) or fewer tone inventories (e.g., Edo, 2 level tones) than Yorùbá. The only language with three contrastive tones, similar to Yorùbá is Bulu. However, the distribution of the three tones in the tonal space is not the same as in Yorùbá. Bulu has two Low tones and one High tone with no Mid tone while Yorùbá has three level tones, High, Mid, and Low.

In sum, while some of the perception and production studies on African tone languages described above were designed to describe the F0 of the tones or the voice quality associated with the tones, others mainly examined the acoustic cues used by native speakers/hearers in the perception of the native tones. The present thesis, on the other hand, is designed to examine whether native and non-native speakers of Yorùbá will be able to perceive the tones accurately or inaccurately due to their knowledge of the tones or due to the inherent features of the tones which may or may not make them easy to perceive. A further question is whether the age and language context of the participants will affect their tone perception accuracy. In order to investigate these questions, the present thesis adopted a series of tasks, AX discrimination, tone identification, lexical recognition, and tone production. The AX discrimination task, in which listeners classify pairs of tones/tone patterns as the same or different, is designed to test the lowest level of tone perception. As tone discrimination does not necessarily require knowledge of Yorùbá, all participants should be equally able to do the task (Bakare, 1995). In contrast, the

identification task requires some knowledge of the three tones, as participants need to categorize the three tones according to the perceived pitch height. Here, knowledge of Yorùbá might be an advantage to participants to carry out this task (Wayland & Guion, 2004). The lexical recognition task requires a deeper knowledge of the language as it examines participants' knowledge of real Yorùbá lexical items that differ in tone patterns. As such, only native speakers are included in this task (Yeh & Lu, 2012). Finally, the production task tests all participants' tone perception and imitation ability. As such, participants' basic experience at pitch manipulation will be investigated in this task (Hao, 2012). Since one major focus of the present thesis is on tone perception, the next section presents some of what has been found in the area of perceptual development with particular reference to tonal development.

1.3 Perceptual development

Infants' language development begins in their mother's womb. Several studies (e.g., De Casper & Spence, 1986; Nazzi, Floccia, & Bertoncini, 1998) have shown that postnatal language preference is related to the prenatal language experience. Newborns show sensitivity to the speech prosody of the language of their immediate surroundings to the extent of being able to differentiate word pitch contours. However, this early sensitivity is not restrictive, as the infants remain sensitive to sound contrasts of any language to which they are exposed at this early stage. This ability is documented in several studies (e.g., Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Chen & Kager, 2012; Liu & Kager, 2012) which observed that infants could discriminate speech tones and non-speech tones at both four and six months of age. Between the age of six to nine months, the child begins to be selectively sensitive. After this stage, the sensitivity to native language contrasts remains or increases while the sensitivity to non-native language contrasts decreases. This is referred to as the perceptual reorganization (PR) stage.

In a context where there are two different language types, for example, a tone (e.g., Yorùbá) and a non-tone language (e.g., English), an English infant, after the perceptual reorganization stage, will have more sensitivity to sound contrasts peculiar to the English language. Several studies of American infants (e.g., Jusczyk, Houston, & Newsome, 1999) have shown that at as early as 7.5

months, the 24 American infants in their study could detect strong/weak-stressed target words in sentential contexts, but they were not as successful in detecting weak/strong-stressed target words. The 24 older infants in the study (10.5 months), also American, showed improvements and could detect weak/strong target words from the different contexts in which they appeared, in a manner characteristic of mature users or hearers of the language. At this stage, exposing these infants to sound contrasts not specific to the English language would result in a diminished perceptual advantage.

After the establishment of the sensitivity stage, infants learning tone languages such as Yorùbá, Thai or Chinese will perceive tones differently than English infants. This was observed by Mattock et al. (2008) who tested 48 Montreal-Canadian four month-, six month-, and nine-month-old English and French infants with Thai Low versus rising lexical tone contrast. The results showed that while the four- and the six-month-old infants discriminated these tones with no decline in discrimination performance across these ages, nine-month-olds in the study failed to discriminate the tones. A related study of perceptual reorganization concerning lexical tone was conducted by Harrison (2000). The study included six Yorùbá and six English infants of six- to eight-months-old. The six Yorùbá children were exposed to Yorùbá language at home while the English infants were not. The study showed that in this age range, only the Yorùbá infants trained successfully on pitch and responded significantly to pitch-differentiated stimuli. One of the Yorùbá infants successfully perceived pitch differences of 40Hz and 20Hz differentials but failed to perceive 10Hz differential. Interestingly, this infant showed significant results at the pitch value of 190 / 210Hz, the same perceptual boundary relevant in adult perception tests. The English infants recorded no success either in pitch training or perception. In addition, the study found that English infants did not respond to isolated syllables with a pitch difference, while the Yorùbá infants did. The Yorùbá infants, like their adult counterparts, discriminated the High tone from other tones in a single nuclear domain. Harrison's findings on Yorùbá infants further supported the assertion that in the first year of life, infants tune into the lexical tones of their native language. The issue remains open as to whether this perceptual ability can be maintained as these infants grow into youth and adulthood, especially when they might have encountered significant L1 input changes. A brief survey of tone acquisition by Yorùbá children could clarify this issue. This follows in the next section.

1.3.1 Tone acquisition by Yorùbá children

Like any child acquiring a mother tongue, Yorùbá children, at the initial stage of acquisition, are also prone to specific production errors which have their roots in Yorùbá typology. Instances of tone simplification, tone absence, vowel denasalization after obstruents and the simplification of labiovelar consonants, have been observed by Orie (1997) and Oyebade (1990). Studies on tonal acquisition involving Yorùbá children have revealed that the three basic tones (High, Mid, Low) surfaced in the child's language as early as 12 months (Ajolore, 1974) and were fully acquired at 16 months (Orie, 2012). At this stage, complex tones (e.g., L**H** and H**L** contours) had not yet been acquired. According to Ajolore (1974), the child, at this stage, did not exhibit the adult-like falling / rising final tones required in the production of these complex tones. Ajolore (1974:272) suggested that "perhaps the transition that a Low-High contour demands, was more than they could cope with at that age". These contour tones became fully acquired by the child between the second and third year of age (Orie, 2012; Ajolore, 1974). As is the case with tone acquisition by Yorùbá children, studies dealing with adult acquisition of Yorùbá tones are also sparse. In what follows, I go through the available literature on Yorùbá tone acquisition by adults.

1.3.2 Adult acquisition of Yorùbá tones

Orie (2006b) is one of the few studies available that focuses on the acquisition of lexical tones by young Yorùbá heritage speakers and adult L2 learners of Yorùbá. Evidence from this study showed that the Yorùbá heritage speakers (K5 - 12) attending Yorùbá Language Summer Camp, recognized the three Yorùbá basic lexical tones better than the older learners (over 18 years of age). This was plausible since these young learners were born of diaspora Yorùbá parents and were acquiring the language as a heritage language.[8] They received input of the heritage language at home from at least one adult. The adult group, on the other hand, was mainly

[8] Heritage Language: L1 acquired at a very young age (0-4 yrs) in an L2-dominant environment; heritage speakers: speakers of a Heritage Language (Valdes, 2001).

learning the language for academic purposes. Their input was mostly through four months of classroom instruction.

Orie (2006b) further observed that the complexity of the phonological and syntactic behaviour of the Mid tone made it a difficult tone to acquire in L2 acquisition. She observed that while the adult Yorùbá L2 learners could differentiate between High and Low tones, they showed a distorted perception of Mid tones. Depending on the context, they identified Mid tones as High tones or as Low tones (Orie, 2006b). Concerning production, participants tended to impose High and Low tones on all nouns, in a pattern reflecting the English stress system. According to the study, production error rates for High and Low tones (12% and 26% respectively) were less than that for Mid tones (97%), indicating that Mid tone was hardly acquired by these learners. These acquisition studies focused on adults learning a new language different from their L1, so the errors observed were not unexpected. Some of the challenges identified by Orie (2006b) included interference from the L1, the adult English speakers' narrow pitch range, and age. At this juncture, the question remains whether what is observed in the acquisition and production of a tone language will necessarily be observed in the perception of that language. This issue is taken up in the next section which takes a cursory look at the tone perception ability by native and non-native tone language speakers. This is an area of research not widely covered in perception studies of African tone languages. As such, the reviews in the following sections will be based primarily on perception studies of Asian tone languages.

1.4 Perception of tone by native tone and non-tone language speakers

Except for Bakare (1995), the domain of the perception of tone of African tone languages by native tone and non-tone language speakers, which is an objective of the present thesis, has received less attention. Meanwhile, the literature in this domain is based mainly on Asian tone languages. Some of these studies on Asian tone languages relate to the objective of the present thesis and will be discussed in the following sections.

One area of investigation regarding tone perception in Asian tone languages focusses on the perception of non-native tones by speakers of tone and non-tone languages. Three diverging

findings have been reported. Some studies observe that native speakers of tone languages perform better than native speakers of non-tone languages (see, e.g., Wayland & Guion, 2004; Lee, Vakoch, & Wurm, 1996), some studies report the opposite result (e.g., Wang, 2006), and other studies find no difference between these groups (e.g., Francis, Ciocca, Ma, & Fenn, 2008; So, 2006). Wayland and Guion (2004) asked native Thai speakers, pre-trained native English and native Mandarin Chinese speakers to discriminate and classify a pair of Thai tones (Mid and Low) known to be difficult to discriminate. The stimuli were presented in triads and participants were asked to listen to the tones of the three words and then decide which of the three words differed in tone from the other two. This was done by clicking a button marked '1', '2', and '3' on the PC. A fourth button marked 'none' was to be clicked if they thought that the three words had the same tone. The procedure involved both discrimination (deciding whether the three tones in a set were the same or different) and identification task (locating the position of the different tone in a set with one different tone). The results showed that while native Thai listeners were better at both discriminating and identifying the two Thai tones, native Chinese listeners performed better than the native English listeners in the tasks both before and after the training. The English participants showed only a little progress in tonal classification. The study concluded that native tone language speakers have an advantage over speakers of a non-tone language in the perception of an L2 tone. The study suggests that in addition to the presence of phonological representations of tones in their long-term memory, Chinese listeners could have transferred the ability to track F0 movement and direction, employed in discriminating their native language, to the discrimination of a foreign tone.

In contrast, Wang (2006) observed that tone language speakers perform worse than non-tone language speakers in tone perception. Wang's (2006) study involved beginning learners who belonged to three different L1s, Hmong (a tone language), Japanese (a pitch-accent language), and English (a stress language). Having been trained before the experiment, the participants were asked to listen to each of the 40 randomized stimuli once and then identify the Mandarin tones by circling the tone of the stimulus they heard on a given answer sheet. The results showed that, while the accuracy rates of the English and the Japanese participants were comparable, the Hmong participants performed significantly worse than the English and the Japanese participants. The study suggests that the tonal system of the Hmong speakers' L1 might have negatively affected their perception of the L2 (Mandarin) tones.

The third finding in this regard includes studies such as Francis, Ciocca, Ma, and Fenn (2008), Hao (2012), and So, (2006), who reported no difference in perception accuracy between tone and non-tone language speakers. Francis, Ciocca, Ma, and Fenn (2008) conducted identification and difference rating task for 10 Mandarin speakers, 10 American English speakers and 12 native Cantonese speakers. On each trial of the identification task, participants listened to a single sentence containing the word bearing one of the six Chinese traditional tone number 1 to 6. They were then asked to identify the target word/tone by clicking a mouse button on the appropriate on-screen symbol. In the difference rating task, participants listened to a single pair of sounds presented in random order. A pair could be of same-token (e.g., 22 – 22, Low-Level followed by Low-Level) or different-token (e.g., pair 21 – 55, Low-Falling followed by High-Level). Each pair was presented once. Then, they were asked to rate the degree of similarity of the two sounds on a scale of 1 – 10. 1 indicated a perfectly similar (identical) pair and 10 indicated a very different pair. The results showed comparable performance between the native tone language speakers and the native English speakers. The only difference observed was in the type of tonal confusion exhibited by the two groups of participants. According to Francis et al. (2008), one reason for the comparable performance by the English-speaking participants was due to the presence of a well-defined English category (i.e. the English question intonation) which was clearly assimilated to the Mandarin lexical tones (Tone 1 [55] and Tone 2 [25].

With respect to tone perception of African tone languages, Connell (2000) reported similar findings. Connell (2000) examined the perception of lexical tones in Mambila, a tone language with four-level lexical tones (High, Mid-High, Mid-Low, and Low), using synthesised tones of ten level frequencies in steps of 5Hz inserted in a carrier phrase. Six native Mambila speakers and six native English speakers were trained for the experiment. The participants' main task was to listen to the stimuli over the headphones and then group them into four categories (High, Mid-High, Mid-Low, and Low) according to the perceived height of each stimulus. The results revealed that while pitch height is the primary important cue to tone perception in the language, native Mambila listeners did not perform significantly better than the English listeners in their sensitivity to pitch contour. The only difference between the two groups was the linguistic experience shown by the Mambila native listeners in their response patterns. They showed

consistency in placing the tones into different perceptual spaces while the English participants did not.

These studies have shown that being a native or non-native speaker of a language may be a factor in processing peculiar features of the language. Native tone language speakers may do better at tone perception because for them, tone perception is generally influenced by the presence of the native tone experience. The phonological representation of tone is always present in the native tone language speakers' long-term memory to be accessed when needed. As for non-tone language speakers, accurate performance at tone perception as shown by some of these studies could be the result of general language experience by these listeners. It could also occur in situations where well-defined native categories are clearly assimilated into non-native lexical tone categories. Even in situations where there is a comparable performance by native and non-native tone language groups, the error patterns (Francis et al., 2008), or the consistency in tonal placement (Connell, 2000) by the two groups sometimes reflects their native language experience. Meanwhile, native tone language speakers may perform worse in L2 tone perception due to the negative effects of the interference from native tones. The following section looks at the literature relating to the subjects' age and how it influences language processing.

1.5 Age and perception

1.5.1 Children vs. adult perception studies

Among the studies that undertake a cross-sectional approach to tone perception is Ciocca and Lui (2003), and Sze (2004). Ciocca and Lui (2003) studied adults (> 18) as well as children of different age groups (4, 6, and 10-year-olds) to test their perception of six Cantonese contrastive tones. They asked the participants to identify the six tones presented in monosyllabic words with the syllable /ji/. The target tones were inserted in a semantically neutral carrier phrase. The participants were presented with two pictures of the minimal pair for the tonal contrasts constructed from the combinations of the six lexical tones with the stimuli /ji/. They were asked to choose the picture that matched with the stimulus syllable they just heard. The results showed that while children of higher age groups performed better than those of lower age groups, the performance of the 10-year-old children was comparable to that of the adult participants. The

study by Sze (2004) used the same research paradigm as Ciocca and Lui (2003) with five age groups (2, 4, 6, and 10-year-olds, and adult controls). Replicating the findings by Ciocca and Lui (2003), the study found that the tone perception ability of 10-year-olds was comparable to that of adult participants. Participants who were less than ten years old were reported to find specific Cantonese tonal pairs easier to identify and others harder to distinguish or identify. What these two studies commonly suggest is that before the age of ten, Cantonese children's tone perception ability, when measured by tone identification accuracy, is still in development, while at the age of ten onwards, it has already been stabilized, hence its comparability to adult perception.

In another contribution, So (2000) conducted both perception and production experiments with three groups of Cantonese participants. The three groups consisted of: 1- those who immigrated to Canada in their teen years (10-15); 2-those who were either born in Canada or immigrated to Canada at a young age (before the age of seven); and 3- adult native Cantonese speakers who immigrated to Canada less than two years from the test. In the perception experiment, forty-eight stimuli of the target words /si/ and /fu/ in citation forms, produced by four experienced Cantonese language instructors, were presented to the participants. They were to listen and then identify the six lexical tones associated with the target words in two sets of stimuli. In the production experiment, participants were asked to read out 16 prepared lists containing the two root-word sets (/si/ and /fu). The target words were read out in three formats: 2 repeated citation forms; 3- repeated phrasal forms (with the target word being the second syllable) and one sentential form (with the target word being in the middle position). The results indicated that while there was a severe confusion of lexical tones in both perception and production for the young Canadian-raised Cantonese (2), the production and perception abilities of teenage Canadian-raised Cantonese (1) were comparable to those of the adult native Cantonese speakers (3), lending support to Ciocca and Lui's (2003) finding. The study suggested that participants from the first and third groups had achieved complete tonal development before immigration to Canada. In contrast, participants who immigrated to Canada at a young age (2), had not yet fully acquired the Cantonese tones before moving to Canada. Additionally, the study pointed to the effects of language contact (English) as a possible factor responsible for the wrong tone placement observed for the young Canadian-raised Cantonese participants; this effect was not observed for the teenage Canadian-raised and the native Cantonese-speaking participants.

1.5.2 Language experience and tone perception

Among the studies that suggest beneficial effects of prior tone language experience include Wayland and Guion (2004) as well as Lee et al. (1996). So and Best (2011) further suggest that experience with a language in which pitch features perform lexical functions may also be helpful in identification of non-native tones. With respect to African tone languages, one of the studies that examined the effect of prior tone language experience on tone perception is Omozuwa (1991), who investigated the perception and production (repetition) of Edo tones by native Edo speakers and non-Edo speakers from five different linguistic groups namely Yorùbá, Ika, Hausa, French, and Chinese speakers. In the experiment, recorded Edo bisyllabic nouns were played back to the non-Edo participants in a listening and repetition task aimed at simulating Edo tones. Then, the repeated tokens were presented to ten Edo native speakers/hearers to analyze, using their knowledge of their native language to assess the acceptability or non-acceptability of the utterances. Among others, the results indicated that participants' L1 influenced their perception of Edo tones. Participants whose L1 shared some affinity with Edo (e.g., Yorùbá and Ika) performed better than those whose L1 did not (e.g., Hausa, French and Chinese). Furthermore, among these participants whose L1 was not closely related to Edo, those with previous training in tone manipulation (e.g., phonetician, musicologist, speech therapist), performed well in pitch perception and repetition, regardless of their L1. Thus, while some of the earlier studies presented suggest that tone language knowledge benefits native speakers in the perception of their native and non-native tones, the studies presented here further suggest that non-native speakers of a tone language may also perceive non-native tones which share some similarities to their native tone system. Moreover, non-tone language speakers may also identify foreign tones if they have native language categories that can be assimilated into the non-native tones.

1.5.3 Bilingualism and tone perception in tone-language majority contexts

Tong, Lee, Lee, and Burnham (2012) investigated the effects of bilingualism and experience on tone perception. The study was conducted in Hong Kong where Cantonese, a tone language is predominantly spoken. It involved three participant groups: Cantonese-English bilingual adults,

Cantonese-English bilingual children, and English monolinguals. They all participated in an AXB experiment to test their Cantonese tone perception as well as English stress perception. They were asked to decide whether the second stimulus, a disyllabic sound was similar to, or the same as, the first or the last stimulus (also disyllabic). The results revealed that bilingual children's performance was comparable to bilingual adults. In addition, there was a comparable performance on both tone and stress perception by bilingual children. A comparison between bilingual and monolingual adults showed that bilingual adults performed better than monolingual adults in tonal contrast perception test as well as in stress perception tests. The results also revealed that, while both bilingual and monolingual adults performed better at perceiving stress than at perceiving tone, the fact that bilingual adults performed better at stress perception (a non-native contrast) than at tone perception (a native contrast) was an interesting one. It suggests that possibly the level of bilingual language experience and age play a role in suprasegmental speech perception. Meanwhile, Tong et al. (2012) represents the effect of bilingualism on tone perception in a situation where the participants reside in an environment where tone language is dominant. What should we expect in a situation where the tone language is the minority? This is discussed in the next section.

1.5.4 Bilingualism and tone perception in tone-language minority contexts

Two potential outcomes can occur in a situation where bilingual participants reside in an environment where the tone language is the minority. The first outcome is attrition which occurs as a result of migration at an older age. As suggested by Paradis (2007), language attrition refers to a particular case of language change which is usually prompted by the decreased frequency of use of a language. It refers to a gradual decline in language use proficiency, which can follow from a decreased overall frequency of use, or decreased use in particular functional domains. Attrition can set in as a result of the normal aging process (non-pathological language loss) or in a bilingual context and language contact situations. Some studies on L1 attrition have indicated that movement to an L2 dominant environment can lead to L1 attrition due to a change in the relative use of the L1 (see Bullock & Gerfen, 2004; Seliger & Vago, 1991). The aspect of the language that is most prone to loss is the lexicon, probably because vocabulary acquisition is a

continuous process (Hutz, 2004). Furthermore, it is hypothesized by Montrul (2008:60) that L1 attrition in a child will be more severe than in an adult undergoing the same experience. This means that the degree of language loss will be more pronounced in children (early bilingualism) than in adults (late bilingualism) (see Montrul, 2015: 2008 for a detailed review). Thus, in all cases, the effects of bilingualism may manifest themselves in the loss (attrition) or forgetfulness of some of the features of the native language previously acquired in childhood.

The other outcome is incomplete acquisition or acquisition without mastery which may occur to child migrants who reside in a host country with a dominant language. These types of migrants are referred to as heritage speakers. Heritage speakers refer to early bilinguals who were either born in the host country or moved to the host country at an early age (before age 4). According to Montrul (2015), heritage speakers may start off learning the two languages simultaneously and may be proficient in both. However, when they become of school age and they begin receiving instructions in the L2, which, in this case is the majority and the official language, the balance in the acquisition of the two languages is typically shifted. They go through reduced input in quantity, and less use, of the L1. In a situation such as this, the heritage language of these speakers may never fully develop, a situation referred to as 'incomplete acquisition', or 'acquisition without mastery'[9] (Montrul, 2015).

In the domain of perceptual attrition with respect to tones, some of the existing scholarly works include Yeh and Lu (2012), Yeh and Tu (2012), as well as Yeh and Lin (2013) and they all focus on attrition by older speakers. According to these studies, the infrequent use of the tone language is the primary cause of tonal loss in speakers. Infrequent use of the L1 occurs mainly due to the dominant influence of the L2, which gradually replaces the L1. Usage of the L1, in this case, is construed as a two-way process - speaking (production) and listening. If one factor (especially production) is missing, attrition is bound to set in. It is plausible to assume that listening to the language alone, in the absence of speaking, will affect the subjects' production and perception patterns, both in quantity and quality. One of the studies that consider the effects of attrition on tone perception is Yeh and Lu (2012). Their approach serves as a frame for the present thesis.

[9]According to Montrul (2015:118) *acquisition without mastery* of several aspects of the heritage grammars occurs when input is not optimal in quantity. Optimal quantity includes input from the immediate family context, the school context and the broader sociolinguistic context. Some participants in the present study may belong to this category.

Yeh and Lu (2012) specifically examined the effects of language attrition on Hakka low-level tone. Their findings showed the effects of language attrition on the tone perception in this language. Their study sought to ascertain whether language attrition could result in significant tonal confusion. The participants included three groups of Hakka speakers: Young Attriters, Young Fluent speakers, and Older Fluent speakers. The Young Attriters (mean age 17.3 years old) used to be Hakka-Mandarin bilinguals but were currently speaking only Mandarin at home. They also had reduced exposure to Hakka (once a month or less often), and they never spoke Hakka at school or work. The Young Fluent speakers (mean age 38.9 years old) spoke Hakka at home almost every day and mostly spoke Mandarin at work. The Older Fluent speakers (mean age 59.1 years old) spoke Hakka all the time with very little exposure to Mandarin.

Yeh and Lu (2012) employed both production and perception tasks to test its hypotheses. The production task (PRO) had the participants read a word list of 40 disyllabic Hakka lexical items (2 syllables x 5 tones x 2-word positions) using rudimentary and intermediate vocabulary. The perception tasks were three in total: the AXB Discrimination task, the Identification task (IDN) and the Lexical task (LEX). The AXB discrimination task required participants to listen to three monosyllabic sounds separately and then decide whether the second sound was more similar to the first sound or the third. In the tone Identification task, participants listened to one monosyllabic sound in each trial and then identified the tonal type of the sound. The Lexical task provided participants with one monosyllabic sound in each trial. Their task was to recognize the word meaning of the sound.

The percent accuracy results of the study were as follows: Older Fluent speakers (%Correct: AXB=98.75; IDN=90.97; PRO=91.67; LEX=94.58); Young Fluent speakers (%Correct: AXB=98.56; IDN=87.02; PRO=89.04; LEX=90.58); and the attriters (%Correct: AXB=97; IDN=86.13; PRO=51.25; LEX=76). Statistical analysis of the results showed that there was no significant difference across the three groups in the AXB discrimination and the Identification tasks while significant difference was posted in the production and the lexical tasks. Further analysis of the production and the lexical results showed no cross-group difference between the young and older fluent speakers. Meanwhile, significant cross-group difference was observed between the attriters and non-attriters in these two tasks. The study found that generally, the fluent speakers displayed fewer tonal errors than attriters in all tasks, confirming the effects of

language attrition on tonal processing. The study further observed that due to the phonetic similarity of pitch height, the low-level tone (T5) was mostly substituted for the low falling tone (T3) in the production task by all participants. In the perception tasks, the low-level tone was substituted mostly for the high-level tone (T1) due to the similarity in pitch contours between the two tones. Thus, the study observed that the tones were some of the language-internal factors that could influence tone perception.

A closer look at these results showed that in the AXB discrimination task, where no knowledge of the language was required, all participant groups performed equally well. In the identification task, where knowledge of the tone language was essential, all participants also performed well. However, in this task, there was a relative drop in accurate performance for all groups (although not significant) when compared to the AXB discrimination. Moreover, the old and young fluent speakers posted a slightly better accurate performance in both the production and lexical tasks than what they posted in the Identification task. In the lexical and the production task where both lexical and tonal knowledge were essential, there was a significant drop in accurate performance for the attriters compared to the non-attriters. The overall performance revealed a stepwise pattern where older fluent speakers were better than the young fluent speakers, who, in turn outperformed the attriters.

Meanwhile, one issue with Yeh and Lu (2012) is the difference in the mean age between the young fluent speakers (mean age = 38.9 years) and the young attriters (mean age = 17.3 years). As mostly teenagers, the age difference could have been an explanation for the young attriters' attrition, hence, the significant difference in accuracy between the two groups. The next section further presents some other phonological factors that can affect perception.

1.6 Some phonological factors that can influence perception

As the present thesis is dealing with lexical tone perception involving speakers from tone and non-tone language backgrounds, it is essential to show that tone perception, like segment perception, is influenced by the phonology of the listener's native language. Thus, non-native listeners may perceive non-native contrasts differently, or, in extreme situations, may not be able to distinguish the foreign sounds.

Gandour (1983, 1984), and Lee, Vakoch and Wurm (1996) highlighted the effects of differences in lexical tone inventory on tone perception. In Gandour (1983, 1984), 50 native speakers from each of the following tone languages (Mandarin, Cantonese, Taiwanese, Thai), as well as 50 native English speakers were asked to make dissimilarity judgments on tonal pairs involving nineteen synthesized F0 stimuli (3 falling-rising, 4 falling, 4 rising, and 5 level tones). The dissimilarity judgements were on an 11-point scale (0 = no difference; 10 = extreme difference). The results showed that the tone language listeners generally rated the tones differently from the English listeners. Different tone language groups also rated the tones differently in part due to differences in their languages' tonal inventories. While the native English listeners employed tone heights in their ratings, Thai listeners were observed to be focussing on the direction of F0 contour (rising or falling) in their ratings. The Cantonese listeners who had four of the six tones in their language with level contours, appeared to employ the dimension of tone levels in their judgement. Similarly, Lee et al. (1996) asked Cantonese, Mandarin (Taiwan and Mainland) and English (US) listeners to discriminate naturally recorded stimuli of Cantonese and Mandarin tones on word and non-word syllables. They found that the Mandarin and Cantonese listeners discriminated the tones in their respective dialects more accurately. While the English-speaking participants were also able to discriminate the tones, what differentiated them from the native tone language speakers was that the native tone language speakers could put the stimuli more accurately into tonal categories than the native English speakers. This was also observed in Connell (2000) in the study of the perception of Mambila tones by native Mambila and native English speakers. The Mambila listeners were more consistent in placing the tones in their perceptual space while the English listeners showed significant variation (see Section 1.3).

While the absence of tone inventories might impede non-native tone language speakers' organization of pitch differences into appropriate language-specific tonal spaces, the advantage of having similar phonological categories was the focus of the study by Nguyen and Ingram (2005). They found that experienced users tend to use cues that are actively applied in their L1 to process prosodic contrasts in the L2. Their study involved native speakers of Vietnamese, a tone language with different prosodic features from English. In a process that reflected a transfer of L1 perception strategies to L2 perception, these Vietnamese speakers could perceive English stress contrasts because their L1 also contain pitch cues which are essential elements of the English stress system.

Another related factor that may influence perception includes the lexical status of the stimuli (whether real words/or non-words). In a study by Fox and Unkefer (1985), eleven Chinese and eleven English (US) participants listened to four continua of synthetic Mandarin high level and rising tone pairs to perform a forced-choice identification task. The forced-choice task required them to identify a stimulus as having either a High-Level or a Rising tone. The first pair had real words at both ends of the continuum while the second pair had non-words at both ends. The third and the fourth pairs had real words at only one end. The results showed that the Chinese (Taiwan) participants showed more significant identification responses for Mandarin word tokens in non-word/word and word/non-word pairs than for non-word/non-word pairs. In contrast, the English participants in the experiment did not show a similar difference in any sequence. In sum, their study suggested that lexical knowledge facilitates tone perception.

At this juncture, it needs to be mentioned that comparing the present thesis to all the studies mentioned so far is not without some reservations. This is because they are largely based on Asian tone languages in which mostly pitch movement and direction are employed as cues to tone identity. These languages are fundamentally different from African tone languages which use pitch height as cues, hence the term level / register tone languages. Secondly, while Yorùbá has three level tones (High Mid, and Low), the Asian tone languages in the reviewed studies have more than three tones (Mandarin =4; Cantonese=6-9; Taiwanese=5-6; Thai=5) which include both level and contour tones. Moreover, most of the studies (e.g., Wayland & Guion, 2004; Wang, 2006; So, 2006) used non-native participants who were L2 learners of the tone languages. In comparison, the non-native participants in the present thesis are naïve listeners with no previous exposure to the tone language. Finally, while there may be similarities between the present thesis and Yeh and Lu (2012) in the types of experiments used, the details and the procedures of the experiments in the present thesis are adapted to fit with the purpose of the investigation, as discussed in section 2.3 below. The next section looks at studies focusing on the effects of the number of syllables in a stimulus on word or tonal processing.

1.6.1 Syllabicity

Studies of the effects of syllable count on word processing in general, and tone processing, in particular, have been consistent in suggesting that disyllabic items contribute to better processing than monosyllabic items. With regards to word processing, Dommergues and Segui (1989) conducted three word-detection experiments using low-frequency, high-frequency, and concrete French words. Each experiment required participants to listen carefully to two lists of linguistic items consisting of an equal number of words and non-words. The first list contained homogeneous monosyllabic items such as *soc, digue, ruche* (real words), *summe, tive, rugue* (non-words) while the second list contained heterogeneous monosyllabic and disyllabic items such as *gaz, soldat, sac, bateau, gosse, garcon* (real words), *fatu, zotte, pito, summe, bachin, gore* (non-words). Just after listening to each list, participants were asked to give a subjective estimate of the proportion of the real words present in the list. The results showed a similar trend across the three experiments. Participants significantly underestimated the number of real words in the homogeneous list of monosyllabic items but did not underestimate them in the mixed list containing monosyllabic and disyllabic items. The study suggests that monosyllabic items "induce a monotony effect" (p. 252) and, as such could affect the encoding process.

Given that Dommergues and Segui (1989) focus on monosyllabic and disyllabic word processing, it is essential to show that just as syllabicity affects word processing, it can also have some effects on tone perception. One study that deals with syllabicity with reference to tone is Perdereau (2013). He conducted an experiment to test the extent to which the speech of a non-native speaker of a tone language was understood (in the absence of visual or other sensory cues). He asked three native Mandarin speakers to recognize monosyllabic and disyllabic words as well as polysyllabic phrases produced by a non-native Mandarin-speaking subject whose Mandarin was characterized as tone-deficient. Upon listening to the stimulus, the listeners were asked to respond by reproducing the tokens they heard along with the tones or the most probable tones in case of ambiguities, within the shortest time. The results showed that while recognition accuracy in the disyllabic task was comparable to that of the polysyllabic task, there were more recognition errors found in tone identification for the monosyllabic task. This led to the observation that word processing in disyllabic words and polysyllabic phrases was less prone to the ambiguities situated in the lexical tone level.

While these studies tend to suggest that tones may be easier to process in disyllables than in monosyllables, there may be other factors or contexts, beyond the scope of the present thesis that make tone processing difficult. All the experimental tasks in the present thesis (except the second identification experiment) are employed to examine the syllabic effects on participants' tone perception and production accuracy. Meanwhile, as the present thesis involves both perceptual and production skills, the relationship between perception and production is discussed briefly in the next section.

1.7 Perception versus Production

Perception and production require different skills. In the present thesis, the perception tasks involve auditory (and in some cases, visual) prompts followed by a perceptual response, while the production task involves auditory prompts followed by a verbal response. One question that comes to mind in this literature is how related are speech production (imitation) and speech perception? Studies involving the perception and production of segments have found that participants' imitation patterns largely resemble their perception patterns. For example, Jia, Strange, Wu, Collado and Guan (2006) found that participants' accuracy in a vowel imitation task was significantly similar to their accuracy in an identification task. Similarly, studies focusing on perception and production of tones have found some similarities between tone perception and production. Elliot (1991) investigated the relationship between the perception and production of Mandarin tones by University students who were Mandarin L2 learners. The participants included both native and non-native English speakers who did not speak or hear Mandarin regularly outside of the classroom setting. The non-native English speakers included Japanese, Cantonese, Vietnamese and speakers of other Chinese dialects. All the participants went through three tone tests involving tone production, tone perception, and perception of their own tone production. The results showed that generally, participants' perception was better than their production and they were aware of their tone production errors. The study pointed at similarities in participants' production, perception, and self-perception errors by indicating that most of the errors in all the three tasks involved Tone 2 while other tones did not show a similar

pattern. However, the study was criticized for not pointing out that the tone language background of fourteen of the participants could have affected their perception and production patterns.

While Elliot (1991) suggests that participants' performance in perception is better than production, Yang and Ankenmann (2007) suggest otherwise. The latter investigated the tone perception and production of American learners of Chinese who were exposed to Mandarin for one or two years. The results suggested some similarities between tone perception and production performance. However, participants' overall performance in tone production was much better than their tone perception. The study further observed different effects for segmental and suprasegmental categories on tone perception and production. It suggested that while tone categories affected perception, the fact that the participants in the study did not exhibit any pattern of difficulty across the tones in production indicated that tone categories did not affect production. As the present thesis involves processing of phonological features by speakers and non-speakers of a specific language, the next section presents some theoretical models related to phonological processing of native and non-native contrasts.

1.8 Theoretical Models of phonological processing of native and non-native contrasts

Some theoretical models attempt to explain the issues involved in the processing of L1 and L2 speech sounds in general. Strange (2011) proposed the Automatic Selective Perception Model (ASP) that hypothesized that listeners process speech sounds in either phonological or phonetic mode. According to the model, when listeners are confronted with a perceptual task, they selectively detect the acoustic-phonetic information necessary and sufficient for the performance of the task at hand. The automatic and attentional processing lies along a continuum of "effort", connected with the selective detection of information relevant to the performance of the task. Moreover, the two modes, conceived as "ways of perceiving", are further influenced by other factors such as the listeners' knowledge, purpose and intention, stimulus complexity, and task demands. The model predicts that a task that involves simple syllables in citation form as stimuli and does not impose a high memory load will see learners more likely to operate in a phonetic mode. Consequently, learners are more likely to operate on a phonological mode when the task imposes a high memory load with the task involving words of longer syllables or phrases. Thus,

in the present thesis, those tasks that do not impose high memory load such as the simple AX discrimination, the Lexical recognition, and the tonal imitation tasks will likely see participants operating at the phonetic level. The identification tasks, which require participants to perceive the tones, then categorize them as one of the three Yorùbá tones, are embedded in disyllabic and multisyllabic stimuli. Therefore, listeners are likely to operate on a phonological mode.

As the present thesis includes native tone and non-tone language speakers, this section looks at the theory that can account for a perception pattern that may be encountered through the non-native tone language speakers in the study. Based on the general findings that speakers of non-tone languages such as English are able to perceive tones to a certain extent, possibly by focusing on pitch differences, theories related to the perception of tones by non-tone language speakers have received considerable attention. Previous work by So (2006) observes the feasibility of assimilating non-native tones to the listeners' native language prosodic (intonational) categories. This is what is later proposed in the Perceptual Assimilation Model (PAM: So & Best, 2008) with suprasegmental extension (PAM-S). In their study, So and Best (2008) asked eighteen Australian native English speakers to categorize 72 tokens of the four Mandarin tones placed on the syllable /fu/ into four English intonation categories (*Flat pitch, Question, Uncertainty, and Statement*). The stimuli were individual presentations randomized during the experiment. They were asked to press any of the four buttons provided, corresponding to the four intonation categories and a fifth button labelled '*Unknown*' was to be pressed when they could not identify a tone into any intonation category. The results showed that the native English participants perceived Mandarin Tone 1 (High-level) as *Flat pitch*, Tone 2 (Mid-Rising pattern) as *Question*, and Tone 4 (High-Falling) as *Statement*. The only exception was that Tone 3 was not significantly perceived as *Uncertainty;* both Tone 2 and Tone 3 were equally perceived as *Uncertainty*. Thus, the study indicated that native English listeners assimilated non-native Mandarin tones to the English intonation categories. The shared phonetic similarities between the pitch contours of English intonations and those of Mandarin tones were the major factors for this assimilation. Subsequent studies by these authors (see, e.g., So & Best, 2010, 2011, 2014) have further confirmed the PAM-S model. Thus, their studies suggest that speakers of non-tone languages (e.g., native speakers of English) can discriminate Mandarin tones adequately. Consequently, this finding has made the suprasegmental dimension of this theory relevant to the present tone perception study. Through the inclusion of non-native tone language

speakers, the study will attempt to see where the theory stands with regard to English-speaking participants' perception of this non-native suprasegmental feature.

Incidentally, most of the scholarly works available in this area of study are on Asian tone languages which have contour tones making them differ from African tone languages which have register tone systems. This is one gap which the present thesis attempts to fill. As a result, the thesis will explore participants' tone perception with respect to the inherent features of a register tone language. The final section of this chapter presents an outline of the remainder of the thesis.

1.9 Outline of the thesis

The thesis is organized as follows. Each of the five experiments is allocated a chapter. At the beginning of each chapter, there are more details about each experiment with respect to the stimuli, the participants, and the procedures. This is followed by a section showing the results and then the discussion of the results. Chapter 2 presents information about the studies comprising the present thesis. It covers the objectives, the participants, the materials, and the individual experiments, including the factors being considered in the thesis. Chapter 3 presents the first experiment, AX Discrimination, in which participants listen to a pair of tones or tone patterns and then decide whether they are the same or different. Chapter 4 covers the second experiment, the first Identification experiment. In this experiment, participants will identify the three Yorùbá tones as either High, Mid, or Low. Chapter 5 covers the second Identification experiment which focusses on participants' identification of the two Yorùbá contour tones. Chapter 6 presents a Lexical experiment, where only Yorùbá-speaking participants identify Yorùbá minimal pairs with the assistance of sounds and pictures. Chapter 7 reports a Production experiment, in which participants listen and then imitate words produced with Yorùbá tones. Chapter 8 presents the analysis of the language environment factors with respect to the young Canadian-based Yorùbá speakers and their influence in all reported experiments. This is followed by a general discussion and the conclusion in Chapter 9. The following chapter outlines the current study and provides general information about the study's participants, the materials, and the hypotheses surrounding the experiments.

Chapter Two: Current Study

Using four perception experiments and one production experiment (e.g., Yeh & Lu, 2012), the present thesis investigates whether the age and tone language experience/context of the participants influence their tone perception and production. The thesis also examines whether language-internal factors such as tone (H, M, L), syllable count and demographic factors such as age at arrival in Canada, months in Canada and place of birth influence tone perception and production. The goals of this thesis are to find out specifically:

(i) Is tone perception and production affected by the participants' usage of Yorùbá (language context)? To this end, I will test three groups of speakers: native Yorùbá speakers living in Nigeria, Yorùbá speakers living in Canada, and speakers of Canadian English who have no knowledge of Yorùbá.

(ii) Does a participant's age affect tone perception and production; and further, does age interact with a participant's language context? To test this, each speaker group included a group of adults and a group of children. As a further question, we ask whether child and adult Yorùbá speakers who live in Canada away from their native language environment, retain the ability to perceive and produce Yorùbá tones to the same extent as speakers who live in their native language environment?

(iii) Are the participants equally able to perceive and produce all Yorùbá tones. To this end, I will use stimuli that consist of the three tones and the two complex tones of Yorùbá.

iv) Finally, the last study reported in chapter 8 generalizes over the individual experiments reported in chapters 3 - 7 and inspects the influence of various demographic factors on the performance of a subset of the participants, namely the Yorùbá-Canadian young group. The demographic factors include the age at time of testing, age at arrival, years[10] in Canada, and place of birth.

[10] Age (in years) was later converted into months for the statistical analysis. Thus, we have age in months, Months in Canada, and Age in months at arrival.

The next section presents more details about the participants for the thesis who were drawn from speakers of both tone and non-tone language backgrounds.

2.1 Participants

A total of 139 participants (summarized in Table 2.1 and described in more detail below) were recruited for the study at locations in Canada and Nigeria. The Canadian recruitment took place in Edmonton, Calgary, and Vegreville, while the Nigerian recruitment took place in Shaki, Oyo State. The recruitment in Edmonton and Vegreville were through the researcher's contacts, contacts with the principal of a private school north of Edmonton, and through the University of Alberta, Department of Linguistics participant pool. The recruitment in Calgary and Shaki, Nigeria occurred through intermediary contacts in both locations. All participants received Letters of Initial Contact as well as the information sheet and consent form (Appendices A2 -A5) from the researcher or the intermediary contacts. The information sheet explained everything the participants needed to know about the research as well as their rights. Once the participants or their parents or guardians had agreed to be part of the study and had signed the attached consent form, they returned it to the researcher or the intermediary contacts before the experiment were run. All participants gave informed consent following the protocol approved by the Research Ethics Board of the University of Alberta (Appendices A6-A7).

At every location, the recruitment targeted participants of two different age groups, young and adult participants. These young and adult participants fall into two main categories: the native Yorùbá and the non-native Yorùbá populations. The native Yorùbá population was further sub-classified into Nigerian-based Yorùbá speakers and Canadian-based Yorùbá speakers. All the young and adult native Yorùbá participants, both Nigeria- and Canada-based, were either born in Oyo-Ibadan region or originally from this geographical region. That is the region whose variety of Yorùbá largely coincides with the Standard Yorùbá. The non-native Yorùbá population comprised of mainly English-language speaking Canadians recruited in Edmonton. Specifics of each population are described in what follows.

2.1.1 Nigerian-based Young Yorùbá speakers (YorNgrYg)

For this group, twenty-eight (28) participants (*Mean age* = 12;8, *SD* = 1;6, Age range = 9 - 16) who were all school-age youth between the ages of 9 and 16 years old were recruited. They were all born to Yorùbá parents and living with their parents in Shaki, a town in Oyo State, in the South-western part of Nigeria. As school children, they learned or were still learning Standard Yorùbá as a subject at school, so they had some basic knowledge of the tones of the language. Additionally, they were using Yorùbá at home with their parents and siblings, as well as at school with their friends and teachers. They also used it outside the school or home when speaking with their acquaintances or any other member of the community. In addition to Yorùbá, they were also learning English at school; it is the language used to teach the other subjects.

2.1.2 Nigerian-based adult Yorùbá speakers (YorNgrAd)

Seventeen (17) adults (18 years and over) were recruited for this participant group (*Mean age* = 43;5, *SD* = 9;2, Age range = 23 - 59). They were all educated English-Yorùbá bilinguals and they included students, workers, and professionals. They were all native Yorùbá speakers and have been exposed to the language since birth. Like their younger counterparts, they have also lived in Shaki for most of their lives. They learned Standard Yorùbá as a subject in both elementary and high school, and they used it frequently. They used the language at home with their siblings, spouses, or children as well as outside the home with their friends and other people in the community. They also used it at work, except occasionally, when they were in official engagements such as office meetings and conferences where they were required to use the English language.

2.1.3 Canadian-based young Yorùbá speakers (YorCanYg)

This group consisted of thirty (30) school-age children (*Mean age* = 11;9, *SD* = 2;4, Age range = 8 - 16). More than half of these participants were from parents who also took part in the study, described in 2.1.4 below. Seventeen of the children (57%) were born in Nigeria and then brought to Canada by their Yorùbá parents. Thirteen of these children (43%) were born outside Nigeria (Canada=5; Kuwait=7; USA=1) to Yorùbá parents. Thirteen children (43%) were either born in Canada or arrived in Canada before age 4; Eight children (27%) arrived at age 4-9 while nine children (30%) arrived at age 9 and over. These children lived in the same house as their parents but had different rates of exposure to the language. Their Yorùbá language use was limited, but

not restricted to, listening to their parents or other adults speaking it within or outside the house, or on occasions organized by the language community. Thus, the majority of the participants from this group know the language passively. They still understand and can respond to simple instructions given by their parents within the home. Six of these participants (20%) were reportedly compelled by their parents to speak Yorùbá at home. Despite speaking it only to their parents, they claimed that they could not engage in longer discussion in Yorùbá and they tended to quickly switch to English while speaking to their siblings and friends. None of them attended a school where the heritage language was taught. Two of these children were withdrawn by their parents during the experiment.

2.1.4 Canadian-based adult Yorùbá speakers (YorCanAd)

This group was made up of fifteen (15) adult Yorùbá native speakers (*Mean age* = 40;8, *SD* = 5;8, Age range = 28 - 48). Some of the members of this group were parents of the young participants described in 2.1.3 above. They all reportedly have learned the Standard variety at school in Nigeria. They used Yorùbá daily at home with their spouses and, in some cases, with their children. Sometimes, they used Yorùbá in communication with other adults from the Yorùbá community. Thus, their L1 use was also limited in quantity and quality. They migrated to Canada as adults and were either working or studying in Canada at the time of the experiment.

2.1.5 Canadian-based young English language speakers (EngCanYg)

This group consisted of twenty-two (22) school-age children currently attending schools in Edmonton, Canada (*Mean age* = 10;9, *SD* = 2;0, Age range = 8 - 16). Seventeen (17) of these children were born in Canada while the remaining five were born outside Canada. Those who were born outside Canada were brought to Canada at relatively young ages of between two to five years of age, so they were all very fluent in English. Consequently, for all the participants in this group, English had been their dominant language, despite having other languages such as Arabic or Urdu as their L1. As students, they had also learned, or were learning French at school. Preliminary screening of the participants revealed that none of them spoke, or was exposed to a tone language before the time of the experiments.

2.1.6 Canadian-based adult English language speakers (EngCanAd)

Members of this group consisted of twenty-seven (27) University of Alberta students who were all registered in different undergraduate Linguistics courses offered by the Department of

Linguistics (*Mean age* = 19;8, *SD* = 6;5, Age range = 17 - 52). They were all Canadians with English as their first language. The majority of them spoke only English language. Two of them reportedly spoke 'a little' Tagalog while another two spoke French in addition to English. A few among them were currently beginner learners of other languages such as French, Spanish or Japanese. None of them spoke or had previous exposure to either Yorùbá or any other tone language before taking part in the experiment.

Based on the descriptions above, two distinct age groups are identified. First, the young (YG) participants who are between 8 to 16-years old, and second, the adult (AD) participants who are all above the age of 18. In addition to age, the classification also highlights participants' language context which defines their level of exposure to the language of the study. The contexts identified so far are three. There are participants who are all native Yorùbá speakers based in Nigeria (YorNgr). Then, there are participants, who are also native Yorùbá speakers but based in Canada (YorCan). Finally, there are participants who are English-speaking, Canada-based with non-tone language background (EngCan). A summary of the classification of participants into age groups and their tone language contexts is presented in Table 2.1.

Table 2. 1 Summary of participants' age groups and levels of tone language contexts

Lang. Context	Age group	Age range	L1	L2	Residence	Tone lang. Context
YorNgr	Young	8-16	Yorùbá	English	Nigeria	Yorùbá-Nigerian
	Adult	18+	Yorùbá	English	Nigeria	Yorùbá-Nigerian
YorCan	Young	8-16	Yorùbá	English	Canada	Yorùbá-Canadian
	Adult	18+	Yorùbá	English	Canada	Yorùbá-Canadian
EngCan	Young	8-16	English	-	Canada	English-Canadian
	Adult	18+	English	-	Canada	English-Canadian

2.2 Materials

All the materials are based on Standard Yorùbá. The stimuli included monosyllabic, disyllabic and trisyllabic nonce words and real words (Experiment 4) bearing all the three Yorùbá tones

used in Standard Yorùbá. The monosyllabic stimuli were of the consonant - (nasal) vowel (CV) syllable type. The CV syllable is the core syllable in Yorùbá, and it is universal in all languages. It has been used successfully in different experiments involving the perception of suprasegmental features of speech in many languages (see, e.g., Harrison, 2000; Yeh & Tu, 2012). The disyllabic stimuli consisted of VCV as well as CVCV syllable types. The trisyllabic stimuli consisted of VCVCV as well as CVCVCV syllable types. As the emphasis was on tone perception, the stimuli used in the experiments were nonce words except for the lexical experiment (Chapter Six, below) which used real words. All nonce words used conformed to Yorùbá phonotactic rules. For instance, Yorùbá vowel harmony rules were observed to ensure that the stimuli didn't include such unacceptable vowel-initial nouns as **obẹ, *ekọ* which violate the rules (see Section 1.1 and Table 2.2). Similarly, the word formation also ensured that the stimuli did not include word systems that were phonologically non-existent in the language such as words beginning with the [u] vowel, or word-initial V-syllable structure bearing the High tone. Efforts were also made to ensure that all the Yorùbá real and nonce words included each of the seven Yorùbá oral vowels [a - e - ẹ - i - o - ọ - u] as well as the nasal vowels [ã - ĩ - ũ - ọ̃], bearing the three tones High, Mid, and Low. While all the Yorùbá consonants were included in the perception experiments, the production experiment, reported in Chapter 7, excluded the two Yorùbá consonants involving double articulation (*k͡p* and *g͡b*) due to the difficulty they could pose for non-native speakers.

A fluent Yorùbá male speaker recorded the stimuli for the experiments. The recording was made in a sound-attenuated booth at the Phonetics Laboratory located in the Department of Linguistics, University of Alberta using the MB QUART-MBK C 800 headset with combined headphones and electret microphone. All utterances were recorded digitally at a sampling rate of 44.1 kHz (16-bit, mono) using the phonetic analysis software PRAAT, software version 5.4.01 (Boersma & Weenink, 2014). Following the recording, each stimulus was edited and individually extracted into a separate sound file. Before uploading the recording onto the experimental software, all the stimuli were double-checked by an independent verifier, a native speaker who did not know about the purpose of the experiment. This was to ensure accurate representation of the sounds and the tones. Following the recommendation of the independent verifier, all recorded stimuli that resulted in ambiguities, or that were deemed unclear or had unstable loudness, were excluded. Such stimuli were re-recorded in a second session using the same procedures.

2.3 The Experiments

The experiments were designed using E-prime software (version 2.0) which allows for the inclusion of text, sound, and pictures as stimuli. After setting up the experiments, each participant went through the trials independently, following the instructions. The experiment began with the participants wearing the headset and then following the instructions on the computer screen. The participants used the specified keys on the computer keyboard to register their answers.

There were four perception experiments and one production experiment in the present thesis. The type and motivation for each experiment was as follows: The first experiment (Chapter Three) is an AX discrimination task modelled after Liu and Kager (2012). It tests participants' basic tone discrimination ability and, as such does not require the knowledge of the language. This was preferred to the AXB discrimination task, which could have been more cognitively challenging for the younger, and possibly, the older subjects. The second and the third experiments (Chapters Four and Five) involve tone identification tasks (e.g., Bakare, 1995; Hao, 2012). These experiments tap further into participants' knowledge of the tones of the language of the study as they are required to identify the three simple tones and the two complex tones. Since identifying a language's tone is a challenge on its own, these two experiments were designed to reduce the challenges that could have arisen for all participants; hence the instructions were for the participants to identify only the last tone of each stimulus. The fourth experiment (Chapter Six) is the lexical task (e.g., Yeh & Lu, 2012), which tests native Yorùbá-speaking participants for their recognition of Yorùbá real lexical items in a task that requires them to recognize the tone /tone pattern of one of the words in a minimal pair. This experiment taps both into participants' knowledge of the tones as well as requires knowledge of the lexical items of the language of the study. As such, real lexical items are used. Pictures for the lexical items were further included to assist participants in their choice. The last experiment (Chapter Seven) involves a production task (e.g., Hao, 2012; Omozuwa, 1991) where participants are asked to imitate the toned stimuli they heard as accurately as possible. Basic experience at pitch manipulation is enough to do this experiment as participants are required to place the tones in their appropriate tonal spaces while producing them. All the five experiments were conducted one after the other in the following order: AX Discrimination, Identification (1), Identification (2), Lexical, and Production. There

was one trial per experiment for each participant, and each trial was preceded by a short practice session to acquaint the participants with the experiment. All participants went through the same set of practice sessions designed for each experiment. More details about the procedure for each experiment are provided at the beginning of the chapter for the experiments. Before getting into the details of the experiments, the following sections present the hypotheses upon which the study is being carried out.

2.4 Age factor

Three factors, age, tone language context, and some language-internal features are considered. The age factor identifies two different age groups, young and adult participants (see Table 2.1). The young participants are all of the post perceptual reorganization (PR) stage (see section 1.3), who have previously established their language orientation earlier on in life. On the one hand, their general perception pattern is not expected to be in any way different from that of the others within the same age group or even adults. However, excluding other factors, young speakers of a language type are expected to be better at perceiving sound contrasts in their native language than other young speakers of the same age group who do not belong to their language type due to the PR established earlier on during infancy. We may also expect these native young speakers to be better than adults who do not belong to their language type. However, in situations where both the youth and the adult groups share the same first language, then, age can play a crucial role in their perception patterns. The broad language experience that the adult subjects possess could translate into better perception ability than the young subjects. In the present thesis, all the five experiments will be pertinent in testing this hypothesis. The AX discrimination task may not highlight the effects of the Age factor as much as the tasks that require the knowledge of the language (e.g., the Identification and the Lexical tasks). While the age factor could be crucial in any perception pattern, another equally important factor is the type of input the subject has of the language being studied. This is discussed in the next section.

2.5 Language Context factor

In Table 2.1, three language context types are identified, Nigerian-based Yorùbá speakers, Canadian-based Yorùbá speakers, and Canadian-based English speakers. Nigerian-based Yorùbá participants use the language of the study daily as their L1. Both the young and the adult members of this group were born and raised within the larger L1 environment and are still residing there. In other words, they continue to have daily access to the tone language both within and outside the home. As such, they are expected to perform well in the perception and production of the L1 contrasts.

Canadian-based Yorùbá participants are exposed to Yorùbá from birth. It is their L1, and in most cases, they were born within the larger L1 environment and raised there. The only difference is that due to immigration, they currently live in a new environment where another language is dominant. Their only exposure to the L1 is within their home. While the adult members of this group do speak the L1 to each other either as married couples or with friends and acquaintances, the younger ones seldom speak it. For these younger ones, their exposure is limited mostly to listening to their parents, visiting family friends, or other domestic relatives. An example of this group is the young Yorùbá L2 learners in Orie (2006b) mentioned in Section 1.3.2 above. The exposure pattern of this category is limited when compared to the Nigerian-based Yorùbá speakers. Limited exposure could result in not fully mastering the L1 features while infrequent use of the L1 could trigger attrition. Thus, participants in this group are expected to record fluctuating performance both in tone perception and production. This group's knowledge of the native language may have become deficient. As such, the two identification tasks as well as the lexical recognition task which test the knowledge of the language might record lower accuracy rates for the young participants. However, the adult members may still perform better in these tasks because they have more access to L1 input than the younger members.

The third type of language context identified in Table 2.1 is the Canadian-based English speakers. These participants belong to another language group different from the language of the study. They do not speak Yorùbá and do not have experience listening to this language. In the present thesis, participants from this group also reside in their larger L1 community and continue to have continuous exposure to their L1 (English). Thus, their performance in perception and

production of a novel L2 contrast (i.e., tone) is expected to be limited, especially for the younger members of this group. Due to the age and the general language experience factor, the adult members may perform at par with the native speakers in tasks that do not require tonal knowledge, such as the AX discrimination. Even in tasks that require the knowledge of the tone language (e.g., Identification) we may witness comparable performance as suggested by Francis et al. (2008) as well as Connell (2000). As in the age factor hypothesis, all the five experiments will be applicable in testing this hypothesis. In addition to the Age and the language context factors, the hypotheses formulation also involves some of the language's internal features as discussed in the next section.

2.6 Tonal features

Each of the five experiments is designed to focus on a different aspect of the tonal factor. For the AX discrimination experiment, the focus is on the effect of same-tone and different-tone pairs on participants' accurate discrimination. In other words, participants are asked to decide whether the two tones in each pair of nonce words are the same or different. The task itself should be doable by all participant groups (see Yeh & Lu, 2012 above; also, Bakare, 1995) independent of their language background, as it does not require knowledge of Yorùbá or knowledge of the three tones used in the experiment. Despite that, Shittu and Tessier (2014) found that same-tone pairs were harder to discriminate by participants than different-tone pairs. Similar outcome may be expected in the present experiment. Experiment 2, the first of two tone identification experiments, focuses on the identification of the three Yorùbá tones High, Mid, and Low. Knowledge of the tones or experience in identifying pitch height is predicted to be helpful for this task. We may expect different identification patterns for participants if we consider the nature of, and previous findings, about the three tones. As mentioned earlier in Section 1.1.1 above, the High tone is characterized with a higher F0 than the Mid and the Low tones, but it mostly shares breathy phonation with the Mid tone. The Low tone occupies the lower register with a lower F0, but with a distinctive creaky phonation that differentiates it from both the High and the Mid tones (Hayward, Watkins & Oyetade, 2003). Additionally, Hombert (1976) suggests that falling F0 is the primary perceptual cue to identifying Yorùbá Low tones in word-final

position. If we consider Bakare (1995) which highlighted the strength of the High tone compared to the other two tones, we may expect the High tone to attract better identification accuracy than the Mid and the Low tones. However, if we consider the findings of Hayward et al. (2003), and Hombert (1976) which characterized the Low tone as of distinctive features which could aid its perception, then we may expect the tone to attract better perception. The tonal space of the Mid tone is situated between the High and the Mid, as in Mambila (Connell, 2000) and is sometimes considered a neutral tone. Thus, we may expect a situation where both the High and the Low tones are identified more accurately than, or are confused with, the Mid tone.

For the second Identification experiment (experiment 3), the tonal factors highlighted are the contour tones, rising High and falling Low. I expect this experiment to be more challenging than the previous one, and thus knowledge of the language is also predicted to be helpful for the task. These complex tones are derived when the High and Low tones are adjacent to each other in a disyllabic or multisyllabic context. Due to the experimental design, Hombert (1976) included only one of the contour tones namely the rising (LH). The perception by native speakers demonstrates that its rising F0 clearly differentiates it from the other final sounds (Hombert, 1976). The other contour, falling (HL) is similar to the Edo HL featured in Omozuwa (2003). Edo native speakers perceived the final Low tone as a High-falling tone, as it originates from the frequency zone of the preceding High tone. Thus, unlike the native speakers, the perception of the contour tones may be harder for the non-tone language speakers. As for the Yorùbá-Canadian young participants, they may find the identification of these contour tones challenging, compared to the non-contour tones, as observed by Shittu and Tessier (2014).

The lexical experiment (Experiment 4) taps into deeper knowledge of the tone language. It examines native speakers' recognition of real Yorùbá words (minimal pairs) that differ in tones or in tone patterns. This experiment presumes lexical knowledge in Yorùbá. As such, only Yorùbá speaking groups were included. Depending on the language context of the native Yorùbá participants, we might expect different accuracy rates in this task. The Nigeria-based Yorùbá speakers, both young and old might outperform the Canadian-based Yorùbá speakers. Due to the adoption of real Yorùbá words for the experiment, we might also expect the Canadian-based Yorùbá speaking adults to perform comparably to the Nigerian-based Yorùbá speakers. The young Canadian-based Yorùbá speakers are expected to perform worse than the other groups,

due to the effect of the dominant L2 on the L1 features. This is similar to what was observed in Yeh and Lu (2012) where the older, as well as the young fluent speakers of Hakka performed significantly better than the attriters in the lexical task.

Finally, the production experiment (Experiment 5) generally investigates all participant groups' ability at tone imitation. It examines participants' ability to place the three tones in their appropriate tonal spaces. While knowledge of the tones would be expected to be an advantage for the native tone language speakers as suggested by Omozuwa (1991), personal experience at tone perception and imitation, might help non-native tone language speakers.

Meanwhile, the language context for the young Canadian Yorùbá speakers indicates reduced tone language exposure which possibly results from attrition or incomplete acquisition. The demographic factors that may reflect on this group's perception and production pattern are discussed next.

2.7 Demographic factors

The demographic factors include the personal information collected from participants for the present experiments. The relevant information includes their date and place of birth, number of siblings, rate of use of Yorùbá language, date of arrival in Canada, and the number of years in Canada. The information was collected especially for the young Canadian-based Yorùbá speakers to be used in analysis reported in Chapter Eight. The analysis examines the possible effects of these factors on this participant group's tone perception ability. Possibly due to incomplete mastery of the features of the L1, or due to infrequent use, the young participants from the Canadian-based Yorùbá group are expected to exhibit reduced accuracy in the perception and production of some of the features of the L1, especially the tone. The general expectation is that this group will not perform as well as their Nigerian-based Yorùbá-speaking counterparts due to incomplete acquisition or due to attrition effects. To explore the effects of these factors on their perception and production ability, demographic data including their age, place of birth, length of stay in Canada, and their level of L1 use at home are collected. Each of these factors is expected to influence this group's perception and production performance. The lexical experiment, presented in Chapter 6, which involves only the Nigerian and the Canadian-

based Yorùbá speakers will be a focal point in testing this hypothesis. In addition, in Chapter 8, the demographic factors identified for the young Yorùbá-Canadian participants will be analyzed for all the experiments to aid in testing the hypothesis.

Chapter Three: Tone Discrimination Experiment

This chapter presents the first in a series of experiments which test whether participants' age group, namely children (8-16-year-old), or adults (18 years and above) influences tone perception. The experiments also investigate whether the language context which defines participants' exposure to a tone language, will influence tone perception. In addition to the age and language context factors, the present chapter further examines the effect of having the same-tone or different-tone pairs of stimuli, as well as the effect of having monosyllabic and disyllabic stimuli on participants' perception accuracy. The main purpose of this experiment is to investigate the tone perception of all participants regardless of their language background as shown by previous studies (e.g., Hao, 2012; Connell, 2000; Francis et al., 2008; So, 2006; Omozuwa, 1991; Wayland & Guion, 2004; Lee, Vakoch, & Wurm, 1996; Wang, 2006) and discussed in Sections 1.2 and 1.3.3. Among these studies, there are those that suggest that knowledge of the tone language is essential for accurate tone perception (e.g., Wayland & Guion, 2004; Lee, Vakoch, & Wurm, 1996; Omozuwa, 1991). Connell's (2000) study showed that it is not essential to be a speaker of a tone language to be able to perceive pitch accurately. Other studies suggest that knowledge of tone language is not essential for accurate tone perception (e.g., Hao, 2012; Francis et al., 2008; So, 2006). For example, Bakare (1995) found, using a discrimination task, that while non-tone language speakers were able to discriminate Yorùbá tones, they did not perform comparably to the native Yorùbá speakers. They only performed comparably to the music students in the study (as summarized in section 1.2). The use of real Yorùbá words in Bakare (1995) essentially gives more advantage to the native speakers. It allows them to access the meanings of the words to aid their perception rather than attending to the tones alone. The present experiment aims to investigate how language context influences tone perception, when no tone language knowledge is assumed. Therefore, to avoid the possible effects of real words on the native speakers as observed in Bakare (1995), I use the AX discrimination task which includes only nonce words. As the thesis progresses, subsequent experiments investigate other levels of tone perception where knowledge of the tone language becomes increasingly important. The previously discussed literature, including Bakare (1995), did not examine the possible effects of participants' age. In other words, it has not been shown

whether young and adult participants doing the same tone perception tasks will perform comparably.

To this end, the present investigation adopted the AX Discrimination task, in which young and older participants decide whether pairs of monosyllabic or disyllabic nonce words have the same or different tones or tone patterns. Phonotactically-licit nonce Yorùbá words, produced by a native speaker, are used in the experiment. Having been pre-informed that they are going to be listening to Yorùbá tones, participants go through a short practice session before the experiment. Both young and adult native tone language speakers as well as non-tone language speakers are expected to be able to do the task. Nevertheless, factors such as age, language context, the type of tone pairs, and the syllable count may affect participants' accuracy in this task.

Two age groups are identified, namely young and adult. Previous studies such as Sze (2004) and Ciocca and Lui (2003) found tone perception abilities of 10-year-old children as comparable to adult standards. Based on these studies, the prediction would be that in a participant population with comparable language context, we should expect no significant difference in the performance of the two Age groups. However, the present thesis considers the general language experience possessed by the adults. Thus, in situations where both the young and the adult groups share the same first language, then, age can play a crucial role in their perception patterns. Similarly, in a situation where the language context for the adult and the young groups are not comparable, adult participants may be better than the young participants at tone perception.

Three categories of language context are identified: Yorùbá-Nigerian, Yorùbá-Canadian, and English-Canadian. The Yorùbá-Nigerian group consists of native speakers of Yorùbá who are still living in the language community. They are exposed to the language daily and more frequently. The Yorùbá-Canadian group, also native speakers of the tone language, currently live in a non-tone language environment. Their exposure to the tone language is not as frequent as that of the Yorùbá-Nigerian group. The English-Canadian group has participants with no previous or present exposure to a tone language. Thus, against the background of studies that suggest that native and non-native tone language listeners apply different brain activation patterns when perceiving tones (see, e.g., Gandour et al., 2000), the following predictions are made. If native speakers of a tone language are better at tone discrimination than non-native

speakers, as claimed by e.g., Wayland and Guion (2004); Lee et al. (1996), then we would expect both adult and young native Yorùbá speakers to perform better than the non-native speakers. If, however, frequent exposure to a tone language leads to improved tone perception accuracy, as suggested by Yeh and Tu (2012), then the native Yorùbá speakers living in Nigeria should outperform the other groups. Against this background, we might further expect a stepwise pattern, Yorùbá-Nigerian > Yorùbá-Canadian > English-Canadian, if exposure to Yorùbá is the most important factor. Some research suggests that under some conditions, non-native speakers may perform on a par with native speakers (Hao, 2012; Francis et al., 2008), or even better (Wang, 2006). It should be noted here that since our first experiment (see below) is a relatively straightforward same-different decision task, and thus it does not necessarily tap into the knowledge of Yorùbá, it is entirely possible that our non-native participants will be able to complete the task equally well. However, even in this case, exposure to, or experience with, tonal phenomena may help, in which case the adult non-native group, who are all university students in Linguistics classes, may outperform the younger ones. As such, all these may be expected due to factors such as age, language experience, and the nature of the tone discrimination task which may not require knowledge of Yorùbá.

As described in Chapter 1, the language of the study - Standard Yorùbá - has three basic tones identified as High, Mid, and Low (Bamgbose, 1966; Akinlabi & Liberman, 2000). Phonological studies of these tones (e.g., Bakare, 1995), find that the High tone has the highest intensity, the highest fundamental frequency, the highest third formant frequency (F3) as well as the shortest duration. Other studies show that the three tones have distinctive characteristics (e.g., Hayward et al., 2003). The high tone possesses a characteristic breathy phonation in addition to its higher F0 while the Low tone is associated with a distinctively lower pitch than the other two tones and is realized with a creaky phonation. The Mid tone is the least marked of the three tones, but it mostly shares a breathy phonation with the High tone. Its F0 range, and consequently its perceptual space, lies between that of the High and the Low tones. Thus, to test the participants' tone discrimination ability, the experiment examines monosyllabic stimuli having two same-tone pairs (e.g., H vs. H; L vs. L; or M vs. M), or two different-tone pairs (e.g., H vs. L; H vs. M; or L vs. M). Disyllabic stimuli are also examined with pairs having the same patterns (e.g., HL vs. HL; LL vs. LL; or LM vs. LM) or different patterns (e.g., HM vs. LL; LH vs. LM; or MM vs. HH). As the experiment is straightforward and should not require the knowledge of Yorùbá, it is

expected that most participants will be able to discriminate the monosyllabic pairs. Disyllabic patterns may pose some challenge, but participants could still use individual strategies to decide whether the patterns are the same or different. In an earlier study, conducted with a group of L1 Yorùbá-speaking children and adults living in Canada (Shittu & Tessier, 2014), different-tone pairs were found to be easier to discriminate than same-tone pairs for both age groups. Given that, the present thesis also expects a discrimination pattern similar to what was suggested by Shittu and Tessier (2014). Moreover, the study also expects the interplay of the two factors of age and language context on participants' discrimination accuracy. With these hypotheses in sight, the experiment asks participants to listen to pairs of monosyllabic and disyllabic stimuli and then decide whether the tones/tone patterns on the two stimuli are the same or different.

3.1. Methods

3.1.1 Participants

A total of one hundred and thirty-nine (139) participants took part in this experiment. They were broken down as follows: There were seventeen (17) adults and twenty-eight (28) young native Yorùbá speakers recruited in Nigeria. The Canadian-based Yorùbá speakers consisted of fifteen (15) adults and thirty (30) young participants. The third group - the non-speakers of Yorùbá- recruited in Edmonton, consisted of twenty-seven (27) adults and twenty-two (22) young participants.

3.1.2 Stimuli

The stimuli for this experiment consisted of twenty-four (24) pairs of monosyllabic (Table 3.1) and 24 pairs of disyllabic (Table 3.2) stimuli produced with the three Yorùbá tones - High, Mid, and Low. The monosyllabic items were all of CV or CV/Ṽ syllable types, the basic syllable system in the language. Despite all efforts to ensure that all the monosyllabic stimuli were nonce words, a few of them are actual words or near-actual words which may sound meaningful to highly-experienced users of the language. However, within the stimuli pairs, there were rare

instances where two actual or near-actual words co-occur. All attempts to avoid some of these actual words fell short because not only are most of the verbs in the language monosyllabic, most of these verbs also have minimal pairs or even minimal triplets. The disyllabic words, in contrast, were all nonce-words (Table 3.2) with all the possible syllable types in the language. They consisted of VCV, VCṼ, CVCV, CṼCV, CVCṼ and CṼCṼ syllable types. The stimuli included all Yorùbá vowels, consonants, and tones that could combine to form phonotactically-licit Yorùbá words. It needs to be mentioned that the design of the present thesis mainly focussed on participants' accuracy in the perception of the tones. As such, it did not consider the possible influence of the preceding consonants (Hombert, 1977) or the effects of the vowel heights on the three tones (Connell, 2002; Hombert, 1976). For a detailed description about the production and recording of the stimuli, see Chapter 2, Section 2.2. The practice stimuli for this experiment are presented in Appendix B1.

Table 3. 1 Monosyllabic pairs for the AX Discrimination Experiment

Same-tone Pairs		**Different-tone Pairs**	
wó	tḭ́	dọ́	jū
lọ́	dṵ́	bú	gī
fé	lọ́	wó	k͡pà
tḭ́	g͡bó	dṵ́	lè
bẹ̀	gè	gè	bú
wẹ̀	gà	mì	fé
lè	dù	k͡pā	g͡bō
gà	mì	dù	kọ̄
tā	tẹ̄	kọ̄	dọ́
tḭ̄	jū	tẹ̄	dṵ́
jū	gī	mī	gà
g͡bō	mī	gī	bẹ̀

Table 3. 2 Pairs of disyllabic stimuli for the AX Discrimination Experiment

Same-tone Pairs		**Different-tone Pairs**			
làdè	dùrù	bòrí	orì̀	fì̀lọ̀	wẹ́lō
ākp͡ā	dēmū	bọ̄rẹ́	ọ̀kà̀	gb͡ọ̀jẹ̀	wērù
fǐ́jú	wéjẹ́	bísé	lúwā	gb͡ọ̄̀lọ̀	wẹ̀tọ́
bìjẹ̀	dù̇tã̄	dọ̀kú	rādé	ākp͡ā	jábò
bòrí	bẹ̀mú	bōlò	mídò	bọ̄rẹ́	ārè
ārè	jẹ̀bà	dùrù	sẹ̀dá	bẹ̀lọ̀	kēdù̇
dīsẹ́	k͡pọ̄̀jọ́	fádí	tīlō	bẹ̀mú	kẹ́jà
kẹ́jà	rújẹ̀	bánī	jólé		
lúwā	k͡pọ́kī				

3.1.3 Procedure

In the AX discrimination task, participants listened to pairs of stimuli, one at a time. They were instructed to decide whether the two stimuli bore the same or different Yorùbá tones. The experiment, programmed with E-prime software (E-prime 2.0; Psychology Software Tools Inc.), had the stimuli presented in two blocks: monosyllabic pairs of words formed the first part while disyllabic pairs formed the second part. To acquaint participants with what was to come in the experiment, there were six practice trials before the monosyllabic part of the experiment and four trials before the disyllabic part. In each trial, participants listened to a pair of words through the headphones. After listening to the word pair, they pressed the 'Y' key on the keyboard if they thought that the tones were the same. If they thought the tones were different, they pressed the 'N' key on the keyboard. There was an inter-stimulus interval of 2 seconds between the members of any given pair and a post-stimulus interval of 3.5 seconds for participants to input their responses. The interval was to allow participants adequate time to process the two stimuli. Once the duration was reached or the participants had input their response through the keyboard, the next pair of stimuli automatically followed after one second. This procedure continued until new instructions appeared on the screen. Participants could ask the experimenter for more clarification about the task immediately after the practice sessions, before proceeding to the

actual experiment. In case no questions were asked, participants proceeded to the actual experiment. The program was set to select the word-pairs for each trial of the experiment randomly. Each participant worked on his/her own to complete the two parts of the experiment which lasted for about 8-10 minutes. The response accuracy was extracted for analysis, and the results of the analysis follow in the next sections.

3.2 Results

Table 3.3 presents the general accuracy results for the discrimination experiments for the six participant groups. The first column lists the group names. The second column indicates the group's accurate response divided by the total number of responses with the percent accurate responses shown in the third column. The fourth column shows the mean accuracy for each group followed by the standard deviation in the fifth column. The total number of participants from each group is presented in the sixth column. The seventh column shows the number of participants whose accurate scores are better than 29 which is the chance score obtained from the exact binomial test for this experiment. The final column shows the percentage of participants with better than the chance score.

Table 3. 3 Accuracy results for the Discrimination Experiment for all participant groups: ACC=accuracy; SD=standard deviation; BTC= Better than Chance; YorNgr= Yorùbá-Nigerian; YorCan= Yorùbá-Canadian; EngCan= English-Canadian; Ad=adults; Yg=young

P/grp	ACC	%ACC	MEAN	SD	N	BTC(29/48)	%BTC
YorNgrAd	607/816	74%	35.71	6.36	17	15	88%
YorNgrYg	938/1344	70%	33.5	10.06	28	16	57%
YorCanAd	498/720	69%	32.67	9.95	15	9	60%
YorCanYg	758/1440	53%	25.27	10.42	30	12	40%
EngCanAd	954/1296	74%	35.33	7.07	27	20	74%
EngCanYg	541/1056	51%	24.59	4.99	22	4	18%

The percent accuracy (second column) shows that all the adult participant groups were among the top performers in this experiment with percent accuracy of 69% to 74%. Meanwhile, only one of the three young participant groups, the young Nigerian-based Yorùbá speakers, performed comparably with the adult groups with a percent accuracy of 70%. In the better than chance (BTC) scores (last column), the adult participant groups had more participants with BTC scores than the young participant groups with about 60% to 88% scoring better than chance. The young Nigerian-based Yorùbá speakers remained the closest to the adult groups with 57% of its participants scoring better than the chance score. The general results were further analyzed statistically. The procedure and the results of the analysis follow in the next sections.

3.2.1 Statistical analysis

The statistical analysis section for each experiment examines participant groups' performance with respect to the research questions stated above. The lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015) was used to perform a generalized linear mixed-effects regression (function glmer with binomial family) analysis to investigate participants' performance accuracy in each experiment with respect to the identified variables.

Beginning with the AX Discrimination Experiment, the first model examined the relationship between participants' tone discrimination accuracy and the four predictors of interest. The first factor was Age with two levels: *Adult* and *Young*. The second factor was Language Context with three levels: *YorNgr*, *YorCan,* and *EngCan*. The third factor was the tone patterns for the stimuli coded simply as Tone with two levels: *Same* and *Different*. The last factor was the Syllable count with two levels: *Monosyllabic* and *Disyllabic*. The response accuracy, which consisted of the individual score for the 48 target stimuli, was the dependent variable. The response was coded as either '1' or '0' with '1' indicating an accurate or correct score while '0' indicated an inaccurate or incorrect score. As random effects, the model had intercepts for the forty-eight stimuli coded as *item*, as well as intercepts for *participant*. The model also had by-participant random slopes for Tone. The model was based on 6672 observations (i.e., a total of 139 participants x 48 individual item scores).

P-values were obtained by likelihood ratio tests of both a main-effects model and a model with interactions, then the models' fit were compared to see which provided the best fit to the data. Model comparison (function anova) revealed that the model with an interaction term between Tone and each of age and language context fit the data more closely (AIC = 7623.1 vs. AIC(interaction) =7595.0). The first analysis had Yorùbá-Nigerian as the reference level while the second analysis had Yorùbá-Canadian as the reference level. For comparison, the simple model without interaction terms is presented in Appendix B2.

Table 3.4 summarizes the linear mixed-effects model of the effects of age, language context, Tone pattern and Syllable count on tone discrimination accuracy. The coefficients for the parameters estimated by the model are in the second column, with the intercept as the first value followed by the coefficient estimates for the independent variables. The third column presents the standard errors for the coefficients while the z-values are presented in the fourth column. The last column lists the p-values along with the asterisks which indicate the level of significance. The significance codes are in the last row of the table.

Table 3. 4 glmer model of the interactions between Tone and Age, and between Tone and Language Context. Random effects: participant and item. N=6672

Ref. Levels: YorNgr; Disyllabic; Different-Tone; Adult				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.6115	0.2342	6.881	5.96e-12***
Monosyllabic	0.2314	0.1507	1.536	0.124643
Sametone	-0.1204	0.1922	-0.627	0.530847
Young	-0.5587	0.1928	-2.898	0.003756**
YorCan	-0.5422	0.2340	-2.317	0.020480*
EngCan	-0.6894	0.2303	-2.994	0.002752**
Sametone x Young	-0.4559	0.1192	-3.824	0.000131***
Sametone x YorCan	-0.3943	0.1462	-2.698	0.006976**
Sametone x EngCan	0.1304	0.1421	0.918	0.358623
Ref: YorCan; Disyllabic; Different-Tone; Adult				
Intercept	1.0693	0.2383	4.488	7.19e-06***
Sametone	-0.5147	0.1921	-2.680	0.007365**
YorNgr	0.5422	0.2340	2.318	0.020476*
EngCan	-0.1472	0.2297	-0.641	0.521675
Sametone x EngCan	0.5247	0.1394	3.764	0.000167***
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.'** 0.1 ' ' 1				

The interaction model suggests that Age significantly affected participants' discrimination accuracy. Younger participants discriminated the tones significantly worse than the adults, who were the reference level. The analysis also compares the discrimination accuracy along the three language context types by indicating a different reference level for the model. Thus, when Nigerian Yorùbá speakers were compared with both Canadian Yorùbá speakers and Canadian English speakers, the analysis revealed that Nigerian-based Yorùbá speakers performed better at tone discrimination than either Canadian-based Yorùbá speakers or Canadian English speakers. A comparison between Canadian Yorùbá speakers and Canadian English speakers revealed that there was no significant difference between the two language context types concerning tone discrimination. With Tone as a fixed effect with two levels namely Same and Different, the analysis revealed that generally, same-tone pairs attracted worse discrimination than different-tone pairs, but this was only significant when Yorùbá-Canadian was specified as the reference level. The introduction of interaction terms between Tone and Age as well as between Tone and Language Context revealed that the young participants were significantly worse at discriminating same-tone pairs than the adults. In addition, further interactions showed that Yorùbá-Nigerian participants, as well as English-Canadian participants, were significantly better at discriminating same-tone pairs than Yorùbá-Canadian participants. Moreover, while Yorùbá-Canadian participants were significantly worse than Yorùbá-Nigerian participants in discriminating same-tone pairs, there was no significant difference between Yorùbá-Nigerian participants and English-Canadian participants in discriminating same-tone pairs of items. The analysis further revealed that there was no effect of Syllable count, suggesting that participants performed equally well with one and two-syllable stimuli.

To visualize what has been reported in Table 3.4, Figure 3.1 presents a visual representation of the interactions between Tone and the two age groups, the adults, and the young participants. The figure shows that the adults were much better than the young participants in discriminating same-tone and different-tone pairs, and that different-tone pairs recorded more accurate discrimination by the two Age groups. Finally, the distance between the lines for the same-tone and different-tone for the two Age groups reveal that when compared to the adults, the younger participants discriminated different-tone pairs more accurately than same-tone pairs.

Figure 3.1 Mean accuracy in Experiment 1 by age group (adult, young) and types of tone pairs (same tone, different tone); DIFF=different

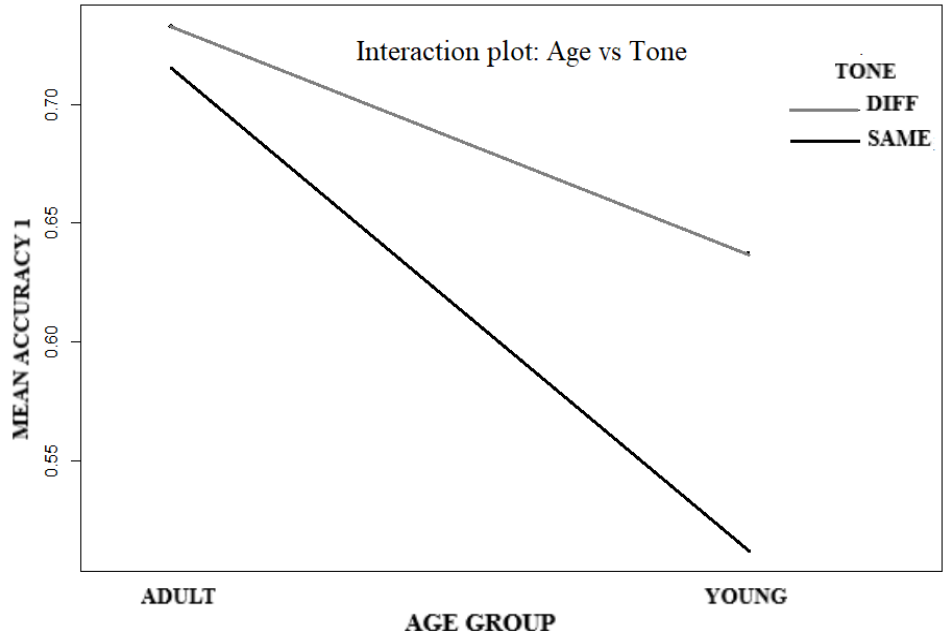
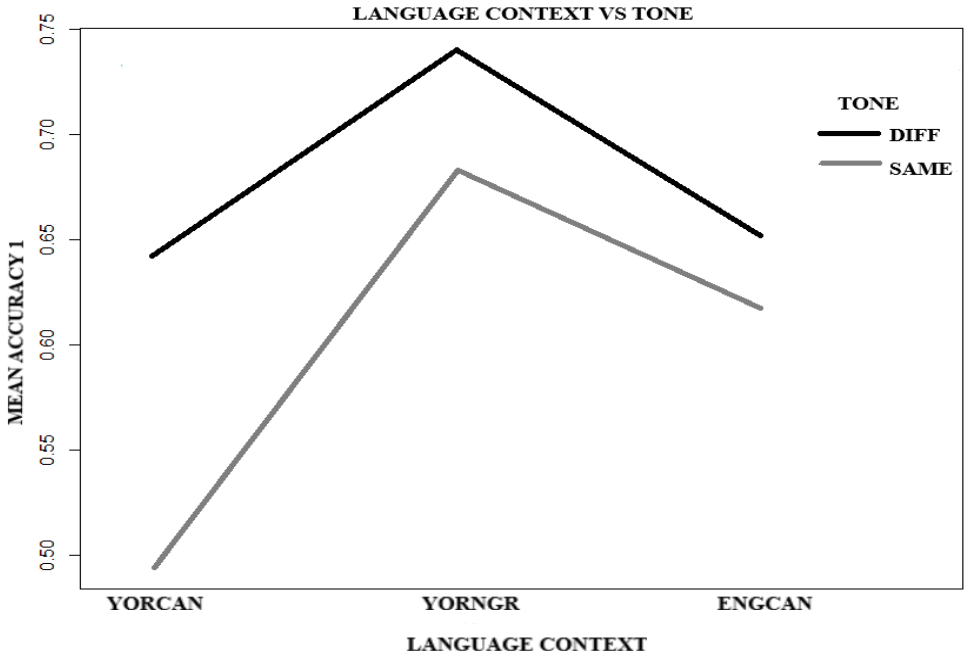


Figure 3.2 presents the plot of the interaction between Tone and Language Context. The figure shows all the three language context types discriminated different-tone stimuli pairs more accurately than same-tone pairs. However, among the three, the Yorùbá-Canadian group discriminated the same-tone pairs compared to different-tone pairs relatively worse than the other two groups. The figure further reveals that the Yorùbá-Nigerian group discriminated both the same-tone and different-tone pairs more accurately than the other two groups. Finally, both the Yorùbá-Nigerian and English-Canadian groups discriminated the two stimuli types more accurately than the Yorùbá-Canadian group. Meanwhile, a model fitted to check for three-way interactions between Tone, Age, and Language Context did not reveal significant results.

Figure 3.2 Mean accuracy in Experiment 1 by language context (Yorùbá-Canadian, Yorùbá-Nigerian, English-Canadian) and types of tone pairs (same tone, different tone)



3.2.2 Tone discrimination by participant groups

This part of the analysis seeks to compare the accurate performance of the participant groups with each other. A simple model was fitted. As fixed effects, the model had Syllable with two levels: Monosyllabic and Disyllabic. Tone had two levels: *Same* and *Different*, while Group had six levels namely *Yorùbá-Nigerian Adult, Yorùbá-Canadian Adult, English-Canadian Adult, Yorùbá-Nigerian Young, Yorùbá-Canadian Young*, and *English-Canadian Young*. Thus, in this model, the 'Group' factor had Age, and Language Context combined into one predictor using contrast coding. In addition to allowing us to examine mean differences between participant groups, contrast coding also allows for testing of the Age, and the Language Context hypotheses between group means. Intercepts were specified for the two random-effect factors: *item* and *participant*. The model fit was checked for by-participant and by-item random slopes, but they did not improve the model's fit. The accuracy score which consisted of the individual score for the 48 target stimuli was the dependent variable. The model was also based on 6672 observations (i.e., a total of 139 participants x 48 individual item score). The presence of an asterisk in the

final column indicates the significance of the results, while the absence of an asterisk indicated no significance.

Table 3. 5 Coefficients of the glmer models of tone discrimination accuracy by participant groups. Random effects: participant and item. N=6672

Ref: Yorùbá-Nigerian adult; Disyllabic; Different-Tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.3914	0.2951	4.716	2.41e-06***
Monosyllabic	0.2559	0.1623	1.577	0.114904
Sametone	-0.4122	0.2097	-1.966	0.049317*
YorCanAd	-0.2052	0.3888	-0.528	0.597666
EngCanAd	-0.1124	0.3367	-0.334	0.738402
YorCanYg	-1.1302	0.3303	-3.422	0.000622***
YorNgrYg	-0.1107	0.3348	-0.330	0.741027
EngCanYg	-1.2263	0.3514	-3.490	0.000484***
Ref: Yorùbá-Canadian adult; Disyllabic; Different-Tone				
Intercept	1.18617	0.31098	3.814	0.000137***
EngCanAd	0.09277	0.36121	0.257	0.797313
YorCanYg	-0.92497	0.34167	-2.707	0.006786**
YorNgrYg	0.09456	0.34943	0.271	0.786690
EngCanYg	-1.02107	0.36031	-2.834	0.004599**
Ref: English-Canadian adult; Disyllabic; Different-Tone				
Intercept	1.278943	0.253769	5.040	4.66e-07***
YorCanYg	-1.017745	0.295201	-3.448	0.000566***
YorNgrYg	0.001785	0.298166	0.006	0.995224
EngCanYg	-1.113851	0.319886	-3.482	0.000498***
Ref: Yorùbá-Nigerian Young; Disyllabic; Different-Tone				
Intercept	1.280723	0.244674	5.234	1.66e-07***
YorCanYg	-1.019522	0.284248	3.587	0.000335***
EngCanYg	-1.115628	0.307690	-0.626	0.000288***
Ref: Yorùbá-Canadian Young; Disyllabic; Different-Tone				
Intercept	0.2612	0.2349	1.112	0.266180
EngCanYg	-0.0961	29850.	-0.322	0.747517

Table 2.5 summarizes the outputs of the model with different reference levels for the group factor. Changing the group reference levels was needed to facilitate the comparisons between and within different groups as demanded by the objectives of the study. To investigate the effect of Age on tone discrimination, the model compared the performance of the young and adult participant groups. Beginning with the two age groups from the same language context, the

analysis in Table 3.5 reveals that within the Yorùbá-Nigerian category, there was no significant difference in tone discrimination accuracy between the young and the adults. As for the Yorùbá-Canadian category, the adults were significantly better at tone discrimination than the young participants. Similarly, for the English-Canadian category, the adults were significantly better than the young ones. Further comparisons between the young and adult participants across language context categories suggests that while the Yorùbá-Nigerian adults did not significantly discriminate the tones better than their younger counterparts from the same language context, they discriminated the tones better than the younger participants from both the Yorùbá-Canadian and the English-Canadian categories. The analysis further suggests that although the Yorùbá-Canadian adults significantly discriminated the tones better than the Yorùbá-Canadian and English-Canadian young participants, there was no significant difference in the discrimination accuracy between this adult group and the young Yorùbá-Nigerian group. Moreover, the third adult group (English-Canadian) also discriminated the tones better than the young Yorùbá-Canadian and the young English-Canadian groups. However, they were not able to discriminate the tones significantly better than the young Yorùbá-Nigerian group. In sum, the models suggested that while all the adult participant groups were better than the younger groups from both the Yorùbá-Canadian and English-Canadian language contexts, they were not better than the young Nigerian-based Yorùbá participants. The comparison between participants of the same Age group but belonging to different language contexts suggests that there was no significant difference in the discrimination accuracy between the adult groups. However, when the young participant groups were specified as the reference levels, a different trend was observed. The young Yorùbá-Nigerian participants, with 70% accuracy, significantly discriminated the tones much more accurately than either the young Yorùbá-Canadian (53%) or the young English-Canadian participants (51%). However, there was no significant difference in the discrimination accuracy between young Yorùbá-Canadian and young English-Canadian participants.

3.2.3 Discrimination of Same-tone vs. Different-tone items

The model whose output is presented in Table 3.6 was fitted to investigate each participant group's accurate discrimination of same-tone versus different-tone items as well as their

discrimination of monosyllabic and disyllabic stimuli. In this model, there was an interaction term introduced between Syllable and Group, as well as between Tone and Group. P-values were obtained by likelihood ratio tests of the model with only the main effects (that is without the interaction term) and another model with the interaction terms. An ANOVA revealed that the model with interaction terms between each of the two factors (Tone and Syllable) and group fitted the data more closely (AIC = 7621.2 vs. AIC(interaction) =7573.1).

Table 3. 6 Coefficients of the glmer models of participant group's accurate discrimination of stimuli pairs along Tones and Syllable count. Reference level: Nigerian Yorùbá adults. Random effects: Participant and Item. N=6672.

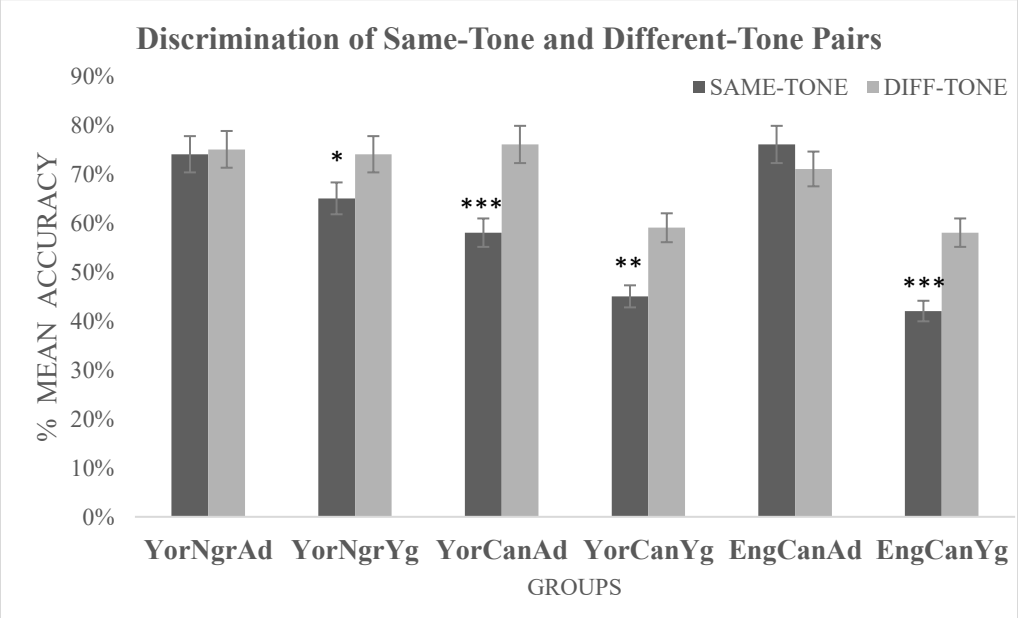
Ref.: YorNgrAd; Disyllabic; DifferentTone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.39481	0.29599	4.712	2.45e-06***
Monosyllabic	-0.20141	0.22088	-0.612	0.361863
YorCanAd	0.06114	0.40515	0.151	0.880048
EngCanAd	-0.35439	0.35024	-1.012	0.311606
YorCanYg	-1.10291	0.34179	-3.227	0.001252**
YorNgrYg	-0.12460	0.35014	-0.356	0.721943
EngCanYg	-1.21889	0.36067	-3.380	0.000726***
Sametone	-0.03870	0.22449	-0.172	0.863130
Monosyllabic x YorCanAd	0.37127	0.25237	1.471	0.141259
Monosyllabic x EngCanAd	0.29936	0.22102	1.354	0.175600
Monosyllabic x YorCanYg	0.57927	0.21148	2.739	0.006159**
Monosyllabic x YorNgrYg	0.48230	0.21978	2.194	0.028205*
Monosyllabic x EngCanYg	0.63353	0.21884	2.895	0.003793**
YorCanAd x Sametone	-102839	0.25388	-4.051	5.11e-05***
EngCanAd x Sametone	0.38629	0.22287	1.733	0.083050 .
YorCanYg x Sametone	-0.70013	0.21252	-3.294	0.000986***
YorNgrYg x Sametone	-0.55231	0.22046	-2.505	0.012235*
EngCanYg x Sametone	-0.72663	0.21975	-3.307	0.000944***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

The analysis compared the discrimination of same-tone pairs versus different-tone pairs of the other participant groups with that of the Yorùbá-Nigerian adult group, which was the reference level. It reveals that compared to the Yorùbá-Nigerian adults, all the three young participant groups, namely, Yorùbá-Nigerian, Yorùbá-Canadian, and English-Canadian, identified same tone-pairs of items significantly worse. Similarly, the Yorùbá-Canadian adult group discriminated same-tone pairs worse than the Yorùbá-Nigerian adults with high statistical significance. The only group to diverge from this trend was the English-Canadian adult, who

discriminated same-tone items better. However, there was no statistically significant difference when compared to the Yorùbá-Nigerian adults, the reference level.

The bar plot in Figure 3.3 further shows each group's discrimination of same-tone versus different-tone pairs. The plot reveals that the Yorùbá-Nigerian adults discriminated same-tone pairs and different-tone pairs at almost equal rates. Other groups appeared to discriminate different tone pairs better than same-tone pairs. Only the English-Canadian adult group discriminated same-tone pairs better than different-tone pairs. All these were not statistically significant as indicated in Table 3.6. However, in comparison to the reference level, four participant groups - Yorùbá-Canadian young and adults, Yorùbá-Nigerian young, and English-Canadian young - discriminated same-tone pairs at worse rates. Thus, the significance stars in Figure 3.3 correspond to the four groups' accurate performance in same-tone pairs in comparison to the Yorùbá-Nigerian adult group.

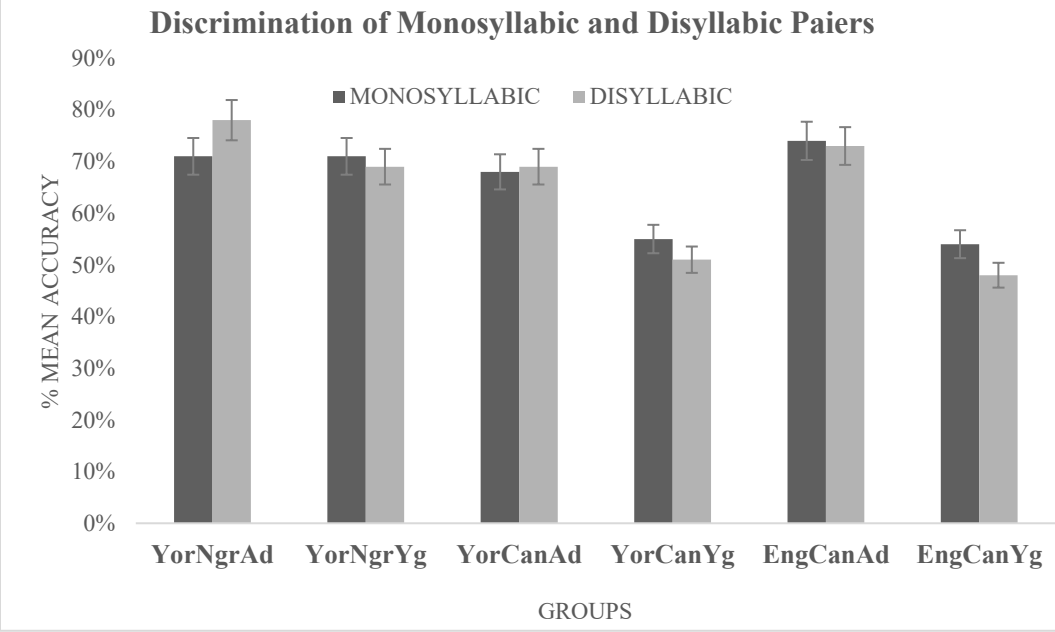
Figure 3. 3 Mean accuracy for the discrimination of same-tone and different-tone pairs by the six participant groups (Error bars=Standard error)



The stimuli for the discrimination experiment also included monosyllabic pairs such as L vs. M, or H vs. H as well as disyllabic pairs such as HH vs. LL, which could be of same-tone or different-tone patterns. As such, the same model compared each group's discrimination of disyllabic and monosyllabic pairs to that of the reference level. Table 3.6 suggests that in comparison to the Yorùbá-Nigerian adult group, the three young groups discriminated

monosyllabic pairs of stimuli better than disyllabic pairs. As for the two adult groups, the analysis does not suggest any significant difference in their discrimination of monosyllabic and disyllabic items, in comparison to the reference level. Furthermore, Figure 3.4 presents each group's discrimination accuracy for monosyllabic and disyllabic pairs. The error bars in the figure does not indicate any significant difference in each group's accurate discrimination of monosyllabic compared to disyllabic pairs.

Figure 3. 4 Mean accuracy for the discrimination of monosyllabic and disyllabic pairs by the six participant groups (Error bars=Standard error)



3.3 Discussion

In this experiment, participants listened to a pair of words bearing two Yorùbá tones. They were to decide whether the tones of the two words were the same or different. The purpose was to investigate their tone perception skills first, as native and non-native speakers of Yorùbá, and second, as young and adult participants. The other purpose was to investigate the language-internal factors that could affect their tone discrimination. The experiment was designed using a simple AX experimental task. As the focus was on tone perception followed by a Yes or No discrimination, this type of task did not require knowledge of the language. As such this type of

task was deemed appropriate. To further prevent language-specific effects, the stimuli were made up of mostly Yorùbá nonce words[11] consisting of both monosyllabic and disyllabic items. Discrimination of monosyllabic pairs formed the first part of the experiment while disyllabic pairs formed the second part. The assumption was that listening to a pair of monosyllabic words bearing the same or different tones would be easier for participants to discriminate than disyllabic pairs, which involved tone patterns. Furthermore, all participants went through the same practice sessions, randomly selected, before the actual experiment.

The results show that globally, both Age and Language Context affected participants' discrimination patterns. The types of tone pairs and Syllable count did not exert significant effects except when comparing accuracy between individual groups. The group analysis further showed that older participants discriminated the tone pairs significantly better than younger ones except within the Yorùbá-Nigerian category where there was no significant difference between the adults and the young participants. Concerning language context, none of the adult groups from the three language context categories discriminated the tones significantly better than the other. In contrast, in the young groups, the Yorùbá-Nigerian young group discriminated the tones significantly better than either of other two young groups while there was no significant difference between the Yorùbá-Canadian and the English-Canadian young groups.

3.3.1 Age and Language Context

Age and Language Context have proved to be crucial factors in this experiment. Considering the previous studies which investigated Age effects on tone perception (e.g., Ciocca & Lui, 2003; Sze, 2004) the expectation is that in typical situations, when young participants (10 years and over) and adults with similar language context are compared, there will be a comparable performance between the two age groups. This suggestion was supported in the present thesis only by the results of the Yorùbá-Nigerian participants, where the adults did not perform significantly better than the young. On the other hand, what was observed in the case of the other two language context categories (Yorùbá-Canadian and English-Canadian), where the adults

[11] As stated earlier, in Section 2.1.2, some of the monosyllabic stimuli were real or near real words due to the difficulty of getting monosyllabic nonce words in the language. The difficulty stems from the fact that more than 90% of the verbs in the language are monosyllabic having minimal pairs or minimal triplets.

significantly outperformed the young participants, shows that other factors, possibly general language experience may be involved.

The three language context types identified include Yorùbá-Nigerian, Yorùbá-Canadian, and English-Canadian. The interaction model (Table 3.4) reveals that specifically, while there was no significant difference between Yorùbá-Nigerian and English-Canadian participants in the accurate discrimination of same-tone versus different-tone stimuli, both the Yorùbá-Nigerian and the English-Canadian participants discriminated same-tone pairs significantly better than the Yorùbá-Canadian participants.

Moreover, when the two factors of Age and Language Context were combined as in the group analysis, the results reveal that, despite the differences in their Language Context, none of the adult groups discriminated the tones significantly better than the other. However, the results of the young groups where the Yorùbá-Nigerian young group outperformed the other young groups, highlight the tone language context factor. This young group appeared to have benefitted from their tone language experience and knowledge. Unlike the young groups, the comparable performance by the three adult groups supports the notion that for the discrimination task, knowledge of the tone language is less important. The adult results appear to conflict with results of other discrimination studies where native tone language speakers outperformed non-native speakers (e.g., Bakare, 1995; Lee et al., 1996; Wayland & Guion, 2004). As for Bakare (1995), it has been mentioned earlier that the use of real Yorùbá words as stimuli could have resulted in more advantage for the Yorùbá speakers, leading to better discrimination. Similar claims could be made about Lee et al. (1996), which also used a mixture of real and nonce words as stimuli. While Wayland and Guion (2004) also used real words as stimuli, in addition to training the participants for the task, the experimental method adopted for the study, namely categorical oddity, involved both discrimination and identification in the same task, making the task more complex than either the AXB, ABX, or the AX discrimination task adopted for the present experiment. It is possible that performing a discrimination task as in Wayland and Guion (2008) might require some knowledge of the language.

Such comparable performance as observed in the present task is possible because the experiment was an AX discrimination task, where listeners simply needed to say whether the two tones (sounds) were the same or not. As such, it did not require knowledge of Yorùbá, which may

explain why Canadian adults performed on a par with the Nigerian adults. The English-Canadian adults could have tapped into their knowledge of pitch and pitch accents such as those used in English to mark lexical (stress shifts), syntactic (e.g., questions) and semantic-pragmatic (information structure) differences, all of which require sensitivity to tonal (pitch, duration, and timing) differences. As such, these adults' experience with dealing with these differences possibly contributed to improved discrimination ability.

In the young age groups, while the young Yorùbá-Nigerian participants discriminated the tones significantly better than either the Yorùbá-Canadian or English-Canadian group, there was no significant difference in the performance between the young Yorùbá-Canadian and the young English-Canadian participants. Thus, while the results of the adult groups does not conclusively support the prediction with respect to language context effects, the results of the young groups appear to support it. Living within the immediate tone language environment, a criterion used to classify Yorùbá-Nigerian participants, contributes to better tone discrimination than not living there. As for the Yorùbá-Canadian young group, some of them may have not fully mastered the tonal features. As for those who have previously mastered the tonal features, movement to, and living in a new language environment, probably have had their tone language perception ability eroded through exposure to a new language system. As for the English-Canadian young group, the fact remains that their unfamiliarity with the new language system puts them at some disadvantage. The English-Canadian adults, despite also not being familiar with the tone language, performed comparably to the other adult groups. It appears that other factors, possibly age (maturity) or general language experience, could explain why the three adult groups posted comparable performance in this task. Perhaps, the experience factor also explains why the literature is abound with evidence that seems to support both comparable and non-comparable performances. Thus, the mixed results observed for both the adult and the young participant groups call for some caution in approaching the effects of language context on tonal perception. Possibly, subsequent experiments in this thesis will further clarify the effects of this factor.

3.3.2 Language-internal Factors

Aspects of the results related to the language-internal features reveal that generally, there was no significant effect for the types of tone pairs on participants' discrimination. However, the interaction model reveals that in comparison to the Yorùbá-Nigerian adult group, four participant groups tended to discriminate same-tone pairs at significantly worse rates. This probably relates to the F0 properties of the three basic tones, each of which has different pitch contours that differentiate them. However, despite having different pitch contours, the High and Mid tones may be confused due to the closeness of their F0. Listeners may misperceive the High tone with a relatively lower pitch than it is, thereby hearing it as a Mid tone. Similarly, they may mistakenly perceive the Mid tone as if it has a slightly higher pitch, thereby being perceived as a High tone. The only tone which may be immune from this confusion is the Low tone. As indicated by the spectrogram in Figure 1.1c, the Low tone is shown with a downward slope while the other two tones were shown with a relatively steady F0 (Fig. 1.1a-b). Hombert (1977) considered the Low tone a falling tone which is an important characteristic of the tone. Connell (2002) also observed that a different physiological mechanism is involved in the production of the Low tones in the four African tone languages he studied. Thus, it is possible that these characteristics may make the Low tone more conspicuous, thus easier to discriminate, when placed beside either the High or the Mid tone. While same-tone pairs might pose some perceptual difficulty, different-tone pairs offer a better chance for correct discrimination because listeners only need to notice the subtle differences in the pitch for discerning monosyllabic pairs, or the pitch patterns for the disyllabic pairs. There is the tendency that tone confusion may occur due to participants' unfamiliarity with the speaker's pitch range. However, the issue of unfamiliarity usually disappears as the experiment progresses. Meanwhile, the confusion between High and Mid tones will be examined more closely in the next experiment which focuses on tone identification.

While studies on the syllabicity effect, such as Dommergues and Segui (1989), observe that monosyllabic items induce a monotony effect, others such as Perdereau (2013) reports that disyllabic items attract better accuracy than monosyllabic items in Mandarin word recognition tasks. However, the present thesis reveals that most participant groups did not show significant superior discrimination accuracy for either monosyllabic or disyllabic words.

In sum, while the results of the adult groups in the present experiment seem to suggest that older Age (i.e., more experience) possibly leads to comparable language processing experience, the results of the younger groups suggest there is a language context effect. The results have also not revealed whether the tones or the syllable count of the stimuli affect tone perception accuracy. Thus, it needs to be investigated through subsequent experiments whether the trend observed here is due to chance rather than a systematic behaviour by these groups. The next chapter covers the next experiment, the first of the two tone-identification experiments, where participants' knowledge of the three Yorùbá tones will be investigated. In the experiment, participants will be required to identify the three Yorùbá tones after listening to individual stimuli. Unlike Experiment 1, this will test domains that tap into their linguistic, and especially Yorùbá knowledge more directly. This will possibly provide more insight on the effects of age and language context on participants tone perception accuracy.

Chapter Four: Tone Identification Experiment 1

The results of the Discrimination experiment presented in Chapter Three suggested that both Age and Language Context affected participants' tone perception accuracy while the type of tone pairs (whether same or different) and Syllable count were shown to be significant only with respect to individual participant groups. The results have shown that generally, Age affects tone discrimination while the results of the young groups have provided evidence that language context affects tone discrimination. However, it has not yet been proved whether the perception of the three Yorùbá tones by these participants is dependent on the exact tone. Thus, the present chapter continues to investigate the effects of the factors related to age, language context, and tone. The first step is to see if the knowledge of the three Yorùbá tones can help disentangle these factors.

Previous studies (e.g., Wayland & Guion, 2004; Lee, Vakoch, & Wurm, 1996; Omozuwa, 1991), have suggested that knowledge of the tone language is essential for accurate tone perception when the task involves processing beyond simple discrimination. In a procedure which involved both discrimination (deciding whether the three tones in a set were the same or different) and identification task (locating the position of the different tone in a set with one different tone), Wayland and Guion (2004) found that native tone language speakers performed better than non-tone language speakers. However, other studies involving a tone identification task (e.g., Hao, 2012; Francis et al., 2008; Connell, 2000) suggest that native and non-native tone language speakers perform comparably at tone identification, suggesting that tone language knowledge may not be essential for tone identification. In the identification task by Hao (2012), participants listened to nonce-word disyllabic Mandarin stimuli and then marked the tone of every syllable they heard with the graphic tonal labels. The results showed comparably accurate identification rates for native Cantonese and the native English speakers. Bakare (1995) also conducted an identification task in which he found that native Yorùbá speakers performed better than non-native speakers. However, the performance by the native Yorùbá speakers possibly relates to the fact that the use of real words for the task allowed the native speakers to attend more to the meanings of the words rather than the to the audio characteristics of the stimulus. Thus, while Hao (2012) used nonce words as stimuli for the identification task and reported a comparable performance between native and non-native tone language speakers, Bakare (1995) used real

words for the same type of task and reported better performance for the native tone language speakers over non-tone language speakers. Finally, our first experiment showed that there was no difference in performance for the adult groups while there was a difference for the young groups.

The present experiment asks whether the role of age and language context are affected by tone language knowledge given a task that we may assume, based on previous studies, would benefit from prior experience with a tone language. Therefore, the tone identification method as in Bakare (1995) and Hao (2012) is adopted, and participants need to identify the three Yorùbá tones as High, Mid or Low. However, in order to reduce the effect of word meaning on native tone language speaker's identification, nonce word stimuli, rather than real words, are adopted. The experiment also investigates participants' age differences and its possible effect on tone perception.

Yorùbá tones are well-defined as having three levels, High , Mid, and Low, each with its own perceptual space distinguished by its F0. As in Mambila (Connell, 2000:168), there is no "overlap or impingement of tones in each other's space" in the production of the three tones by native speakers. In addition, the production of the High and the Low tones are characterized by other features such as voice quality, and phonation type (Hayward et al., 2003). The question is whether the non-impingement of the tones in each other's space attested in production will also be observed in perception. In addition, it is interesting to examine whether the inherent features of the three tones will affect participants' perception. To investigate this, a tone identification experiment was designed in which listeners identify the last tone of disyllabic and trisyllabic nonce words as either High, Mid, or Low Yorùbá tone. The identification method adopted in this experiment is suitable for the purpose because it is easier for participants to focus on a single target tone rather than focussing on all the tones in the disyllabic and trisyllabic utterance. The experiment also used single words instead of longer phrases or sentences to minimize distraction. These stimuli were produced by a native speaker to make the stimuli as natural as possible.

The present task is unlike the previous tone discrimination experiment where the knowledge of the language is deemed not essential. In this experiment, different participant groups may record variabilities in accuracy based on their knowledge of the tone language. Thus, this experiment will reveal, among others, whether non-tone language speakers will perform differently from what was observed in Experiment 1, where lack of knowledge of the language did not impede

their discrimination accuracy. The results will also reveal whether these non-tone language speakers will apply other skills to compensate for their lack of knowledge of the tone language in question. Finally, the results will reveal whether the two Canadian-based younger participants, will apply similar skills as their adult counterparts in their tone perception patterns.

While studies such as Pulleyblank (1986) have argued that Yorùbá has two register tones, upper (High tone) and lower (Mid and Low), an earlier study by Ward (1952:33) described the similarities between the High and the Mid tones, making them distinct from the Low tone thus: "When one hears a word with two level tones which are not felt to be near the bottom of the speaking voice, it is usually impossible to say whether they are two high or two mid tones". Other studies observed features that differentiate these three tones from each other. For example, Connell and Ladd (1990) observed that the Low tone contained a feature that differentiates it from the High and Mid tones. They noted that F0 of the Low tone may fall over the course of an utterance whereas the F0 for the High and Mid tones remain more or less constant. Similarly, Hombert (1976) noted that the lower pitch level, in addition to a falling F0 contour distinguish Low from Mid tones in final position. In a similar vein, Hayward, Watkins and Oyetade (2003) found that the Yorùbá Low tone is typically marked by a distinctive voice quality, which involves more prominent low frequency and a somewhat creakier phonation type while the High tone is marked by a more prominent high frequency and a breathier phonation type which also characterizes the least-marked Mid tone. The creakier phonation distinguishes the Low from both the High and the Mid tones. Furthermore, phonological studies of Yorùbá tones suggest that the High tone possesses more strength and stability while the Mid tone is the weakest and most unstable (e.g., Akinlabi, 1985; Pulleyblank, 1986; Orie, 1997; Akinlabi & Liberman, 2000).

The perception study by Bakare (1995) further found that the High tone had the highest F0, the highest third formant frequency, the highest intensity while the Low tone had the lowest of all these acoustic cues. Moreover, the High tone attracted the best perception by participants and that the hardest tone to perceive for native Yorùbá speakers was the Mid tone. As for non-native Yorùbá participants, the Low tone was the hardest to perceive. In sum, while the three tones are differentiated by F0, the High tone is also marked by greater intensity (Bakare, 1995) while the Low tone is typically marked by a distinctive voice quality (Hayward et al., 2003; Connell & Ladd, 1990; Hombert, 1976). The Mid tone is the least marked of the three tones. Thus, the

present perception study seeks to examine the extent to which listeners will identify these Yorùbá tones based on pitch height as well as the other tonal features described above.

Thus, if, according to Bakare (1995), High tone has the highest F0, the highest F3, and the highest intensity, then we should expect the tone to attract better identification accuracy than the other two tones, as observed in his study. However, if we consider the findings of Hayward et al. (2003), Hombert (1976, 1977), and Connell and Ladd (1990) about the characteristic quality of the Low tone that distinguishes it from the other two tones, then we should also expect the tone to attract relatively more accurate identification. Thus, Mid tone, which remains the least marked of the three tones, may attract the worst identification accuracy. Nevertheless, the mid-range position of the tone, making it closer to both the High and the Low tones may also be a factor in the identification pattern for inexperienced, non-native tone language listeners. This factor may either lead to confusion between Mid and the other two tones, or to the misidentification of the other two tones as Mid.

However, across individual participant groups, there may be a deviation from this general prediction. The Yorùbá-Nigerian young and adult groups, and the Yorùbá-Canadian adult group, may not record many discrepancies in their perception of the three tones. As native speakers, their exposure to the L1 could lead to similar accuracy rates for the three tones. As for the remaining three groups, the High-tone's higher F0 and intensity are expected to contribute to better perception accuracy. However, despite these factors, the closeness of its F0 to that of the Mid tone is predicted to lead to tonal confusion in some listeners, especially when the speaker's voice is also factored in. The Low tone, on the other hand, has a distinctive characteristic, namely creakiness, which expectedly may contribute to better perception accuracy by some listeners. One final observation is that the stress pattern of English, which is the L1 for the English-Canadian groups, may work in favour of these participants. They may compare the High tone with the stressed syllable of English, thereby leading to a somewhat improved perception of the tone.

4.1 Methods

4.1.1 Participants

A total of one hundred and thirty-six (136) participants took part in this experiment. They were the same set of participants as in Experiment 1. There were seventeen (17) adults and twenty-seven (27) young participants in the Yorùbá-Nigerian group, fifteen (15) adults and twenty-eight (28) young participants in the Yorùbá-Canadian group, and twenty-seven (27) adults and twenty-two (22) young participants in the English-Canadian group.

4.1.2 Stimuli

The stimuli consisted of 24 disyllabic nonce words bearing the three Yorùbá tones (Table 4.1) and 24 trisyllabic nonce words also bearing the three Yorùbá tones (Table 4.2). The disyllabic stimuli contained all the possible syllable types in the language such as the VCV, VCṼ, CVCV, CṼCV, CVCṼ and CṼCṼ syllable types. Similarly, the trisyllabic stimuli consisted of the VCVCV, CVCVCV, CVCVCṼ, CVCṼCV, and CVCṼCṼ syllable types. The stimuli included Yorùbá vowels, consonants, and tones that could combine to form phonotactically-licit Yorùbá nonce words.

Table 4. 1 Disyllabic stimuli for the Identification Experiment (1): L=Low tone; (M)=Mid; (H)=High

Stimuli (L)	Stimuli (M)	Stimuli (H)
bìlù	kárē	ìkọ́
kấg͡bò	k͡pètā	òlé
ọ̄mì	kī̄rī	g͡béjú
rẹ̄dà	lọ́g͡bẹ̄	sàtọ́
tìkọ̀	dùsō	sī̄yẹ́
kātẹ̀	ẹ̄dū̃	múdḯ
túdè	díwọ̄	ọ̀sá
wódà̀	ròlū	jẹ̄tí

Table 4. 2 Trisyllabic stimuli for the Identification Experiment (1)

Stimuli (L)	Stimuli (M)	Stimuli (H)
Ādèbà	ālórī	g͡bȉgúré
fádīhṵ̀	àk͡pọ̀tọ̄	ìlūkó
fákọ́tì	àrẹ́tã̄	múrēfọ́
fọ̄sằmù	rōkàlọ̄	k͡pārēdí
k͡péfọ̄lẹ̀	gìdígb̄ō	àsák͡pẹ́
ìfálẹ̀	dīrēk͡pā	défṹjá
k͡pētṹbọ̀	ìlúrē	bọ́lẹ̀mú
ìkōlò	mímúdẹ	sālégb̄ọ́

4.1.3 Procedure

The identification experiments required participants to identify individual tones after listening to a sound segment over the headphone. The participants were asked to identify the tones of the last syllables in the nonce-word tokens. A practice session (see Appendix C1), consisting of six trials and involving both disyllabic and trisyllabic stimuli, preceded the actual experiment. During the practice session, the software played each sound stimulus once while the participant listened carefully for the last segment (syllable) which carried the target tone to be identified. This segment bore either the High, Mid, or Low Yorùbá tone. After listening to the stimulus, they pressed the H key on the keyboard for the High tone, the L key for the Low tone, and the M key for the Mid tone. After playing a sound stimulus, the program waited for input from the participants for up to 2.5s before playing the next sound. Once the inter-stimulus interval of 2.5s was reached or when participants had inputted their response, the program automatically terminated the present stimulus and played the next one. The program was set to select the stimuli for each trial randomly. Once the practice session was completed, and the participants did not require any more clarification from the experimenter, the experimental trials followed. Participants worked on their own until the end of the two-part experiment. The first part of the experiment included only the disyllabic stimuli while the second part included the trisyllabic stimuli with a short break between the two parts. The whole experiment lasted about 8-10

minutes. The response accuracy recorded by the software was extracted and analyzed; the results are presented in the next section.

4.2 Results

Table 4.3 presents the general accuracy results for the Identification 1 experiment. The second column indicates each group's number of accurate scores with the accuracy percentage in the third column. The mean score and the standard deviation are presented in the fourth and the fifth columns respectively. The total numbers of participants are in the sixth column. For this experiment, the chance level score is 24. The last two columns indicate the number and the percentage of participants that have a better than chance (BTC) score.

Table 4. 3 Percent identification accuracy for all participant groups: ACC=accurate scores; SD=standard deviation; BTC= Better than Chance; YorNgr= Yorùbá-Nigerian; YorCan= Yorùbá-Canadian; EngCan= English-Canadian; Ad=adults; Yg=young

	ACC	%	MEAN	SD	N	BTC	%BTC
YorNgrAd	541/816	66%	31.82	14.46	17	13	76.5%
YorNgrYg	691/1296	53%	25.59	13.09	27	15	55.5%
YorCanAd	348/720	48%	23.2	12.7	15	6	40%
YorCanYg	506/1344	38%	18.07	7.33	28	7	25%
EngCanAd	805/1296	62%	29.81	5.88	27	25	93%
EngCanYg	367/1056	35%	16.68	3.86	22	3	14%

With accuracy rate ranging from 35% to 66%, this experiment proved to be low-scoring compared to the tone discrimination experiment. However, all the adult groups in each language context category scored higher than the young groups. The Yorùbá-Nigerian adults accurately identified the tones at a rate of 66% compared to the 53% recorded by the Yorùbá-Nigerian young group. The Yorùbá-Canadian adults also identified the tones more accurately (48%), compared to the rate recorded by their younger counterparts (38%). Similarly, in the English-Canadian category, the adults recorded a 62% accurate identification rate, compared to 35% recorded by the younger participants. The low-scoring nature observed in the experiment was

also reflected in the number of participants with BTC scores. The English-Canadian adults recorded the highest number of participants with BTC scores, followed by the Yorùbá-Nigerian adults and then the Yorùbá-Nigerian young group. The Yorùbá-Canadian adult, Yorùbá-Canadian young, and English-Canadian young groups recorded the lowest number of participants with BTC scores at 40%, 25%, and 14% respectively. Further analysis of the results follows in subsequent sections.

4.2.1 Statistical analysis

Using the lme4 package (Bates, Maechler, Bolker & Walker 2015) in R (R Core Team, 2015), a generalized linear mixed effects regression analysis (glmer) was performed to examine the relationship between participants' tone identification accuracy and four fixed effects, namely syllable count, tone, participants' age, and tone language context. Syllable, as the first fixed effect had two levels - *Disyllabic* and *Trisyllabic*. Tone had three levels: *High, Mid*, and *Low*, while Age had two levels: *Adults* and *Young*. Language Context was divided into three levels namely *Yorùbá-Nigerian, Yorùbá-Canadian*, and *English-Canadian*. All four fixed effects were included in the model. The accuracy score, which consisted of the individual score for the 48 target stimuli, was the dependent variable. The response was coded as either '1' or '0' with '1' indicating an accurate or correct score while '0' indicated inaccurate or incorrect score. The model had intercepts for the two random effects of *Participant* and *Item*. The model was based on 6528 observations (i.e., a total of 136 participants x 48 individual item score).

P-values were obtained by likelihood ratio tests of a main-effects model, a model with an interaction between Age and Language Context, and a third model with an interaction between Age and Tone. A model testing the three-way interaction between Age, Language Context, and Tone did not converge. An ANOVA test comparing the models to see which provided the best fit to the data revealed that the model with the interaction term between Tone and Age (AIC(interaction) = 7782.1) fitted the data more closely than either of the main-effects model (AIC = 7784.8) (see Appendix C2), or the model with an interaction between Age and Language Context (AIC(interaction) =7786.6). The addition of by-participant random slopes for tone did not improve the models' fit. The output of the best fitting model is summarized in Table 4.4.

In Table 4.4, the first value in the second column shows the intercepts and underneath the intercepts are the coefficient estimates for the fixed effects. The third column presents the standard errors for the coefficients, while the fourth column indicates the z-values. The last column lists the p-values; the asterisks indicate the level of significance. The results of the main effects model (Appendix C2) suggest that Age and Language Context significantly affected participants' identification accuracy, while Tone and Syllable count did not. The negative coefficient estimates for Age indicate that younger participants performed worse than the adult participants in tone identification.

Table 4. 4 Coefficients of the glmer model of the interaction between Age and Tone. Random effects: participant and item. N=6672

Ref: Yorùbá-Nigerian; HighTone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.28957	0.25033	5.152	2.58e-07***
Trisyllabic	-0.12861	0.13528	-0.951	0.34179
LowTone	0.02287	0.18515	0.124	0.90170
MidTone	-0.40775	0.18427	-2.213	0.02691*
Young	-1.04505	0.21016	-4.973	6.60e-07***
YorCan	-0.88994	0.23672	-3.759	0.00017***
EngCan	-0.69637	0.23122	-3.012	0.00260**
LowTone x Young	0.10226	0.14072	0.727	0.46743
MidTone x Young	0.35295	0.13958	2.529	0.01145*
Ref: YorCan; HighTone				
Intercept	0.39961	0.25266	1.582	0.1134
EngCan	0.19358	0.23178	0.835	0.40362
Ref: YorNgr; Low Tone				
HighTone	-0.02286	0.18513	-0.123	0.90173
MidTone	-0.43061	0.18449	-2.334	0.01959*
Young	-0.94277	0.21026	-4.484	7.33e-06***
High x Young	-0.10227	0.14071	-0.727	0.46734
Mid x Young	0.25068	0.13983	1.793	0.07301.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

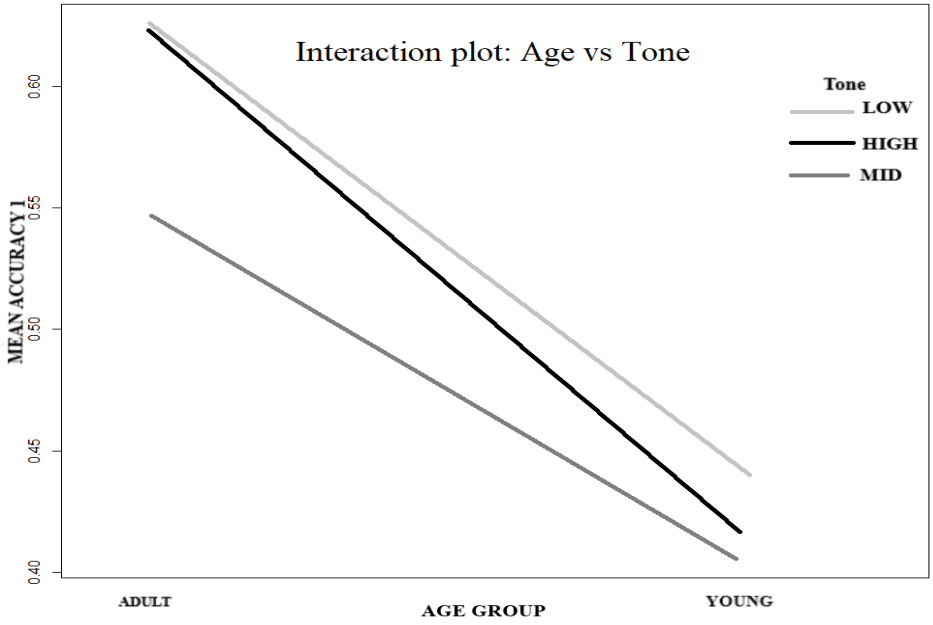
While the interaction model in Table 4.4 reveals the same results for the Age factor, it further reveals that Tone had a moderate effect on the participants' tone identification accuracy. The positive coefficient estimates for Low tone indicates that Low tones were identified better than High tones. However, the effect was not statistically significant. In contrast, the comparison between Mid and High tones suggests that the Mid tone was significantly identified at a worse

rate than High tones. The analysis further suggests that Yorùbá-Nigerian listeners performed better in tone identification than either Yorùbá-Canadian or English-Canadian listeners. However, Yorùbá-Canadian was not significantly better than English-Canadian in tone identification accuracy. Finally, the analysis reveals that compared to the adults, the younger participants identified Low tones better than High tones (statistically insignificant), and they identified Mid tones more accurately than High tones with statistical significance.

When the Low tone was indicated as the reference level to compare the identification accuracy of Low and Mid tones, the results suggests that the High tone was identified at a worse rate than Low tone, although without statistical significance. Moreover, Mid tone was identified at a significantly worse rate than Low tone. Finally, a third Model that examined the interactions between Age and Language Context did not reveal any significant effects.

The interaction between Tone and Age is plotted in Figure 4.1. The non-parallel lines indicate that there were interactions between participants' age groups and their accuracy in identifying the three basic tones. First, the plot indicates that generally, adult participants identified all the three tones more accurately than the younger participants. The plot further indicates that the Mid tone recorded the lowest mean accuracy for both the adults and the young participants, while the Low tone recorded the highest mean accuracy for both age groups. The close distance between the plots for both the High and Low tones for the adults indicates that there were no significant effects in their identification of these two tones. In contrast, the distance between the High- and the Mid-tone lines, and the Low- and Mid-tone lines indicate that the adults tended to identify the High and Low tones better than the Mid tones. In comparison, the plots for the young participants indicate a wider margin between the Low tone and the other two tones, with a wider margin between Low and Mid tones. Thus, the young participants tended to identify Low tones better than either the High or the Mid tone.

Figure 4.1 Mean accuracy in Experiment 2 by age group (adult, young) and tone (High, Mid, Low)

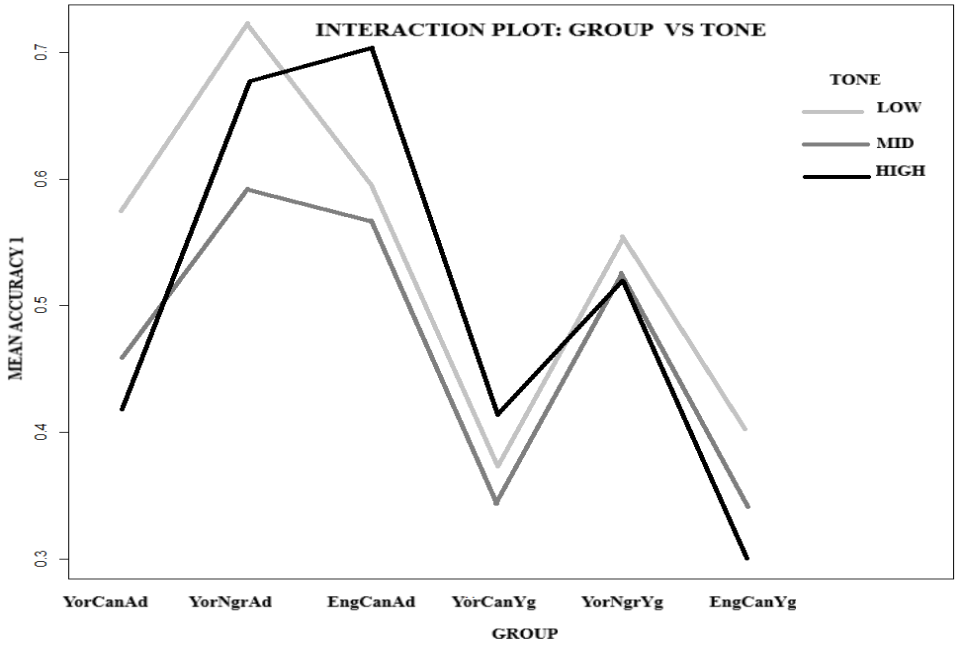


On the effects of Syllable count on participants' identification accuracy, the two models suggested that Syllable, as a factor had no significant effect on participants' accuracy despite the syllabic variability of stimuli.

In sum, the analysis revealed that while Age and Language Context are possible predictors for tone identification accuracy, the effects of the other two predictors, Tone, and Syllable count, are yet to be observed. To do this, there is the need to answer the question about whether all participant groups recorded comparable performance regarding the three tones. This will be the subject of the next section.

4.2.2 Tone identification by the six participant groups

Figure 4.2 Mean accuracy in Experiment 2 by participant groups (YorCanAd, YorNgrAd, EngCanAd, YorCanYg, YorNgrYg,EngCanYg) and tone (High, Mid, Low)



The interaction plot in Figure 4.2 reveals that the High tone recorded the highest identification rate by two groups namely English-Canadian adult and Yorùbá-Canadian young groups while the Low tone recorded the highest identification rate by the remaining four groups. However, there was no instance wherein the Mid tone recorded a higher rate of identification than either the Low or the High tone. The plot further reveals that in comparison to the other two adult groups, the Yorùbá-Nigerian adult group tended to identify Low tones better than the other two tones, while the English-Canadian group tended to identify High tones more accurately than the other two tones. Among the young groups, the Yorùbá-Nigerian young identified the tones more accurately than the other two groups. Nevertheless, the plot does not reveal any identification tendency towards any of the three tones for this group. Thus, the primary revelation of the plot is that the six participant groups did not appear to record similar accurate identification pattern for all the three tones. The next section introduces the statistical analysis to investigate the tone identification patterns of the participant groups further. The first part of the analysis compares the accurate performances of the six participant groups while the second part investigates the groups' identification pattern for the three tones.

To investigate the comparative performance of participant groups, a simple model with Syllable (two levels: *Disyllabic* and *Trisyllabic*), Tone (three levels: *High; Mid; Low*), and group (six levels: *Yorùbá-Nigerian Adult, Yorùbá-Canadian Adult, English-Canadian Adult, Yorùbá-Nigerian Young, Yorùbá-Canadian Young,* and *English-Canadian Young*) was fitted without any interaction term between the fixed effects. As random effects, the model had intercepts for *Participant* and *Item*, as well as by-participant random slopes for the effect of Tone. The accuracy score that consisted of the individual score for the 48 target stimuli remained the dependent variable.

P-values were obtained by likelihood ratio tests of two models: one model with only the three main effects (without random slopes) and another model with by-participant random slopes for the effect of Tone. An ANOVA revealed that the model with random slopes (AIC =7676.9) fitted the model more closely than the model without the random slope (AIC = 7786.6). The model was based on 6528 observations (i.e., a total of 136 participants X 48 individual item score). The output of the model is summarized in Table 4.5.

Table 4. 5 Coefficients of the glmer model of the comparative performance of participants of the same or different Age groups and Language Context. Random Effects: participant and item. N=6672

Ref: Yorùbá-Nigerian Adult				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.1982	0.3177	3.771	0.000163***
YorCanAd	-1.1536	0.4104	-2.811	0.004946**
EngCanAd	-0.4777	0.3674	-1.300	0.193469
YorCanYg	-1.7050	0.3578	-4.765	1.89e-06***
YorNgrYg	-0.8859	0.3590	-2.468	0.013592*
EngCanYg	-1.8677	0.3720	-5.020	5.16e-07***
Ref: Yorùbá-Canadian Adult				
Intercept	0.04465	0.33689	0.132	0.89455
EngCanAd	0.67582	0.39495	1.711	0.08705.
YorCanYg	-0.55140	0.37910	-1.454	0.14581
YorNgrYg	0.26767	0.37231	0.719	0.47217
EngCanYg	-0.71411	0.37999	-1.879	0.06020.
Ref: English-Canadian Adult				
Intercept	0.7205	0.2682	2.687	0.00722**
YorCanYg	-1.2272	0.3055	-4.018	5.88e-05***
YorNgrYg	-0.4082	0.3191	-1.279	0.20092
EngCanYg	-1.3899	0.3432	-4.050	5.12e-05***

Ref: Yorùbá-Nigerian Young				
Intercept	0.3123	0.2620	1.192	0.23320
YorCanYg	-0.8191	0.3079	-2.660	0.00781**
EngCanYg	-0.9818	0.3250	-3.021	0.00252**
Ref: Yorùbá-Canadian Young				
Intercept	-0.5068	0.2574	-1.969	0.04900*
EngCanYg	-0.1627	0.3287	-0.495	0.62066
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

The model in Table 4.5 specified different participant groups as reference levels to make comparisons among the various groups. The analysis suggests that Yorùbá-Nigerian adults identified the tones significantly better than their younger counterparts, including the younger age groups from the other language context categories. For the Yorùbá-Canadian adult group, there was no significant difference in tone identification accuracy between this adult group and the young participants from the same language context as well as those from the other two language context types. The English-Canadian adult group identified the tones significantly better than their younger counterparts as well as the younger group from the Yorùbá-Canadian category. Meanwhile, there was comparable identification rate between this adult group and the Yorùbá-Nigerian young group.

Furthermore, when the three adult groups were compared, the analysis suggests that Yorùbá-Nigerian adults identified the tones significantly better than the Yorùbá-Canadian adults while there was no significant difference between the Yorùbá-Nigerian and English-Canadian adult groups. Moreover, the positive coefficient estimate suggests that the English-Canadian adults were inclined to identify the tones better than Yorùbá-Canadian adults although without statistical significance. As for the young participants, the analysis suggests that the Yorùbá-Nigerian young group identified the tones significantly better than either the Yorùbá-Canadian or the English-Canadian young group. Meanwhile, there was no significant difference between the Yorùbá-Canadian and the English-Canadian young groups. The next section will examine the accurate identification of the three tones by each of the six participant groups.

4.2.3 Individual group's identification of the three tones

The previous section revealed that globally, the Yorùbá-Nigerian adults significantly identified the tones more accurately than all the other groups except the English-Canadian adults. However, this analysis has not answered the question about each groups' identification pattern of the three basic Yorùbá tones, High, Mid, and Low. To answer this, another model was fitted with interaction term introduced between Tone and Group. P-values were obtained by likelihood ratio tests of both a model with only interactions and a second model with interactions, as well as by-participant random slopes for the effect of Tone. An ANOVA revealed that the model with by-participant random slopes for the effect of Tone (AIC=7674.1) fitted the data more closely than the model with the only the interaction terms (AIC = 7757.5). The outputs for the best fitting model are presented in Tables 4.6 and 4.7. The Tables present the outputs of the model fitted to examine participant groups' accurate identification of the three basic tones and the effects of the Syllable count on accurate identification of the three tones. The model chose Yorùbá-Nigerian adult group as the reference level for the Group factor, while 'High' was the default reference levels for 'Tone'.

Table 4. 6 Coefficients of the glmer model of participant groups' identification of the tones. Ref: YorNgrAd; High Tone. Random effects: participant and item. N=6672

Ref: YorNgrAd; High tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.14489	0.34325	3.335	0.000852***
Trisyllabic	-0.13686	0.14208	-0.963	0.335429
LowTone	0.38652	0.33504	1.154	0.248643
YorCanAd	-1.48942	0.45585	-3.267	0.001086**
EngCanAd	-0.07726	0.39792	-0.194	0.846051
YorCanYg	-1.50569	0.39460	-3.816	0.000136***
YorNgrYg	-0.92262	0.39868	-2.314	0.020656*
EngCanYg	-2.03017	0.41424	-4.901	9.54e-07***
LowTone x YorCanAd	0.51885	0.40879	1.269	0.204362
MidTone x YorCanAd	0.72955	0.48762	1.496	0.134618
LowTone x EngCanAd	-0.94025	0.35727	-2.632	0.008495**
MidTone x EngCanAd	-0.25214	0.42295	-0.596	0.551079
LowTone x YorCanYg	-0.57266	0.35700	-1.604	0.108697
MidTone x YorCanYg	0.08294	0.42222	0.196	0.844265
LowTone x YorNgrYg	-0.17179	0.35957	-0.478	0.632817
MidTone x YorNgrYg	0.52298	0.42826	1.221	0.222025
LowTone x EngCanYg	0.10153	0.37123	0.274	0.784463
MidTone x EngCanYg	0.64585	0.44006	1.468	0.142203

The analysis in Table 4.6 suggests that Yorùbá-Canadian adults were significantly different from the Yorùbá-Nigerian adults in identifying disyllabic High-tone stimuli. The same trend was observed for the all the young groups. The negative coefficient estimates suggested that the three young participant groups were more inclined to identifying High tones at a significantly worse rate than the Yorùbá-Nigerian adult group. The English-Canadian adult group was the only group not to register a significant difference in High-tone identification compared to the Yorùbá-Nigerian adults. Following the group comparisons, the analysis made further comparisons of Low and Mid-tone identification to that of High tone. The comparisons did not yield any significant difference except for the English-Canadian adults who remained the only group to record a significant difference in their identification of Low tones versus High tones. The negative coefficient estimate suggested that, compared to the group reference level, this group tended to identify Low tones at a significantly worse rate than High tones.

Table 4.7 presents the output of the same model with Low and Mid tones as the reference levels for Tone. The analysis suggests that Mid tone attracted a significantly lower identification accuracy rate than Low tone, while no significant difference was recorded for High versus Low tone. The group comparison suggested that all the other five groups were significantly different from Yorùbá-Nigerian group in Low-tone identification. As in the High-tone identification cited above, the negative coefficient estimates suggested that these groups tended to identify Low tones at a significantly worse rate than the Yorùbá-Nigerian adult group. In addition to the Low tone, the English-Canadian adult group further recorded significant differences in their identification of High tones. Thus, in comparison to the Yorùbá-Nigerian adults, they identified High tones better than Low tones as revealed by the positive coefficient estimate.

Table 4. 7 Coefficients of the glmer model of participant groups' identification of the tones. Ref: YorNgrAd; Low and Mid. Random effects: participant and item. N=6672

Ref: YorNgrAd; LowTone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.5314	0.3589	4.267	1.98e-05***
MidTone	-0.8203	0.3658	-2.243	0.02493*
YorCanAd	-0.9706	0.4718	-2.057	0.03968*
EngCanAd	-1.0175	0.4124	-2.467	0.01361*
YorCanYg	-2.0784	0.4119	-5.045	4.53e-07***
YorNgrYg	-1.0944	0.4144	-2.641	0.00827**
EngCanYg	-1.9286	0.4292	-4.494	7.00e-06***
Mid x YorCanAd	0.2107	0.4579	0.460	0.64537
Mid x EngCanAd	0.6881	0.3980	1.729	0.08380.
Mid x YorCanYg	0.6556	0.4007	1.636	0.10177
Mid x YorNgrYg	0.6948	0.4043	1.718	0.08573.
Mid x EngCanYg	0.5443	0.4138	1.316	0.18833
Ref: YorNgrAd; MidTone				
Intercept	0.71113	0.39068	1.820	0.06873.
EngCanAd	-0.32942	0.45945	-0.717	0.47338
YorCanYg	-1.42277	0.45883	-3.101	0.00193**
YorNgrYg	-0.39967	0.46391	-0.862	0.38895
EngCanYg	-1.38435	0.47869	-2.892	0.00383**
Low x YorCanAd	-0.21069	0.45757	-0.460	0.64519
Low x EngCanAd	-0.68810	0.39751	-1.731	0.08345.
Low x YorCanYg	-0.65559	0.40021	-1.638	0.10140
Low x YorNgrYg	-0.69475	0.40387	-1.720	0.08539.
Low x EngCanYg	-0.54430	0.41338	-1.317	0.18794

With Mid tone as the reference level for tone, the group comparison reveals that only two groups, the Yorùbá-Canadian and English-Canadian young groups, recorded a significant difference in their identification of Mid tones compared to the Yorùbá-Nigerian adults. Meanwhile, further comparison of the reference level tone with the other tones, for the participant groups, does not yield significant difference for one tone over the other.

In sum, this analysis reveals that Yorùbá-Nigerian adults, the reference level for Group, identified High tones significantly better than the other groups except for the English-Canadian adult group. The group also identified Low tones significantly better than all the other groups, while their identification of Mid tone was significantly better than the young Yorùbá-Canadian and English-Canadian groups. Furthermore, the analysis suggests that generally, there was no significant difference in the identification of High versus Low tones. However, High tones were

identified significantly better than Mid tones, and Low tones were identified significantly better than Mid tones as revealed by the interaction model in Table 4.4. Having observed that the participant groups did not record similar rates of accurate identification for the three basic tones, the next section will analyze participants' misidentification patterns to suggest if any factors may influence participants in identifying one tone as another.

4.3 Tone misidentification by participant groups

In this section of the analysis, participants' actual responses inputted through the keyboard were extracted from the experimental software for analysis. The analysis was undertaken to reveal what was otherwise not obtainable through the statistical linear mixed-effect regression analysis. The analysis unveiled how participants incorrectly identified the three tones. In addition, it revealed the tone that the target tone was most commonly misidentified as. The tone misidentification data for all participants are presented in the pie charts in Figure 4.3. The Mid tone attracted about two-thirds (67%) of all High-tone misidentification, while Low tone attracted the remaining one-third (33%). Both the Low and the High tones attracted almost equal rates of Mid-tone misidentification. Finally, the Mid tone attracted two-thirds (71%) of all Low-tone misidentification with the remaining one-third (29%) going to the High tone.

Figure 4. 3 Misidentification of the three basic tones

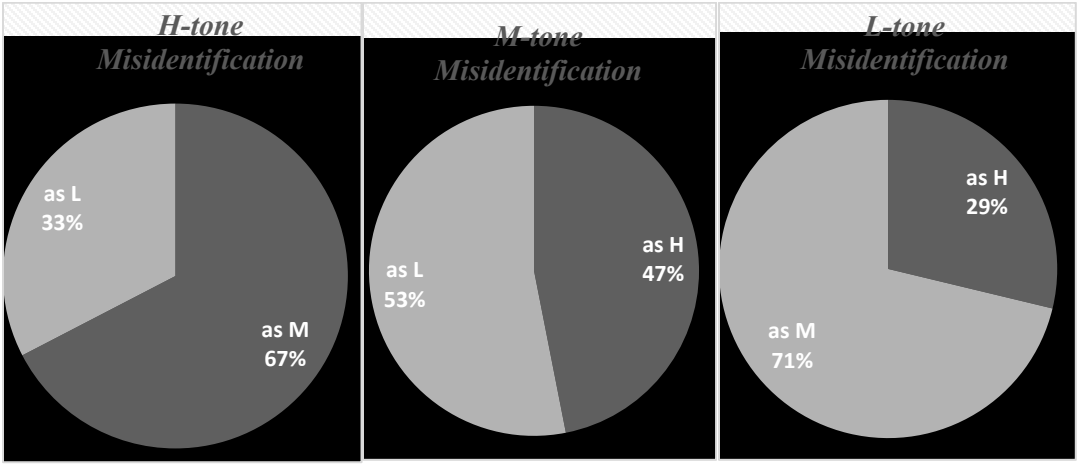


Table 4.8 further presents the results of the misidentification analysis for each of the six participant groups. It shows the numbers for the incorrect identification of each of the three

target tones for each participant group. Beneath each number, the percentage of the total for each tone's misidentification is included in brackets. To examine the relation between tones and their correct and incorrect identification by participant groups, chi-square tests of independence were performed for the results of each participant group. The test results show that the relation between tones and participants' identification patterns were highly significant for all groups: YorCanAd (X^2 (2) = 110.37, p = < .00001); YorNgrAd (X^2 (2) = 73.629, p = < .00001); EngCanAd (X^2 (2) = 354.264, p = < .00001); YorCanYg (X^2 (2) = 331.895, p = < .00001); YorNgrYg (X^2 (2) = 149.52, p = < .00001); EngCanYg (X^2 (2) = 291.777, p = < .00001).

Table 4. 8 Incorrect identification of the three tones by the six participant groups

	YorCanAd		**YorNgrAd**		**EngCanAd**		**YorCanYg**		**YorNgrYg**		**EngCanYg**	
H as	M	L	M	L	M	L	M	L	M	L	M	L
	52	26	34	16	101	15	121	63	74	39	114	81
%	**(67)**	(33)	**(68)**	(32)	**(87)**	(13)	**(66)**	(34)	**(65)**	(35)	**(58)**	(42)
M as	H	L	H	L	H	L	H	L	H	L	H	L
	35	34	25	23	57	111	100	108	52	38	95	98
%	(51)	(49)	(52)	(48)	(34)	**(66)**	(48)	(52)	**(58)**	(42)	(49)	(51)
L as	H	M	H	M	H	M	H	M	H	M	H	M
	28	18	17	23	12	150	49	145	38	45	55	112
%	**(61)**	(39)	(43)	**(58)**	(7)	**(93)**	(25)	**(75)**	(46)	**(54)**	(33)	**(67)**

The analysis reveals that all participant groups misidentified High tones as Mid tones more than as Low tones. Across all groups, the percent rate of High tone misidentification as Mid tone ranged from 58% (English-Canadian young) to 87% (English-Canadian adult). The remaining groups ranged from 65% to 68%. Incorrect identification of High tones as Low was not as prolific as Mid tone, except for the English-Canadian young group. They misidentified High tones as Low in 42% of their total High tone misidentification. All the other groups identified them as Low at rates between 13 and 31%. It needs to be noted that the English-Canadian adult group recorded the highest rate of High-tone misidentification as Mid tone (87%) as well as the lowest rate of High-tone misidentification as Low (13%). The numbers suggested that High tone

was often identified as Mid by all groups and, except for the English-Canadian young group, the tone was less likely to be identified as a Low tone.

Mid-tone misidentification data revealed that four out of the six groups, namely Yorùbá-Nigerian and Yorùbá-Canadian adults, and Yorùbá-Canadian and English-Canadian young, misrepresented Mid tones either as Low or as High at almost equal rates. Out of the remaining two groups, the English-Canadian adult misidentified the Mid tone mostly as Low (66%), while the Yorùbá-Nigerian young group misidentified the tone mostly as High (58%). Overall, the numbers showed that, except for these two groups, all the other groups misidentified Mid tones either as High or as Low at almost equal rates. This misidentification trend observed for the tone here, as well as in the overall misidentification in Figure 4.3, suggested that the Mid tone presented more challenges to the participants than the other tones.

The data for Low-tone misidentification (Table 4.8) reveal that five of the six participant groups, namely the Yorùbá-Nigerian and Yorùbá-Canadian adults, as well as all the three young groups, misidentified Low tones as Mid tones at a percent rate ranging from 54% to 93%. Only the Yorùbá-Canadian adult group misidentified the tone mostly as High at a rate of 61%. Both Yorùbá-Nigerian adult and young groups further recorded a considerable rate for misidentification of Low tones as Mid with 43% and 46% respectively. As observed in their High-tone misidentification, the English-Canadian adult group's misidentification of Low-tones as Mid (93%) was also the highest rate. They also had the lowest rate of Low-tone misidentification as High (7%). Overall, the data suggested that most groups chose to identify Low tones mostly as Mid tones.

4.4. Discussion

In this experiment, participants listened to nonce words produced with Yorùbá lexical tones and then identified the last tone of each word as either High, Mid, or Low. The purpose was to investigate tone identification patterns of both native and non-native tone language speakers. In addition to examining whether Age and Language Context affect participants' tone identification ability, the experiment also investigates whether the acoustic features of each of the three tones will influence participants' accurate identification of the tones. The experiment was designed

using disyllabic and trisyllabic nonce words. It did not include monosyllabic words due to the expectation that monosyllabic stimuli might result in more confusion for listeners and therefore could have distorted the results. Thus, the polysyllabic stimuli provided listeners with a penultimate tone to contrast with the target final tone. Moreover, nonce words were adopted to eradicate language-specific effects, especially for the native speakers. All participants went through the same pre-experiment practice sessions involving randomly selected stimuli.

Similar to what was observed in Bakare (1995), and Yeh and Lu (2012), this experiment recorded a lower mean score in comparison to the AX Discrimination experiment, possibly due to the relative difficulty of the tone identification task or to factors related to participants' differences. The task required participants to listen to a word and then identify the tone of the last syllable as either High, Mid, or Low. It is primarily a mental exercise which can be described as involving phonological encoding (Strange, 2011; Hao & Jong, 2016) due to its high memory demand and polysyllabic stimuli. It is comparatively more difficult than the AX Discrimination task, which is relatively straightforward, requiring participants to employ simple discriminatory skills to decide whether two stimuli bear the same tone or not.

The analysis pertaining to the general age and language context factors suggests that, concerning Age, adults were globally better than young participants at tone identification. This may indicate that Age affords these adults a broader experience at general language or tone processing. It may also indicate that the adults have greater familiarity, or facility with experimental tasks of this nature. Concerning the language context factor, no significant effect was observed in the identification accuracy between the Yorùbá-Nigerian and the English-Canadian adults. Meanwhile, both the adult and the young participants from the Yorùbá-Canadian group and the young English-Canadian group recorded comparable performances.

Perhaps it is more appropriate here to compare the findings of Hao (2012) and Bakare (1995) to the findings of the present experiment. As mentioned earlier, the identification tasks in these studies and that of the present experiment are similar. However, while Hao (2012) shares more affinity with the present experiment in the use of nonce words, Bakare (1995) differs by using real words as stimuli. The fact that the results of the present experiment, where there was no significant difference between the adult native Yorùbá speakers and the native English speakers, supports the findings of Hao (2012) should not be surprising. It indicates that generally, the

adults, regardless of language background could identify the tones accurately. As has been mentioned earlier, it is possible that the use of real words in Bakare (1995) benefitted the native Yorùbá speakers, allowing them to attend to the word meaning as well as to the tone of the words, leading to improved identification.

The results of the present experiment also reflect what was found in Connell (2000) and Francis et al. (2008). In Connell (2000) native Mambila speakers as well as English speakers showed equal sensitivity to the pitch-based Mambila stimuli. In the case of Francis et al. (2008), native tone language speakers (Mandarin Chinese, and Cantonese Chinese) as well as native American English speakers were able to identify the tones as well as successfully do the difference rating tasks. The results of the two studies did not report any differences between the native and non-native tone language speakers. The comparative performance reported in the two studies is similar to what was observed in the present study, between the Yorùbá-Nigerian and the English-Canadian adults. It needs to be mentioned that, similar to the present study as well as Hao (2012), Francis et al. (2008) and Connell (2000) respectively used nonce words and synthetic tokens as stimuli. The only difference between these two studies and the present study is the participants' training before the experiments.

The increasing importance of the knowledge of the tone language is highlighted by the results of the young participant groups than the adult groups. The Yorùbá-Nigerian young group identified the tones significantly better than their Yorùbá-Canadian and the English-Canadian counterparts. Unlike the Yorùbá-Canadian young group who rarely used the tone language, the Yorùbá-Nigerian young group benefitted from the language experience acquired through frequent use of the native tone language. The performance of the English-Canadian young group is expected, due to their lack of any previous exposure to the tone language.

The identification task could be more demanding especially for the English-Canadian participants who were naïve to lexical tones. The only information they had about the language was from the short practice session that preceded the experiment. Thus, for adult non-native speakers to successfully identify this novel feature, it requires some general language experience. As university students taking various linguistics courses, perhaps these participants were better prepared for this task than some of the other participant groups in the study. Unlike the English-Canadian participants, the task was not expected to be as hard for the Yorùbá-Nigerian

participants who have acquired the tonal information since birth and were accessing it daily through L1 usage. Second, these participants had the privilege of learning the Yorùbá tonal system at school, and they also went through the short practice sessions preceding the experiments. By comparison, the other native-speaking groups, especially the Yorùbá-Canadian young participants did not possess most of these privileges. Although most of the younger members were born in Nigeria, they arrived in Canada at a relatively young age. Therefore, they did not have the advantage of learning the Yorùbá tonal system within the school system in Nigeria. Thus, infrequent usage of the L1 may account for their poor performance in this task. Moreover, continuous exposure to, and greater mastery of the L2 may also lead to gradual displacement of the L1 features, as suggested by Major (1992).

Despite the arduousness of the task, the results show that the English-Canadian adults identified the tones with 62% accuracy, a rate closer to what was recorded by the Yorùbá-Nigerian adults (66%). Nevertheless, closer scrutiny of the accurate results of individual participants reveals that some of the top performers among the Yorùbá L1 participants scored between 90 – 100% accuracy, while the highest scorers among the English L1 participants scored 80%. This is an indication that the identification accuracy of the native tone language participants was not random chance; they were consciously aware of the minute tonal differences. However, this is not to dispute the fact that some participants with a tone language background also found the task difficult. For some of the native Yorùbá speakers, both young and adult, it is possible that the use of nonce words affected them negatively. They might expect to hear real words as stimuli. Such expectation might prompt them to tap into what they independently knew rather than focusing on the task. In contrast, the non-tone language speakers, especially the adults, did not have real-word expectations about the tones. They could focus on the task at hand, tapping into their language perception skill, in addition to the experience they obtained from the short practice session. The young non-tone language speakers might not be able to do this possibly due to their lack of experience compared to the adults. It is also reasonable to suggest that level tones in languages such as Yorùbá are more perceivable to non-native tone language speakers, because the tones are differentiated mainly by pitch height. In that case, the success recorded by the English-Canadian adults may suggest the use of pitch height as cues to identify the tones. This aspect of the results is similar to what Connell (2000) observed in the identification pattern of native and non-native speakers for Mambila tones. He observed that while the native speakers

were not in any significant way better at pitch perception than non-native speakers, they were more consistent than the English speakers in placing the four Mambila tones more accurately in their tonal spaces. The English speakers had more variation in their placement of the tones. We may also attribute the performance in Connell (2000) as well as in the present experiment, to the fact that the English speakers, due to the lack of lexical tone in their language, adopt individual strategies in the identification or categorization of the stimuli.

Another possible explanation for the relative success for the English-Canadian adults in this experiment resides in PAM-S model as proposed by So and Best (2010; 2011, 2014). The model highlights the critical role played by the similarities or dissimilarities between the tonal features of the listeners' L1 and the target language in non-native tone perception. As native English speakers and linguistics students, the English-Canadian adults might have applied their advanced language experience and knowledge of the English stress system to notice the similarity between the Yorùbá High tones and English primary stress. The analysis of the results, which indicated that the English-Canadian adult group was the only group to identify High tones at a significantly higher accuracy rate than either the Mid or the Low tones, supported this suggestion. The association of a breathy phonation with the Yorùbá High tone, and the possibility that English speakers are sensitive to it (Hayward et al., 2003), may also have informed the identification pattern of the English-Canadian adult group regarding the High tone. Concerning the Yorùbá Mid and Low tones, it is possible that they either considered the two tones as unstressed or they could not find an English equivalent for the two tones. Hence the lack of significant difference in their identification rates for Low versus Mid.

4.4.1 The three basic tones

The results of the statistical analysis (Table 4.4) reveals that generally, there was no significant difference in the identification of High tones compared to Low tones, even though the coefficient estimates indicate that Low tones attracted better identification accuracy than High tones. The analysis further suggests that Mid tones were identified significantly worse than either the High or the Low tone. Thus, among the three basic tones, Mid tones recorded the lowest accuracy rates overall.

Considering what has been discussed about the effects of the inherent features of the three tones on participants' perceptual accuracy, in addition to the finding by Bakare (1995), the present thesis predicted that the High tone should record better identification rates. Although the High tone was identified more accurately than Mid tone, there was no significant difference in identification accuracy rates between the the High and the Low tone. Thus, the prediction that the inherent features of the Low tone which distinguish it from the other two tones, namely the falling F0 (Hombert, 1976, 1977) and the creakiness (Hayward et al., 2003) likely contributed to the improved identification accuracy and is hereby attested by the present result. It is, however, possible that this experiment is not sensitive enough to find differences between High and Low tone identification. Meanwhile, the present result concurs with Bakare (1995) that the hardest tone for native Yorùbá participants to perceive is the Mid tone. It further adds that it is also the hardest to perceive for both native and non-native Yorùbá speakers. The other prediction was that the 'default' Mid tone, which is the least marked (Hayward et al., 2003), and with less tonal space (similar to the two Mambila Mid tones (Connell, 2000), may be confused with the High tone due to the closeness of their F0 and a somewhat similar phonation type. As such, it would attract lower rates of accurate identification. The result of the present experiment, where the Mid tone was identified significantly worse than either the High or the Low tone supports this part of the prediction.

Similar to the results in the present analysis, the second experiment in Hombert (1976) suggests instances where the Low tone may be identified as Mid. Having observed the fact that falling F0 is an important feature for the identification of Low tone in word-final position, Hombert (1976) asked native speakers to identify LL patterns in which the final Low tones had been modified with either a steady F0 or with a steady F0 and an increased duration and amplitude. The results revealed that the modified final Low tones were significantly identified as Mid. This is an indication that we may have instances where participants identify Low tones as Mid especially if they could not perceive the falling F0 which characterizes the production of the Low tone, and which is equally essential for its accurate identification.

4.4.2 Tone Misidentification

Evidence from the analysis of all participants' misidentification data (Figure 4.3) indicates that both High and Low tones were misidentified as Mid tones in about two-thirds of the total misidentification (High as Mid = 67%; Low as Mid = 71%). We might not be able to say categorically what prompted participants to misidentify High and Low tones as Mid, but the following line of reasoning may be suggested. In this experiment, participants had to choose between three tones where two of these tones were at either ends of the language's tonal space namely top and bottom. The third was situated in-between. In such a situation misperception could occur mainly due to lack of knowledge of the language or due to participants' naivety to pitch perception. Using the F0 factor (Hombert, 1976), it is necessary that participants perceive these tones with the appropriate frequency range as well as notice the direction of change of the tone as either falling, steady, or rising (Hombert, 1976). Probably due to lack of knowledge or experience of the tones, listeners fail to perceive the tones based on these criteria, which may lead to misperception. If listeners fail to notice the subtle changes in tone height, then they may resort to perceiving the tones at one pitch range, mostly Mid. Meanwhile, it is possible that certain contexts such as the preceding consonant, the vowel types, or the speaker's voice may affect the tones, thus leading some listeners to incorrectly represent High and Low as Mid. Although these contexts are not expected to significantly affect the results of the present experiment, the possibility should be considered while viewing the results. In contrast, the misidentification of High tones as Low, and misidentification of Low tones as High occurred in only one-third of the total misidentification. This observation supports the suggestion that both High and Low tones possess inherent features that make them more distinguishable to listeners than the Mid tone. The next section will discuss misidentification patterns with respect to individual participant group.

4.4.3 Misidentification by individual group

4.4.3.1 High vs. Low

The data in Table 4.8 showed a trend where almost all participant groups recorded lower rates of misidentification of High tones as Low and vice versa. The English-Canadian adult group

strongly represented this trend. This group had the lowest rates for these two types of misidentification. Two factors, namely the falling F0 of the Low tone as well as the English stress system, can explain this occurrence. In addition to the falling F0 (Hombert, 1976; 1977), creakiness is another feature that typically accompanies the production of the Yorùbá Low tone in many environments as observed in Hayward et al. (2003). This feature always makes the tone conspicuously identifiable to a native Yorùbá listener, and, as the results show, to experienced non-native Yorùbá listeners as well. Moreover, the English-Canadian adults possibly employed the English stress system, to their advantage. This group, who were all undergraduates and mostly English-speaking adults, could readily access and positively apply, their acquired knowledge and experience of the English stressed and unstressed syllable system in the identification of the High and Low tones. The High tone is comparable to the primary stressed syllable while the Low tone is similar to the unstressed syllable. The two factors suggested here may also have been applied by experienced listeners from the other participant groups to help them reduce the misidentification of High tones as Low or Low tones as High. Finally, the larger F0 difference between High and Low tone (42Hz), coupled with the difference in phonation type between the two tones (breathy versus creaky) are factors that may have caused listeners to distinguish these two tones clearly and then perceive them more accurately.

4.4.3.2 High vs. Mid

It has become clear that the bulk of the High tones were misrepresented as Mid tones by all the participant groups. However, while High tones were mostly misidentified as Mid, the data showed that Mid tones were not identified mostly as High except by only one participant group: the Yorùbá-Nigerian young group. Misidentification of High tones as Mid is possibly caused by speaker's voice, which may affect all participant groups. Depending on the speaker's pitch level, listeners may misperceive a High tone as Mid, especially in monosyllabic words where there are no other tones to make the final tone more discernible. In addition to the speaker's voice, the closeness of F0 (19Hz) between these two tones may also lead to confusion among participants. Perhaps, participants preconceived the High tone as a distinct and easily identifiable tone with a higher pitch that stood it out from the other tones. It is also possible that due to the context and probably due to the speaker's voice factor, listeners may resort to misidentifying them as Mid, or in some cases, as Low tones.

The misidentification rate of Mid tones was almost equally split between High and Low tones; The trend observed here may be related to the issue of speaker's F0 range and to Mid – High tone confusion due to less space for the Mid tone. In the tonal space, the Mid tone lies in intermediary position between the High and Low tones, and it is considered as the least marked of the three Yorùbá tones (Hayward et al., 2003). The voice measurement for the native speaker who recorded the stimuli shows that the F0 difference between the Mid and the High tone (19Hz) is not much unlike what is observed between Mid and Low (23Hz). In comparison, the F0 difference between Low and High tone is 42Hz. Thus, it is possible that listeners did not perceive the Mid tones as distinct enough to classify them categorically into Mid tones; hence the resulting misidentification of the tones as High or Low at almost equal rates.

Finally, the English-Canadian group's low misidentification of Mid tones as High may still be due to the English stress factor. With their linguistic experience, this group might not have perceived the Mid tones as high-pitched enough to compare to English primary stress. As such, they did not often misidentify them as High tones.

4.4.3.3 Mid vs. Low

In the preceding paragraph, misidentification of Mid tones has been observed to be equally split between High and Low tones; F0 closeness has been cited as a possible cause of this misidentification pattern. One aspect of the data unaccounted for is the fact that unlike the other groups where the Mid tone misidentification rates were equally split between High and Low tones, the English-Canadian adult group misidentified Mid tones mostly as Low tones with a 66% misidentification rate. With a different outlook, the same factor of English stress system may also account for this. While the F0 of the Mid tone is not easily comparable to the English primary stress, the flatness of the Mid tone (with F0 contour lower than that of High tone), makes it more comparable to an unstressed syllable than a stressed one. Hence, there were relatively higher rates of misidentification of Mid tones as Low. By comparison, the conspicuously lower pitch of the Low tone makes it discernable from Mid tones to trained listeners such as the young and adult Yorùbá-Nigerian groups. Thus, all the other groups except the Yorùbá-Canadian adults, misidentified more Low tones as Mid and less as High tone. The issue of F0 distance remains a factor that could have informed the perception pattern observed here. According to Hombert (1976), the falling F0 of the Low tone was used to judge the L-final

patterns (ML; LL). It is possible that due to the F0 distance factor, these participants perceive the Low tones as a steady tone, which, according to Hombert (1976) was the criteria used by Yorùbá native speakers to judge the M-final patterns (MM and LM). In his second experiment, when the final Low tones in LL pattern was modified with a steady-state F0 or with a steady F0 plus increased duration and amplitude, participants identified the final Low tone as Mid. The results suggest that falling F0 is an important cue for the identification of Low tone in word-final position.

To further support the F0 distance factor, the lowest rate of misidentification of Low tones as High (7%) was recorded by the English-Canadian adult group. The misidentification data suggest that when the High tone was the target tone to be identified, it was mostly misrepresented as Mid by all groups. When the Mid was the target tone, the direction of its misrepresentation was split equally between High and Low. When the Low was the target tone, most groups (five out of six) misrepresented it mostly as Mid. Thus, Mid tone emerges as the tone which mostly attracts misidentification of the other tones. The data further suggests that for most groups, it is uncommon to misrepresent High tones mostly as Low, neither is it common to misrepresent Low tones mostly as High. Some of the reasons suggested for the higher rate of misidentification of High as Mid include the shorter F0 distance between the two tones which may constitute confusion to listeners. In addition to the F0 distance, speaker voice factor may further compound this High-Mid confusion.

The fact that High tone is not often misidentified as Low and that Low tone is not often misidentified as High may also be linked to the F0 distance between these two tones. A higher F0 distance between the two tones gives each tone its distinct quality and features that allow it to either be identified correctly or, if incorrectly, as a tone with relatively shorter F0 distance (Mid tone). One of the causes for the lower rates of misidentification of High tone as Low is the creakiness feature that usually accompanies the production of Low tone. It is such a prominent feature that both trained and untrained listeners are unlikely to miss. Finally, for the English-Canadian adults, the English stress system is a factor that may have contributed to more accurate identification of High and Low tones. As native English speakers as well as being university students, they may resort to their experience of the English stress system as one of their

strategies to accomplish the identification task which taps into the knowledge of the tone language more than simple tone discrimination.

The F0 distance between Mid and Low tones is also relatively small. This fact probably explains what is observed in the misidentification of Low tones as Mid and the participant groups' lack of preference for either High or Low tone in Mid tone misidentification data. Further evidence is found in Hombert (1976), where the final Low tones in the LL and ML patterns were identified as M (i.e. L → M). The explanation given was that this could occur in cases where the participants judged the final Low tone as having a steady F0, or as having a steady F0 with increased duration and amplitude. In the case of the present experiment, it may occur when participants fail to perceive the final Low tones with the appropriate falling frequency or when they fail to notice the direction of change of the tone.

In sum, the present experiment has further indicated that there is no similarity between the participant groups in their accurate identification of the three tones. Yorùbá-Nigerian adults identified the tones significantly better than only the Yorùbá-Canadian adults while they did not identify them better than the English-Canadian adults. The Yorùbá-Nigerian young participants identified the tones significantly better than the other two young groups. Furthermore, while the English-Canadian adults performed comparably to the other adult groups, they demonstrated a peculiar identification pattern different from all the other participant groups. Could all these observations be attributed to participants' language context or some language-internal factors? This will continue to be a major focus of investigation in the next experiment.

Chapter Five: Tone Identification Experiment 2

The aim of the present experiment is to investigate participants' identification of only two tones, High and Low, which are the two tones involved in the formation of the Yorùbá complex or contour tones. This makes the present experiment different from the previous identification task in which all the three tones were targeted. By targeting these two tones, I can observe, (i) whether the High tone attracts better identification rates due to its acoustic cues (Bakare, 1995), (ii) whether the inherent complexity of the contour tones leads to reduced perception rate, and (iii) whether there is a difference in the perception rates between the two contour tones, H (in LH context) and L (in HL context). As in the previous experiments, all results are considered within the context of participants' age and language context.

As in the previous identification experiment, this experiment examines another level of tone perception where the knowledge of the tone language becomes increasingly important. Instead of identifying simple tones by themselves, participants in the present experiment are required to identify these tones in the context of a preceding tone, in effect identifying two Yorùbá complex tones. Previous studies on the perception of Yorùbá tones (e.g., Hombert, 1976; Bakare, 1995) have focused mainly on the perception of the three basic tones. The use of only monosyllabic stimuli by Bakare (1995) indicated that only the three basic tones were targeted for the identification task. Targeting complex tone patterns can only be achieved using Yorùbá disyllabic patterns. Hombert (1976) used disyllabic stimuli to examine how native listeners perceive the tones in bisyllabic contexts, allowing him to include one of the Yorùbá complex tones. The study by Shittu and Tessier (2014), observed, using the identification task, that young Yorùbá-Canadian participants found the two complex tones more difficult to identify than the simple tones. However, there has been no study focusing solely on how native and non-native tone language listeners perceive these two tones depending on their complexity, as described in Section 1.1.2. Thus, for the present experiment, I adopt the identification task using disyllabic nonce words as stimuli. The present experiment also examines the effects of participant differences with respect to age, a factor, which was not considered in the previous perception studies.

As in the preceding Identification experiment (Chapter 4), participants are asked to identify the final tones of the stimuli as either High, Mid, or Low, when these tones are preceded by either of the three tones in the first syllable. Doing this allows me to investigate how these two complex tones are being perceived and misrepresented. Participants are expected to identify or misidentify the tones either based on the complexity of the tones or based on their knowledge, or lack of it thereof, of the tone language. The procedure is the same as in the first identification experiment.

Among the studies that examined the local effects exerted by the three Yorùbá tones on each other are Connell and Ladd (1990), and Laniran and Clements (2003). These studies suggested that of all the tones, neighbouring H and L are found to exert local influences on each other. Moreover, Hombert (1976) distinguished one of the two contour tones (LH pattern) as a rising contour, making it different from the falling and level tones. The other contour tone (HL pattern; falling) was not included in Hombert (1976) because Yorùbá does not have VCV nouns with the HL tone pattern. However, the HL pattern was featured in Omozuwa (1991) who distinguished the Edo HL pattern as a falling tone similar to the Yorùbá pattern. The Low tone in the second syllable has its source from the F0 of the preceding High tone. In addition, the F0 measurements for the complex tones (see Chapter 1; Table 1.3 and Figure 1.2) further indicate that they are too distant from the simple tones at onset, middle, and offset points, and are liable to confuse non-trained listeners. Features such as these have led to the expectation that the contour tones may be difficult to perceive to untrained listeners.

An earlier study (Shittu & Tessier, 2014) reported relatively lower perception accuracy rates for the Yorùbá contour tones, compared to the simple tones. Thus, the present thesis predicts, that based on the participants' L1 background (knowledge of the tone language or lack of it), there will be discrepancies in their perception for each of the two Yorùbá contour tones. In typical situations, native tone language speakers will not face difficulty in identifying the complex tones. On the other hand, the English monolingual speakers may find the complex tones more difficult. However, experienced English monolingual adults may apply their knowledge of the English prosodic system (especially the iambic and trochaic system) to help them process the complex tones, thereby leading to improved performance. This prediction is based on a study by Nguyen and Ingram (2005) which suggests that experienced users tend to use cues that are actively applied in their L1 to process prosodic contrasts in the L2. Moreover, a recent study by

Chrabaszcz, Winn, Lin, and Idsardi (2014) suggests that when cues are contrastive across syllables, they are stronger when they are in the iambic contour than when they are in the trochaic contour. Thus, this study may be adduced to support the relevance of the English prosodic system to the English monolingual speakers in the present perception study. These hypotheses will be revisited in the light of the results obtained in subsequent sections.

5.1 Methods

5.1.1 Participants

A total of one hundred and seventeen (117) participants took part in this experiment. They were the same participants as in the previous experiments. However, some participants were not available to take part in this aditional experiment; hence the different total. They were broken down as follows: seventeen (17) adults from the Yorùbá-Nigerian group and twenty-seven (27) from their younger counterparts; eleven (11) participants from the Yorùbá-Canadian adult group and thirteen (13) participants from the Yorùbá-Canadian young group. The English-Canadian group had twenty-seven (27) adults and twenty-two (22) young participants.

5.1.2 Stimuli

This experiment was designed with the disyllabic stimuli in Table 5.1. The 48 disyllabic nonce words were bearing the three Yorùbá tones. The stimuli included a mixture of contour as well as regular tonal sequences. As in the other experiments the stimuli included all the Yorùbá vowels as well as consonants that could combine with the vowels to form phonotactically-licit Yorùbá nonce words. The stimuli contained all the possible disyllabic syllable types in the language such as VCV (e.g., ọmi), VCṼ (e.g., ọrin), CVCV (e.g., jabo), CṼCV (e.g., finju), CVCṼ (e.g., wodan) and CṼCṼ (e.g., sanmun) syllable types. Each nonce word was recorded once.

Table 5. 1 Disyllabic Stimuli for the Identification Experiment 2

LH	HL	MH	ML	HH	LL
àpá	jábò	jẹ̄tí	ọ̄mì	múdǐ	bìlù
ìlú	kấgbò	ōsá	rẹ̄dà	gbéjú	tìkọ̀
ìkọ́	túdè	sǐ̄yẹ́	kātè	fíjú	ọ̀kọ̀
òlé	wódầ	dīsẹ́	ārè	wéjẹ́	ìlù
sàtọ́	kẹ́jà	pọ̄jọ́	jẹ̄bà	bísé	dùrù
bórí	rújẹ̀	bọ̄rẹ́	bōlò	fádí	ọ̀rǐ̀
bẹ̀mú	mídò	rādé	wērǔ̀	Jólé	ọ̀kầ
dọ̀kú	fálè	īkǔ	kēdǔ̀	gúré	sầmǔ̀

5.1.3 Procedure

All participant groups took part in this experiment. The experiment followed the same procedure as stated in Identification 1 above. Participants listened to the stimuli one at a time and then identified the last syllable as having a High, Mid, or Low tone, by pressing either the H, M, or L key on the keyboard. As in the previous experiments, a short practice session, consisting of the six tonal patterns, preceded the actual experiment. The experimental software randomly presented the stimuli. The stimuli for the practice session are presented in Appendix D1.

5.2 Results

The general results presented in Table 5.2 show an improved performance by most participant groups compared to the previous Identification Experiment. The lowest percent accuracy was 32% (EngCanYg) while the highest percent accuracy was 80% (YorNgrAd). In each language context category, the adult participants scored higher than the younger ones. As in Identification 1, the better-than-chance (BTC) score for this experiment was 21 out of 48. The last two columns show the number and the percentages of participants with BTC scores. The Yorùbá-Nigerian and English-Canadian adults recorded the highest percent scores with 82% and 96% respectively, while the English-Canadian young participants recorded the lowest percent score with 18%. The

Yorùbá-Nigerian young participants scored 67% while the Yorùbá-Canadian adult and young groups respectively had 54.5% and 61.5% of their participants with BTC scores.

Table 5. 2 Participants' percent accuracy in the Identification Experiment 2

	ACC	**%**	**MEAN**	**SD**	***N***	**BTC**	**%BTC**
YorNgrAd	653/816	80%	38.41	13.27	*17*	14	82%
YorNgrYg	844/1296	65%	31.26	15.09	*27*	18	67%
YorCanAd	288/528	55%	26.18	14.22	*11*	6	54.5%
YorCanYg	279/624	45%	21.46	10.99	*13*	8	61.5%
EngCanAd	815/1296	63%	30.19	4.75	*27*	26	96%
EngCanYg	336/1056	32%	15.27	5.71	*22*	4	18%

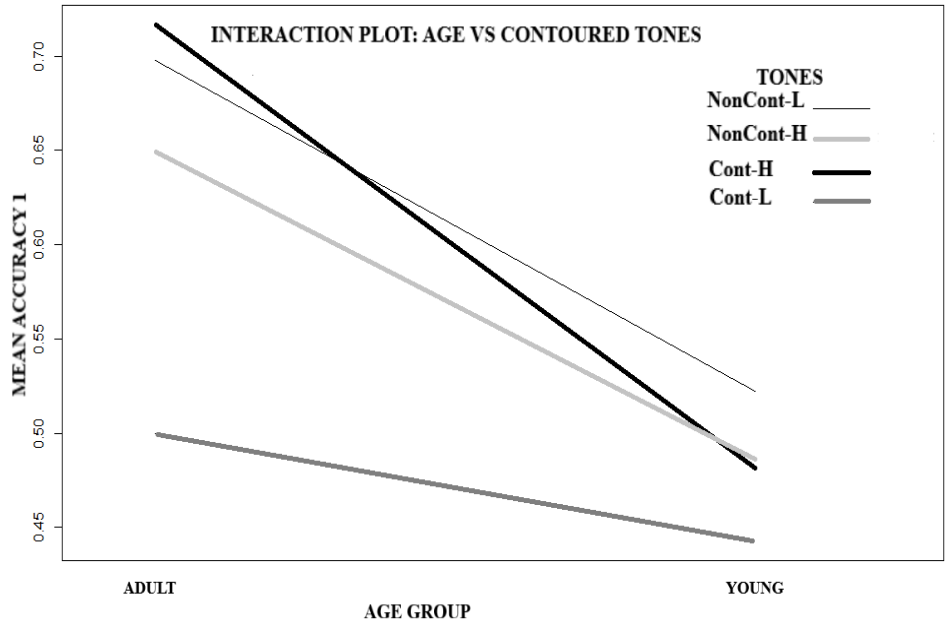
5.2.1 Statistical analysis

Lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015) was used to perform a generalized linear mixed effects regression analysis (glmer) to examine the relationship between participants' tone identification accuracy and three fixed effects, Age, Language Context, as well as the type of tonal contour. In the models fitted, Age had two levels, *Adult* and *Young*, while Language Context had three levels, *Yorùbá-Nigerian, Yorùbá-Canadian*, and *English-Canadian*. The third fixed effect, Contour, had four levels: *Contour Low* tone*, Contour High* tone*, Non-Contour Low* tone, and *Non-Contour High* tone. The model also had intercepts for the two random effects, *item* and *participant*. The accuracy score, which consisted of the individual score for the 48 target stimuli, was the dependent variable. It was coded as either '1' or '0' with '1' indicating an accurate or correct score while '0' indicated inaccurate or incorrect score. The model was based on 5616 observations (a total of 117 participants x 48 individual item score=5616).

P-values were obtained by likelihood ratio tests of both a main-effects model and another model with the interactions between contour and each of the other two main effects, Age, and Language Context. An ANOVA revealed that the model with interaction terms between contour and each

of the other two predictors (AIC(interaction) = 5957.2) fitted the model more closely than the model without interaction terms (AIC = 5985.5). The introduction of by-participant and by-item random slopes did not improve the model's fit. The outputs of the model are summarized in Tables 5.3 to 5.5 below. Meanwhile, before moving on to the tables, a visualization of the interaction plot between Age and the four contour levels is presented in Figure 5.1.

Figure 5.1 Mean identification accuracy of the contour levels by the two Age groups (adult, young): Cont-H=contour High; Cont-L=contour Low; NonCont-H=non-contour High; NonCont-L = non-contour Low



The plot reveals some interactions between accurate identification of the contour levels and participants' age groups. The adult and young groups identified contour Low worse than the other three contour levels with the adult group showing more significant effects than the young group. Significant interactions were further observed between the contour Low and the other three contour levels. These interactions were more pronounced for the adult age groups more than for the younger age groups. The parallel lines for non-contour Low and non-contour High indicate that there was no significant Age group effect in the identification of these two contour types. The next plot in Figure 5.2 further presents the interaction between Language Context and accurate identification of the contour items.

Figure 5.2 Mean identification accuracy of the contour levels along the three Language Contexts. Cont-H=contour High; Cont-L=contour Low; NonCont-H=non-contour High; NonCont-L = non-contour Low

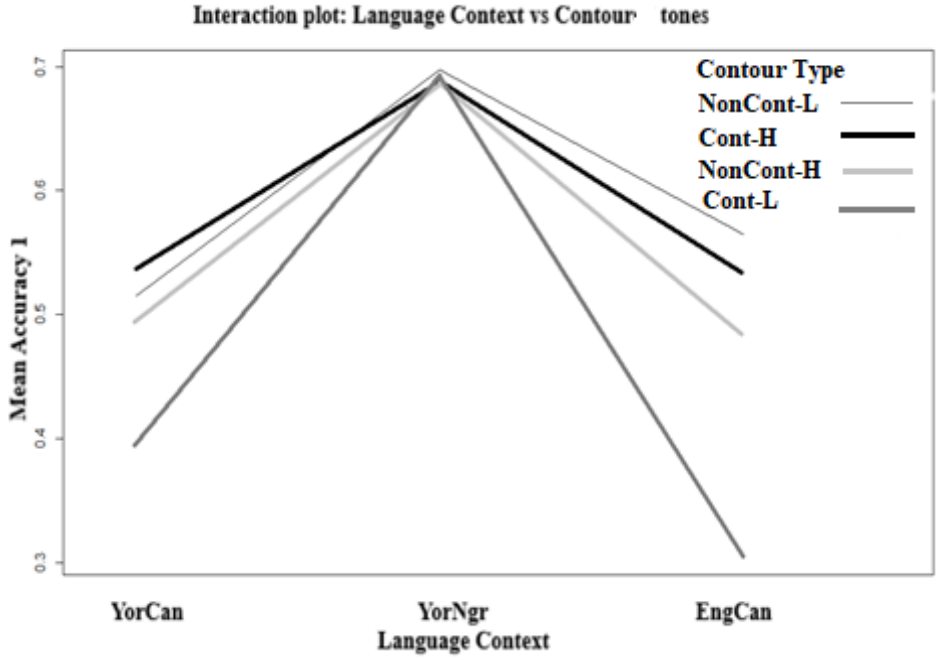


Figure 5.2 reveals that the Yorùbá-Nigerian participants identified all the four contour types more accurately than either the Yorùbá-Canadian or the English-Canadian category. The identification accuracy of the four contour levels did not indicate a significant difference like what was observed for the other two Language Context types. Both the Yorùbá-Canadians and the English-Canadians showed a significant effect for the identification of the contour Low compared to the other three levels of the contour factor. The English-Canadians identified the contour Low less accurately than the other contour levels, especially the non-contour Low and the contour High. While the Yorùbá-Canadians also recorded significant effects in identifying contour Low, there was no significant effect in their identification accuracy of non-contour Low and non-contour High when compared to the other contour levels. The two plots in Figures 5.1 and 5.2 have revealed that generally, participants exhibited tone identification discrepancy along the factors of Age and Language Context. To highlight which of these factors are significant, this observation is explored further in the statistical analysis presented in Table 5.3 which summarizes the linear mixed-effects model of the effects of Age, Language Context, and Contour type on participants' tone identification accuracy.

Table 5. 3 Coefficients of the glmer model of the effects of the interactions between contour type, Age, and Language Context. Random effects: participant and item. N=5616.

Ref: YorNgr; Contour-High tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	2.3899	0.3796	6.296	3.06e-10***
Cont-L	-1.28178	0.36863	-3.477	0.000507***
NonCont-H	-0.2813	0.3092	-0.910	0.362900
NonCont-L	-0.1596	0.3098	-0.515	0.606448
Young	-1.4934	0.3194	-4.676	2.92e-06***
YorCan	-1.2866	0.4346	-2.960	0.003075**
EngCan	-1.5690	0.3635	-4.317	1.58e-05***
Cont-L x Young	0.8454	0.2372	3.564	0.000365***
NonCont-H x Young	0.4137	0.2019	2.049	0.040489*
NonCont-L x Young	0.3616	0.2027	1.784	0.074405.
Cont-L x YorCan	-0.7546	0.3236	-2.332	0.019717*
NonCont-H x YorCan	-0.1909	0.2784	-0.686	0.492809
NonCont-L x YorCan	-0.1678	0.2791	-0.601	0.547626
Cont-L x EngCan	-0.9607	0.2759	-3.482	0.000498***
NonCont-H x EngCan	-0.1525	0.2360	-0.646	0.518325
NonCont-L x EngCan	0.1635	0.2364	0.691	0.489257
Ref.: YorCan; Contour High tone				
Intercept	1.10333	0.41929	2.631	0.008503**
Cont-L	-0.5272	0.3572	-1.476	0.139984
NonCont-H	-0.47226	0.31596	-1.495	0.134998
NonCont-L	-0.32743	0.31644	-1.035	0.300788
EngCan	-0.28244	0.41475	-0.681	0.495880
Cont-L x EngCan	-0.20609	0.29719	-0.693	0.488014
NonCont-H x EngCan	0.03846	0.25215	0.153	0.878770
NonCont-L x EngCan	0.33131	0.25280	1.311	0.190021

The analysis suggests that Age and Language Context had significant effects on participants' identification accuracy. The analysis further suggests that only one of the four levels of contour, contour Low (CL) showed a significant effect on participants' accuracy, as it was identified significantly worse than the contour High. Meanwhile, there was no significant difference between contour High and the other two levels, non-contour High, and non-contour Low. Moreover, as revealed in Table 5.3 and Figure 5.1 above, younger participants performed significantly worse than adult participants in identifying the other tones versus the contour High tone. The model further suggests that globally, the Yorùbá-Nigerian participants performed better than either the Yorùbá-Canadian or the English-Canadian participants in identifying the

other levels of the contour tones versus the contour High tone (see also Figure 5.2). With Yorùbá-Canadian as the reference level, the model suggests that there was no significant difference between Yorùbá-Canadian and English-Canadian participants in the identification of the contour tones.

The second part of the analysis examines the effects of the interactions between the contour factor and the other two predictors, Age and Language Context. With contour High as the default reference level in Table 5.3, the analysis reveals that, in comparison to the adults, younger participants tended to identify contour Low and non-contour High tones significantly better than the default contour High tone, while there was no significant difference in their identification of contour High versus non-contour Low. As for the interaction between contour type and Language Context, the analysis suggests that compared to the Yorùbá-Nigerian participants, contour Low tones appeared to pose more difficulty for the Yorùbá-Canadian and English-Canadian participants than the contour High tones. These two Language Context categories identified the contour Low tones significantly less accurately than the contour High tones. In contrast, there was no significant difference in their accurate identification of contour High tones and the other two contour tone levels, non-contour High and non-contour Low.

When contour Low was indicated as the reference level in Table 5.4, the analysis reveals that, overall, there was no significant difference in the identification rates of contour Low compared to the other three contour levels (contour-H, nonCont-H, and nonCont-L), even though the positive coefficient estimates pointed to better identification for these three levels. Moreover, the younger participants misidentified all the other three levels more often than contour Low compared to the adults. Finally, the Yorùbá-Canadian and English-Canadian participants identified the both non-contour High and non-contour Low significantly better than the contour-Low when compared to the Yorùbá-Nigerian category.

Table 5. 4 Coefficients of the glmer model of the effects of the interactions between contour type, Age, and Language Context with contour Low as the reference level

Ref: YorNgr; Contour-Low tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.8627	0.3774	4.935	8.00e-07 ***
NonCont-H	0.2458	0.3080	0.798	0.424703
NonCont-L	0.3676	0.3087	1.191	0.233813
Young	-0.6480	0.3203	-2.023	0.043080 *
YorCan	-2.0411	0.4356	-4.686	2.79e-06 ***
EngCan	-2.5297	0.3650	-6.931	4.18e-12 ***
NonCont-H x Young	-0.4316	0.2045	-2.111	0.034802 *
NonCont-L x Young	-0.4837	0.2056	-2.353	0.018618 *
NonCont-H x YorCan	0.5637	0.2799	2.014	0.044003 *
NonCont-L x YorCan	0.5868	0.2808	2.090	0.036639 *
NonCont-H x EngCan	0.8082	0.2384	3.390	0.000698 ***
NonCont-L x EngCan	1.1242	0.2391	4.702	2.57e-06 ***

5.2.2 Tone Identification by participant groups

Figure 5. 3 Mean accuracy for the identification of contour levels by the six participant groups: Cont-H=contour High; Cont-L=contour Low; NonCont-H=non-contour High; NonCont-L = non-contour Low

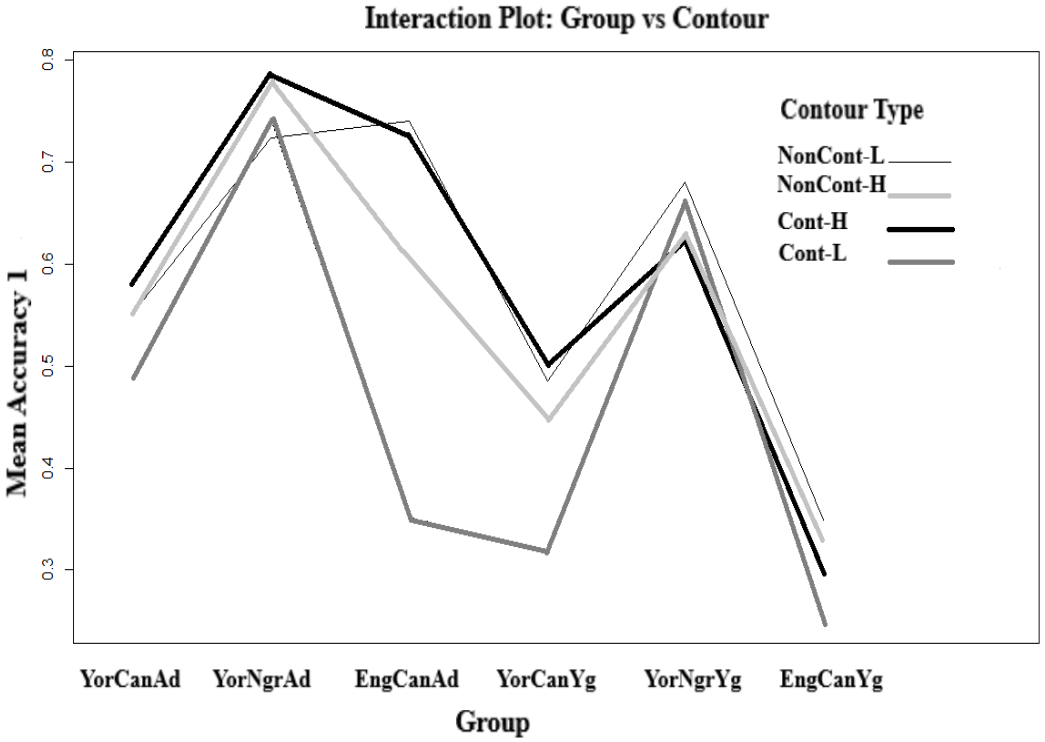


Figure 5.3 shows the identification of contour tone for the six participant groups. The plot reveals that contour Low was the worst identified tone by four of the six participant groups, namely the English-Canadian and Yorùbá-Canadian young and adult groups. The plots further revealed that unlike the other groups, the English-Canadian adult and the Yorùbá-Canadian young groups posted wider distances between the points for the four contour levels, indicating significant effects between their accurate identification of contour Low and all the other three contour levels. This observation will be explored further through statistical analysis to highlight which of the factors are significant with respect to the participant groups.

A model with Contour and Group as the two factors was fitted to compare pairwise the performance of the participant groups with one another. Contour had four levels, *Contour High, Contour Low, Non-Contour Low*, and *Non-Contour High*. As before, Group had six levels: *Yorùbá-Nigerian Adult, Yorùbá-Nigerian Young, Yorùbá-Canadian Adult, Yorùbá-Canadian Young, English-Canadian Adult,* and *English-Canadian Young*. The model had intercepts for the two random effects, namely *item* and *participant*. The dependent variable was the accuracy score and consisted of the individual score for the 48 target stimuli. The model was based on 5616 observations (i.e., a total of 117 participants x 48 individual item score=5616). To undertake a pairwise comparison of the groups, the model specified different reference levels for Group. The outputs of the model with different group reference levels are presented in Table 5.5.

Table 5. 5 Coefficients of the glmer model of the comparative performance in tone identification by participant groups of the same or different Age and Language Context. Random Effects: participant and item. N=6672.

Ref: Yorùbá-Nigerian Adult; Contour High tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	2.14018	0.41765	5.124	2.99e-07***
Cont-L	-0.72794	0.26237	-2.774	0.00553**
NonCont-H	-0.17484	0.22688	-0.771	0.44093
NonCont-L	0.08173	0.22706	0.360	0.71889
YorCanAd	-1.59544	0.58159	-2.743	0.00608**
EngCanAd	-1.39167	0.46804	-2.973	0.00295**
YorCanYg	-2.27426	0.55198	-4.120	3.79e-05***
YorNgrYg	-0.82127	0.47218	-1.739	0.08198 .
EngCanYg	-2.87268	0.48767	-5.891	3.85e-09***
Ref: Yorùbá-Canadian Adult				
Intercept	0.54474	0.48055	1.134	0.25698
EngCanAd	0.20377	0.52512	0.388	0.69798
YorCanYg	-0.67882	0.60100	-1.129	0.25870
YorNgrYg	0.77418	0.52998	1.461	0.14408
EngCanYg	-1.27723	0.54238	-2.355	0.01853*
Ref: English-Canadian Adult				
Intercept	0.74851	0.33403	2.241	0.025036*
YorCanYg	-0.88260	0.49183	-1.795	0.072733.
YorNgrYg	0.57040	0.40251	1.417	0.156447
EngCanYg	-1.48101	0.41821	-3.541	0.000398***
Ref: Yorùbá-Nigerian Young				
Intercept	1.31891	0.34221	3.854	0.000116***
YorCanYg	-1.45298	0.49732	-2.922	0.003482**
EngCanYg	-2.05139	0.42484	-4.829	1.37e-06***
Ref: Yorùbá-Canadian Young				
Intercept	-0.13409	0.44379	-0.302	0.76255
EngCanYg	-0.59841	0.51023	-1.173	0.24086

The table shows comparisons between the two Age groups in each Language Context category as well as comparisons between participants who are in the same age group but different Language Context. The table reveals that while there was no significant difference in the identification rate between adult and young participants from Yorùbá-Nigerian Language Context, the Yorùbá-Nigerian adults identified the tones significantly better than the young groups from the other two Language Contexts. With Yorùbá-Canadian adults as the reference level, there was no significant

difference between this adult group and their younger counterparts, and the young Yorùbá-Nigerian group. However, this adult group identified the tones significantly better than the English-Canadian young group. With English-Canadian adult as the reference level, this adult group was the only group to identify the tones significantly better than their younger counterparts. However, there was no significant difference between this group and the younger groups from both the Yorùbá-Nigerian and Yorùbá-Canadian categories.

The second comparison examines participants' performance across Language Context types. It considers participants of the same age group (adult vs. adult; young vs. young) but with different Language Context. In the adult groups, the analysis indicates that the Yorùbá-Nigerian adult group was significantly better in this task than either the Yorùbá-Canadian or English-Canadian adults. Meanwhile, the analysis does not indicate any significant difference in the accurate identification rate between the Yorùbá-Canadian and the English-Canadian adults.

In the younger age groups, a similar trend was observed. The young Yorùbá-Nigerian participants posted a significantly higher accurate identification rate than either the Yorùbá-Canadian or the English-Canadian younger ones. Meanwhile, there was no significant difference in accurate identification rates between the Yorùbá-Canadian and the English-Canadian young participants.

5.2.3 Identification of High vs. Low tones

With the stimuli ending in either a High or Low tone, this experiment provides a suitable environment to further investigate participants' accurate identification rate for the two tones. It should be noted that the High tones in this experiment contained simple High tones (preceded by either Mid or High tone in a disyllabic context, e.g., M.**H**; H.**H**) and complex High tones (preceded only by Low tone, i.e., L.**H**). Similarly, the Low tones also contained both the simple Low tones (preceded by either Mid or Low tone as in M.**L** and L.**L)** as well as complex Low tones (preceded only by High tone, i.e., H.**L**).

Figure 5.4 Mean identification accuracy for High versus Low tones by the six participant groups.

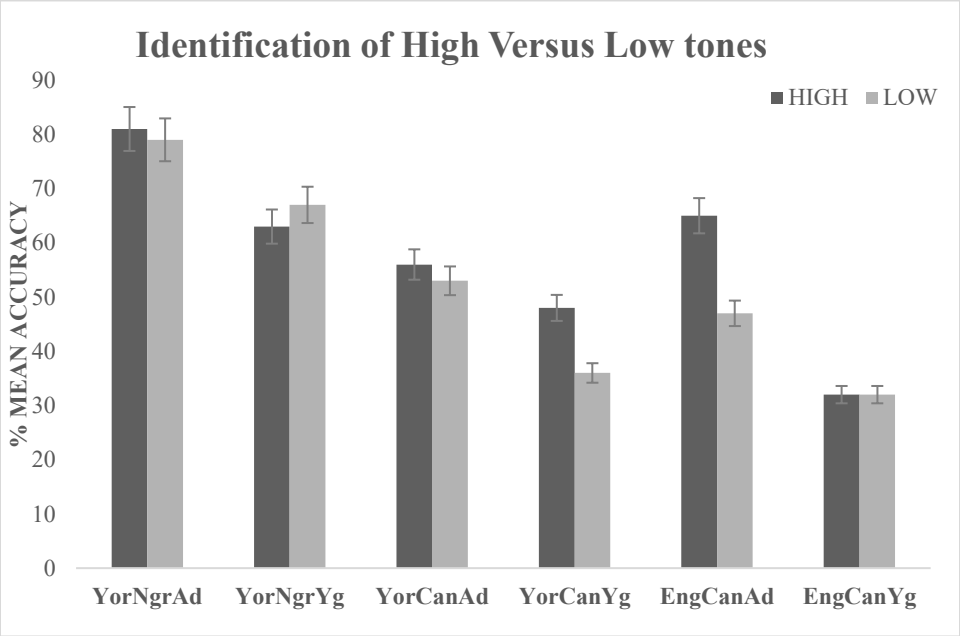


Figure 5.4 reveals that the Yorùbá-Nigerian adults performed better than the other groups in the identification of the High and the Low tones with 81% and 79% respectively. The Yorùbá-Nigerian young and the English-Canadian adult groups also recorded comparatively higher rates in the identification of the two tones than the other groups. Two of the young participant groups, Yorùbá-Canadian and English-Canadian, recorded the lowest identification rates for High and Low tones. The figure also reveals that the trend for the other groups was either to identify High tones slightly better than Low tones, or to record equal rate of identification for the two tones.

To investigate the statistical significance of what is observed in Figure 5.4, a model was fitted to investigate participant groups' identification of High and Low tones. The model included two fixed effects, namely Tone, with two levels, *High* and *Low*, and Group, with six levels, with the interaction term between the two factors. As in the previous models, this model had intercepts for the two random-effect factors, *item*, and *participant*. The dependent variable was the accuracy score that consisted of the individual score for the 48 target stimuli. The model was based on 5616 observations (a total of 117 participants x 48 individual item score=5616).

P-values were obtained by likelihood ratio tests of the model with the interaction term between the two main effects but without random slopes, and another model with by-participant random slopes for the effect of Tone. An ANOVA revealed that the there was no significant difference

between the two models (AIC = 5949.6 vs. AIC (random slope) = 5949.4). Thus, the addition of by-participant random slopes for Tone effect on accuracy did not improve the model's fit. The output of the model is summarized in Table 5.6.

Table 5. 6 Coefficients of the glmer model of the identification of High versus Low tones by participant groups. Random Effects: participant and item. N=5616

Ref: Yorùbá-Nigerian Adult; High tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	2.090259	0.403046	5.186	2.15e-07***
Lowtone	-0.114491	0.365076	-0.314	0.75382
YorCanAd	-1.635991	0.591090	-2.768	0.00564**
EngCanAd	-1.409103	0.477133	-2.953	0.00314**
YorCanYg	-2.299069	0.561918	-4.091	4.29e-05***
YorNgrYg	-1.150438	0.481560	-2.389	0.01690*
EngCanYg	-2.974231	0.497527	-5.978	2.26e-09***
LowTone x YorCanAd	0.008832	0.459923	0.019	0.98468
LowTone x EngCanAd	-0.016743	0.379530	-0.044	0.96481
LowTone x YorCanYg	-0.041246	0.437747	-0.094	0.92493
LowTone x YorNgrYg	0.680080	0.388363	1.751	0.07992.
LowTone x EngCanYg	0.063819	0.396998	0.161	0.87229

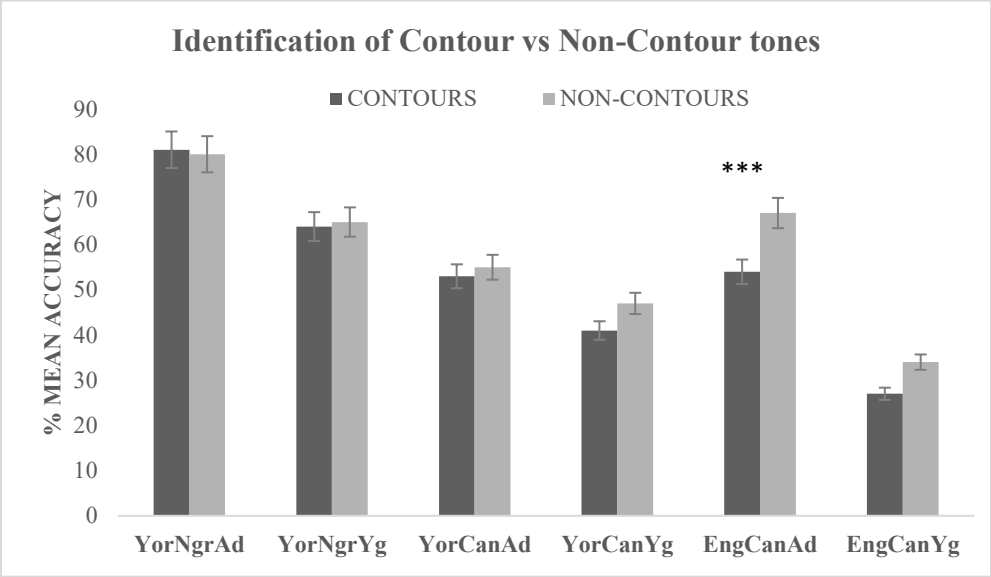
Table 5.6 suggests that there was no significant effect of Tone in this experiment. It further suggests that all the remaining five groups were significantly different (worse) than Yorùbá-Nigerian adult group in the identification of the High tone. The Yorùbá-Nigerian young participants were the only group that recorded some significance in their identification of Low tones compared to High tones when the Yorùbá-Nigerian adult was the group reference level. The other groups did not record any significant differences.

5.2.4. Identification of the contour tones by participant groups

This section investigates the identification of all-contour and all-non-contour stimuli in this experiment. The percent identification accuracy by the six participant groups is presented in the graphs in Figure 5.5. The figure shows that the Yorùbá-Nigerian adults and their younger counterparts performed better than the other groups. Moreover, the figure reveals no significant difference in the identification of the two different tones by five of the six participant groups. The only group to record significant effect in the identification of all-contour tones versus all

non-contour tones was the English-Canadian adult group (All contour=54%; all non-contour=67%). The significance stars indicate the group's identification of contour and non-contour tones when compared to the Yorùbá-Nigerian adults.

Figure 5. 5 Mean identification accuracy for all-contour and all-non-contour tones by the six participant groups



To further examine what was revealed by the plot, a statistical model was fitted to analyze the individual group's identification rates for the contour and non-contour stimuli in this experiment. The model had Contour as a fixed effect with four levels: *Contour High, Contour Low, Non-Contour High*, and *Non-Contour Low*. As in the previous models, it also had *Group* as a fixed effect with six levels (*YorCanAd; YorNgrAd; EngCanAd; YorCanYg; YorNgrYg,* and *EngCanYg*). An interaction term was introduced between these two main effects. As random-effect factors, the model included *item* and *participant* and had intercepts for them. The dependent variable was the accuracy score which consisted of the individual score for the 48 target stimuli. The model was based on 5616 observations (a total of 117 participants x 48 individual item score=5616).

P-values were obtained by likelihood ratio tests of a model with only the interaction term and another model with the interaction term as well as by-participant random slopes for the effect of Tone. An ANOVA revealed that the model with by-participant random slopes for the effect of

Tone (AIC=5885.1) fitted the data more closely than the model with the only the interaction terms (AIC = 5942.2). However, another model with by-participant as well as by-item random slopes for the effect of Tone did not converge. The output for the best fitting model is presented Tables 5.7 to 5.9. Yorùbá-Nigerian adult was the chosen reference level for the group factor in all the summaries.

Table 5. 7 Coefficients of the glmer model of the identification of different contour tones by participant groups. Random Effects: participant and item. N=5616. YG = young; AD=adult Cont=contour; NonCont=non-contour; H=High tone; L=Low tone

Ref: Yorùbá-Nigerian Adult; Contour-High				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	2.60853	0.59374	4.393	1.12e-05 ***
Cont-L	-0.54064	0.56928	-0.950	0.34227
NonCont-H	-0.66539	0.48743	-1.365	0.17222
NonCont-L	-0.62828	0.53056	-1.184	0.23634
YorCanAd	-1.89719	0.82980	-2.286	0.02224 *
EngCanAd	-1.43040	0.67837	-2.109	0.03498 *
YorCanYg	-2.66415	0.79213	-3.363	0.00077 ***
YorNgrYg	-1.46219	0.68150	-2.146	0.03191 *
EngCanYg	-3.68159	0.71049	-5.182	2.20e-07 ***
Cont-L x YorCanAd	-0.12080	0.68210	-0.177	0.85943
NonCont-H x YorCanAd	0.32093	0.58555	0.548	0.58363
NonCont-L x YorCanAd	0.40564	0.65115	0.623	0.53331
Cont-L x EngCanAd	**-1.37148**	**0.57738**	**-2.375**	**0.01753 ***
NonCont-H x EngCanAd	-0.02635	0.49148	-0.054	0.95725
NonCont-L x EngCanAd	0.75800	0.54778	1.384	0.16643
Cont-L x YorCanYg	-0.39962	0.66105	-0.605	0.54549
NonCont-H x YorCanYg	0.43192	0.56591	0.763	0.44532
NonCont-L x YorCanYg	0.62343	0.62433	0.999	0.31801
Cont-L x YorNgrYg	0.73670	0.57080	1.291	0.19682
NonCont-H x YorNgrYg	0.40320	0.49346	0.817	0.41388
NonCont-L x YorNgrYg	**1.10423**	**0.55336**	**1.996**	**0.04599 ***
Cont-L x EngCanYg	0.34656	0.60586	0.572	0.56731
NonCont-H x EngCanYg	0.92227	0.51921	1.776	0.07569 .
NonCont-L x EngCanYg	0.92878	0.57402	1.618	0.10565

Table 5.7 presents the summary of the model where Yorùbá-Nigerian adult was the reference level for the group factor while contour High was the reference level for the contour factor. The first part of the analysis compared the three contour levels to contour High. The results suggest that contour Low, non-contour High, as well as non-contour Low, were all identified at worse rates than contour High, although the difference was not statistically significant. In the group

comparison section, the summary suggests that all the other five participant groups were significantly worse than Yorùbá-Nigerian adults in the identification of contour High Tone. In the last section of the analysis, where each group was compared to the Yorùbá-Nigerian adults in their identification of contour High, the analysis reveals that the English-Canadian adult group identified contour Low at a significantly worse rate than contour High. It also reveals that the Yorùbá-Nigerian young group identified non-contour Low tones significantly better than contour High, while the English-Canadian young group leaned towards identifying non-contour High tones better than contour High. The next paragraph compares contour Low to the other contour levels as shown in Table 5.8.

Table 5. 8 Coefficients of the glmer model of the identification of different contour tones by participant groups with contour-Low tone (Cont-L) as the reference level

Ref: Yorùbá-Nigerian Adult; Contour-Low tone				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	2.06791	0.54990	3.760	0.00017***
NonCont-H	-0.12477	0.50652	-0.246	0.80543
NonCont-L	-0.08768	0.48475	-0.181	0.85647
YorCanAd	-2.01800	0.76583	-2.635	0.00841 **
EngCanAd	-2.80191	0.62837	-4.459	8.23e-06 ***
YorCanYg	-3.06382	0.73404	-4.174	2.99e-05 ***
YorNgrYg	-0.72550	0.62691	-1.157	0.24716
EngCanYg	-3.33505	0.65406	-5.099	3.41e-07 ***
NonCont-H x YorCanAd	0.44176	0.63607	0.695	0.48736
NonCont-L x YorCanAd	0.52648	0.58234	0.904	0.36595
NonCont-H x EngCanAd	**1.34515**	**0.53089**	**2.534**	**0.01128 ***
NonCont-L x EngCanAd	**2.12951**	**0.49449**	**4.306**	**1.66e-05 *****
NonCont-H x YorCanYg	0.83158	0.61338	1.356	0.17518
NonCont-L x YorCanYg	1.02310	0.56254	1.819	0.06896 .
NonCont-H x YorNgrYg	-0.33348	0.53022	-0.629	0.52939
NonCont-L x YorNgrYg	0.36756	0.49031	0.750	0.45346
NonCont-H x EngCanYg	0.57573	0.55312	1.041	0.29794
NonCont-L x EngCanYg	0.58226	0.51264	1.136	0.25604

To compare contour Low with the other levels, the model summary in Table 5.8 specified contour Low as the reference level. The analysis suggests that overall, contour High was better identified than contour Low, while both non-contour High and non-contour Low attracted worse identification rates than contour Low, although all these were without statistical significance. The group comparison section suggests that, except for the Yorùbá-Nigerian young group, all the

other groups significantly identified contour Low less accurately than the Yorùbá-Nigerian adult group. In the final part of the analysis, the English-Canadian adult group (bolded) was the only group to identify non-contour High, and non-contour Low significantly better than contour Low when compared to the reference level (YorNgrAd). All the remaining groups did not record any significant difference in their identification of contour Low versus other contour levels when compared to the Yorùbá-Nigerian adult group. To compare the other levels of the contour factor, non-contour Low was specified as the reference level in Table 5.9.

Table 5. 9 Coefficients of the glmer model of the identification of different contour tones by participant groups with non-contour High as the reference level

Ref: Yorùbá-Nigerian Adult; Non-Contour High				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	1.943139	0.396779	4.897	9.72e-07***
Cont-L	0.124771	0.506526	0.246	0.80543
NonCont-L	0.037112	0.434081	0.085	0.93187
YorCanAd	-1.576247	0.570943	-2.761	0.00577**
EngCanAd	-1.456753	0.461531	-3.156	0.00160**
YorCanYg	-2.232231	0.544085	-4.103	4.08e-05***
YorNgrYg	-1.058978	0.466382	-2.271	0.02317*
EngCanYg	-2.759330	0.482014	-5.725	1.04e-08***
Cont-L x YorCanAd	-0.441764	0.636378	-0.694	0.48757
NonCont-L x YorCanAd	0.084700	0.558448	0.152	0.87945
Cont-L x EngCanAd	**-1.345161**	**0.531144**	**-2.533**	**0.01132***
NonCont-L x EngCanAd	0.784336	0.462788	1.695	0.09011.
Cont-L x YorCanYg	-0.831584	0.613655	-1.355	0.17538
NonCont-L x YorCanYg	0.191486	0.532071	0.360	0.71893
Cont-L x YorNgrYg	0.333481	0.530204	0.629	0.52937
NonCont-L x YorNgrYg	0.701026	0.472602	1.483	0.13799
Cont-L x EngCanYg	-0.575739	0.553243	-1.041	0.29803
NonCont-L x EngCanYg	0.006504	0.481092	0.014	0.98921

Table 5.9 indicates that the other three levels were all identified better than non-contour High, although the differences were not statistically significant. Meanwhile, the group comparison suggests that all the other groups identified non-contour High significantly worse than the Yorùbá-Nigerian adult group. The last part of the analysis further suggests that only the English-Canadian adult group (bolded) accurately identified contour Low significantly less than the reference-level non-contour High compared to the Yorùbá-Nigerian adult group.

To sum up, the analysis reveals that overall, there was no significant difference in the identification of all the four levels of the contour factor (contour High, contour Low, non-contour High, and non-contour Low). The analysis further suggests that Yorùbá-Nigerian adults, as the group reference level, performed significantly better than all the other groups in all the contour level comparisons. However, there was no significant difference between this group and their younger counterparts in the identification of contour Low (also see Table 5.8). Finally, the analysis highlights the identification patterns of the English-Canadian adult group with respect to the contour-Low tone when compared to the reference level. This group consistently identified the contour Low tone less accurately than each of the other three levels of the contour factor.

5.3 Misidentification of High and Low Tones by individual groups

In this section of the analysis, participants' responses were extracted and analyzed with the objectives of revealing participants' overall misidentification pattern for the contour and the non-contour tones in this experiment. Additionally, it will reveal the tones that were mostly misidentified, and which tones they were commonly mistaken for.

Figure 5. 6 Misidentification of High tones as Mid or Low by the six participant groups.

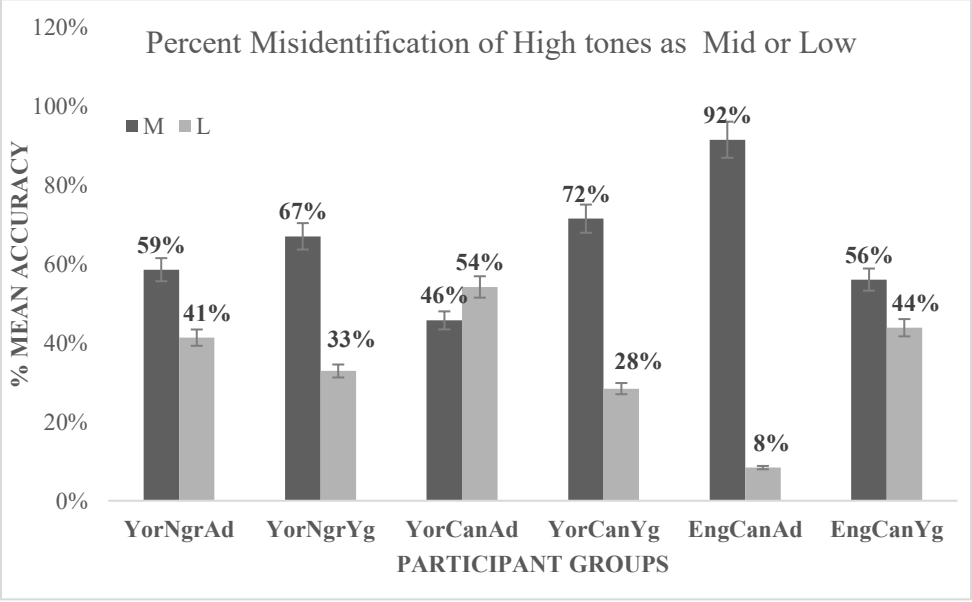


Figure 5.6 represents the data in the last two rows of Table 5.10 which presents the total and the percent rates of High-tone misidentification by each of the six participant groups. It indicates that except for the Yorùbá-Canadian adults, all the other groups misidentified High tone as Mid with percent ranging from 56% to 92%. The figure further reveals that two groups, namely English-Canadian adults, as well as Yorùbá-Canadian young group, led the participants in the perception of High tones as Mid, with 92% and 72% misidentification rates respectively. They also recorded the lowest percentages of misidentification of High tone as Low (YorCanYg, 28%; EngCanAd, 8%). The remaining three groups, the Yorùbá-Nigerian adults, Yorùbá-Nigerian young group, and English-Canadian young group, misidentified High tones as Mid with percent rates of 59%, 67%, and 56% respectively.

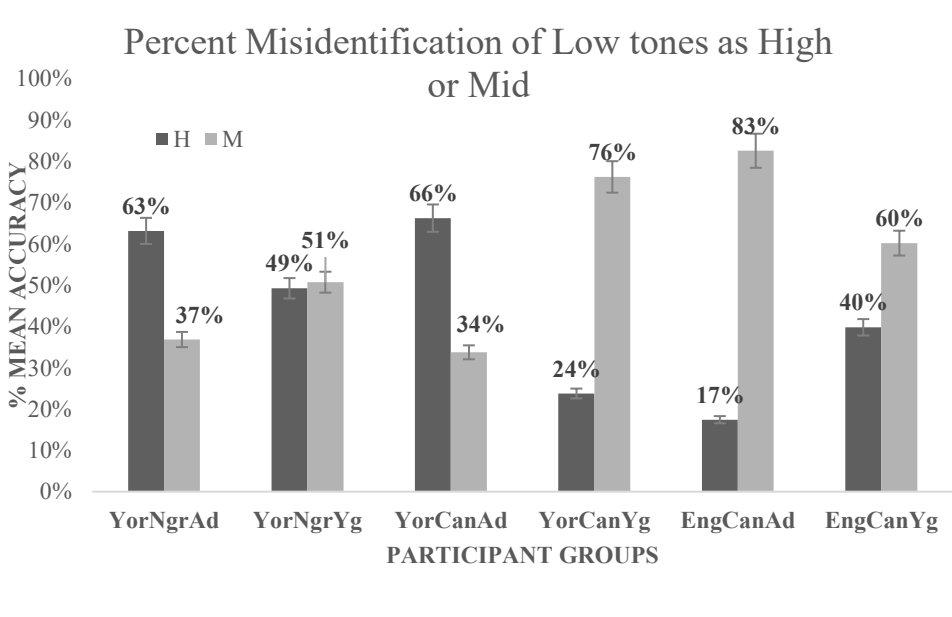
Table 5. 10 High-tone misidentification in MH, HH, and LH patterns

P/GR	**YorNgrAd**		**YorNgrYg**		**YorCanAd**		**YorCanYg**		**EngCanAd**		**EngCanYg**	
As→	MID	LOW	MID	LOW	MID	LOW	MID	LOW	MID	LOW	MID	LOW
MH	19	7	36	11	7	11	31	6	66	5	55	41
	(73%)	(27%)	(77%)	(23%)	(39%)	(61%)	(84%)	(16%)	(93%)	(7%)	(57%)	(43%)
HH	10	5	41	17	12	10	26	11	89	7	53	52
	(67%)	(33%)	(71%)	(29%)	(55%)	(45%)	(70%)	(30%)	(93%)	(7%)	(50%)	(50%)
LH	5	12	31	25	8	11	21	14	52	7	62	40
	(29%)	(71%)	(55%)	(45%)	(42%)	(58%)	(60%)	(40%)	(88%)	(12%)	(61%)	(39%)
TOT	34	24	108	53	27	32	78	31	207	19	170	133
	(59%)	(41%)	(67%)	(33%)	(46%)	(54%)	(72%)	(28%)	(92%)	(8%)	(56%)	(44%)

Table 5.10 further presents the direction of the misidentification observed for the six tonal patterns, by each group. It reveals the rate at which the last tone in each pattern was misidentified as other tones. This analysis seeks to highlight which of the three basic tones attracts the most misrepresentations. The table reveals that in addition to the English-Canadian Adults, all the three young participant groups misidentified High tones mostly as Mid in all the three tonal patterns. Out of the remaining two groups, the Yorùbá-Nigerian adults misidentified High tones mostly as Mid in the MH and HH patterns, and mostly as Low in LH pattern. The Yorùbá-Canadian adults, on the other hand, misidentified the High tones as Low in MH and LH patterns and as Mid in HH pattern. Thus, the High tone was misidentified mostly as Mid in HH pattern

universally by all groups, while the High tone in MH pattern was misidentified mostly as Mid by all groups except the Yorùbá-Canadian adult group. As for the LH pattern, except for the Yorùbá-Nigerian and Yorùbá-Canadian adult groups, all the other groups misidentified the High tone there primarily as Mid. In sum, Mid tone recorded higher rates (percent) than the Low tone in 14 out of 18 possible instances, emerging as the tone that the majority of the participant groups misidentified High tones as. The Low-tone misidentification data is presented in the last two rows of Table 5.11 and the percent rate of misidentification is represented by Figure 5.7.

Figure 5. 7 Misidentification of Low tones as High or Mid by the six participant groups



Three groups - English-Canadian adults, Yorùbá-Canadian young group, and English-Canadian young group - misidentified Low tones as Mid at the rates of 83%, 76%, and 60% respectively. For two groups, the Yorùbá-Canadian adults, and Yorùbá-Nigerian adults, about two-thirds of their Low-tone misidentification (66% and 63% respectively) were misrepresented as High. The only remaining group, namely the Yorùbá-Nigerian young group, misidentified Low tones almost equally as High or as Mid tone (49% and 51% respectively). Thus, Low tones were misidentified mostly as Mid by three groups. Two groups misidentified them mostly as High tones, while one group oscillated between identifying them as either Mid or High tone.

Table 5. 11 L-tone misidentification in ML, LL, and HL patterns

P/grp	**YorNgrAd**		**YorNgrYg**		**YorCanAd**		**YorCanYg**		**EngCanAd**		**EngCanYg**	
As→	HIGH	MID	HIGH	MID	HIGH	MID	HIGH	MID	HIGH	MID	HIGH	MID
ML	13	8	21	30	15	12	12	29	5	55	41	59
%	(62)	(38)	(41)	(59)	(52)	(41)	(30)	(73)	(8)	(92)	(41)	(59)
LL	11	6	16	18	18	8	4	27	3	44	33	60
%	(65)	(35)	(47)	(53)	(69)	(31)	(13)	(87)	(6)	(94)	(35)	(65)
HL	12	7	30	21	20	7	13	37	34	100	47	64
%	(63)	(37)	(59)	(41)	(74)	(26)	(26)	(74)	(25)	(75)	(42)	(58)
TOT	36	21	67	69	53	27	29	93	42	199	121	183
%	(63)	(37)	(49)	(51)	(66)	(34)	(24)	(76)	(17)	(83)	(40)	(60)

The misidentification data (Table 5.11) further reveals that in the ML pattern, all the young participant groups, and the English-Canadian adults, misidentified Low tones mostly as Mid, while both the Yorùbá-Nigerian and Yorùbá-Canadian adult groups misidentified them mostly as High. For the LL pattern, both the English-Canadian young and adult groups and the Yorùbá-Canadian young group misidentified the final low tones mostly as Mid while both the Yorùbá-Nigerian and Yorùbá-Canadian adult groups misidentified them mostly as High. For the HL pattern, the same three groups, previously mentioned, also misidentified the Low tones in this pattern as Mid, while the Yorùbá-Nigerian young group joined both the Yorùbá-Nigerian and Yorùbá-Canadian adult groups in misidentifying the Low tones mostly as High.

To sum up, the misidentification of High tone reveals a trend where the tone was misidentified mostly as Mid by almost all participant groups across the three tonal patterns. The only group to show a departure from the trend was the Yorùbá-Canadian adult group. In contrast, the misidentification of Low tones did not yield a universal trend. It was almost equally divided between those who misidentified the tones mostly as High (i.e., Yorùbá-Nigerian and Yorùbá-Canadian adults; and to some extent Yorùbá-Nigerian young group) and those who misidentified them mostly as Mid (i.e., English-Canadian young and adult groups; Yorùbá-Canadian young group). Thus, the Low tone misidentification data did not support any conclusive statement about whether High or Mid tone attracted the most misidentification of Low tones.

5.4 Discussion

In this experiment, the aim was to investigate participants' identification of only two of the three basic Yorùbá tones. This makes it different from the previous identification task in which all the three tones were targeted. By targeting these two tones, we can observe, (i) whether the High tone attracts better identification rates due to its higher F0, F3 and intensity (Bakare, 1995), (ii) whether the inherent complexity of the contour tones leads to reduced accurate perception rate, and (iii) whether there is a difference in the perception rates between the two contour tones, H (in LH context) and L (in HL context). As in the previous experiments, all these are viewed first within the context of participants' age and language context.

An improved performance recorded by most participant groups in this experiment suggests that the task is comparatively more accessible than the previous Identification Experiment. Among the Yorùbá-Nigerian and Yorùbá-Canadian groups, there are participants with accurate scores of 100 %, while the highest scores range between 70% and 80% among the English-Canadian groups. One possible reason for this improved performance is that the disyllabic stimuli lead to better processing as suggested by Perdereau (2013). The other reason may be that because participants were previously exposed to disyllabic items and the three basic tones in the two preceding experiments, they now had some experience, which they carried over into the present experiment. The improved results further suggest that participants got better at the tasks as they progressed. As for the native speakers, especially those with the highest accuracy scores, they positively applied their knowledge of the tones to do the task. However, it may also be that these native speakers, as well as other good performers, quickly noticed that the ranges of choices were getting smaller as they progressed in the experiment.

The overall accuracy rate of 63% recorded by English-Canadian adult group puts them as comparable to the native-Yorùbá-speaking Yorùbá-Nigerian young group (65%) and better than the Yorùbá-Canadian adult group (55%) who are also native Yorùbá speakers. The performance of these adult non-native tone language speakers as well as their performance in the previous two experiments suggests that, as university students of Linguistics, these participants might have greater familiarity with experimental tasks of this nature.

Language context is possibly the only factor to cite for the discrepancy between the two native-speaking adult groups. Regarding the performance of the adult native Yorùbá versus non-Yorùbá speakers, one possible interpretation for this discrepancy relates to the use of nonce words and its relationship with being either a native or a non-native speaker of a language. The absence of real words is arguably one of the main reasons that make the task harder for some native speakers of Yorùbá. However, the higher percent accuracy recorded by Yorùbá-Nigerian adult group (80%) suggests that they were not affected by the absence of real words. As for the remaining native Yorùbá-speaking groups, the lower percent accuracy may indicate that their knowledge of real Yorùbá words interfered with their performance. In other words, they had some expectations, based on their experience. They were possibly thinking of what they expected to hear, thereby causing them problems while carrying out the task. Some stimuli that could have caused interference were picked out during the analysis of participants' response. [Bọ̄rẹ́ MH], a nonce word, has a minimal pair in the real word [bọ̄rẹ̀ ML] 'to worship an idol'. Similarly, the nonce word [jẹ̄bà ML] has real-word minimal pairs in [Jẹ́bà HL], 'a city in Central Nigeria', and [jẹ̀bà LL], 'to eat cassava porridge'. Another such stimulus is the nonce word [ọ̀kọ́ LH], which has minimal triplets in [ọ̄kọ́ MH] 'a hoe'; [ọ̄kọ̀ ML] 'a car'; and [ọ̄kọ̄ MM] 'a husband'. All the native Yorùbá-speaking groups recorded higher misidentification rates on these words. Thus, when those participants heard those equivalent nonce words, they were bound to be influenced by the tones of the real words that have an affinity to the nonce words they heard. Unlike the native tone-language speakers, the English-Canadian groups had no real words to distract them; they had no pre-conceived notion of what tones the segments should have. This could explain the noticeably improved performance of the English-Canadian adults in this experiment.

Another aspect highlighted by the results is the general effects of Age and Language Context on tone identification. The analysis suggests that globally, adults are significantly better at tone identification than young participants. This occurs despite all participants having benefitted from previous, although limited, exposure to the tones through the practice sessions and the two preceding experiments. This indicates that older age may be interpreted as more experience at general language or tone processing.

Concerning the Language Context factor, the analysis suggests that generally, Yorùbá-Nigerian significantly leads to better identification than either Yorùbá-Canadian or English-Canadian.

This factor is highlighted by the results of both the adult and the young groups. The Yorùbá-Nigerian adults identified the tones significantly better than both the Yorùbá-Canadian and the English-Canadian adults. Similarly, the Yorùbá-Nigerian young group identified the tones significantly better than the Yorùbá-Canadian and the English-Canadian young groups. When compared to the previous two experiments (AX discrimination and Identification 1), it is observed that the Yorùbá-Nigerian young group have consistently highlighted the importance of the knowledge of the tone language, by outperforming their Yorùbá-Canadian and English-Canadian counterparts. This is the first experiment in which the Yorùbá-Nigerian adult group highlights the importance of this factor. The experiment shows that perhaps due to their complexity, the contour tones require more than sensitivity to pitch height to process accurately.

In this experiment, knowledge of the tone language, acquired through frequent use, leads to better performance. This is in line with studies such as Bakare (1995), Omozuwa (1991), and Wayland and Guion (2004) mentioned earlier, which found that native tone language speakers tended to perform better than non-tone language speakers at tone identification. The use of real words as stimuli in these studies, in addition to the pre-test training provided to participants in Wayland and Guion (2004) are the main differences that set these studies apart from the present study. In contrast, the other groups of studies (e.g., Connell, 2000; Hao, 2012; and Francis et al., 2008) suggest that knowledge of the tone language is not required for accurate identification of the tones. While the use of nonce words as stimuli makes these studies similar to the present study, the comparable performance reported for native and non-native tone language listeners at tone identification is a finding not attested in the present experiment. It is possible that the type of the identification experiments (categorization and difference rating) coupled with the training sessions provided to participants in Connell (2000) and Francis et al. (2008) contributed to the comparable results observed for the native and non-native groups.

The results of the three experiments conducted so far have revealed a trend which indicates the increasing importance of knowledge and experience of the tone language. The first experiment (AX discrimination), which requires no knowledge of the tone language, revealed no difference between the three adult groups. The second experiment (Identification 1), a task which demands knowledge of the tones, showed some differences in the adults' performance. The Yorùbá-Nigerian adult group outperformed the Yorùbá-Canadian adult but not the English-Canadian

adult group. The present experiment revealed that the Yorùbá-Nigerian adults significantly outperformed the other two adult groups. While this result shows that the present task was more demanding, it further reveals that accurate identification of the complex tones benefits from adequate knowledge of the tone language.

More experience with the tone language predisposes participants to tap into their acquired tonal knowledge, thereby leading to improved tone identification. Native Yorùbá speakers may use their in-depth knowledge of the language to re-analyze the nonce lexical items into a more meaningful one. One instance would be the stimulus [bòrí LH]. Although this stimulus is supposedly a nonce word, not meaningful in Yorùbá, it has minimal triplets in [bōrí MH] 'be victorious', and in another word contracted from the verb [bò L] 'to cover' and the noun [ōrí MH] 'head', meaning 'to cover the head'. One would have expected interference from these equivalent real words to impede participants' identification process. However, for proficient users, it is not an interference issue; rather they can reprocess the stimuli, leading to better identification. As a native Yorùbá speaker listening to this type of stimulus, there is still the expectation that (some of) the words were Yorùbá words. Thus, there is the tendency for such listeners to resort to their knowledge of the language to arrive at the assumption that [bòrí LH] may comprise of two words contracted into one as is commonly the case with the language. Thus, [bo L] 'to cover' as a verb, combines with the other word [òrí LH] 'shea butter' to form [bòrí LH], meaning 'to cover the shea butter'. The same process can be applied to other stimuli in this Experiment. This is an advantage possessed by only the Yorùbá-Nigerian participants. This description explains why the Yorùbá-Nigerian adults recorded a percent accuracy of more than 80% for most of the stimuli in this Experiment, with 59% as the lowest percent accuracy for a stimulus.

Meanwhile, as observed in the previous experiments, the performance in the present experiment, for both the Yorùbá-Canadian adult and the young groups, was comparable to their English-Canadian counterparts. The fact that generally, there was no significant difference between these two participant groups further suggests that these two language context types are somewhat similar. They either lack tone language knowledge and experience (in the case of English-Canadian group) or, the knowledge of the tone language they have is not optimal; hence they are

unable to tap into this acquired tonal knowledge to identify the contour tones (in the case of Yorùbá-Canadian group).

The overall accuracy results reveal that participants did better as they progressed in the experiments. This further suggests that they positively utilized their experience from the two preceding experiments to their advantage. When the performances of the two age groups with the same Language Context are compared, the results suggest comparable performance for the Yorùbá-Nigerian and Yorùbá-Canadian categories; the adults did not significantly outperform their younger counterparts. In contrast, the English-Canadian adult group performed significantly better than their younger counterparts. Thus, in addition to this group's noticeable improvement as the experiments progressed, these adults, because of their age, also possess more language experience which gives them some edge over their younger counterparts. The younger participants either had not positively benefitted from the previous tonal experiments or do not have the experience necessary to transfer previous tone perception training and experience into subsequent tasks.

The outcome from the comparison of participants of the same age groups but belonging to different Language Context reflects what is obtained in the general Language Context effects. The revelation that Yorùbá-Nigerian young and adult groups were significantly better than their counterparts from Yorùbá-Canadian and English-Canadian further suggests the positive effects of tonal exposure on tone identification. In contrast, no significant effects were noted when comparing the Yorùbá-Canadian and English-Canadian adults. Although these two adult groups belong to different Language Contexts, neither group was significantly better than the other in this experiment. The comparable performance may be the result of either attrition, in the case of the Yorùbá-Canadian category, or lack of tone language experience, in the case of the English-Canadian category.

5.4.1 Language-internal factors

In support of the findings of the preceding Identification Experiment, the present Identification Experiment further suggests that there is no difference in the rate of identification for High and Low tones. Thus, while the issue of the acoustic cues of the High-tone (F0, F3 and intensity), as

observed by Bakare (1995), may contribute to more accurate identification of High tone, the lower-pitch, coupled with the distinctive phonation type may also have informed the comparable identification rate observed for the Low tone. In addition to these peculiar features associated with High and Low tones, the knowledge and application of the English stressed, and unstressed syllable system might be an additional tool for the adult English-speaking participants.

Out of the two complex tones, one was easier to identify than the other as suggested by the statistical models in Tables 5.3 and 5.5. One possible explanation for this finding resides in the pitch contour measurements for these two complex sets in (see Chapter 1, Table 1.3 and Figure 1.2 above). The contour Low tone recorded the highest onset measurement (169.3Hz), with a difference of between 66 – 71 Hz between it and the non-contour Low tones (ML = 102.7; LL = 97.3). At mid-point, the differences were reduced to about 39.7- 42.5 Hz. Meanwhile, the offset measurement recorded by the contour Low tone (98.26Hz) was, in fact, more than what was recorded for LL onset (97.3Hz), and almost close to the measurement for ML onset (102.7Hz). This fact possibly explains why half of the participant groups, especially the English-Canadian young and adult groups and to some extent, the Yorùbá-Canadian young group mostly misidentified Low tones, especially the contour Low, as Mid tones. It is interesting that these groups recorded lower rates of misperception of Low tones as High. They might intuitively know that the contour Low was not a High tone (comparable to English primary stress) and had therefore decided to identify it as Mid tone instead.

In contrast, the other participant groups (Yorùbá-Nigerian young and adults as well as the Yorùbá-Canadian adults) mostly misperceived the Low tones as High. We may also cite the same argument of the offset measurements for the contour Low tone as a possible reason for these groups' misidentification pattern. Being all native Yorùbá speakers, it is possible that they could not perceive creakiness that is usually associated with the Low tone. The production of the falling contour usually sets off from the preceding High tone, making the realization of the creakiness nearly impossible at the offset. Another look at the tone measurements in Table 1.3 may support this assertion. Except for the contour High offset (212.7), the onset measurement for the contour Low (169.3) was higher than all the measurements recorded for the H-final patterns at onset, middle and offset (Min =113.8Hz; Max=157.5Hz). Thus, whether participants pay attention to the direction of change of tone of the second syllabic peak or to its end point

(Hombert, 1976), it may still lead to confusion due to the wide difference in F0 onset between the contour Low and the other Low tones. Similar to the Edo HL pattern, with voiced consonants, the Low tone becomes assimilated to the level of the preceding High tone before the descending glide (Omozuwa, 1991). It is possible that the F0 offset of the Low tone did not reach the level that makes it perceivable to most listeners or did not contain relevant acoustic cues which could have assisted listeners towards more accurate perception. The downward glide of the Low tone may even be more difficult with voiceless consonants where the initial F0 value of the second syllable is usually higher than the final F0 value of the first syllable as observed by Omozuwa (1991). All these factors may lead to inaccurate perception of the Yorùbá LH pattern.

The measurements for the contour High tone (L**H**) indicate a lower measurement at onset (113.8Hz), which was between 39 to 43Hz lower than non-contour MH and HH. There was not much difference at mid-point, but at the offset (212Hz), it was about 55 and 56Hz higher than the High tone in MH (157.5Hz) and HH (156.7Hz) respectively. This spike in pitch height at the offset position for contour High expectedly resulted in better perception for listeners. It is possible that in the LH pattern, participants paid more attention to direction of change of the tone of the second syllabic peak, or to its end point (Hombert, 1976), hence the improved accuracy rate for the contour High. This probably explains the significantly better accuracy obtained for contour High (LH) versus contour Low (HL) perception. While we may cite confusion between the High and Mid tone as one reason for the global misperception of High tone as Mid, the fact that individual groups displayed a peculiar identification pattern requires a second look. In HH patterns, all groups misidentified the High tone mostly as Mid due to possible High-Mid confusion. Successive High tones in a disyllabic context are easily confused as MM to both trained and untrained listeners. In the MH pattern, all groups except the Yorùbá-Canadian adults misperceived Mid as High. Moreover, misperception of the High tones as Low in this pattern by the Yorùbá-Canadian adults was possibly due to confusion or due to the effect of the preceding Mid tone. In the complex LH pattern, the Yorùbá-Canadian and the Yorùbá-Nigerian adult groups often misperceived High tones as Low. This could be due to the preceding-tone effect. The lingering effect of the lower onset of the contour High, caused by the preceding Low tone, continues until the mid-point before spiking at the offset. This might have hindered the listeners' proper identification of the tones as shown by the misidentification pattern of the Yorùbá-Nigerian and Yorùbá-Canadian adult groups.

The findings of Hombert (1976) and Omozuwa (1991) need to be viewed while bearing in mind the differences between these studies and the present thesis in terms of objective and the method. The two previous studies share a similar objective in terms of investigating the acoustic cues used by native speakers / listeners of the respective languages in identifying the tones of their languages. Thus, the experimental methods, in which listeners rated the perceptual distances between two Yorùbá tone patterns (Hombert, 1976) or assessed the acceptability or non-acceptability of Edo bisyllabic patterns (Omozuwa, 1991), were suitable for the objectives of the studies. In contrast, the present thesis did not investigate the role of specific acoustic cues. It examined native and non-native participants' accurate / inaccurate perception of the Yorùbá contour tones while considering participants' age, and language context, as well as the effects of the inherent characteristics of the two Yorùbá complex and the three simple tones. Additionally, it investigated the possible effects of participants' L1 on their perception patterns.

Moreover, the use of nonce words as stimuli ensured that the native speakers should not have had considerable lexical advantage over the non-tone language speakers. Using nonce words as in the present experiment also has an added advantage when we compare the present experiment to Hombert (1976) which used real bisyllabic nouns. It allows for the inclusion of all possible disyllabic patterns in Yorùbá, a feat which would have been impossible with real nouns. Three of the nine Yorùbá disyllabic patterns were not included in Hombert (1976) because there were no real Yorùbá nouns for these excluded patterns. With respect to Omozuwa (1991), one other reservation to consider is the fact that Edo has only two tones compared to the three tones in Yorùbá. In the present thesis, the instruction for participants to identify the tones as either High, Mid, or Low was to provide us with more data to investigate participants' misidentification patterns, especially the extent to which the two tones were identified either as each other or as the third tone, Mid.

5.4.2 Individual groups vs. language-internal factors

The analysis in Table 5.7 indicates that when compared to the Yorùbá-Nigerian adults, only the English-Canadian adults identified contour Low tones significantly worse than contour High. Secondly, this same group also identified contour Low tones at significantly worse rates than

either the non-contour Low or non-contour High (Table 5.9). It is possible to cite the peculiar complexity of this contour tone as a factor which might have influenced the poor identification rate by the English-Canadian adult group. At the same time, we can also consider the fact that the language experience possessed by this mostly English-speaking group means that they could tap into some of the commonalities between English and Yorùbá while undertaking this perception experiment. Despite belonging to different language typology with different suprasegmental features, the two languages share one common feature: F0 contour is the primary cue for tones in Yorùbá (Bakare, 1995) and English stress (Beckman, 1986). Secondly, the contexts of the two Yorùbá complex tones are comparable to the alternations of stressed and unstressed syllables in English prosody. While LH is closely related to the iambic, which indicates a weak stress followed by a strong stress, HL shares close affinity to the trochaic, which refers to a succession of strong and weak stresses in a disyllabic or multisyllabic context. Meanwhile, a recent study by Chrabaszcz, Winn, Lin, and Idsardi (2014:1) suggested that "cues were stronger when they were in the iambic contour than when they were in the trochaic contour." Thus, based on this suggestion, we can say that despite their non-exposure to a tone language, the English-Canadian adult group, found the stronger-cued iambic-like LH pattern easier to identify than the trochaic-like HL pattern, which supposedly has 'weaker cues'.

Summing up, the difference in tone identification accuracy between the Yorùbá-Nigerian participant groups and the other participant groups, as observed in the present experiment, suggests strong positive effects of tone Language Context. More exposure to the tone language leads to better tone identification. Furthermore, the analysis does not suggest any significant difference in the identification accuracy for High versus Low tones. Thus, the hypothesis that the inherent features of the High tone may lead to better perceptual accuracy is not being supported by the results. Instead, the results suggest that like the High tone, the Low tone is equally identifiable due to its own distinctive perceptual characteristics (Hombert, 1977), namely the falling contour and in some cases the creaky phonation which is typically associated with production of the tone. Moreover, this experiment does not suggest a significant difference in perceptual accuracy between Yorùbá contour and non-contour tones, except at individual group level. The English-Canadian adult group was the only group to identify non-contour tones better than contour ones. Similarly, among the two contour tones, the contour High tone (in LH) attracted better identification than the contour Low (in HL). Individual group analysis further

suggests that the English-Canadian adult group was the only group to identify contour High significantly better. The improved accuracy in the contour High (in LH pattern) for this group is suspected to be linked to either the affinity between the High tone and the English primary stress or the affinity between the LH pattern and the English iambic meter. Furthermore, the misidentification analysis suggests that the HL pattern was misidentified more than any other patterns. It further suggests that Mid tone attracted the most misidentification of the other tones. It was the tone to which most participants resort as a substitute for the other tones. This observation possibly supports the intermediary position occupied by this tone in Yorùbá tonal space.

Chapter Six: Lexical Recognition

The previous chapter suggests that more exposure to a tone language leads to better tone identification accuracy than less exposure. However, as we have seen in the previous chapters as well as in prior studies (e.g., Francis et al., 2008; So, 2006), participants may not need to have prior knowledge of the tone language in question in order to be quite successful in identifying the tone contrasts and types.

The results of the second identification experiment (complex tones, Chapter 5) as well as the overall better results of the Yorùbá-Nigerian young group for all the previous experiments, when compared to the other young groups, provide some evidence that knowledge of the tone language is an added advantage to tone identification/perception. These results are in line with previous studies which suggest that knowledge of the tone language facilitates more accurate perception of tones (e.g., Wayland & Guion, 2004; Lee, Vakoch, & Wurm, 1996; Bakare, 1995, and Omozuwa, 1991). In particular, e.g., Yeh and Lu (2012) carried out a lexical task, in addition to discrimination and identification, which focused on the perception of low-level Hakka tones. The study involved older fluent speakers, young fluent speakers, and young attriters. In Yeh and Lu (2012), participants were provided with one monosyllabic sound in each trial. Participants were asked to recognize the word meaning of the sound. The results found that both the old and the young fluent speakers significantly outperformed the young attriters in the lexical task, while no difference was observed in either the discrimination or the identification tasks. Thus, more experience with the language leads to better performance in a lexical perception task involving tones.

With respect to the present study, it is still not clear the extent to which knowledge of Yorùbá plays a role in these tasks. Following Yeh and Lu (2012), the present experiment investigates whether the use of real words leads to improved tone perception accuracy compared to the use of non-words. Unlike Yeh and Lu (2012), the present task included pictures of the minimal pair to assist listeners in deciding which item of the pair they hear. Especially interesting is whether native speakers who are based in Nigeria will outperform those who are based in Canada. It is also interesting to see how the performance of older speakers in these two environments compares to that of the younger speakers. Thus, the age and language context factors involve two

categories from each factor (Age: adult vs. young; Language Context: Yorùbá-Nigerian vs. Yorùbá-Canadian) instead of three categories included in Yeh and Lu (2012).

This investigation involves two population groups belonging to the language of study but with different language contexts at the time of the experiment. One of the groups continues to live in the immediate tone-language (L1) environment while the other group has moved away from the immediate tone-language environment to a new environment where a non-tone language (L2) is dominant. According to Bullock and Gerfen (2004), as well as Seliger and Vago (1991), movement to an L2-dominant environment can lead to L1 attrition due to a change in the relative use of the L1. Other studies (e.g., Hutz, 2004) suggest that the aspect of the language that is most prone to loss is the lexicon, probably since vocabulary acquisition is considered a continuous process. In another relevant study, Montrul (2008) observes that if L1 attrition were to occur in children, it would be more severe than L1 attrition in adults.

As in the previous experiments, each population group also consists of two age groups: young and adult. Thus, the hypothesis builds on both the Age and the Language Context factor tested in all the other experiments. In other words, the chapter investigates whether Age and Language Context predict the tonal accuracy of participants with real words (as opposed to non-words). Thus, the present experiment tests the prediction that real words will lead to more accurate tone perception for all participants. Considering several studies (e.g., Lee et al., 1996; Wayland & Guion, 2004; Yeh & Lu, 2012) that suggest that prior tone language experience provides an advantage for subjects in tone perception tasks, this experiment predicts that participants with more exposure to the tone language will outperform those with less exposure. Furthermore, based on the assumption that participants tend to have more exposure as a result of older age, adult native tone language speakers are expected to do better than the young ones. Similarly, participants with richer tone language contexts are expected to perform better than those with poorer tone language contexts. To test these predictions, the participants will be presented with minimal pairs of existing Yorùbá words instead of identifying tones in non-word strings.

Minimal pairs for Yorùbá verbs and nouns differing in tone patterns are used as stimuli. Pictures depicting the meaning of the words are also provided to eliminate confusion. Using real words as is the case here, is expected to highlight the effect that higher-level knowledge of the language will have on participants' perception. Frequent usage of the tone language which leads to richer

exposure, is expected to lead to better identification than infrequent use. Participants who know the language well are expected to apply the knowledge in recognizing the appropriate tone pattern for one of the minimal pairs they hear and whose picture is also shown on the screen.

The use of real Yorùbá words as stimuli for this experiment as well as the decision to use minimal pairs is similar to Hombert (1976). Despite this similarity, the purpose as well as the method of investigation for the two experiments diverge. Hombert (1976) examined the perceptual cues used by native speakers to distinguish these nouns that differ in their tonal patterns. As such, the task required participants to rate the perceptual distance between two patterns on a scale of 0-10. The approach of the present thesis takes Hombert (1976) a step further. Rather than evaluating the perceptual distance, it examines whether knowledge of the tones will influence native speakers' perceptual accuracy in recognizing one of the minimal pairs being presented.

6.1 Methods

6.1.1 Participants

Eighty-six (86) participants from only the tone-language-exposed groups, namely the Yorùbá-Nigerian and the Yorùbá-Canadian groups, took part in this experiment. English-Canadian participants were not included in this experiment because it was a lexical experiment. The participants were broken down as follows: Yorùbá-Canadian groups had fifteen (15) adults and twenty-eight (28) young participants; Yorùbá-Nigerian groups had sixteen (16) adult and twenty-seven (27) young participants.

6.1.2 Stimuli

The stimuli contained twelve (12) monosyllabic minimal pairs with tonal contrasts (Table 6.1) that were all real Yorùbá verbs. All the verbs were monosyllabic since disyllabic verbs are very few in Yorùbá. Moreover, the few Yorùbá disyllabic verbs could not yield the minimal pairs required for this experiment. All the selected verbs were frequently-occurring ones involving

everyday actions. They included verbs denoting actions related to eating, clothing, personal hygiene, farming, and general personal development, among others.

Table 6. 1 Monosyllabic minimal pairs for the Lexical Experiment

Tonal contrast	Pic/Word 1	Gloss	Pic/Word 2	Gloss
H- L	bó	'to peel'	bò	'to cover'
	fá	'to shave'	fà	'to pull'
	fọ́	'to break'	fọ̀	'to wash'
	ká	'to pluck'	kà	'to read'
H – M	jẹ́	'to answer'	jẹ̄	'to eat'
	kí	'to greet'	kī	'be thick'
	kọ́	'to build'	kọ̄	'to write'
	mú	'to pick'	mū	'to drink'
L – M	rò	'to think'	rō	'to plough'
	sè	'to cook'	sē	'to do'
	sò	'be loose'	sō	'to tie'
	sṳ̀	'to sleep'	sṳ̄	'to roast'

The stimuli also contained ten (10) disyllabic minimal pairs with tonal contrasts (Table 6.2), all of which belonged to real Yorùbá nouns. They were comprised mostly of items within the immediate surroundings of a typical Yorùbá community. They included nouns denoting body-parts, foods, household utensils, animals, transport, entertainment, and items relating to nature or the environment. Furthermore, due to the absence of monosyllabic nouns in Yorùbá, all the nouns used in the experiment were disyllabic, and they are considered high-frequency words. Each stimulus was recorded once. Pictures representing the minimal pairs were also included in this experiment as shown in the samples in Appendix E1. To ensure that all the minimal pairs were used, the experiment had two versions. One version included the sounds for one set of the two words in all the minimal pairs while the second version included the other set. Both versions included the same set of two pictures for the minimal pairs. The positioning of the pictures either to the left or right of the screen was not the same for the two versions.

Table 6. 2 Disyllabic minimal pairs for the Lexical Experiment

Tonal contrast	Word 1	Gloss	Word 2	Gloss
	ōdò	'river'	ōdó	'mortar'
	ọ̄wọ̀	'broom'	ọ̄wọ́	'hand'
ML - MH	īkà	'eggplant'	īká	'termite'
	īkù	'belly'	īkú	'squirrel'
	ọ̄rù	'neck'	ọ̀rū	'sky'
ML – LM	ọ̄bẹ̀	'soup'	ọ̀bẹ̄	'knife'
	āg͡bọ̀	'basket'	àg͡bọ̄	'coconut'
LL – ML	ọ̀kọ̀	'spear'	ọ̄kọ̀	'car'
LH – LL	ìlú	'town'	ìlù	'drum'
LH – MH	àk͡pá	'scar'	āk͡pá	'arm'

6.1.3 Procedure

The main task was for participants to recognize Yorùbá lexical items based on the sound stimuli they listened to. Participants viewed the text of the minimal pair displayed on the screen. The text was written in Yorùbá orthography but without the tone markings. Next, the program played the sound prompt representing one of the lexical minimal pairs. In the case of monosyllabic lexical items, only one tone differentiated the meaning of the minimal pairs. In disyllabic lexical items, participants were required to pay attention to the two tones accompanying the two syllables as the two tones together determined the difference in meaning. Along with the sound prompt, two pictures appeared on the screen representing the two lexical items. The picture on the left was labelled "1" while the one on the right was labelled "2". The participants' task was to identify which of the two pictures (1 or 2) matched the word they heard. Based on the instruction, they were to press the "1" key on the keyboard for picture 1 and the "2" key for picture 2. After the sound stimulus was played, the program waited for an input from the participants for up to 3.5s before displaying the next stimulus. Once they had selected their answer, the picture stimuli disappeared from the screen, and the next stimuli followed. The experimental software automatically recorded their response. The main experiment was preceded by four practice trials, after which the participants had a short period in which to ask questions

before moving on to the actual experiment. The main experiment was divided into two blocks, the first block included the monosyllabic verbs while the second block included the disyllabic nouns. The whole experiment lasted for about 8 minutes. The stimuli for the practice session for this experiment is presented in Appendix E2.

6.2 Results

Table 6.3 presents the general results of the Lexical experiment. The results reveal that the highest percent accuracy rate by participants in this experiment was 95% while the lowest was 57%. The exact binomial test sets the better than chance score (BTC) score for this experiment at 15. As shown in the last column of Table 6.3, three of the four participant groups, the young and adults from the Yorùbá-Nigerian category and the Yorùbá-Canadian adult group, had all their participants scoring significantly better than chance. The last participant group, the Yorùbá-Canadian young participants, had 64% of its members scoring better than chance.

Table 6. 3 Participants' percent accuracy in the Lexical Experiment

	ACC	**%**	**MEAN**	**SD**	**N**	**BTC**	**%BTC**
YorNgrAd	335/352	95%	20.94	1	16	16	100%
YorNgrYg	521/594	88%	19.3	2.27	27	27	100%
YorCanAd	303/330	92%	20.2	3.71	15	15	100%
YorCanYg	354/616	57%	12.64	4.18	28	18	64%

6.2.1 Statistical Analysis

Lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015) was used to perform a generalized linear mixed-effects regression (glmer) analysis to investigate the general effects of Syllable count, Age, and Language Context on their lexical recognition patterns. Syllable, as a fixed effect, had two levels: *Monosyllabic* and *Disyllabic*. Age had two levels: *Adult* and *Young*. Finally, Language Context also had two levels: *Yorùbá-Nigerian* and *Yorùbá-Canadian*. The accuracy score, which consisted of the individual score for the 22 target

stimuli, was the dependent variable. Ther accuracy was coded as either '1' or '0' with '1' indicating an accurate or correct score while '0' indicated inaccurate or incorrect score. As random effects, the model had intercepts for *item* and *participant*. The model was based on 1892 observations (a total of 86 participants x 22 individual item score).

P-values were obtained by likelihood ratio tests of both main-effects model and another model with interactions. An ANOVA revealed that the model with the interaction term between Syllable and Language Context (AICinteraction = 1417.5) fitted the data more closely than the model without an interaction term (AIC = 1422.6). The addition of by participant random slopes for the effect of Syllable count did not improve the model's fit. The model specified Yorùbá-Nigerian as the reference level for Language Context. The output of the simple model is presented in Appendix E3 while the output for the interaction model is summarized in Table 6.4.

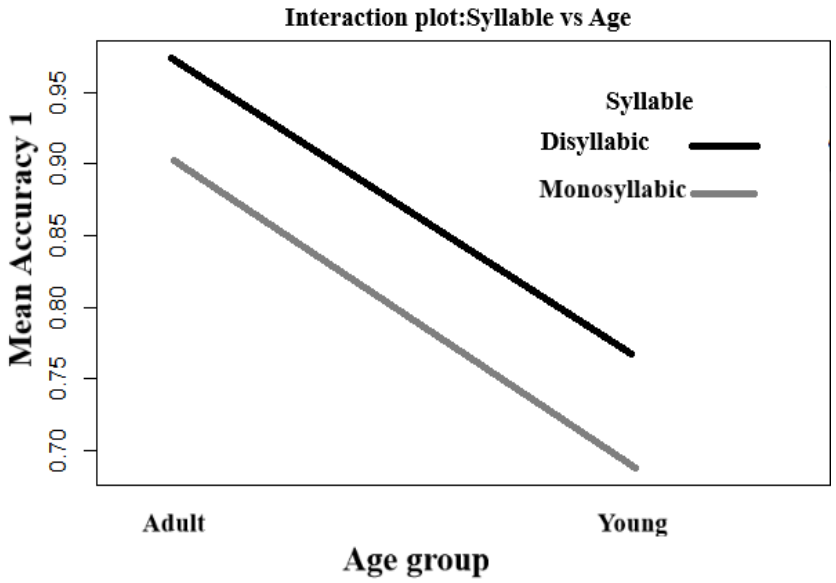
The analysis suggests that all three fixed effects, Age, Language Context, and the number of syllables in the stimuli, had significant effects on participants' lexical recognition accuracy. Younger participants recognized the lexical items significantly less accurately than the adults. Monosyllabic verbs were recognized significantly less accurately than disyllabic nouns. The model further compares the two Language Context types in this experiment and the analysis suggested that generally, Yorùbá-Canadian participants performed significantly worse in lexical recognition accuracy than Yorùbá-Nigerian participants. With respect to the interaction between Syllable and the Language Context factor, the analysis indicates that compared to Yorùbá-Nigerians, Yorùbá-Canadians did not record as much difference in accurate recognition of monosyllabic and disyllabic lexical items (see Figure 6.2).

Table 6. 4 Coefficients of the glmer model of the effects of Age and an interaction between Syllable count and Language Context on lexical recognition. Random effects: Participant and Item. N=1892. Reference level: YorNgr; Disyllabic

Ref: YorNgr; Adult; Disyllabic				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	5.9415	0.5457	10.887	< 2e-16***
Young	-2.2212	0.3332	-6.666	2.62e-11***
Monosyllabic	-1.8183	0.5126	-6.547	0.00039***
YorCan	-2.8384	0.4582	-6.195	5.85e-10***
Monosyllabic x YorCan	1.1569	0.4463	2.592	0.00954**

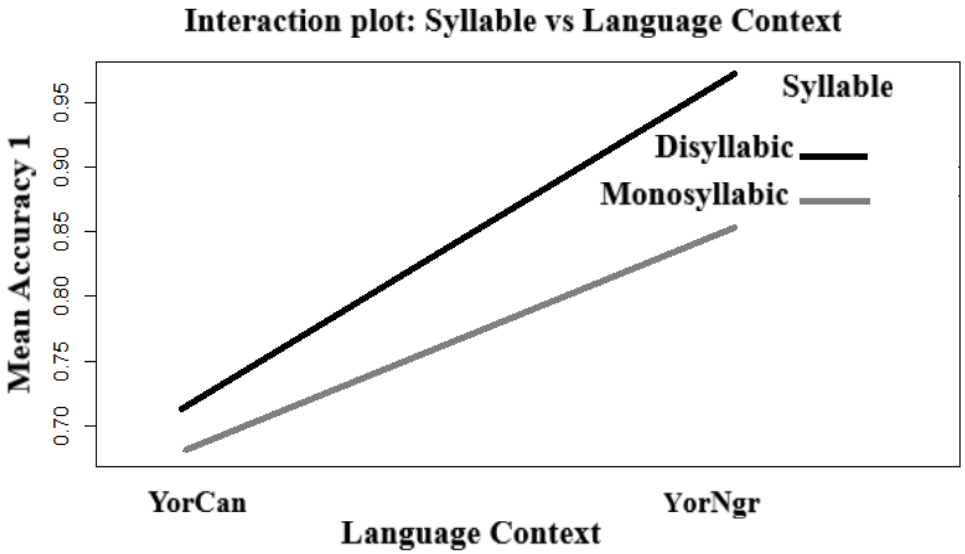
To further explore the nature of the interaction between the main effects with respect to accuracy, Figure 6.1 reveals that the two main effects of Age and Language Context were crucial. Adult participants were more accurate than the younger ones, while Disyllabic stimuli attracted more accurate responses than Monosyllabic stimuli. However, the parallel lines of the plots indicate that the two main effects did not exert any special synergistic effect on accuracy.

Figure 6.1 Mean accuracy in Experiment 4 by syllable count (monosyllabic, disyllabic) and age group (adult, young)



The plot in Figure 6.2, also reveals the prominence of the two main effects of syllabicity and Language Context. Disyllabic stimuli had higher accuracy than monosyllabic stimuli while Yorùbá-Nigerians were more accurate at lexical recognition than Yorùbá-Canadians. The figure further indicates the presence of an interaction between the two main effects. It reveals a more significant effect of syllabicity for Yorùbá-Nigerian than for Yorùbá-Canadian participants.

Figure 6.2 Mean accuracy in Experiment 4 by syllable count (monosyllabic, disyllabic) and Language Context (Yorùbá-Canadian, Yorùbá-Nigerian)



6.2.2 Lexical recognition by participant groups

This section investigates the lexical recognition accuracy by the four participant groups. It compares the accuracy of the four groups with one another and then investigates the groups' lexical recognition accuracy with respect to the monosyllabic and disyllabic stimuli. First a visual representation of the interaction between the two fixed effects is presented in Figure 6.3.

Figure 6. 3 Mean accuracy in Experiment 4 by syllable count (monosyllabic, disyllabic) and the four participant groups

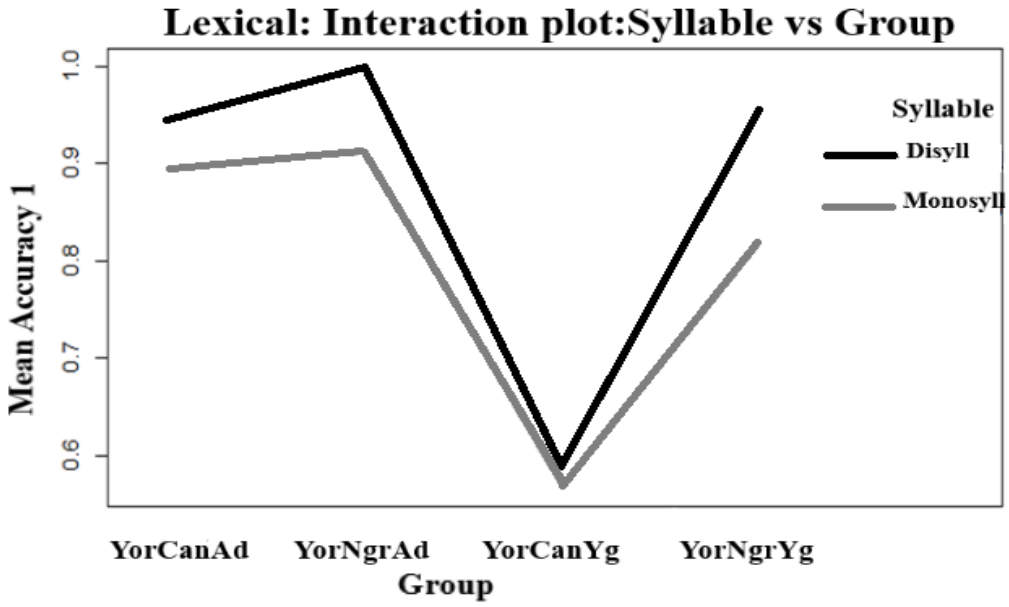


Figure 6.3 indicates that there were interactions between the two main effects, Syllable and Group. First, it reveals that the Yorùbá-Nigerian adult group recorded higher accuracy than all the other groups in both the monosyllabic and disyllabic stimuli. It also reveals that the Yorùbá-Canadian young group performed worse than the other groups in both the monosyllabic and disyllabic stimuli. To examine the statistical significance of the difference between these groups, a model was fitted to undertake a comparison within and between the four participant groups. The model had 'Syllable' as a fixed effect with two levels: *Monosyllabic* and *Disyllabic*. The model also had 'Group' as a fixed effect with four levels which included the *Yorùbá-Nigerian Young* and *Adult* groups as well as the *Yorùbá-Canadian Young* and *Adult* groups. As random-effects, the model included intercepts for *participant* and *item*. The dependent variable was the accuracy score which consisted of the individual score for the 22 target stimuli. The model was based on 1892 observations (a total of 86 participants x 22 individual item score =1892). P-values were obtained by likelihood ratio tests of the model with the two main effects and another model with the main effects and by-participant random slopes. An ANOVA revealed that the simple model with only the main effects fitted the data more closely than the model with random slopes (AIC = 1419.2 vs. AIC (random slopes) =1420.3). The summary of the model is presented in Table 6.5.

Table 6. 5 Coefficients of the glmer model comparing the lexical recognition accuracy of the four groups. Random effects: Participant and Item. N=1892

Ref: YorNgrAd				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	4.8191	0.5632	8.557	< 2e-16***
YorCanAd	-0.9379	0.5876	-1.596	0.11044
YorCanYg	**-3.8012**	**0.5044**	**-7.537**	**4.82e-14*** **
YorNgrYg	**-1.3350**	**0.4822**	**-2.768**	**0.00563****
Ref: YorCanAd				
Intercept	3.8812	0.5359	7.242	4.43e-13***
YorCanYg	**-2.8633**	**0.4481**	**-6.390**	**1.66e-10*** **
YorNgrYg	-0.3971	0.4857	-0.818	0.4136
Ref: YorNgrYg				
Intercept	3.4842	0.4401	7.917	2.43e-15***
YorCanYg	-2.4662	0.3727	-6.617	3.67e-11***
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.'** 0.1 ' ' 1				

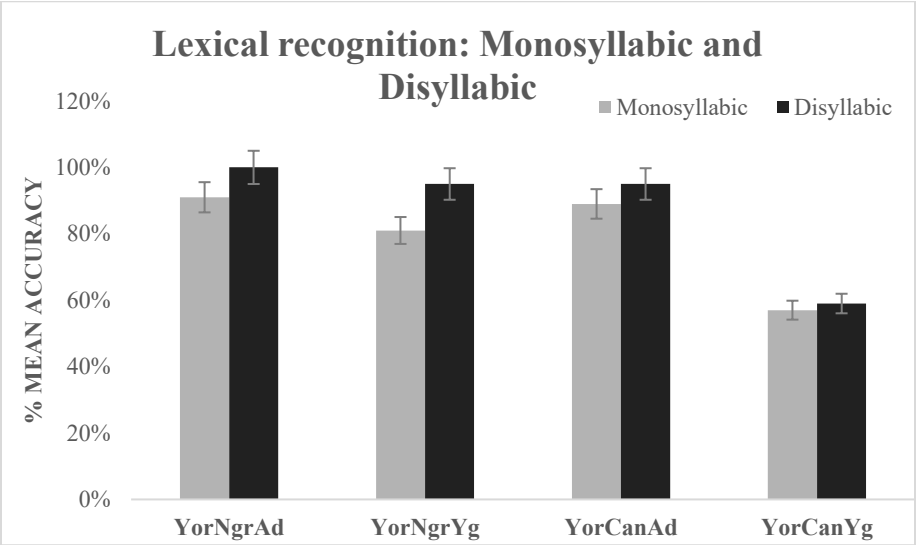
Table 5.5 summarizes the outputs of the model with different reference levels for the group factor. The results suggested that within the Yorùbá-Nigerian category, the adult members were significantly better than the young members in Lexical recognition accuracy. The rows in bold black print indicates that the Yorùbá-Nigerian adults were also significantly better than the Yorùbá-Canadian young participants. Furthermore, when Yorùbá-Canadian adult group was specified as the reference level, this adult group was significantly more accurate at lexical recognition than the younger members from the same Language Context type. However, there was no significant difference between this adult group and the Yorùbá-Nigerian young group.

When participant groups belonging to the same Age group but from different Language Context were compared, the analysis revealed that there was no significant difference in the lexical recognition rate between Yorùbá-Nigerian and the Yorùbá-Canadian adults. In contrast, with Yorùbá-Nigerian young as the reference level, the analysis suggested that the young Yorùbá-Nigerians recorded a significantly better lexical recognition accuracy than the young Yorùbá-Canadian participants.

6.2.3 Recognition of Monosyllabic and Disyllabic items

Figure 6.3 above appears to suggest that disyllabic stimuli recorded higher accuracy rates than monosyllabic ones for the Yorùbá-Nigerian groups than for the Yorùbá-Canadian groups. To get a clearer picture, Figure 6.4 presents the percent accuracy rates for the four groups in monosyllabic and disyllabic stimuli.

Figure 6.4 Mean accuracy for the recognition of monosyllabic and disyllabic lexical items by the four participant groups



The graph in Figure 6.4 reveals that the Yorùbá-Nigerian adults recorded a 100% recognition rate for the disyllabic items while recognizing the monosyllabic items at a rate of 91%. Their younger counterpart recorded a 95% recognition rate for the disyllabic nouns compared to 81% for monosyllabic verbs. For the Yorùbá-Canadian participants, the difference was not so pronounced. The adults recorded 95% recognition rate for disyllabic items compared to 89% for monosyllabic ones while the younger participants recorded almost equal rate of recognition for disyllabic (59%) and monosyllabic (57%) lexical items. However, for all the four groups, the significance of the nature of the interaction, especially the differences between the accuracy in disyllabic and monosyllabic stimuli is yet to be determined. This will be revealed by subsequent statistical analyses that follow.

To investigate each group's lexical recognition accuracy with respect to the number of syllables of the stimuli, a model with an interaction term between Syllable and Group was fitted. The introduction of by-participant random slopes for the effect of the Syllable count on accuracy did not improve this model's fit. The output of the model summarized in Table 6.6 suggests that when the other groups were compared to the Yorùbá-Nigerian adult group, there was no significant difference in the accurate recognition rate for monosyllabic and disyllabic lexical items across all participant groups.

Table 6. 6 Coefficients of the glmer model of the interaction between Syllable count and Group. Random effects: Participant and Item. N=1892

Ref: Yorùbá-Nigerian Adult; Disyllabic;				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	19.00	30.28	0.627	0.530
Monosyllabic	-15.72	30.28	-0.519	0.604
YorCanAd	-14.92	30.28	-0.493	0.622
YorCanYg	-18.35	30.28	-0.606	0.544
YorNgrYg	-15.26	30.28	-0.504	0.614
Monosyllabic x YorCanAd	14.30	30.28	0.472	0.637
Monosyllabic x YorCanYg	15.20	30.28	0.502	0.616
Monosyllabic x YorNgrYg	14.14	30.28	0.467	0.641

6.3 Discussion

The Lexical experiment examines the native speakers' accurate recognition of real-word Yorùbá minimal pairs whose meaning differs due to a change of tone or tone patterns. The experiment also examines the effect of Age on participants' accuracy, in addition to the effect of participants' Language Context on their accurate recognition of the real Yorùbá monosyllabic and disyllabic words. Native Yorùbá-speaking participants were asked to listen to real Yorùbá words while looking at two pictures labelled 1 and 2 on the screen. Their task was to choose the picture that matched the word they heard. The pictures on the screen represented the two words in the minimal pair, while the word they heard was for one of the minimal pairs. Because the experiment differs from the previous experiment both in terms of method (only real Yorùbá words), and types of participants (only native speakers), the findings related to comparability or incomparability in the performance of native and non-native tone language speakers will not be alluded to in this discussion. The experiment shares some affinity with Yeh and Lu (2012) by involving only native speakers of the tone language of the study. As in Yeh and Lu (2012), the study investigates whether native speakers' knowledge and experience of the tone language will be reflected in the recognition of the lexical items.

The relative ease of the present experiment is reflected in the accuracy results where three of the four participant groups, the Yorùbá-Nigerian young and adults and the Yorùbá-Canadian adults, performed excellently with accuracy scores of between 88% and 95%. All participants from these three groups also scored better than the chance score. One main reason for this success may be the use of real words as opposed to nonce words (see, e.g., Fox & Unkefer, 1985; Lee et al., 1996). Using real words induces more engagement with the stimuli, and it allows participants, who are also of tone language background, to retrieve the word meaning and thus process the tone more accurately. In addition, the experience acquired from the three previous tasks, coupled with the pictures, which added more information to the stimulus, assist in the accurate processing of the lexical items.

The results of the statistical analysis in Table 6.4 highlight the Age factor, suggesting that generally, adults are significantly better than the young participants. Even though all the participants in this experiment are native Yorùbá speakers, the mean age for the two adult groups was far higher than the mean age of the young participant groups (YorNgrAd: $M_{age} = 43.5$; YorCanAd: $M_{age} = 40.8$; YorNgrYg: $M_{age} = 12.8$; YorCanYg: $M_{age} = 11.9$). This age gap means that the adults, being more experienced, possibly possess more lexical repertoire than the young participants. The adults may also be more familiar with experimental tasks of this nature than the children. The Age factor is also manifested when the Yorùbá-Nigerian adult group was compared to their younger counterparts. These are two groups with similar Language Context. The adults, with their experience due to their older age, were significantly better in this experiment than the young. This finding is in line with Hirsh, Morrison, Gaset, and Carnicer (2003) who suggest that for both monolinguals and bilinguals, the age of acquisition of a vocabulary item is a significant predictor of picture naming ability. Moreover, according to Herschensohn (2007:196), “since age of learning for any given item determines length of experience with that item, the variable is not maturational, but experiential”. Thus, in the absence of other factors, one may say that both Age and more experience with the lexical items are at work in tandem; older age cumulatively translates into more or better experience.

It is also possible to apply a similar assumption of Age and Experience as the main factors responsible for the highly significant performance by the Yorùbá-Canadian adults over their

younger counterparts in this experiment. However, in the present comparison, the poor performance by the younger participants may be due to age as well as incomplete acquisition of the L1, a description that applies to some members of this participant group. It may also be due to infrequent use of the L1 which can induce an attrition effect. Being away from the native language environment, the young Yorùbá-Canadian participants, more than the adults, are more prone to attrition effects (Montrul, 2008). This is further suggested by Yeh (2011) that if the frequency of use of a dominated language decreases, it lowers the probability of occurrence of lexical representations. This, in turn, increases the mismatch between acoustic signals and their corresponding lexical representations. Moreover, some members of this young age group have not fully acquired enough native language vocabulary before migration. For those who have acquired it, the native language vocabulary is systematically being displaced by another vocabulary due to infrequent usage of the L1.

Despite belonging to the same Age group, none of the two adult groups in this experiment, Yorùbá-Nigerian and Yorùbá-Canadian adults, posted significant differences from each other. This is possible because as adult speakers, they have comparable lexical experience. This aspect of the result is in line with Yeh and Lu (2012) where the older and the younger fluent speakers performed comparably at the lexical task. Both groups have acquired the native language along with the lexical items and have been using them since infancy. Even though the Yorùbá-Canadian adults have since migrated away from a native-language speaking environment, they reportedly continue using the language daily with their spouses and friends. Thus, the fact that they are away from the native language environment has not resulted in extensive loss of the previously-acquired lexical items. Their comparable performance in this task is evidence that the acquired lexical items are still intact and accessible whenever they are needed. Additionally, the fact that the stimuli were comprised of commonly-used lexical items, coupled with the inclusion of pictures, may also have assisted the Yorùbá-Canadian adults to perform comparatively well.

In contrast, the results for the two young participant groups in this experiment are not comparable. The Yorùbá-Nigerian young participants were significantly better than the Yorùbá-Canadian. The different results posted by these young groups underscore the fact that frequent use of the native language is more beneficial than infrequent use. The results further highlight the

importance of knowledge and experience of the tone language, a trend that has been outlined consistently by the young groups throughout this dissertation. Being Yorùbá-Canadian may have resulted in the lack of, or loss of, native language lexical knowledge. As such, even the presence of pictures might not have helped these young participants when confronted with those items they did not know or could not recollect. Moreover, this part of the results reflects the results of the young fluent speakers and the young attriters in Yeh and Lu (2012). The young fluent speakers performed better than the young attriters in the lexical task. Meanwhile, in comparing the young groups in Yeh and Lu (2012) to those in the present study, one issue to be considered is the difference in the mean age between the two groups. While the mean age for the young groups in the present study are close (YorNgrYg=12;8 yrs); YorCanYg=11;9 yrs), the mean age for the young groups in Yeh and Lu (2012) is not as close (young fluent speakers = 38;9 yrs; young attriters = 17.3 yrs). As mostly teenagers, the age difference could have informed the young attriters' attrition, hence, the significant difference in accuracy between the two groups.

Regarding the syllabicity factor, previous studies such as Perdereau (2013) tend to suggest that word processing in disyllabic words and polysyllabic phrases is less prone to the ambiguities that are situated in the lexical tone level. Thus, more recognition errors are to be expected in tone recognition for monosyllabic stimuli than disyllabic and polysyllabic stimuli. Despite these findings, the results of the present experiment do not suggest significant differences in accurate lexical recognition between monosyllabic and disyllabic items by the participant groups.

In sum, the results of the present experiment have shown that in typical situations, knowledge of the language acquired through more exposure to a language are crucial in the accurate recognition of the real-word stimuli of the language. Age probably proves to be more crucial in the case of the Yorùbá-Canadian adults whose performance is comparable to the Yorùbá-Nigerian adults and better than their younger counterparts. For this group, one may say that higher Age provides more lexical experience. Meanwhile, the observation that the Yorùbá-Canadian young participants posted the lowest accuracy scores among the four groups requires more study. This will be taken up later in Chapter Eight that looks at the group's demographic data which can give us more clues as to the reason for the lower accuracy rates posted by this group in this and in the previous experiments.

Chapter Seven: Tone Production

This chapter presents the tone production[12] experiment designed to test all participants' accurate imitation of the Yorùbá tones. Thus, in addition to the hypotheses about Age and Language Context, the experiment also tests participants' production patterns for the three simple Yorùbá tones and the two complex tones. The previous experiments (1-4) have investigated participants' perception of the Yorùbá tones. As the experiments progress, the tasks have revealed the increasing importance of knowledge of the tone language in tone perception. The present experiment shifts the focus from perception to production.

A previous study by Hao (2012), focussing on production of tones involving native tone language and non-tone language speakers, found that native tone language knowledge was not required for accurate imitation of tones. The study found no significant difference between native English and native Chinese speakers in imitating Mandarin tones. The study further found that participants were significantly more accurate at the imitation task than at identification or reading tasks. However, it is important to note that the non-tone language participants in Hao (2012) were all L2 learners of the tone language with significant amount of training in the tone language being studied. Another study by Omozuwa (1991) involved native and non-native speakers in a repetition task involving real words with Edo tones. The results showed that non-Edo speakers who were also tone language speakers and especially those whose tone languages were related to Edo recorded better performance than non-tone language speakers. Moreover, non-tone language speakers with previous training in tone manipulation (e.g., phoneticians, musicologists) also performed better than those who did not receive such training. A similar finding was reported in Yeh and Lu (2012) where older and younger fluent speakers of Hakka, with more knowledge and experience of the native tones, were significantly better than the young attriters in the reading task involving the Hakka low-level tones. Thus, similar to Hao (2012) and Omozuwa (1991), the present experiment adopted the repetition/imitation task.

[12] In the present study, production is used synonymously with repetition. It involves imitating from audio stimuli. This differentiates it from other forms of production experiments such as read-aloud and picture description which use orthographic (written) and pictorial (pictures) stimuli respectively.

The purpose of the present experiment is to investigate all listeners' accuracy at imitating Yorùbá tones. To this end, participants listen to recordings of monosyllabic and disyllabic nonce Yorùbá words made by a male native Yorùbá speaker and then repeat what they hear. The repeated tokens were recorded and later presented to two evaluators to identify them as either High, Mid or Low Yorùbá tones in the case of the monosyllabic words. The disyllabic patterns were to be identified as any of the nine possible Yorùbá disyllabic tone patterns. Secondly, the experiment seeks to investigate how the participants misrepresent the three tones as one another. This will shed more light on the tones which attract most misrepresentation of the other tones. Unlike the reading aloud method or production though picture description, which exerts much demand on the participants, this experiment only requires participants to listen and then repeat what they hear. As in the other experiments, the use of nonce words is adopted due to the presence of both native Yorùbá speakers as well as and non-tone language speakers among the participants. The evaluation method relies on the judgement of the two evaluators and where they disagree, it rests on me to decide on the tones or the tone patterns nearest to my own perceived judgement. The evaluators' judgement is further used as the basis for assessing participants' accuracy and for further analysis of the participants' utterances.

The acoustic difference between the three basic tones may influence their imitation pattern. Although the High tone is characterized as having the Highest F0, the highest F3 as well as the highest intensity (Bakare, 1995), available studies on the production of these tones suggest that both the High and the Low tones are produced better than the Mid tone by L2 learners. In her study of young and adult learners of Yorùbá as a second language (L2), Orie (2006b) observed that at the initial stage of acquisition, the pattern shown by L2 learners was to impose the High and the Low tones on all nouns based on the English stress pattern (Orie, 2006b:122). In the same study, these learners also identified utterance or word-initial Mid tones as High tones while utterance or word-final Mid tones were identified as Low tones. The study suggested that at that stage of their learning, participants possibly did not perceive the Mid tone as a distinct tone; hence, it was not manifested in their production (see examples in Chapter 1, Section 1.1.1).

The expectation that native Yorùbá speakers will perform well in this task is based on evidence from several studies (e.g., Fowler et al., 2003; Shockley et al., 2004) that suggest that L1 speakers of a language should be able to imitate the subtle phonetic differences that are not

contrastive in their native language. Despite the scarcity of research on the production of Yorùbá tones by non-native tone language speakers, it is nonetheless possible to predict the overall production outcome for native tone and non-tone language speakers. For instance, as previously noted in Omozuwa (1991), eight tone and non-tone language speakers were asked to repeat Edo bisyllabic nouns recorded by a female Edo speaker. The results revealed that Yorùbá and Ika speakers were more accurate at producing Edo tones than the French speakers (non-tone language speakers) or Chinese speakers who were also tone language speakers but whose language was not closely related to Edo. The study also revealed that some of the French and Chinese speakers who had previous professional training in pitch manipulation (phonetician, musicologist, speech therapist) performed better than those who did not have any training. Li (1995) also investigated native and non-native speakers of tone languages. The study found that native speakers of a tone language were better at producing the tones of another tone language than speakers of non-tone languages. The study further suggested that tone production difficulties were not the same for native speakers of tone and non-tone languages. Speakers of non-tone languages who have no previous tonal experience, might find it more difficult to instinctively categorize and then imitate the tones accurately. In another study comparing the accurate performance in perception versus production, Yang and Ankenmann (2007) compared the production and perception accuracy of American learners of Mandarin Chinese. They observed that the participants' tone production performance was much more accurate than their tone perception.

In sum, this experiment will examine the production pattern for the participants in the light of the hypotheses mentioned above. Additionally, the production error pattern will be analyzed to shed more light on whether participants follow certain patterns in their misrepresentation of the three basic and the two complex tones.

7.1 Methods

7.1.1 Participants

A total of one hundred and sixteen (116) participants took part in this experiment. Eleven (11) participants were from the Yorùbá-Canadian adult group and twenty-seven (27) participants

were from their younger counterparts. Sixteen (16) adults and twenty-seven (27) young participants were from the Nigerian-based Yorùbá-Nigerian group. Twenty-seven (27) adults and twenty- two (22) young participants were from the English-Canadian group.

7.1.2 Stimuli

The production stimuli (Table 7.1) consisted of 21 monosyllabic and 27 disyllabic items. In the monosyllabic items, each tone was represented seven (7) times, while in the disyllabic items, each of the nine possible tonal patterns was represented three times in the stimuli. All the stimuli were nonce words, and each stimulus was recorded once. The stimuli included all Yorùbá vowels excluding the nasal vowels. The stimuli also included all Yorùbá consonants except the double articulation /k͡p/ and /g͡b/ which could have posed some difficulty for non-tone language speaking participants.

Table 7. 1 Production stimuli: Monosyllabic and disyllabic items

Monosyllabic			Disyllabic		
H	M	L	HH-LH-MH	HM-LM-MM	HL-LL-ML
gọ́	dẹ̄	dọ̀	dúfẹ́	lékō	múdè
ló	fọ̄	fù	fọ́rú	rágī	sálò
sẹ́	gī	gà	tíkó	séyọ̄	yọ́tì
sú	rē	gò	ọ̀dẹ́	ẹ̀yọ̄	dèlù
tá	sā	jẹ̀	tìkó	ìdē	ìlọ̀
tí	tō	jì	wùsá	mọ̀kā	wàgò
yé	wū	lè	rẹ̄kú	jẹ̄mī	jẹ̄tì
			sāwé	kōdẹ̄	kōdù
			tōfú	wārē	sāgè

7.1.3 Procedure

In this production task, participants listened to a word and then repeated what they heard. Their repetition was spoken into a headset microphone connected to a Sony ICD-PX333 digital

recorder. During the practice session, two monosyllabic words, as well as two disyllabic words, were played for participants to imitate. After completing the practice session, the main task began once participants did not have further questions to ask the experimenter. All participants went through the two blocks of the experiment in which monosyllabic stimuli formed the first block of the experiment while disyllabic stimuli formed the second block.

Two fluent native Yorùbá speakers assisted in evaluating the recordings of the participants. They were instructed to identify the tone of each syllable as either High, Mid, or Low, in the case of the monosyllabic items, and to identify the tonal patterns, which could be any of the nine possible tonal combinations, in the case of the disyllabic items. The two evaluators worked independently of each other. The rates of agreement between the evaluators for the participant groups' production stimuli were as follows: Yorùbá-Nigerian adult (98%); Yorùbá-Nigerian young (95%); Yorùbá-Canadian adults (95%); Yorùbá-Canadian young (90%); English-Canadian adult (87%); English-Canadian young (85%). In cases of disagreement between the two raters, I listened to all the recordings to select the rating that was the same or the closest to my own judgement.

7.2 Results

Table 7.2 presents the results of the production task for the six participant groups. It shows their accuracy in both the monosyllabic (MS) and disyllabic (DS) stimuli. The first row lists the six participant groups with the number of participants from each group in brackets in the second row. The accuracy scores (ACC) in the Monosyllabic stimuli are listed in the third row followed by the percent accuracy in brackets in the fourth. In the fifth row, the accuracy scores for the disyllabic stimuli are listed with their percent accuracy in brackets in the sixth row. The seventh and the eighth rows present the total accuracy scores and the percent accuracy (in brackets) respectively. The results show that three participant groups, the Yorùbá-Nigerian young, Yorùbá-Nigerian adult, and the Yorùbá-Canadian adult, recorded perfect imitation for the three tones (between 97% to 100%), while the other three groups recorded percent accurate production rate ranging from 66% to 79%.

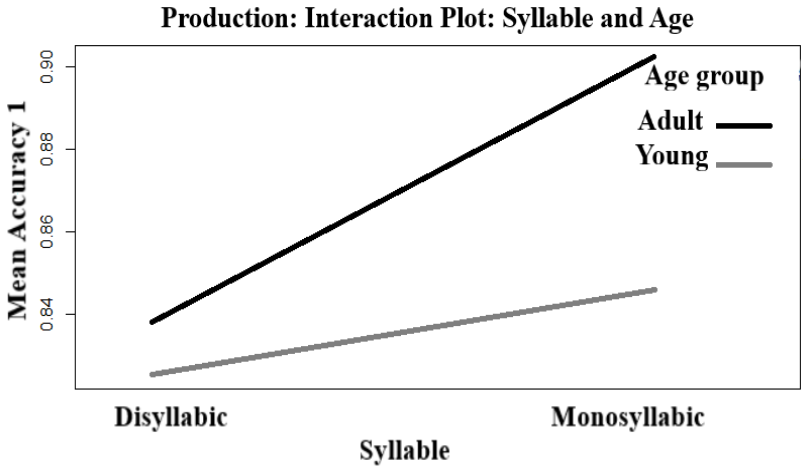
Table 7. 2 Accuracy results for the production tasks for all participant groups: MS=Monosyllabic; DS=Disyllabic; ACC=Accuracy Score; N= No. of participants

		YorNgrAd	YorNgrYg	YorCanAd	YorCanYg	EngCanAd	EngCanYg
		(N=16)	(N=27)	(N=11)	(N=13)	(N=27)	(N=22)
MS	ACC	334/336	562/567	224/231	220/273	465/567	316/462
	%	(100)	(99)	(97)	(81)	(82)	(68)
DS	ACC	431/432	726/729	290/297	270/351	501/729	384/594
	%	(100)	(100)	(98)	(77)	(69)	(65)
TOTAL	ACC	765/768	1288/1296	514/528	490/624	966/1296	700/1056
	%	(100)	(99)	(97)	(79)	(75)	(66)

7.2.1 Statistical Analysis

Figure 7.1 presents the plot of the interaction between Syllable count and Age on tone production. The plot indicates that adults were more accurate than the young groups in the production of both monosyllabic and disyllabic stimuli. However, the interaction is such that the young participants were significantly worse than the adults in producing monosyllabic stimuli when compared to their production of disyllabic stimuli. To get the statistical significance of the effects observed in Figure 7.1, the next paragraphs present the analysis.

Figure 7. 1 Mean accuracy in Experiment 5 by Age group (adult, young) and Syllable count (monosyllabic, disyllabic)



As in the other experiments, Lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015) in R was used to create regression models to check for the general effects of some identified fixed effects on participants' production accuracy. The first model examined the relationship between participants' tone production accuracy and the three factors of Age, Language Context, and Syllable count. As fixed effects, this model included Age with two levels: *Adult* and *Young*; Language Context had three levels: *Yorùbá-Nigerian, Yorùbá-Canadian*, and *English-Canadian* and Syllable had two levels: *Monosyllabic* and *Disyllabic*. The accuracy response, which consisted of the individual score for the 48 target stimuli was the dependent variable. As random effects, the model had intercepts for the forty-eight stimuli coded as *Item*, as well as intercepts for the young and adult subjects coded as *Participant*. The model was based on 5568 observations (a total of 116 participants x 48 individual item score).

P-values were obtained by likelihood ratio tests of both the main-effects model and another model with an interaction between the two factors of Syllable and Age. An ANOVA revealed that the model with an interaction term between Syllable and Age fitted the data more closely (AIC = 3292 vs. AIC(interaction: Syll: Age = 3285). A model with an interaction between Syllable and Language Context yielded the same fitting as the simple main effects model (AIC=Interaction: Syll: Lang.Context =3292.1). The introduction of by-participant random slopes for Syllable improved the model's fit.

To allow for comparisons to be made across the three Language Context levels, the main-effects model had the reference level for the Language Context factor alternated, resulting in two comparable analyses. The first analysis had Yorùbá-Nigerian as the reference level, while the second analysis had Yorùbá-Canadian as the reference level by default. The outputs of the model are summarized in Table 7.3.

Table 7. 3 Coefficients of the glmer models of the effects of the interaction between Syllable count and Age on tone production. Random effects: Participant and Item. N=5568

Ref: YorNgr				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	6.5799	0.4781	13.764	< 2e-16***
Monosyllabic	0.9144	0.3476	2.631	0.00852**
Young	-0.7174	0.2666	-2.691	0.00713**
YorCan	-3.4421	0.4545	-7.574	3.62e-14***
EngCan	-5.3156	0.4258	-12.483	< 2e-16***
Monosyllabic x Young	-0.5903	0.2503	-2.358	0.01837*
Ref: YorCan				
Intercept	3.1378	0.3638	8.625	< 2e-16***
EngCan	-1.8735	0.3069	-6.105	1.03e-09***

The analysis in Table 7.3 suggests that younger participants produced the tones significantly less accurately than the adults. Moreover, the results suggest that globally, the monosyllabic stimuli recorded slightly better tone production accuracy than disyllabic stimuli. On the general effects of Language Context, the analysis reveals that when Yorùbá-Canadian and English-Canadian were compared to Yorùbá-Nigerian, both Language Context types resulted in significantly worse tone production accuracy than Yorùbá-Nigerian. Similarly, a comparison between Yorùbá-Canadian and English-Canadian reveals that English-Canadians had significantly worse tone production accuracy than Yorùbá-Canadian. Additionally, the results reveal that Monosyllabic stimuli attracted better production accuracy than disyllabic stimuli. As suggested by the results, the nature of the interaction between Syllable and Age was such that younger participants were significantly worse at producing monosyllabic stimuli than their production of disyllabic stimuli when compared to the adult participants.

7.2.2 Tone production by individual groups

Another model was fitted to compare the performance of the participant groups. This model enabled the comparison between individual groups belonging to the same or different Age groups or Language Contexts. The model had two fixed effects, Syllable, with two levels, *Monosyllabic* and *Disyllabic*; and Group with six levels, *Yorùbá-Canadian Adult, Yorùbá-*

Nigerian Adult, English-Canadian Adult, Yorùbá-Canadian Young, Yorùbá-Nigerian Young, and *English-Canadian Young*. The model had random intercepts for *Participant* and *Item*. The dependent variable was the accuracy score consisting of the individual score for the 48 target stimuli. The model was based on 5568 observations (a total of 116 participants x 48 individual item score=5568). P-values were obtained by likelihood ratio tests of the simple model with only the main effects and another model with by-participant random slopes for the effect of Syllable on tone production accuracy. An ANOVA revealed that the model with by-participant random slopes with Syllable (AIC=3281.2) fitted the data more closely than the model without random slopes (AIC = 3292.1). Table 7.4 presents the summary of the model's output.

Table 7. 4 Coefficients of the glmer models of the tone production accuracy by individual participant groups. Random effects: Participant and Item. N=5568

Ref: YorNgrAd				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	6.1660	0.6741	9.147	< 2e-16***
YorCanAd	-2.0304	0.7609	-2.668	0.00762**
EngCanAd	-5.0641	0.6804	-7.443	9.82e-14***
YorCanYg	-4.4050	0.7040	-6.257	3.92e-10***
YorNgrYg	-0.2868	0.7740	-0.370	0.71103
EngCanYg	-5.4435	0.6783	-8.025	1.02e-15***
Ref: YorCanAd				
Intercept	4.1356	0.4620	8.952	< 2e-16***
EngCanAd	-3.0337	0.4717	-6.431	1.27e-10***
YorCanYg	-2.3746	0.5087	-4.668	3.04e-06***
YorNgrYg	1.7437	0.6044	2.885	0.00391**
EngCanYg	-3.4130	0.4695	-7.270	3.60e-13***
Ref: EngCanAd				
Intercept	1.1019	0.2824	3.902	9.55e-05***
YorCanYg	0.6591	0.3614	1.824	0.0682.
YorNgrYg	4.7774	0.5013	9.530	< 2e-16***
EngCanYg	-0.3793	0.3109	-1.220	0.2224
Ref: YorNgrYg				
Intercept	5.8793	0.4914	11.963	< 2e-16***
YorCanYg	-4.1183	0.5309	-7.757	8.68e-15***
EngCanYg	-5.1567	0.4967	-10.381	< 2e-16***
Ref: YorCanYg				
Intercept	1.7610	0.3526	4.994	5.91e-07***
EngCanYg	-1.0384	0.3668	-2.831	0.00464**

In Table 7.4, the first comparison examines the performance of young and adult participants in the three Language Context categories. With Yorùbá-Nigerian adult as the reference level, the analysis suggests that while there was no significant difference in tone production accuracy between Yorùbá-Nigerian adult and young participants (also in Table 7.4, Ref: Yorùbá-Nigerian young), the Yorùbá-Nigerian adults produced the tones significantly better than the young members from both the Yorùbá-Canadian and English-Canadian categories. When the Yorùbá-Canadian adult group was specified as the group reference level, the analysis suggests a significant difference between Yorùbá-Canadian adults and their younger counterparts. Moreover, while the Yorùbá-Canadian adults produced the tones significantly better than their younger counterparts as well as the young group from the English-Canadian category, the young Yorùbá-Nigerian participants produced the tones significantly better than the Yorùbá-Canadian adults. With English-Canadian adult as the reference level, the results indicate that there was no significant difference in tone production accuracy between the English-Canadian adults and their younger counterparts. Neither was there any significant difference in tone production accuracy between this adult group and the Yorùbá-Canadian young group. However, young participants from the Yorùbá-Nigerian category produced the tones significantly better than the English-Canadian adults.

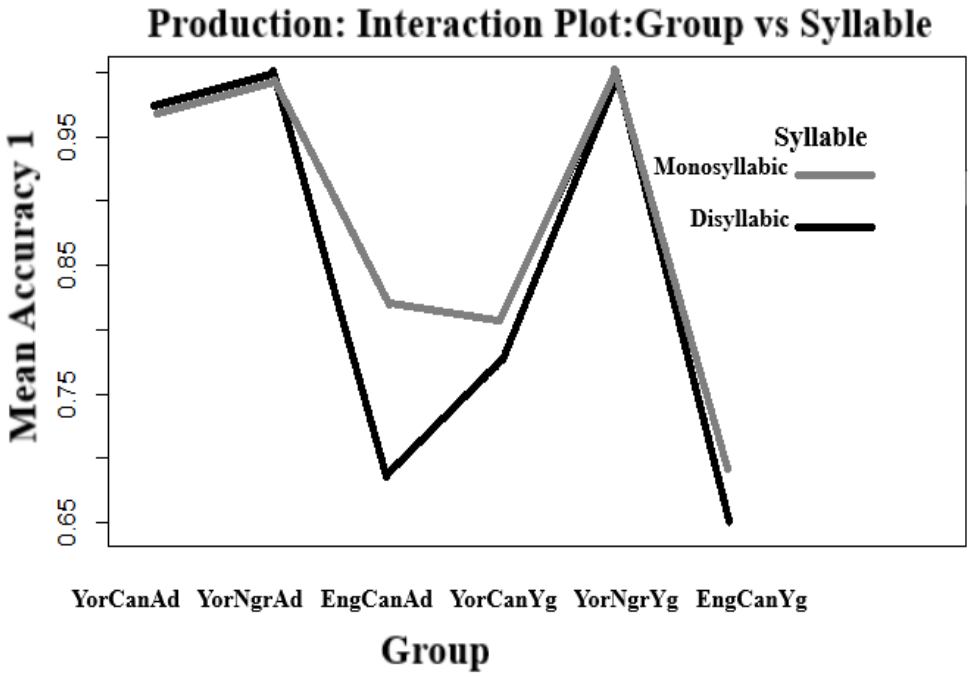
The second comparison examines the groups' performances across Language Context types. It compared participants belonging to the same age group (e.g., adult vs. adult; young vs. young) but with different Language Context. In the adult groups, when Yorùbá-Nigerian adult group was the reference level, the analysis suggests that the Yorùbá-Nigerian adult group was significantly better at tone production accuracy than either the Yorùbá-Canadian or English-Canadian adults. With Yorùbá-Canadian adults as the reference level, the analysis suggests that the English-Canadian adults produced the tones significantly worse than the Yorùbá-Canadian adults.

In the younger age groups, the analysis reveals a similar trend to what was observed for the adults. With Yorùbá-Nigerian young as the reference level, the analysis reveals that young Yorùbá-Nigerian participants were significantly better in tone production accuracy than either the Yorùbá-Canadian or English-Canadian young participants. Meanwhile, when the Yorùbá-

Canadian young group was the reference level, these young participants were significantly better at the tone production task compared to their English-Canadian counterparts.

7.2.3 Production of Monosyllabic and disyllabic items by individual groups

Figure 7. 2 Mean accuracy in Experiment 5 by syllable count (monosyllabic, disyllabic) and the six participant groups



The interaction between Syllable count and participant groups regarding production accuracy is further plotted in Figure 7.2. The plot indicates that the accurate performance of three groups, namely the Yorùbá-Nigerian adult, Yorùbá-Nigerian young, and the Yorùbá-Canadian adult is better than the remaining three groups in the production of both the monosyllabic and disyllabic stimuli. The figure further reveals that only two groups, the English-Canadian adults and the Yorùbá-Canadian young appeared to have effects for Syllable count while the other four groups did not. In what follows, a statistical model that looks at the interaction between group and syllable count will explore whether the effects shown in Figure 7.2 are significant or not.

The model analyzed each participant group's production accuracy with respect to monosyllabic and disyllabic items. This was done by introducing interaction term between the Syllable factor and the Group factor. In addition to the intercepts for the two random effects of *participant* and *item*, the model also had by-participant random slopes for the effect of the Syllable count on production accuracy. The introduction of the random slopes improved the model's fit (AIC=3285.8 vs. AIC (random slopes) = 3280.6). The summary of the results of the best model is presented in Table 7.5.

Table 7. 5 Coefficients of the glmer models of the production of monosyllabic and disyllabic items by participant groups. Random effects: Participant and Item. N=5568

Ref: Yorùbá-Nigerian Adult; Disyllabic				
	Estimate	Std. Error	Z value	Pr(>|z|)
Intercept	6.7924	1.0487	6.477	9.37e-11***
Monosyllabic	-0.6898	1.2769	-0.540	0.5891
YorCanAd	-2.4331	1.1383	-2.137	0.0326 *
EngCanAd	-5.7657	1.0493	-5.495	3.92e-08***
YorCanYg	-5.0033	1.0713	-4.670	3.01e-06***
YorNgrYg	-0.5531	1.1966	-0.462	0.6439
EngCanYg	-6.0145	1.0537	-5.708	1.14e-08***
Monosyllabic x YorCanAd	0.6415	1.3717	0.468	0.6400
Monosyllabic x EngCanAd	1.7394	1.2537	1.387	0.1653
Monosyllabic x YorCanYg	1.1670	1.2718	0.918	0.3588
Monosyllabic x YorNgrYg	0.4304	1.4632	0.294	0.7687
Monosyllabic x EngCanYg	0.9940	1.2578	0.790	0.4294

The analysis reveals that generally, there was no significant effect for Syllable count on participants' production accuracy, even though the positive coefficient estimates signified that monosyllabic stimuli tended to attract better production accuracy than disyllabic stimuli. Thus, Table 7.5 indicates that the Syllable count effects shown for the English-Canadian adult and the Yorùbá-Canadian young groups in Figure 7.2 were not significant. The next section will present the analysis of participants' response to further reveal their tone production patterns.

7.3 Response Analysis

7.3.1 Monosyllabic stimuli

Table 7. 6 Percent accurate production of monosyllabic stimuli by the six participant groups

	High	Mid	Low	Total	N
YorCanAd	97%	97%	98%	97%	11
YorNgrAd	99%	100%	100%	100%	16
EngCanAd	72%	75%	98%	82%	27
YorCanYg	60%	95%	87%	81%	13
YorNgrYg	99%	98%	100%	99%	27
EngCanYg	40%	77%	88%	68%	22
Total	78%	90%	95%	88%	116

Table 7.6 presents each group's accurate production of the monosyllabic stimuli. It presents the accuracy percentage in each of the three tones. The numbers indicate that the Low tone recorded the highest rate of accurate production overall. The table also reveals that the three tones attracted similar rates of production accuracy by the Yorùbá-Nigerian young and adult groups as well as the Yorùbá-Canadian adult group. The Low tone recorded a higher rate of accuracy than either the High or the Mid tone by both the young and adult English-Canadian groups. The Yorùbá-Canadian young group represented the only instance where the Mid tone was produced more accurately than either the Low or the High tone. Moreover, both Mid and High tones were produced at almost equal rates by most participant groups (Yorùbá-Nigerian young and adult groups, Yorùbá-Canadian adult, and English-Canadian adult groups), but Mid tones were produced more accurately than High tone by the young participant groups from the Yorùbá-Canadian and English-Canadian categories. Both the young and adult groups from English-Canadian category produced the Low tones more accurately than either the High or Mid tones. However, the lowest percent accuracy in this task (40%) was in the production of the High-tone recorded by the English-Canadian young group.

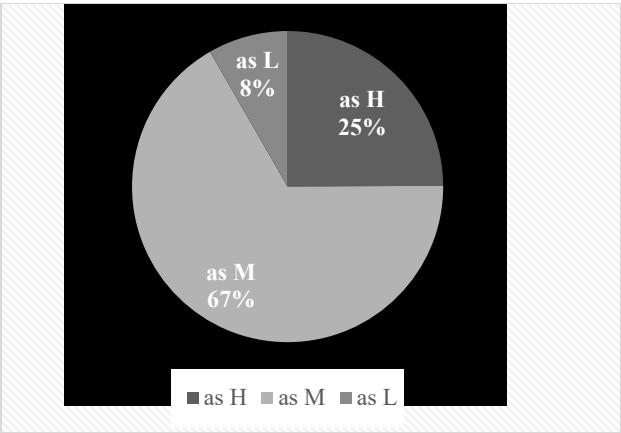
At this juncture, the analysis of the results will focus on only three participant groups, namely the Yorùbá-Canadian young group, the English-Canadian adults, and the English-Canadian young group. This is because the other three participant groups (Yorùbá-Nigerian young and adults and

Yorùbá-Canadian adults) have displayed optimal or near-optimal tone production accuracy, leaving us to look elsewhere for hints about participants' production behaviour. In what follows, the study will analyze the production errors for the three groups to identify each group's tone production pattern.

7.3.2 Monosyllabic Production errors

Monosyllabic items provide one of the perfect avenues to observe for production errors due to the absence of a penultimate syllable that could affect the production of the target syllable. Figure 7.3 shows that overall, a total of 301 monosyllabic stimuli were misrepresented by the three focus groups. Out of these, 201 items (67%) were misrepresented as Mid tones, 75 items (25%) were misrepresented as High tones, and 25 items (8%) were misrepresented as Low tones, as shown in Figure 7.3.

Figure 7. 3 Misrepresentation of Monosyllabic tones by the three participant groups



7.3.3 Individual group's tonal misrepresentation patterns

Table 7.7 presents the misrepresentation patterns for the monosyllabic items by the three participant groups. The table presents the number of items misrepresented by the participant groups along with the direction of their misrepresentation. The first column indicates the target tones while the second row indicates the tones which the target tones were misrepresented as.

The numbers misrepresented are presented in the table while the numbers in brackets represent the percentage of the total misrepresentation.

Table 7. 7 Monosyllabic Production: Confusion Matrix: H=High; M=Mid; L=Low

	YorCanYg			**EngCanYg**			**EngCanAd**		
as→	H	M	L	H	M	L	H	M	L
H	-	35	1	-	86	7	-	49	3
%		(97)	(3)		(92)	(8)		(94)	(6)
M	4	-	1	26	-	9	43	-	4
%	(80)		(20)	(74)		(26)	(83)		(7)
L	0	12	-	0	18	-	2	1	-
%		(100)			(100)		(67)	(33)	
TOTAL	4	47	2	26	104	16	45	50	7
%	(8)	(89)	(4)	(18)	(71)	(11)	(44)	(49)	(7)

The Yorùbá-Canadian young group misrepresented a total of 53 tones (18% of all misrepresentations). Out of these, 47 tones (89%) were misrepresented as Mid tone, while four (8%) were misrepresented as High tones. Only two tones (4%) were misrepresented as Low. It should be noted that 35 out of the 47 items misrepresented as Mid tones (74%) originated from High tones, while all the four remaining items misrepresented as High tone originated from Mid tones. The English-Canadian young group misrepresented a total of 146 tones (48% of the total misrepresentations), out of which 104 (71%) were misrepresented as Mid tones. 26 items (18%) were misrepresented as High tones, while 16 items (11%) were misrepresented as Low tones. Interestingly, 83% (86 items) of the items misrepresented as Mid tones were High tones, while all the 26 items (100%) misrepresented as High tones were Mid tones. As the only adult participant group in this analysis, the English-Canadian adults misrepresented a total of 102 items (34% of all misrepresentations). 50 of these items (49%) were misrepresented as Mid tone. 45 items (44%) were misrepresented as High tone while seven items (7%) were misrepresented as Low tone. Of all the 50 items misrepresented as Mid tone, 49 (98%) were High tones. In a similar vein, 43 of all the items misrepresented as High tone (i.e., 95%) were Mid tones. In sum, the data revealed that the three groups mostly misrepresented High tones as Mid. Conversely, Mid tones were mostly misrepresented as High by the two English-Canadian groups, while the

Low tone, with the lowest misrepresentation rate, was mostly misrepresented as Mid by the Yorùbá-Canadian and English-Canadian young groups. The next sections of the analysis examine participants' production and misrepresentation of the disyllabic stimuli.

7.3.4 Disyllabic tonal patterns

Table 7.8 presents the percent accurate production of disyllabic tonal patterns by all the six groups. Three groups, the Yorùbá-Nigerian young, Yorùbá-Nigerian adult, and Yorùbá-Canadian adult groups, recorded the highest percent accuracy with 100%, 100%, and 98% respectively. The other three groups, the Yorùbá-Canadian young group, English-Canadian adult, and English-Canadian young groups scored 77%, 69%, and 65% respectively. This is quite similar to the trend observed for the monosyllabic stimuli. Thus, similarly, subsequent analysis of the disyllabic results will include only the three latter participant groups.

Table 7. 8 Percent accurate production of disyllabic stimuli by the six participant groups

P/GRPS	***N***	**TONAL PATTERNS**									**TOTAL**
		HM	HL	HH	MH	ML	MM	LH	LM	LL	
YorNgrAd	*16*	100%	100%	100%	100%	100%	100%	100%	98%	100%	100%
YorCanAd	*11*	100%	100%	94%	97%	100%	94%	97%	97%	100%	98%
YorNgrYg	*27*	100%	100%	99%	100%	100%	99%	100%	99%	100%	100%
Total	*54*	100%	100%	98%	99%	100%	98%	99%	98%	100%	99%
EngCanAd	*27*	75%	51%	52%	70%	81%	62%	98%	54%	75%	69%
YorCanYg	*13*	82%	59%	67%	82%	87%	72%	79%	82%	82%	77%
EngCanYg	*22*	56%	38%	35%	76%	77%	68%	94%	76%	62%	65%
Total	*62*	70%	48%	49%	75%	81%	66%	92%	68%	72%	70%

The analysis begins with the overall production accuracy for the nine disyllabic tonal patterns for the three participant groups presented in the second part of Table 7.8 (in the last four rows). The data indicate that globally, the three participant groups recorded the highest rate of production accuracy in the LH pattern (92%) while the lowest rates were recorded in the pattern containing

the Yorùbá falling contour (HL) (48%) and the pattern with the High-High tonal sequence (HH) (49%). The six remaining patterns recorded accuracy rates of between 66% and 81%. Further examination of the results reveals that the HH and HL patterns attracted the lowest accuracy rates for all the three groups while LH pattern attracted the highest accuracy rates for the two English-Canadian groups. The Yorùbá-Canadian young group performed well in all the remaining patterns with ML (87%) as their best-produced pattern. The data also reveals the patterns that the participants found more challenging to produce. They are represented by those patterns with the lowest rates of accurate production in the table. Thus, for all three groups, HL and HH were the most misrepresented patterns. However, among the three groups, the production error rates for the English-Canadian young group for the two patterns were higher (HL=62%; HH=65%) than what was recorded by the English-Canadian adult group (HL=49%; HH=48%), or the Yorùbá-Canadian young group (HL=41%; HH=33%). Further analysis of participants' error rates follows in the next section.

7.3.5 Disyllabic Production error analysis

Table 7. 9 Disyllabic Production: Error Matrix

as→	**HM**	**HL**	**HH**	**MH**	**ML**	**MM**	**LH**	**LM**	**LL**	**TOTAL**
HM		0	2	3	2	41	6	1	2	57
%		(0)	(4)	(5)	(4)	(72)	(11)	(2)	(4)	(11)
HL	13		0	7	63	7	1	3	3	97
%	(13)		(0)	(7)	(65)	(7)	(1)	(3)	(3)	**(19)**
HH	4	0		38	2	34	14	2	1	95
%	(4)	(0)		(40)	(2)	(36)	(15)	(2)	(1)	**(18)**
MH	3	0	5		0	16	20	3	0	47
%	(6)	(0)	(11)		(0)	(34)	(43)	(6)	(0)	(9)
ML	3	17	0	2		7	1	0	4	34
%	(9)	(50)	(0)	(6)		(21)	(3)	(0)	(12)	(7)
MM	16	1	4	27	0		11	4	0	63
%	(25)	(2)	(6)	(43)	(0)		(17)	(6)	(0)	(12)
LH	1	0	0	7	1	1		4	0	14
%	(7)	(0)	(0)	(50)	(7)	(7)		(29)	(0)	(3)
LM	1	0	0	1	3	10	44		1	60
%	(2)	(0)	(0)	(2)	(5)	(17)	(73)		(2)	(12)
LL	4	1	0	1	26	2	5	13		52
%	(8)	(2)	(0)	(2)	(50)	(4)	(10)	(25)		(10)
TOTAL	45	19	11	86	97	118	102	30	11	519
%	(9)	(4)	(2)	**(17)**	**(19)**	**(23)**	**(20)**	(6)	(2)	(100)

Table 7.9 presents the results of the analysis of the disyllabic production errors. It shows the number and the percentages (in brackets underneath) of the misrepresentation of the disyllabic patterns. In the table, the first column indicates the target patterns while the first row indicates the patterns which the target patterns were produced as. The last column indicates the total number and the percentages (in brackets underneath) of the misrepresentation of the target patterns. The last row indicates the total number and the percentages (in brackets, underneath) of the patterns attracting most misrepresentations of the target patterns.

In the last column (in bold), the table reveals that the HL and the HH patterns were the most misrepresented at the rates of 19% and 18% respectively. The LH pattern was the least misrepresented with only 3% of total misrepresentation. Other patterns were misrepresented at rates of between 7% and 12%. Meanwhile, in the last row (in bold fonts), four patterns, MM (23%), LH (20%), ML (19%), and MH (17%) mostly appear to attract the misrepresentation of the target patterns. In other words, those were the patterns that the target patterns were mostly produced as. Three patterns, LL (2%), HH (2%), and HL (4%) attracted the lowest rates of misrepresentation of the target patterns. Table 7.10 examines how the production error patterns apply to the three participant groups.

Table 7. 10 Patterns attracting most misrepresentations of the other patterns

GROUPS	**N**	**TONAL PATTERNS MISREPRESENTATION AS:**								
		HM	HL	HH	MH	ML	MM	LH	LM	LL
YorCanYg	81	15	2	4	16	12	16	6	8	2
%	(100)	(19)	**(2)**	**(5)**	(20)	(15)	(20)	(7)	(10)	**(2)**
EngCanYg	210	15	6	3	40	34	74	20	12	6
%	(100)	(7)	**(3)**	**(1)**	(19)	(16)	(35)	(10)	(6)	**(3)**
EngCanAd	228	15	11	4	30	51	28	76	10	3
%	(100)	(7)	**(5)**	**(2)**	(13)	(22)	(12)	(33)	(4)	**(1)**
Total	519	45	**19**	**11**	86	97	118	102	30	**11**
%	(100)	(9)	**(4)**	**(2)**	(17)	(19)	(23)	(20)	(6)	**(2)**

Table 7.10 presents a breakdown of the production error patterns with respect to each of the three groups. It shows how the three participant groups misrepresented the target patterns. The second row indicates the patterns which the target stimuli were misrepresented as. The number of stimuli misrepresented are presented along with the percentage of total misrepresentation in brackets underneath. The table reveals that while the Yorùbá-Canadian young group misrepresented other patterns mostly as MM (20%), MH (20%), or HM (19%), the English-Canadian young group misrepresented other patterns mostly as MM (35%) and MH (19%). The English-Canadian adult group differed from these other two groups. They misrepresented other patterns mostly as LH (33%) and ML (22%). Thus, the results for the three participant groups revealed that most participants resorted to four patterns, MM, LH, ML, and MH, whenever they faced confusion regarding the tones of some of the disyllabic stimuli. Among the four patterns, MM was the most prominent, followed by LH. On the other end of the scale, HL, and the double-toned HH and LL, attracted the lowest rates of misrepresentation.

In sum, participants' tonal misrepresentation patterns revealed the following trends: First, other patterns were misrepresented as MM and MH by both the Yorùbá-Canadian and English-Canadian young groups. In addition to these two patterns, Yorùbá-Canadian young group also resorted to HM as an alternative. Second, for the three groups, ML also attracted substantial misrepresentation of other patterns. Moreover, the English-Canadian adult group, more than any other group, resorted mostly to LH pattern to represent other patterns. Last, in addition to the complex HL patterns, the two double-toned patterns (HH and LL) attracted the fewest misrepresentation of the other patterns across all three participant groups.

7.4 Discussion

In this experiment, participants listened to Yorùbá-toned monosyllabic and disyllabic nonce-word stimuli and then repeated what they heard. This is a simple imitation task which requires auditory perception followed by motor production. This makes it different from other types of production tasks such as a reading aloud task, which would have required more skills (Hao & Jong, 2016). The purpose is to investigate the effects of Age and tone Language Context on accurate tone production. In addition, the experiment examines participants' tone production

pattern with respect to the tones they most often imitated correctly or incorrectly. Finally, the experiment investigates how participants chose to represent the tones that they often produced incorrectly.

The relatively better performance recorded by all participants in this task indicates that, irrespective of their language background, they all possess the skills required to produce Yorùbá tones. The results further support Hao and Jong (2016) as well as Hao (2012) which observed that tonal production (imitation) tasks appear to be easier than tasks that require phonological categorization (identification).The results also support Connell's (2000) finding that pitch discrimination is better than tone categorization in non-tone language speakers.

Despite the observation above, analysis of their performance along the Age factor suggests that adults are significantly better than the young participants (Table 7.3). In this, as well as in the previous experiments, maturity, which affords the adults more language experience, will continue to be cited for the performance. On Language Context effects, the analysis suggests that Yorùbá-Nigerians are more accurate at tone production than Yorùbá-Canadians or English-Canadians, while the Yorùbá-Canadians participants are more accurate than the English-Canadians. These findings highlight the fact that frequent use of the tone language leads to improved tone production accuracy compared to more infrequent (or no) use. This finding is in line with Omozuwa (1991) where native speakers of the target tone language as well as speakers of other tone languages similar to the target tone language outperformed non-native speakers. In the same study, non-native speakers with experience in tone manipulation also performed better than other non-native speakers. The finding of the present study also reflects Yeh and Lu (2012) where native speakers of the tone language with more experience and knowledge of the tone language outperformed those with less knowledge and experience. Meanwhile while comparing the results with those of Omozuwa (1991) and Yeh and Lu (2012), there is the need to mention that the two studies used real words as stimuli in contrast to nonce words used for the present task. Additionally, Yeh and Lu (2012) adopted the reading task, which could have been more difficult than the repetition task adopted in the present study.

An aspect of the result of the present study where native speakers of the tone language outperformed the non-native speakers (also Omozuwa, 1991) does not fall in line with Hao (2012), where there was comparable performance between the two groups. This might be

expected if we consider the difference between the non-native participants in the two studies. Unlike the present study (also Omozuwa, 1991), where the non-native participants did not undergo any training before the test, the participants in Hao (2012) were all university students learning Mandarin as an L2 in the United States. Some of them had been learning Mandarin for at least 2.5 years or more prior to the time of the study. This is a factor which might have contributed to the comparable performance observed in the study.

Within the Yorùbá-Nigerian and English-Canadian groups, the analysis does not reveal any significant difference between the young and the adult members, suggesting that there is no Age factor effect. This points to the fact that, not only does the production task impose less memory demand on all participants, but also that the levels of tone language experience for the young and the adult members of these groups are somewhat similar especially for the English-Canadian group. As for the Yorùbá-Nigerian group, while not disputing the fact that the adults, by definition, have more cumulative exposure to tones, we can also say that the young participants have received the amount of input that is necessary for correct use of tones. The comparable performance by the Yorùbá-Nigerian young group may also indicate that children are more adept at mimicry. In contrast, the Yorùbá-Canadian adult group showed significantly better performance than their younger counterparts. While one cannot entirely rule out the effects of the Age factor, the Language Context effects appear to be a major factor for the discrepancy displayed by these two age groups. Despite belonging to the same Language Context characterized by limited and infrequent use of the native tone language, the tone language levels of the two age groups are not that similar. The adult members are educated professionals, including some parents of the younger participants. Except for the workplace where they use the dominant language, these adults reportedly still use the native tone language daily at home in conversation with their spouses and, in some cases, their children and family friends. In comparison, the young members have diverse tone language contexts. Seventeen of the children were born in Nigeria while thirteen were born outside Nigeria (Canada = 5; Kuwait = 7; USA = 1) but were brought to Canada by their parents. In all these cases, most of the participants rarely used the language willingly at home, even though they were exposed to it daily through listening to their parents. Thus, the fact that the adults still used the native language daily while the younger ones did not may explain the significant effect recorded for this Language Context category.

The analysis suggests that Language Context is an essential factor. For both the adult and the young groups, the analysis reveals that there is gradient effect of Language Context on participants' performance in the production task. It suggests that an enhanced tone Language Context contributes to better tone production accuracy. Yorùbá-Nigerian is better than Yorùbá-Canadian while Yorùbá-Canadian is better than English-Canadian. Finally, the analysis for the general syllabicity factor on participants' production reveals no significant effect, suggesting that the Syllable count for the stimuli did not affect participants' production accuracy.

7.4.1 Monosyllabic production

Analysis of the monosyllabic production results involving the three basic tones reveals that uniformly, High tones recorded the lowest percent accuracy by the three participant groups (Table 7.6). The results posted by these three participant groups give an initial indication that production of Yorùbá tones may not be the same as the perception of the tones. It further reveals that the production of the High tone in a monosyllabic context is not the same as producing it in a disyllabic context. For participants with a non-tone language background, equating a High tone with the English primary stress may not be the answer. This is because the Yorùbá High tone may require a breathier phonation or a pitch height more than that required to produce the English primary stress. Not reaching that pitch height during its articulation may render such production as Mid or even as Low tone. Another observation is that the English-Canadian adult and young groups produced the Low tones more accurately (with 98% and 88% respectively) than the Mid tones (75% and 77% respectively) (Table 7.6). Yorùbá Low tone typically has a characteristic voice quality, according to Hayward et al. (2003), and a more distinctive perceptual characteristic (Hombert, 1977). To produce the tone, speakers need to descend to a distinct vocal register otherwise described as creaky phonation. If this fact is true for Yorùbá Low tones, we may assume that the non-native listeners in the present thesis were able to perceive and produce this distinctive voice quality associated with the tone, leading to more accurate repetition of the tone. A plausible explanation for this production pattern relates to the pitch height of the Mid tone, which lies between High and Low. Its production requires raising the pitch level above the level for the Low tone but not as high as for the High tone. This possibly proves difficult for the two non-tone language groups due to the absence of an

equivalent neutral system in the prosodic system of their L1. By comparison, the production of the Low tone is relatively less difficult for all participants as it involves only lowering the pitch as much as they can.

While the two English-Canadian groups faced some difficulty producing the Mid tone, the Yorùbá-Canadian young group produced the tone more accurately than the Low tone (95% vs. 87%). This suggests that as native speakers, even with their infrequent tone language usage, they still accorded the correct pitch height to this seemingly 'neutral' tone. Thus, some members of this group still possess the tone production skill required. In what follows, the analysis of the error patterns will reveal how participants misrepresented each of the three tones.

7.4.2 Monosyllabic production errors

As indicated earlier in Figure 7.1, Mid tones globally attracted up to two-thirds (67%) of the total tonal misrepresentation across the three participant groups, while High tone attracted a quarter (25%) of the misrepresentation. Low tone attracted only about 8% of these misrepresentations. Meanwhile, across the three groups, the target tones were often produced as Mid tones (Table 7.7). The production error data highlights a trend wherein the Mid tone emerges as the tone attracting the most misrepresentation of the other tones. Among others, this trend points to an earlier suggestion that participants' performance in perception does not reflect their production performance. The fact that a tone is difficult to perceive may not necessarily translate into similar difficulty in the imitation or production of that tone. Second, evidence from Hayward et al. (2003) has shown that both the High and the Mid tones are different from the Low tone in their phonation characteristics. They are mostly breathier while the Low tone is mainly creakier. The fact that these two tones share similar phonation characteristics, may have led to confusion, leading some participants to produce the two tones in the same pitch range. The Mid tone's intermediary position between the two extreme tones, High and Low, in addition to the closeness of the tone's F0 to those of the two tones are other factors that could have led to participants' resort to the Mid tone in their production of the High and the Low tones. Moreover, due to the narrow pitch range ascribed to listeners with no tone language background as suggested by Orie (2006), most of these participants found it challenging to attain the pitch contour level required

to produce the High tone or even possibly the Low tone. Such untrained listeners could only resort to the nearest available pitch level namely, Mid, in their production of these two tones. Thus, High tones were often produced as Mid by all the three participant groups, while Low tones were produced as Mid in about a quarter (26%) of the time by the Yorùbá-Canadian young group. The young and adult English-Canadian groups respectively produced 17% and 2% of their Low tones as Mid.

The suggestion that perception is not the same as production is further highlighted by the English-Canadian adult group's production performance with respect to the High tones. In the perception experiment, this was the only group to identify the High tones significantly more accurately (Chapter 3; Table 3.10). One would have expected the group to record similar success in High-tone production. Comparing the relatively high-pitched Yorùbá High tone to the English primary stress system probably worked for the perception task but not for production. Instead, when a High tone should have been produced, a Mid tone was produced. This trend indicates that the two independent evaluators judged the pitch levels for the High tones produced by these participants as not high enough to be the Yorùbá High tone; hence their classifying them as Mid tones. Nevertheless, the fact that these participants are accustomed to the English primary stress system did not affect them to the extent of producing the High tones as Low. The results indicate that they could perceive the vast difference between the pitches of the High and Low tones. Thus, for the three groups, there were only a few instances of High tones being produced as Low.

The rate at which participants misrepresented other tones as High tones was not as high as that for Mid tones. About a quarter (25%) of the other tones were misrepresented as High. Misrepresentation as High by each of the three participant groups stands at 8%, 18%, and 44% respectively for the Yorùbá-Canadian young, English-Canadian young, and English-Canadian adult groups. Interestingly, almost all the tones misrepresented as High were Mid tones (Table 7.7). This indicates that participants rarely or never misrepresented Low tones as High. Mid tones being produced as High tones does not contradict the observation that the High tone is more difficult to produce. Rather, the low rate of such occurrence is a clear testimony to that. Perhaps what is being observed here is the result of tonal confusion between Mid and High, possibly caused by a similar breathier phonation type that differentiates them from the Low tone

(Hayward et al., 2003). Such confusion may ultimately lead to perceptual and production confusion for listeners. Another look at the production error rates for the three participant groups further supports this point. The Yorùbá-Canadian young group recorded the lowest rate (8%) for tones misrepresented as High, all of which originated from Mid tones. This result is an indication that some members of this group could perceive and reproduce Mid tones. The English-Canadian young group, on the other hand, recorded a higher error rate of 26%, all of which also originated from Mid tones. This group, due to their lack of tone language experience, exhibited more confusion in Mid-tone production than the Yorùbá-Canadian young group. However, their lack of experience probably worked more in their favour than for the English-Canadian adult group who recorded the highest rate for tones misrepresented as High (44%), most of which also originated from Mid tones (96%).

The results observed in this part of production data highlight the Mid–High confusion regardless of the experience factor. The results of the English-Canadian adult group highlight this fact better because, for this group, the rate of misrepresentation of High tones as Mid (49%) was not that different from the rate recorded for their misrepresentation of Mid tones as High (44%). While most of the tones misrepresented as High were from Mid tones, none, or just a negligible few Low tones, were produced as High. Two factors may be cited to account for this observation. First, in a monosyllabic context, the absence of a preceding or a following tone makes the Low tone more conspicuous to the auditory system. Second, a low-pitch contour, which usually results in some creakiness, characterizes the production of the Low tone. Together, these two factors may facilitate better perception and subsequent production of the Low tone. Unlike the High or Mid tone which may confuse listeners with little tone language experience, listeners perceiving the creakiness of the Low tone will strive not to raise their pitch level more than necessary while imitating it. Essentially, these same factors may account for the low rates recorded for the misrepresentation of both High and Mid tones as Low by the three participant groups.

7.4.3 Disyllabic production

The disyllabic production analysis in Table 7.8 reveals that the LH pattern, which contains a contour tone, attracted higher production accuracy compared to the other patterns. One possible reason for this high accuracy seems to be the iambic nature of the pattern which could have made it easier for most participants especially the English-Canadian groups. In contrast, the other contour pattern, HL, recorded lower production accuracy by participants. Thus, the LH pattern was produced more accurately than the HL pattern. This part of the production results reflects what was observed in perception (Identification 2 experiment) where the contour H (in LH) recorded a significantly higher accuracy than the contour Low (in HL). Although the two patterns are near equivalent in the English prosodic system, but due to reasons yet unknown, participants appeared to be doing better at the iambic-like LH than the HL which has the trochaic semblance. In addition, it is also possible that due to narrow pitch range (Orie, 2006), the two English-Canadian groups found it especially difficult to initiate the high pitch of the initial High tone of this pattern. Although the High tone may be comparable to the English primary stress in its use of pitch as cues, the two features are different. The realization of the English primary stress does not require such a high pitch as that required to produce the Yorùbá High tone.

If the production of the initial High tone in the HL pattern poses some challenges for non-tone language speakers, the production of two High tones in succession is expected to be equally challenging for these groups. This observation is attested to by the relatively lower accuracy rates observed for the HH and HL patterns by all three groups. Yet, in the production of these two patterns, the Yorùbá-Canadian young group, possibly due to their tone language background, still performed slightly better than the two English-Canadian groups. In sum, HL and HH emerged as the most wrongly-produced of all the nine Yorùbá disyllabic patterns. Considering what has been observed so far, the next section will suggest some reasons why some patterns may attract higher rates of misrepresentation of the other patterns.

7.4.4 Suggested Reasons for patterns attracting misrepresentation of other patterns

In Table 7.10, five patterns are involved in attracting the most misrepresentation of other patterns across the three participant groups. MH and MM are the patterns mostly preferred by the English-Canadian young group as well as the Yorùbá-Canadian young group with the latter additionally favouring the HM pattern. The English-Canadian adult group misrepresented other patterns mostly as either ML or LH. The following reasons may be suggested as to why these patterns were often chosen to represent other patterns.

Tonal confusion: High – Mid confusion may lead to the misrepresentation of High tones as Mid or Mid tones as High in a disyllabic pattern containing any of these two tones. Such misrepresentation may occur at any syllable position in the disyllabic word. Such words end up being produced mostly as MM or MH, and not as either HL, HM, or HH due to possible difficulty in producing the High tone at the initial syllable as will be shown next.

Production difficulty: Due to the difficulty encountered in the production of some tones, participants may substitute easier tones for other tones, resulting in a different pattern entirely. The production of the initial High tone in a disyllabic word is one good example of this. Thus, H-initial patterns such as HM, HL, and HH may end up being misrepresented as ML, MM, or MH.

Prosodic similarity: The LH pattern may attract misrepresentations due to its similarity to the English iambic meter. As a pattern with a closer affinity to an existing contrast in the prosodic system of some participants' L1, it becomes readily available to certain speakers of such language as a substitute for any pattern that sounds close to it.

Salient acoustic features: Some patterns may have tones inherent with features easily perceivable to the listener. Such features are, at the same time, not difficult to reproduce. Patterns such as these will still be misrepresented but not significantly. However, such patterns may rarely be resorted to for misrepresenting the other patterns. For example, the Low tone whose creakiness makes it easily perceivable to the ears and is easily reproduced does not attract considerable confusion.

In what follows, the analysis will examine the disyllabic production error matrix presented in Table 7.9 above to highlight how the reasons suggested above apply to one or more of the

misrepresentation patterns observed. In all the H-initial disyllabic patterns (HM, HL, HH), High tones were mostly produced as Mid. This indicates that perhaps participants usually resorted to the Mid tone when they could not attain the necessary pitch height required to produce the High tones. Moreover, in the HH pattern, where both the initial and the final tones are High tones, participants produced more than a third (36%) of this pattern as MM, an indication that the final High tone was also produced as Mid. This observation further lends support to the suggestion that generally, participants faced difficulty in the production of initial High tones and that the production of the final High tone could also be difficult if the penultimate tone is another High tone (i.e., in the HH pattern). The perception of a similar pattern in Edo required that the initial H tone be produced in the same frequency level as the final H tone (Omozuwa, 1991). In cases where the F0 of the initial H tone did not reach the required pitch level, the pattern was perceived as LH. When this occurred in the Yorùbá HH, the pattern was perceived as MH and rarely as LH. This is expected due to the difference between Yorùbá and Edo tone inventories. Edo has only two tones (High and Low) while Yorùbá has three tones (High, Mid and Low). Thus, in Yorùbá tonal space, the nearest tone to the High tone is the Mid tone, hence the perception of HH as MH.

The M-initial patterns reveal that in the production of MH as LH, the Mid tone was represented as Low. To speakers with little or no experience of a tone language, the Mid tone may sound like a neutral tone, comparable to the unstressed syllable in English. Thus, a disyllabic pattern involving a Mid tone followed by a High tone may be perceived as the English iambic meter. Thus, it is not unusual to see some participants producing this pattern as LH. Moreover, about a third of participants also produced the final High tone in the MH pattern as Mid, resulting in MM. This is possibly due to the confusion between High and Mid tones or the general difficulty experienced in producing the High tones.

The neutral nature of the Mid tone is also detected in participants' production of the ML pattern as HL, which shares some similarity with the English trochaic meter. A Mid tone has no equivalence in the prosodic system of the English-speaking participants in the present analysis. This means that the tone may be perceived either as a neutral or as a High tone. It may not be perceived as a Low tone because a Low tone is already present in the disyllabic pattern and is conspicuous enough to contrast with the neutral-sounding Mid tone. When such a 'neutral' tone

is followed by a Low tone, which is comparable to the unstressed syllable in English, this may lead participants to produce the resulting combination closer to the nearest match in the English prosodic system, the trochaic meter. The other observation that the ML pattern was also produced as MM in about one-fifth of participants' misrepresentation also requires some explanation. Possibly due to participants' unfamiliarity with the pattern, the attempt to correctly produce the Mid tone led them to maintain the pitch height of the Mid tone and then stretch it across the production of the two syllables in the ML pattern. As such, the lower pitch essential in producing the Low tone in the second syllable was compromised.

While MM has been shown as the alternative whenever participants are faced with a difficult pattern, it is interesting to note that the pattern itself was misrepresented by a sizeable number of participants. Equally interesting is the fact that both the initial and the final Mid tone of this pattern were mostly misrepresented as High. This unidirectional path for Mid tone misrepresentation further highlights the Mid-High confusion observed so far in this study.

For L-initial patterns, the initial Low tones are mostly misrepresented as Mid when they are followed by High and Low. Thus, participants produced both LH and LL as MH and ML respectively in about half (50%) of the misrepresentation of these patterns. In the cases where LH was produced as MH, participants' attempts to produce the pattern did not begin with the appropriate low pitch before moving on to produce the following High tone. While the evaluators correctly judged the final High tone as High, producing the preceding Low tone without adequate low pitch makes it sound like a Mid tone to trained listeners. Apparently, the evaluators did not perceive it as low enough, thereby judging it as a Mid tone. The LH pattern was also produced as LM in about a third of its misrepresentations. In this case, participants produced the final High tone as Mid, suggesting another instance of High-Mid confusion.

High-Mid confusion can further be invoked to explain the production of LM as LH in three-quarters (73%) of the misrepresentations. The final Mid tone in LM is frequently produced as High. In addition to the High-Mid confusion, another factor that may account for this misrepresentation is the similarity of the LH pattern to the commonly used English iambic meter. Moving from a Low tone to an apparently 'neutral' tone midway between a Low and a High tone may have given rise to the resulting LH pattern. Participants who produced LM as LH possibly found it difficult to produce a 'neutral' tone immediately following a Low tone. Thus, based on

their experience with the English iambic, the other tone that can naturally follow a Low tone will be a High tone.

In the production of the LL pattern, participants were faced with the difficulty of producing two low-pitched syllables in succession. This is a feat that they are possibly not used to doing in their L1, which in most cases, is English. Only native speakers of the tone language as well as experienced language users can accomplish this feat without some difficulty. Thus, when they misrepresented LL as ML, these participants appeared to produce the initial Low tone as a Mid tone as a way out, resulting in the ML pattern. In another revelation of the difficulty of producing successive Low tones in a disyllabic context, about a quarter (25%) of LL patterns were produced as LM. This indicates that in the latter case, participants chose to substitute the final Low tone as Mid instead of the initial one sacrificed in the former case. This effect was observed in Hombert (1976) with the modified stimuli. Hombert (1976) found that a Low tone in word-final position is correctly perceived if it has a falling F0. The presence of a steady F0 in this position causes it to be perceived as a Mid tone, but rarely as a High tone. In contrast, if this happens in Edo, with two-tone inventory, the LL pattern is usually perceived as LH (Omozuwa, 1991). Thus, as in the HH pattern, the language's tonal distance and tonal inventory are important factors.

The objective of Hombert (1976) was to investigate the acoustic cues necessary of the perception of Yorùbá tones by native speakers. Thus, in the second experiment within his study, he used modified stimuli. Low tones modified with steady F0 resulted in significant identification of the modified Low tone as Mid. The modified stimuli led to the findings that falling F0, rather than duration and amplitude was the most essential cue to identifying the Yorùbá Low tones. In short, the purpose of the experiment required the use of the modified stimuli.

Omozuwa (1991) cited here also refers to the second part of the study which looked at the acceptability / unacceptability judgements of the repetitions of the non-Edo speakers in the experiment. For the analysis, ten Edo speakers were asked to listen to the repetitions by the non-Edo speakers and then assess which of the utterances were acceptable or non-acceptable utterances in Edo. This stage of the experiment was similar to the present production experiment in terms of procedure and evaluation method. Meanwhile, the use of real Edo words was deemed appropriate for the purpose as all participants in the repetition exercise were non-Edo speakers.

This is unlike the present thesis which comprised of both Yorùbá and non-Yorùbá language speakers.

Despite these similarities, the misrepresentation patterns encountered in Edo tonal repetition need to be viewed with some caution when compared to the present thesis. This is in part, due to the main differences in the tonal inventory between Yorùbá and Edo. Yorùbá has three tones (High, Mid, Low) while Edo has two (High and Low). Thus, while native Yorùbá evaluators could judge the utterances as either of the three tones, Edo evaluators could only judge the non-native utterances as either High or Low. For instance, in Yorùbá disyllabic production error matrix, the HH pattern was assessed as either MH (40%), MM (36%), or as LH (15%), while for the same HH pattern in Edo, native Edo speakers could judge the error in the pattern only as LH.

In sum, the production error analysis has so far revealed that participants misrepresented the tones due to High – Mid confusion or due to the difficulty of producing each of the three tones in specific disyllabic contexts. The analysis further suggests that when faced with a difficult pattern, participants often resort to the MM pattern or the familiar iambic-like LH pattern as substitutes. Additionally, the analysis suggests that there are fewer instances of initial High tones being produced as Low. Likewise, there are fewer instances of initial Low tones being produced as High, possibly due to the distance between the two tones in the language's tonal space. A comparison of the production of the HH and the LL patterns with similar patterns in Edo suggests that a language's tonal inventory influences the perception of the two patterns by native speakers. Finally, in tone production, the Yorùbá-Canadian young group performed comparably to the English-speaking groups rather than performing comparably to the other native tone-language speaking groups. The next chapter explores some of the reasons why this group's tone production (and perception) and abilities are not comparable to the other native tone language speakers.

Chapter Eight: Demographic survey for the Yorùbá-Canadian Young group

The present chapter examines some of the demographic factors that could shed more light on why the Yorùbá-Canadian young participants did not perform comparably to the other native Yorùbá-speaking participant groups, especially their Yorùbá-Nigerian counterparts. Some members of the Yorùbá-Canadian young group fits what is usually referred to as Heritage Language children by being born in Canada or moving to Canada at young age and for having exposure to their L1 through their parents and possibly grandparents. However, on moving to another language environment, they are exposed to an L2 which eventually becomes their dominant language (Valdes, 2001). Heritage Language children often attain native-like proficiency in the L2, but their L1 may undergo some deficiency. According to Montrul (2008), their L1 can exhibit deficiencies due to incomplete acquisition[13] or due to attrition. These deficiencies also tend to be more severe in early than in late bilinguals.

Several studies have identified some predictors as essential for Heritage Language development and maintenance. These include the Age of Arrival in the immigrant country and the Language Environment. Montrul (2008) linked low level of Heritage Language proficiency in pronunciation, the lexicon and morphosyntax to younger Age of Arrival. Other studies (e.g., Jia, 2008) suggest that the quantity and the quality of Heritage Language input are equally essential for the development and maintenance of the Heritage Language. Moreover, studies such as Jia and Aaronson (2003) have reported a shift from the L1 to a dominant L2 among immigrant population who arrived in the L2 country at an early age. Jia and Aaronson (2003) studied Chinese who migrated to USA before the age of nine. They found that within one year of English immersion, these young immigrants shifted their preferred language from Chinese to English.

In the domain of tonal attrition, Yeh and Lu (2012), Yeh and Tu (2012), as well as Yeh and Lin (2013), have observed that, the infrequent use of the tonal language is a primary cause of tonal loss in speakers (see Section 1.6 for more details on Yeh & Lu, 2012). Infrequent use of the L1 occurs mainly due to the dominant influence of the L2, which gradually replaces the L1. Usage of the L1, in this case, is construed as a two-way process - speaking (production) and listening. If

[13] Incomplete acquisition: When Heritage Language speakers do not attain the proficiency levels appropriate for L1 children of the same age groups in grammatical structures and other areas of the L1 (more details in Section 1.7).

one factor (especially production) is missing, attrition is bound to set in. It is plausible to assume that listening to the language alone, in the absence of speaking, will affect the subjects' production and perception patterns, both in quantity and quality.

For the present analysis, the demographic factors were obtained through a pre-experiment survey called the Alberta Language Environment Questionnaire (ALEQ) developed by the CHESL Center Project (Paradis, Emmerzael, & Sorenson Duncan, 2010) adapted for use with this participant group (see the adapted ALEQ in Appendix A1). It was distributed to the parents as well as to the participants to collect background information and relevant demographic data. Information obtainable from the questionnaire included the date and place of birth of the children, number of siblings, the year of arrival in Canada, as well as the rate and frequency of Yorùbá usage at home. An exposure to the native tone language informed the hypothesis that the Yorùbá-Canadian young group should demonstrate a tone processing advantage, especially over their English-Canadian counterparts (Wayland & Guion, 2004). However, despite this tone Language Context, the analyses suggest that most members of this group continued to face tonal processing difficulties.

The Yorùbá-Canadian young group consisted of thirty (30) school-age children (Mean age = 11;9, *SD* = 2;4, Age range = 8 - 16) who were either brought to Canada by their Yorùbá parents or born in Canada to Yorùbá parents. They lived with their parents in the same house but had different rates of exposure to the language. Their Yorùbá language use was limited, but not restricted to listening to their parents or other adults speaking it within or outside the house, or in occasions organized by the language community. They neither spoke it willingly with their siblings nor used it in conversations with their friends, including those who were also from the same language community. However, a few of the children from this group who had recently migrated to Canada reportedly spoke the language, although not frequently.

An earlier study (Shittu & Tessier, 2014) compared the tone perception abilities of this participant group with their adult counterparts. The results revealed that the younger participants performed worse than the adults in AX discrimination (adults =74%, young =51%); Tone Identification (adults =54%, young =37%); and Lexical recognition (adults = 96%, young = 54%). The largest young versus adult difference in accuracy score was recorded in the Lexical recognition task. Moreover, in the present thesis, group analyses for all the five experiments

reveal that among the four groups with tone language background (Yorùbá-Nigerian adults and young groups; Yorùbá-Canadian adult and young groups), the Yorùbá-Canadian young group recorded the worst accuracy performance in all the experiments, including the last two experiments, the Lexical recognition and the tone Production. When their performance was compared to their peers from the Yorùbá-Nigerian group, the analyses indicate that they performed worse than the Yorùbá-Nigerian young group in the AX Discrimination, the first Identification Experiment, the second Identification Experiment, the Lexical Experiment, and the Production Experiment. In the Lexical recognition task, the other three groups with tone language background posted high accuracy with rates ranging from 88% to 95%, while the Yorùbá-Canadian young group posted an accuracy rate of 57%. Similarly, in the Production experiment, the other three groups recorded percent accuracy ranging from 97% to 100% while the Yorùbá-Canadian young group recorded accuracy rate comparable only to the English-Canadian group who are non-tone language speakers. While they predictably performed worse than their Yorùbá-Nigerian counterparts, the analyses indicate that there was no significant difference in the performance between the Yorùbá-Canadian young group and their English-Canadian peers either in the AX Discrimination, the first Identification, or the second Identification Experiment. It was only in the Production task that they demonstrated an accurate performance significantly better than their English-Canadian counterparts. Meanwhile, evidence from participants' response analyses in the second tone Identification Experiment and the Production Experiment have further demonstrated some affinity in the way the Yorùbá-Canadian young group and the English-Canadian adult and young groups process the tones. In the second Identification Experiment, the three groups often identified the High and the Low tones in the tone patterns as Mid. Similarly, in the Production Experiment, the three participant groups recorded the lowest accuracy rates in the two tone patterns 'HL' and 'HH' in addition to misrepresenting other patterns mostly as either HL, HH, or LL. Thus, the present thesis seeks to investigate the factors which may have contributed to this performance. Data from the adapted ALEQ survey which provided valuable demographic information about the groups are consulted. The information obtained from the survey is summarized briefly in the next section.

8.1 The adapted ALEQ survey

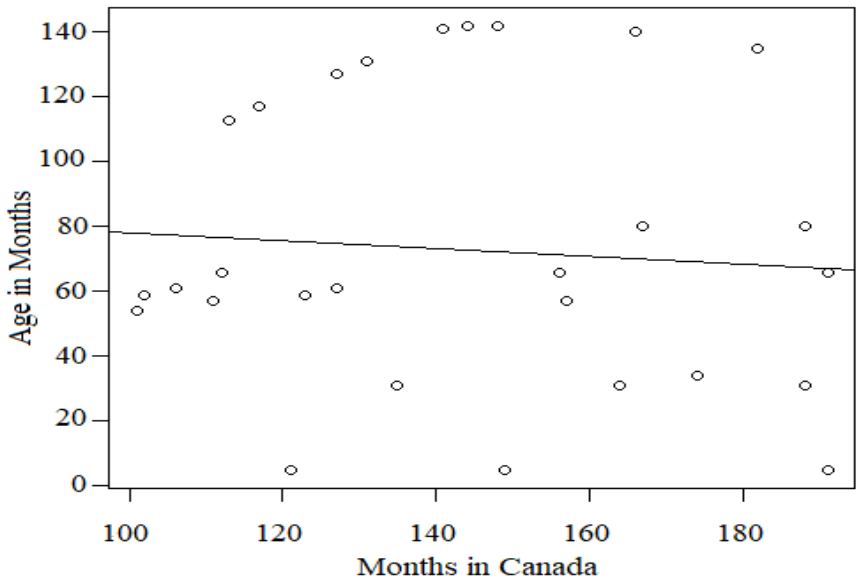
The adapted pre-experiment survey was distributed to the parents as well as to the participants to collect background information and relevant demographic data. The information obtained from the questionnaire included the date and place of birth of the children, number of siblings, the year of arrival in Canada, as well as the rate and frequency of Yorùbá usage at home. A simple initial analysis of the parents' responses to the questionnaire yields the following observations. All the participants received native language input within the house, through listening to, or hearing their parents conversing with each other daily. More than half of them also received the L1 input by listening to their parents using the language in conversation with their siblings or family friends within the house, or in a telephone conversation with other relatives. The analysis also shows that for about half of the participants, their parents sometimes spoke to them in the native language. However, two-thirds of these participants reportedly did not speak the native language to their parents in return. The remaining one-third did speak it, but not all the time. The questionnaire also considers the participants' native language use with their siblings and other relatives (if any) within the home. The results show that less than one-fifth of the participants always used the native language with their siblings, while none of the participants had a relative residing permanently in the home. Information from the survey were prepared for statistical analysis presented in the following section.

8.2 Statistical Analysis

Among the demographic factors provided by the survey, four measurements were chosen for inclusion in the subsequent analysis. They included the following four major variables: participant's age at the time of testing, length of stay in Canada, age at arrival in Canada and their place of birth. The language environment measurement was not included because a preliminary analysis which included all the predictors did not show any significance for this factor. All the variables, except the place of birth, were converted into months to facilitate the analysis.

Lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Development Team, 2015) in R was used to conduct a generalized linear mixed effect regression (glmer) analysis on the effects of each of the four variables on participants' accuracy in each of the five experiments. To facilitate the analysis in R, three of the four variables were converted into numerical data. The fourth variable, 'place of birth' remained a factor with two levels: *Nigerian (NGR)* and *Non-Nigerian (NNGR)*. Before performing the linear mixed-effect analysis, correlation tests were performed for the three numeric fixed effects, and the tests reveal a very weak correlation between age (in months) and time in Canada (in months) (r = -0.1186822; p-value = 4.746e-16). The tests also reveal a medium correlation between age (in months) and age at arrival (r = 0.5688306; p – value < 0.2.2e-16). Finally, the tests reveal a strong correlation between age at arrival and months in Canada (r = -0.8841519; p- value < 2.2e-16). The correlation plots are shown in Figures 8.1 to 8.3.

Figure 8. 1 Correlation plot: Age in months vs. Months in Canada



The correlation plot for 'Age in months' and 'Months in Canada' in Figure 8.1 shows that the points between these two factors do not converge, indicating no correlation. Thus, older age at the time of testing does not necessarily translate to a longer stay in Canada. For instance, participants who arrived a few months before the test will have fewer months in Canada than those who arrived earlier, even though they may be older. In contrast, participants who were born in Canada may have more or fewer total number of months in Canada depending on when they were born. In other words, participants with more months in Canada can be of any age.

Figure 8.2 presents the correlation plot for the two factors “Age in months” and “Age at arrival”. The medium correlation shown in the figure indicates that participants’ Age at arrival is predictably less than their actual Age. The relationship is such that even if participants were born here in Canada, which puts their Age at arrival at zero months, their actual Age at time of testing could be no less than eight years (96 months), which is the minimum age to participate in the study for this participant group.

Figure 8. 2 Correlation plot: Age at arrival vs. Age in months

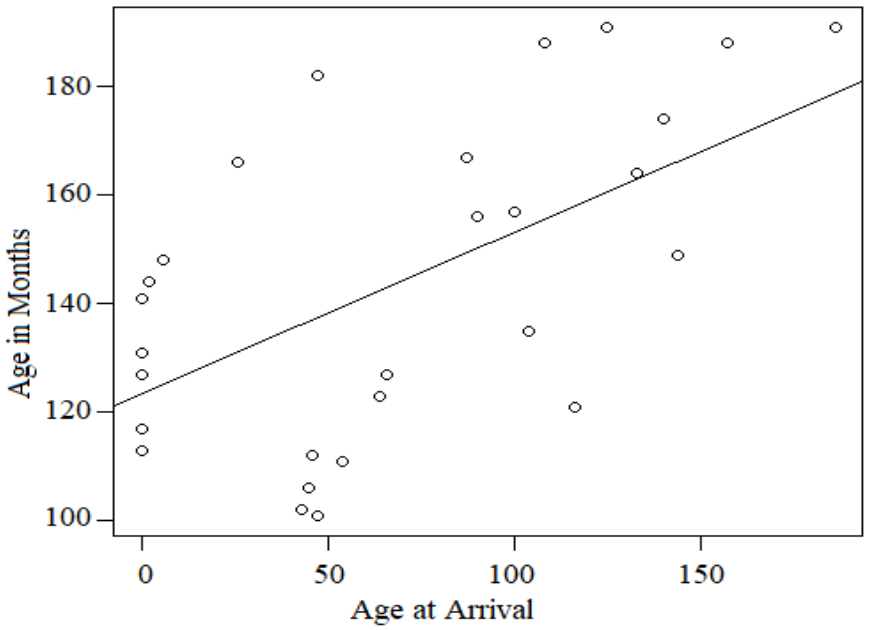
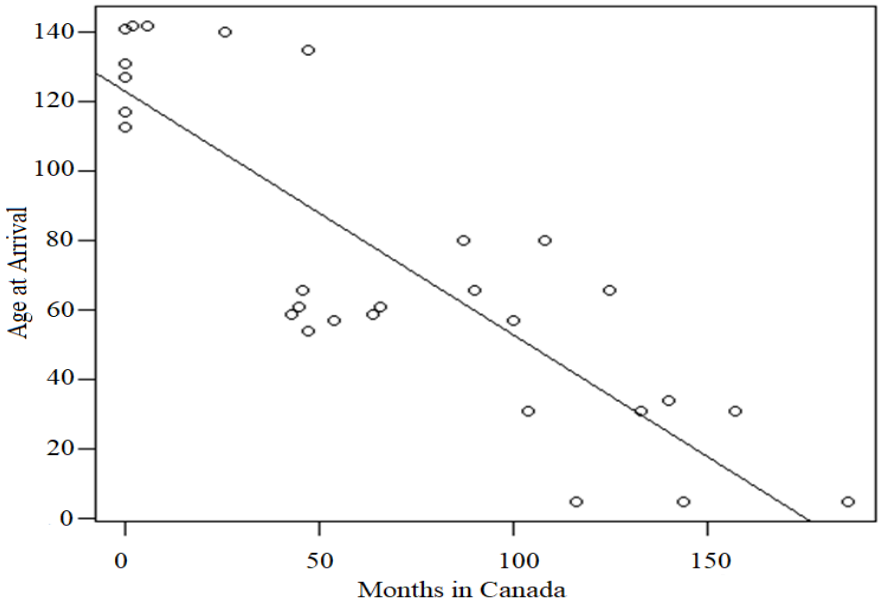


Figure 8.3 presents the correlation plot for Age at Arrival and Months in Canada. As the figure shows, there is a strong correlation between Age at arrival and Months in Canada. The Age at arrival is predictably higher than the Months in Canada. This indicates that most of the participants with lower Age at arrival were either brought to Canada young or were born here in Canada. Canadian-born participants and those brought to Canada at a younger age are likely to record more months in Canada than those who were brought here at an older age.

Figure 8. 3 Correlation plot: Age at arrival vs. Months in Canada



To attest this collinearity between Age at arrival and Months in Canada, a full model was run that included all the three factors. The result of the full model returned the same coefficient estimates, the same standard errors, z-values as well as p-values for Age at arrival and Months in Canada. Given this correlation, the Age at arrival factor was excluded from subsequent linear mixed-effect regression analyses due to its high collinearity with Months in Canada as well as its medium correlation with the other variable Age in months. Thus, for each of the five experiments, the analysis included two fixed effects, namely *Age in months* and *Months in Canada* as two numerical variables. The third fixed effect was the participants' place of birth, a factor with two levels, *Nigerian-born (NGR*) and *Non-Nigerian-born (NNGR),* with Nigerian-born as the reference level. The participants' accuracy scores, which consisted of the individual score for each experiment's target stimuli, was the dependent variable. As a random-effect factor, the stimuli were coded as *Item*, while the young participants were coded as *Participant*.

In the tables below, since the two numeric factors have no reference level, the minus symbol (-) before the coefficient estimate indicates a negative effect of this factor on overall accuracy in the experiment, while the absence of a symbol indicates a positive effect of the factor on accuracy in the experiment. As for the third factor, place of birth (POB), which is non-numeric, a minus symbol indicates that the Non-Nigerian-born participants performed worse in that experiment

than the Nigerian-born participants, the reference level. The absence of a minus symbol signifies a better accuracy than the reference level (NGR). The asterisks indicate the significance level of the results, while the absence of an asterisk indicates no significance. In what follows, the results of the generalized linear mixed-effects models for each of the five experiments is presented, beginning with the AX Discrimination Experiment.

8.2.1 Demographic factors vs. the AX Discrimination Experiment

Table 8.1 presents the model showing the effects of the demographic factors on tone discrimination. The model was based on 1440 observations (a total of 30 participants x 48 individual item score = 1440). The analysis reveals that older age at the time of testing significantly led to better accuracy in tone discrimination. However, a longer stay in Canada resulted in significantly worse tone discrimination accuracy. The analysis further reveals that there was no significant difference in tone discrimination accuracy between participants who were born in Nigeria and those born outside Nigeria. The interaction between Age in months and place of birth reveals no significance even though the negative estimate seemed to indicate that older age, coupled with being born outside Nigeria, might lead to reduced tone discrimination accuracy.

Table 8. 1 Coefficient estimates of the glmer analysis model showing the effects of the demographic factors on tone discrimination by the Yorùbá-Canadian young group

	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	-1.808856	1.020241	-1.773	0.076234**.**
Age in Months	0.015972	0.006650	2.402	0.016305 *
Months in Canada	-0.014385	0.003838	-3.748	0.000179 ***
POB_NNGR	3.150426	1.683948	1.871	0.061365**.**
Age in Months x POBNNGR	-0.011181	0.011481	-0.974	0.330127

8.2.2 Demographic factors vs. the Tone Identification Experiment 1

Table 8.2 presents the results of the model showing the effects of the demographic factors on the first Tone Identification Experiment. The model was based on 1344 observations (a total of 28 participants X 48 individual item score=1344). Unlike the AX Discrimination Experiment above, participants' age at the time of testing and their place of birth did not have any significant effect on their tone identification accuracy. The other factor, namely 'months in Canada' showed significant effects similar to what was observed in the AX Discrimination Experiment. Thus, a longer stay in Canada led to reduced tone identification accuracy. The interaction between participant Age and their place of birth did not produce any significant effect on tone identification accuracy.

Table 8. 2 Coefficient estimates of the glmer analysis model showing the effects of the demographic factors on tone Identification by the Yorùbá-Canadian young group

	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	-1.805450	0.885644	-2.039	0.04149 *
Age in Months	0.010913	0.005625	1.940	0.05235.
Months in Canada	-0.010153	0.002765	-3.673	0.00024 ***
POB_NNGR	1.912154	1.320267	1.448	0.14753
Age in Months x POBNNGR	-0.008139	0.008786	-0.926	0.35426

8.2.3 Demographic factors vs. the Tone Identification Experiment 2

The model showing the effects of the demographic factors on the second Tone Identification Experiment is presented in Table 8.3. It was based on 624 observations (a total of 13 participants x 48 individual item score = 624). All the three factors recorded highly significant effects in this experiment. Older age at the time of testing significantly contributed to better identification of the tones, while a longer stay in Canada significantly reduced tone identification accuracy. In a

similar vein, those who were born outside Nigeria performed significantly better than those who were born in Nigeria. However, the interaction between Age and Place of birth suggested that older Age, coupled with being born outside Nigeria, might significantly reduce participants' tone identification accuracy.

Table 8. 3 Coefficient estimates of the glmer analysis model showing the effects of the demographic factors on tone Identification (2) by the Yorùbá-Canadian young group

	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	-6.15833	1.76867	-3.482	0.000498 ***
Age in Months	0.03812	0.01121	3.401	0.000671 ***
Months in Canada	-0.01522	0.00355	-4.288	1.81e-05 ***
POB_NNGR	6.07369	2.28116	2.663	0.007755 *
Age in Months x POB_NNGR	-0.02879	0.01433	-2.009	0.044511 *

8.2.4 Demographic factors vs. the Lexical Experiment

Table 8.4 presents the results of the model showing the effects of the demographic factors on Lexical recognition. The model was based on 616 observations (a total of 28 participants x 22 individual item score=616). The results indicate that the Age effect, with positive estimates, was a significant factor in this experiment. Older age resulted in better lexical recognition. Similarly, the place of birth factor indicated that participants born outside Nigeria were better at lexical recognition than those who were born in Nigeria. However, just as in the Identification 2 Experiment, the results of the interaction between Age and Place of birth suggested that older participants who were born outside Nigeria tended to have significantly lower lexical recognition accuracy than those born in Nigeria. Finally, as in all the previous experiments, the 'Months in Canada' factor posted high significance, indicating that longer stay in Canada resulted in lower lexical recognition accuracy.

Table 8. 4 Coefficient estimates of the glmer analysis models showing the effects of the demographic factors on Lexical recognition by the Yorùbá-Canadian young group

	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	-1.673914	0.931978	-1.796	0.07248.
Age in Months	0.018506	0.006017	3.076	0.00210 **
Months in Canada	-0.012231	0.002955	-4.139	3.48e-05 ***
POB_NNGR	3.726181	1.400385	2.6661	0.00779 **
Age in Months x POB_NNGR	-0.023966	0.009347	-2.564	0.01035 *

8.2.5 Demographic factors vs. the Tone Production Experiment

Table 8.5 presents the results of the model showing the effects of the demographic factors on the participants' tone production. The model was based on 624 observations (a total of 13 participants x 48 individual item score=624). While this experiment also did not record significant effects for the Age factor, a longer stay in Canada resulted in low production accuracy with a very high statistical significance. There was also no significant difference in tone production accuracy between participants born in Nigeria and those born outside Nigeria. Finally, the interaction between Age and Place of birth did not yield any significant effect.

Table 8. 5 Coefficient estimates of the glmer analysis models showing the effects of the demographic factors on tone Production by the Yorùbá-Canadian young group

	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	80018253	3.901318	2.055	0.039853 *
Age in Months	-0.021075	0.023567	-0.894	0.371187
Months in Canada	-0.018643	0.004944	-3.771	0.000163 ***
POB_NNGR	-8.090198	4.416074	-1.832	0.066953.
Age in Months x POB_NNGR	0.041929	0.026718	1.569	0.116575

8.3 Discussion

This section is devoted mainly to one participant group, the Yorùbá-Canadian young group. Demographic measurements for the group were collected through an ALEQ survey adapted for this experiment. The survey produced information about participant's age (in months) at the time of testing, Age at arrival in Canada, Months in Canada, and Place of birth, among others. After performing correlation tests on all these factors, 'Age at arrival' was excluded as a fixed effect due to its correlation with both 'Age in months' and 'Months in Canada'. Thus, only the three remaining factors were included as fixed effects in subsequent analysis. These demographic factors were considered with respect to their effects on participants' accuracy in each of the five experiments; the results of the statistical analysis are discussed in what follows.

8.3.1 Months in Canada

All five experiments recorded negative estimates for Months in Canada as well as highly significant values for the effect of this factor on participants' accuracy. This result indicates that the longer the participants stay in Canada, the more likely it is that they will perform worse in each of the four tone perception as well as the tone production experiments. This is highly likely because some of these young participants have moved away from the native language community into a new community where a new language is dominant, leading to infrequent use (Paradis, 2007) or, in some cases a total lack of use of the tone language. In the case of those participants who were born outside Nigeria, it is possible that a case of incomplete acquisition manifests in their tone perception and production ability despite living in a tone language environment. As for the children who were born in Nigeria and then migrated to Canada, living in an L2 dominant environment means that attrition is bound to set in (Bullock & Gerfen, 2004; Seliger & Vago, 1991). The results of this young participant group reflect Montrul's (2008) observation that language loss in children is usually more severe than in adults. Moreover, the present thesis, as well as studies such as Yeh and Tu (2012) have shown that attrition can affect the suprasegmental features of a language. The likely effects of participants' age on tone perception and production is discussed next.

8.3.2 Age in months

Positive coefficient estimates were recorded for 'Age in months' in the AX discrimination, the Identification (1), Identification (2), and the Lexical experiments, while the only negative estimate was recorded in the Production experiment. This suggests that older Age at the time of testing is a possible contributing factor to better performance in the perception experiments as opposed to the Production experiment. Meanwhile, statistical significance was attained in the AX discrimination, the second Identification, and the Lexical experiments, while no significance was attained in the tone Identification (1) and the tone Production experiments.

Older age could indicate more tone language experience for the participants and consequently translate into better performance. Older age, on the other hand could also indicate longer stay in the L2 environment resulting in more loss of the tone perception skill and then, less accurate performance. However, the correlation test and the plot in Figure 8.1 suggest that older age at the time of testing did not necessarily translate into a longer stay in Canada for these participants. Thus, the significant results observed for the Age factor in three of the five experiments highlights the positive effect of this factor. The lack of significance observed in the first Identification and in the tone Production experiments may be due to the fact that in the First Identification experiment, the accuracy score for this participant group was at the chance level, while in tone Production, the group recorded its best performance. Therefore, the Age factor may not have significant effect. Meanwhile, the interaction of the Age factor with the Place of Birth reveals that despite the significance obtained for the Age factor in the three experiments, older children who were not born in Nigeria were likely to perform significantly worse in the identification of the contour tones and in Lexical recognition. This indicates that these older participants who were not born in Nigeria might have spent more time in the L2 environment since birth than those who were born in Nigeria. Perhaps, the low perceptual accuracy rates recorded for the contour tones and the Lexical items are signs of incomplete acquisition of these two aspects of the language. While the complexity of the contour tones makes them the last to be acquired by Yorùbá children (between 2 – 3 years) (Orie, 2012; Ajolore, 1974), the fact that these older participants were not born in Nigeria, coupled with living in a L2-dominant environment may further hinder their acquisition of the lexical items. The next section discusses likely effects of native speakers being born in or outside the country of the native language.

8.3.3 Place of Birth (POB)

The place of birth factor recorded significant effects in only two of the five experiments namely the Tone Identification (2) and the Lexical experiments with positive estimates. One interesting aspect of this result is the suggestion that the Non-Nigerian-born participants (NNGR) performed significantly better than the Nigerian-born participants in these two experiments while there was no significant difference between the two categories of participants in tone Discrimination, Identification of the three basic tones and in tone Production. The absence of significant difference between the two categories of participants in tone discrimination might be due to the fact that the task required little or no knowledge of Yorùbá vocabulary. In the identification of the three basic tones, the task required the knowledge and possibly frequent use of the lexical tones which both the Nigerian- and the non-Nigerian-born participants might not adequately have, hence the comparable performance. It should be mentioned that this was the task in which the Yorùbá-Canadian group performed significantly worse than in the other tasks. The production experiment was only about articulating the tones appropriately following a prompt. This was purely a repetition skill which any of the groups could perform comparably well.

The revelation that the children born outside Nigeria performed significantly better than those born in Nigeria in the identification of the contour tones and the Lexical recognition may be due to an interplay of several factors. First, being born outside Nigeria does not mean that these participants were not exposed to the native tone language. These participants, just like their Nigerian-born counterparts, have also undergone the perceptual reorganization stage at infancy (Mattock et al., 2008). Given that, further reference to the ALEQ survey data reveals that some parents of the non-Nigerian-born participants usually spoke the native tone language to their children and had their children speak the language in return. Although these children did not always speak the native language, the exposure they got from hearing the language almost daily from their parents could translate into improved tone perception experience. On the other hand, the results may further allude to the fact that after immigration to Canada, the Nigerian-born participants might have experienced a rapid shift from L1 to L2 as suggested by Jia and Aaronson (2003). Some of the parents of these children may be less demanding than others in making the children speak the L1 since they are in a country where English, a lingua franca and

the official language in Nigeria, is spoken. Thus, the 'displacing effect' effected by the dominant language might inform this performance. Despite these explanations, the results above suggest that being born in a native tone language speaking environment may not necessarily offer more advantage over not being born there. The fact that these participants of similar age groups have been removed from the native tone-language environment into an environment where another language is dominant, affects the two categories of participants equally. This point is possibly highlighted by the analysis of the interaction between Age and the place of birth factors (Tables 8.3 and 8.4). The analysis reveals that in the two experiments where the non-Nigerian-born participants performed significantly better than the Nigerian-born participants, older children who were not born in Nigeria were likely to perform worse. This aspect of the results possibly relates to another demographic factor namely months in Canada. Thus, the Age and Place of birth interaction results could indicate that the older these participants, the longer their stay in Canada. Meanwhile, it has been observed in all the experiments that longer stay in Canada invariably led to less accurate performance in both tone perception and production.

Chapter Nine: General Discussion and Conclusion

The present thesis investigates the perception of lexical tones by three groups of speakers using four perception experiments (AX Discrimination, Identification 1, Identification 2, and Lexical) and one Production experiment. The AX discrimination was modelled after Liu and Kager (2012) and was designed to examine participants' tone discrimination ability. The two Identification experiments were similar to Bakare (1995) and Hao (2012). While the first Identification experiment investigated participants' identification of the three basic tones, the second Identification experiment investigated their identification of the two complex Yorùbá tones. The Lexical experiment was modelled along Yeh and Lu (2012). It examined the recognition of real Yorùbá lexical items by participants who are native tone language speakers. The Production experiment was similar to the mimicry task in Hao (2012) and it examined all participants' ability to imitate lexical tones.

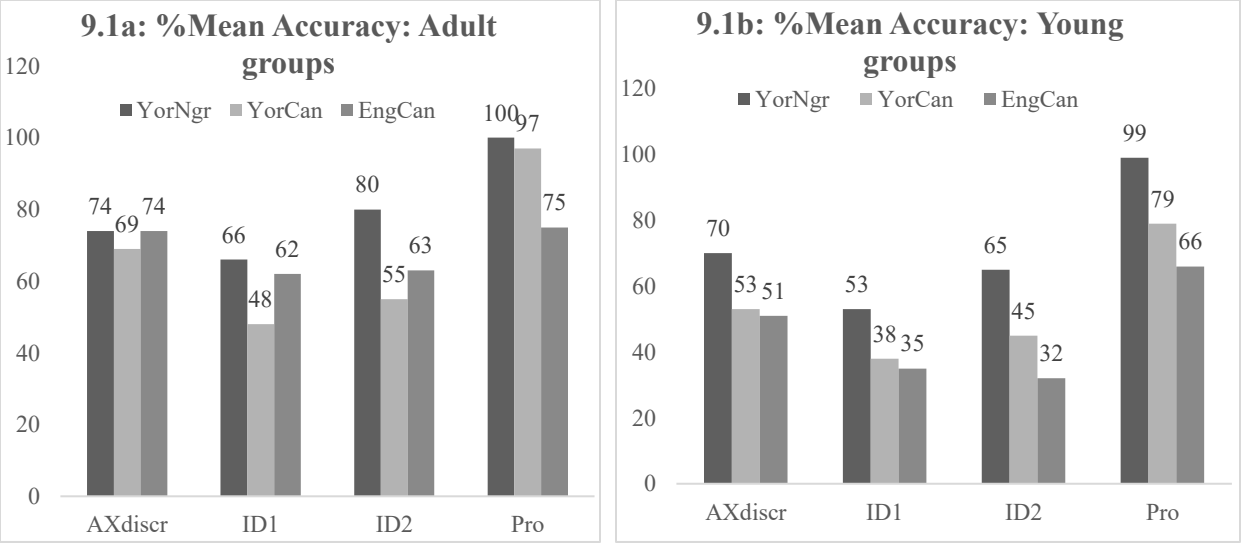
Two of the three participant groups have varying degrees of exposure to a tone language. One group, the Yorùbá-Nigerian, has participants with continuous experience through residence in an environment where the tone language is spoken. The second group, Yorùbá-Canadian, contains participants who were exposed to the tone language at birth but are not currently living in an immediate environment where the language is spoken. Thus, they currently experience infrequent use of the tone language. The third group, labelled English-Canadian, consists of participants with no previous or present exposure to a tone language. Each group contains both young and old sub-groups. To examine the tone perception and production performance of the Yorùbá-Canadian young participant group, an adapted Alberta Language Environment Questionnaire (ALEQ) was administered to this group. Demographic data obtained from the questionnaire was used in the analysis reported in Chapter 8.

9.1 The Experiments

There were five experiments in the study and each experiment had a different motivation. The AX Discrimination experiment examined participants' tone discriminatory skills. The First and the Second Identification experiments tested participants' knowledge of the three tones and the

two complex tones. They investigated the possible effects of the tone features on perception. The Lexical experiment investigated the advantage of a deeper knowledge of the tone language by using real words as opposed to nonce words. The Production experiment examined participants' perception and production ability of the Yorùbá tones. The overall results for the five experiments indicate that some of the experiments posed more difficulty than the others. This section will attempt to explore why this might be the case.

Figure 9.1 Mean accuracy for the adult and young groups in the experiments



Figures 9.1a and b present the mean accuracy for adult and young groups, respectively, for the experiments in which all participant groups participated, excluding the lexical experiment (reported in Chapter 6). The mean percent accuracy shown in Figure 9.1a indicates that except for the English-Canadian adults, the other adult groups recorded lower mean accuracy in the first Identification experiment (ID1) compared to the other three experiments. In the young groups (9.1b), while the English-Canadian young group recorded comparable mean percent accuracy in the two Identification experiments, the other young groups recorded lower mean accuracy in the first Identification experiment (ID1) compared to what they recorded in the other three experiments. Moreover, while the mean accuracy results of the Yorùbá-Nigerian young group consistently indicate the benefits of tone language knowledge in all the experiments, the Yorùbá-Nigerian adults indicate this advantage only in the second Identification (ID2) and the Production (Pro) experiments. The overall performance of the Yorùbá-Canadian adults compared to the other two adult groups will be discussed later in section 9.4 below.

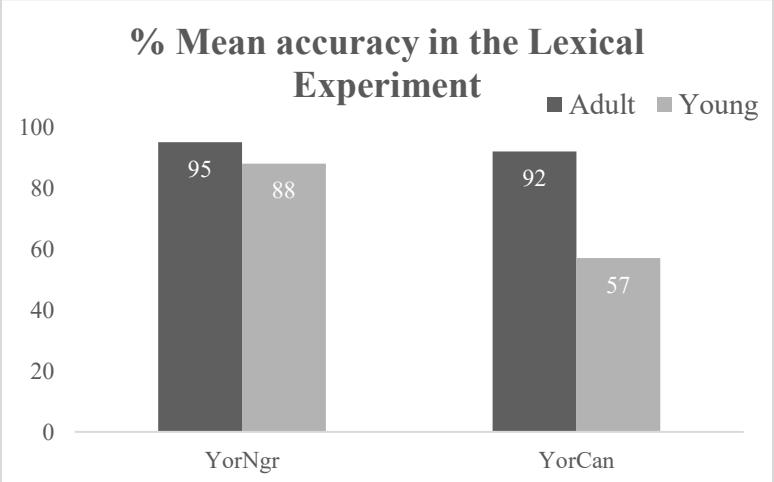
The difficulty posed by the Identification experiments may lie in the processing mechanism involved. As hypothesized by Strange (2011) in the Automatic selective perception (ASP) model, learners process speech sounds either in a phonological or phonetic mode, depending on the interactions of factors such as the listeners' knowledge, purpose and intentions, stimulus complexity and task demands. The Identification tasks in the present thesis imposed higher memory loads on participants. They were required to identify the tones of the last syllables in disyllabic and trisyllabic words as one of the three possible tones. This task requires auditory perception as well as phonological encoding. In addition, the context was such that the High and Mid tones may be confused with each other, as demonstrated by the results. Thus, the difficulty experienced by participants in this task may emanate from the phonological encoding of tones and the consequent application of the tone labels. In contrast, the AX discrimination task did not impose comparable memory demands and, as such, enabled an acoustic/phonetic level of processing (Dupoux et al., 1997; Halle, Chang, & Best, 2004). Similarly, the Lexical recognition task required listeners to "detect only the acoustic-phonetic information about tone sequences that is sufficient to distinguish one word from another" (Strange, 2011: 460). As for the Production task, the ASP model predicts that listeners will use the phonetic mode as they may not need to relate the sounds/tones with linguistic categories to be able to imitate them. Thus, according to the results of the present experiments, the Identification task, a task that likely requires phonological categorization, attracted reduced accuracy. Similar results were reported in Hao and Jong (2016), and Hao (2012). In the case of the English-Canadian adults in the present thesis, they likely make use of their phonological knowledge of the English stress system in the Identification tasks.

On the difference between discrimination and identification, Hombert (1977:184) writes that "In contrast with our amazing ability to decide if two tones presented successively are similar or not, our identification ability, i.e., our ability to identify and name sounds (or tones), is rather poor." Perhaps the lower score for the Identification task also emanates from the usage of nonce words as stimuli. Although the experiment used nonce words in order not to be biased against non-tone language speakers, the possible effect of this factor on the performance of the Yorùbá-Nigerian groups must also be considered. The expectation of real Yorùbá words as stimuli might interfere with their tone processing. In contrast, the English-Canadian groups, especially the adults, harboured no such expectation and therefore no such interference was expected.

It is interesting to observe that the Production task recorded the highest percent mean compared to the four perception experiments. This result supports Hao and Jong's (2016:164) suggestion that "when the targets are tone, the task that requires only perception and production, i.e., imitation, appears to be easier than those that require phonological categorization". One reason for this trend is the fact that the tone production task in the present thesis was not cognitively challenging. The production targets were lexical tones, which, unlike other sounds (e.g., consonants), typically have perceptible acoustic cues that are stable and clearly noticeable (Burnham & Mattock, 2007). The task involved mainly the manipulation of laryngeal posture to produce the required tone (Hao & Jong, 2016).

In the Lexical experiment, I made use of only native Yorùbá speakers as participants as well as real Yorùbá words as stimuli which differentiate the Lexical experiment from the other four experiments. The mean percent accuracy scores presented in Figure 9.2 show higher averages for all the native participant groups except the Canadian-based young participants. These higher averages possibly show that knowledge of the language, and possibly experience, benefit the other native Yorùbá-speaking groups. The other Yorùbá participants, including the Yorùbá-Canadian adults have learned Yorùbá in Nigeria. In contrast, most of the Yorùbá-Canadian young participants have learned Yorùbá out of Nigeria (e.g., Canada, USA, Kuwait). Those who have learned Yorùbá in Nigeria are presently exposed to a more dominant L2. In other words, due to either incomplete acquisition (e.g., Montrul, 2015) or infrequent use (e.g., Paradis, 2007) they do not have the same quantity or quality of input as the other native Yorùbá participants. The improved performance in the Lexical recognition task further supports studies that suggest that there are positive effects of real words on tone perception (e.g., Lee et al., 1996). Real words allow listeners to tap into lexical information, which helps listeners access otherwise fading tone information. Nonce words provide listeners with no such information.

Figure 9.2 Mean accuracy for the adult and young native Yorùbá speakers in the Lexical experiment



9.2 Comparing Asian and African Languages

Due to the sparse literature on perception of African tone languages, the present thesis made substantial reference to works performed on Asian tone languages. Meanwhile, the results reported from such Asian tone language studies should be observed with caution when comparing them with the results obtained in the present thesis due to the following reasons. First, the Asian tone languages (e.g., Mandarin, Thai, Cantonese) are typologically different from the African tone languages (e.g., Yorùbá, Edo, Mambila, Ika, Bulu, Igbo) with respect to the acoustic cues for the perception of these tones. While the Asian tone languages use pitch movement and the direction of pitch movement for tone identity, African register tone languages use pitch height. Second, in Yorùbá, neither amplitude nor duration have been shown to be relevant in tone perception (Hombert, 1976), whereas in Mandarin for instance, these two acoustic cues are relevant for the perception of its tones. Finally, Asian tone languages typically have a larger inventory size and make greater use of other cues, such as phonation, than African tone languages.

While the comparison between the tonal systems of African and Asian tone languages may not affect the results of the present thesis regarding such factors as Age and Language context, it may affect the results with respect to the language-internal features especially in the language

identification tasks where individual tones are categorized. If we focus on the differences between Asian and African tone systems and then assume that one tone system is more complex or diverse than the other, then we may expect some fundamental differences in the perception of these two different tone systems. Depending on the tone system, native listeners may focus on pitch height or the direction of F0 contour or the direction of tone levels. However, if we consider the fact that as tone languages, the F0 control mechanisms are similar (although not in all details), then there is no reason for their non-comparability in situations where similar experimental methods are used.

Furthermore, while comparing the results of the present thesis with the results obtained from perception studies of African languages (e.g., Mambila, Bulu, Edo), it is necessary to consider the differences between those studies and the present thesis in terms of purpose and the methodology. The present thesis examined tone perception and production ability of native Yorùbá speakers and non-tone language speakers while those studies investigated the frequency range used by native speakers of these languages (and non-native speakers as in Connell, 2000; Omozuwa, 1991) in the perception of the native tones. Moreover, while the present thesis adopted tone discrimination and identification tasks, those studies adopted methods ranging from categorization of synthetic stimuli into the language's tonal inventory (Connell, 2000; Hombert, 1988), assessing tonal acceptability or unacceptability of non-native utterances (Omozuwa, 1991), or rating the perceptual distance between two real-word patterns (Hombert, 1976).

Another area of consideration while looking at the results is the difference between these African languages and Yorùbá in terms of tonal inventory. While Edo has two contrastive tones (High and Low), Bulu has three contrastive tones (two Low and one High). Mambila has four contrastive tones (High, Upper Mid, Lower Mid and Low). Thus, while Edo and Bulu have no Mid tones, Mambila has two. The presence or absence of an extra tone in a language's inventory respectively indicates more, or less crowded tonal space. This may have greater or lesser implication for tone perception or misperception of such a language as indicated by the example of the comparative data for the misperception of Yorùbá and Edo HH patterns in Section 7.4.4. above. Such differences call for caution in comparing the perception results obtained from these languages with those of the present thesis.

9.3 Production vs. Perception

The findings of the present thesis indicate a lack of symmetry in tone perception and production. In the perception experiments, the results were inconclusive as to whether High or Low tones attracted better perception accuracy. In contrast, the Production experiment highlighted the Low tone as the best-imitated tone by participants. This is perhaps due to the observation that the Low tone has easily noticeable auditory cues. The only instance in the results where there was perception – production symmetry was in the identification of contour tones. The results were similar in the sense that in both perception and production, the contour High (in LH) generally attracted better accuracy than the contour Low (in HL) (see Tables 5.3, 5.5, and 7.8).

The findings on tone production by Orie (2006b) suggest that the Mid tone was hardly acquired by Yorùbá L2 learners who reportedly substituted High and Low tones for the Mid in their production. Similarly, the findings of the present thesis concerning Mid-tone production patterns indicate that participants resorted to it especially when they found it hard to imitate the acoustic cues necessary for the production of the other tones. This is not to be interpreted as indicating that the participants are consciously doing this with adequate knowledge of the tones. To a large degree, it is more likely that the evaluators could not detect the pitch height necessary for the realization of High and Low tones in the non-native participants' production. Hence the classification of such tokens as Mid. In an example for a mis-produced Low tone, if it is produced with a relatively higher F0 or without the falling contour which characterizes the production of the tone, it will be similar to what Hombert (1976) observed where the Low tone, modified with a steady F0 was perceived as Mid by native participants.

The English-Canadian adults, with more experience, represent this trend when they recorded similar rates for both the misrepresentation of High tones as Mid and the misrepresentation of Mid tones as High (Table 7.7). As for the Yorùbá-Canadian young, and the English-Canadian adult and young groups they mis-produced other tones (mostly High tones) as Mid. This trend is also linked to the possibility that the Mid tone is closer relatively to the High tone than to the Low tone in the language's tone space and in terms of F0 measurement. Additionally, the production of the High tone shares more characteristics with the production of the Mid tone, mostly in terms of breathiness (Hayward et al., 2003).

However, there is a fundamental difference between Orie (2006b) and the present thesis. Orie (2006b) was a longitudinal study with data collected over a period of five years in a classroom environment. Her results were based on the analysis of long-term data, which involved reading aloud, naming objects, or picture descriptions, all of which are more demanding than the imitation task in the present thesis. The imitation task in the present thesis was a controlled experiment lasting a few minutes. Moreover, participants were pre-informed of the three Yorùbá tones and were aware of the task required of them. Thus, the task was much easier because participants only needed to listen to nonce-word stimuli and repeat them. In addition, all participants undertook three or four tone perception tasks first before the production task; all of which had practice sessions before the actual experiments. The preceding experiments possibly served as enough exposure for these participants to be able to differentiate and imitate the tones better than if they were to produce them in a void.

9.4 Age and Language Context

One aim of the present thesis was to examine the general effect of Age on tone perception and production. The mixed-effects analyses suggest that adults recorded better accuracy in all the tone perception experiments (AX Discrimination, Tone Identification 1 and 2, Lexical recognition), and the tone Production experiment. It is possible that the adults' age, combined with the level of general language experience contributed to their better performance in these experiments, as suggested by Tong et al. (2015) and Ciocca and Lui (2003).

Another aim of the study was to investigate whether participants' tone language context will generally affect their tone perception and production accuracy. The results for the adult groups suggest no language context effects in the AX Discrimination and the first Identification experiments. In the two experiments, the Yorùbá-Nigerian adults performed comparably to their English-Canadian peers. They also compared to the Yorùbá-Canadian adults in AX Discrimination while performing better than them in the first Identification experiment. Meanwhile, the second Identification and the Production experiments indicate significant effects for language context between the adult groups. The Yorùbá-Nigerian adults were more accurate than their Yorùbá-Canadian and English-Canadian peers. Similarly, the Lexical experiment

which compared the two Yorùbá-speaking groups, suggests that a richer tone language environment, which affords participants continuous exposure to the language, leads to better tone perception accuracy than a poorer one. Unlike the adult groups, the results of the young groups continuously highlight the positive effects of the tone language context on tone perception and production accuracy. The Yorùbá-Nigerian young participants were significantly more accurate than their peers from the Yorùbá-Canadian and the English-Canadian young groups in all the four perception experiments and in the production experiment. Meanwhile, in all the perception experiments, there was no significant difference between the young and adult Yorùbá-Canadian groups and their English-Canadian counterparts. These disparate results call for different interpretations in what follows.

In the AX Discrimination experiment, the effect of the knowledge of the tone language is manifested in the performance of the young groups rather than the adult groups. The two adult groups with tone language background did not perform significantly better than the adult group without tone language background. This possibly relates to the fact that the discrimination task does not require the knowledge of the language. The results of the three adult groups support Connell (2000), Hao (2012) and Francis et al. (2008), who suggest no difference in tone perception performance between tone and non-tone language speakers. It appeared that, coupled with the fact that the task does not require the knowledge of the tone language, other factors, possibly general language experience afforded by the age factor contribute to the comparable performance observed for the adult groups in this experiment. In contrast, the results for the young participant groups highlight the language context effect. Knowledge of the tone language benefitted the young Yorùbá-Nigerian group to outperform the other young groups while there was no such effect in the results of the adult groups. Thus, at least for the young age groups, the result supports the view that exposure to a tone language tends to result in better tone perception accuracy than non-exposure, as found in Wayland and Guion (2004) and Lee et al. (1996). Furthermore, the fact that the Yorùbá-Nigerian young participants performed significantly better than the Yorùbá-Canadian young group, replicates the general findings on attrition effects by Yeh and Lu (2012) that showed that infrequent use of the tone language leads to loss of tone perception accuracy, or to tonal confusion for the attriters.

The first Identification experiment, which tested the knowledge of the three basic tones, presented the most conflicting results. As in the first experiment, the result of the young groups is consistent in highlighting the effects of tone language knowledge. The Yorùbá-Nigerian young group outperformed the other two young groups, whereas there was no difference between the Yorùbá-Canadian and the English-Canadian young groups. Meanwhile, the results of the adult groups did not reflect what was suggested in Bakare (1995) in which native tone language speakers performed better than non-tone language speakers in a similar identification task. It has been suggested earlier that likely the use of real words benefitted the native speakers in Bakare (1995). Instead, the present results support Hao (2012) which shares more affinity with the present experiment in the use of nonce words. The results further reflect the findings of Connell (2000) and Francis et al. (2008) which suggest that both native and non-native tone language speakers can identify lexical tones accurately. These two studies adopted synthesized stimuli (Connell, 2000) or semantically-neutral stimuli (Francis et al., 2008).

Another explanation for this aspect of the results is found in comparing the mean age of the adult participant groups. The English-Canadian group had a much lower mean age of 19;8, compared to 40;8 and 43;5 for the Yorùbá-Canadian adults and Yorùbá-Nigerian adults respectively. Thus, it is possible that the English-Canadian adult group who were all young adults, employed that as an advantage towards general language processing. Second, all the English-Canadian adult participants were students taking courses in Linguistics at the time of the test. Although none of these participants spoke a tone language, most of them had learned, or were learning, a second (non-tone) language. About one-fifth of these participants were reportedly bilinguals, speaking a second (non-tone) language. Thus, being bilingual, whether fully or partially, is an added experience possessed by many of these participants. It is an experience that they may have positively applied, especially in the development of strategies for the perception of alien linguistic features. This is what is suggested by Tong et al. (2015), where English monolinguals, as well as Cantonese-English bilingual adults, performed well in the perception of non-native contrasts (Cantonese tones and English stress).

The effects of attrition and possibly incomplete acquisition are highlighted when the Yorùbá-Canadian young group did not perform comparably to their Yorùbá-Nigerian counterparts, while there was no significant difference between the Yorùbá-Canadian young and their English-

Canadian counterparts. As mentioned earlier, the Yorùbá-Nigerian young group represents the positive effects of tone language knowledge in perceiving lexical tones.

The results of the by-group comparisons for the three remaining experiments are similar. The results show that as the tasks progress, they begin to highlight the increasing importance of the knowledge of the tone language. In the second Identification Experiment which also examined the knowledge of the Yorùbá tones (in this case the complex tones), both the Yorùbá-Nigerian adult and young participants identified the tones significantly better than their Yorùbá-Canadian and English-Canadian counterparts. Thus, the positive effects of tone language context are highlighted, in line with Wayland and Guion (2004), Bakare (1995), and Omozuwa (1991).

Moreover, the aspect of the results which showed that the Yorùbá-Canadian adult and young groups did not perform comparably to their Yorùbá-Nigerian counterparts while performing comparably to the English-Canadian group, depicts the negative effects of attrition on tone perception (Yeh & Lu, 2012). Concerning the performance of the Yorùbá-Canadian adults, the group performed less accurately than the Yorùbá-Nigerian Adults while performing comparably to the English-Canadian adults in the perception experiments (Figure 9.1a). This below-par performance may be due to the relatively limited input of the native language available to these adult participants when compared to the Yorùbá-Nigerian adults. The Yorùbá-Canadian adults comprised of participants who are either the parents of the younger participants or couples residing in Canada. The typical native language usage for these participants is likely restricted to within the home, where the L1 usage usually involves only the parents or the couples, except occasionally with family friends through telephone calls or when they go to Nigeria on a visit. Secondly, the discussion between the couples likely does not cover a wide array of topics. The breadth of the discussion likely covers personal issues and rarely extends to broader discussions of other topics such as political, sports, social and scientific issues. The limited scope in the discourse variety may have contributed to less diverse input.

Moreover, in cases where there are children in the house, the quantity of the L1 input is often reduced further as most of the parent-children communication is in the dominant L2. Thus, infrequent use could lead to limited input in both quality and quantity, or to attrition of some of the features of the L1. This is possibly reflected in their tone perception performance as observed in the present thesis. Meanwhile, the results show that their tone production, as well as their

lexical recognition skills, were not as prone to attrition as their tone discrimination and identification skills. Tone production, as in the present thesis was a simple imitation task while lexical recognition task included pictures and it involved frequently-used lexical items (nouns and verbs) in the language. Additionally, as mentioned in sections 4.4 and 9.5.3, the use of nonce words for the perception experiments may also have contributed to the performance observed for these adult participants. As native Yorùbá speakers, the expectation of real Yorùbá words could have led to wrong analysis of the tones of the stimuli.

The poor performance by the Yorùbá-Canadian young participants has been discussed in different sections of the thesis as possibly the effect of either incomplete acquisition or attrition (see, e.g., section 6.3, 7.4, 8.3, 9.1, and 9.6). Incomplete acquisition by a section of this participant group was possibly aggravated by the home environment from where they acquired their L1. As described above, the L1 usage status of the parents has been characterized by limited input both in quantity and quality. Acquiring the L1 from such an environment is bound to be incomplete. Specific aspects of the tone language will not be fully mastered. Regarding the other members of this group especially the older children who have acquired the L1 from Nigeria, more extended stay in Canada and infrequent use of the L1 in the home, coupled with the strong displacement effect by the dominant L2 contribute to a gradual attrition of those tonal features previously acquired. Meanwhile, the comparative performance between Yorùbá-Canadian and English-Canadian participants could have been compared to studies such as Francis et al. (2008) and Connell (2000) which found no difference in native and non-native tone language speakers' perception of tones. However, the infrequent use of the tone language which characterizes the Yorùbá-Canadian groups in the present study makes them not comparable to the native tone language participants in the studies mentioned.

The results of the Lexical Experiment underscore the importance of tone language knowledge and experience on tone perception. There was no significant difference between the two adult groups: Yorùbá-Nigerian and Yorùbá-Canadian. This is possibly because as adult speakers, they have comparable lexical experience. The results indicate that the lexical items previously acquired, are still intact and accessible whenever they are needed. Additionally, the tasks were made less difficult by the inclusion of pictures. The comparable performance reflects the study of Yeh and Lu (2012) where the older and the younger fluent speakers performed comparably at the

Lexical task. In contrast, the Yorùbá-Nigerian young group recorded significantly better results than their peers from Yorùbá-Canadian category. The results suggest that frequent use of the native language is more beneficial than infrequent use as suggested by Yeh and Lu (2012). The results further highlight the importance of knowledge and experience of the tone language, a trend that has been outlined consistently by the young groups throughout this thesis. The performance of the young Yorùbá-Canadian participants may be due to the lack of use or loss of the native language lexical knowledge. It may also be due to the reduced quantity and quality of the native language input available to these participants.

Meanwhile, the results of the Production Experiment suggets that knowledge and experience of the tone language leads to accurate perception and production of lexical tones. Both the young and adult participants from Yorùbá-Nigerian and Yorùbá-Canadian groups perceived, and consequently produced, the tones significantly better than their counterparts from the English-Canadian group. This result is in line with Yeh and Lu (2012) as well as Omozuwa (1991). The result does not support Hao (2012) which found comparable production ability for native and non-native tone language participants. Additionally, the production result, like the lexical, also highlights the effects of incomplete acquisition and attrition as both the young and adult Yorùbá-Canadian participants still performed significantly worse than their Yorùbá-Nigerian counterparts.

It is important to point out that much of the work cited regarding the native and non-native tone perception (e.g., Wayland & Guion, 2004; Francis et al., 2008; Wang, 2006; Hao, 2012; Orie, 2006) involved non-native speakers who were L2 learners of the tone language. The present thesis differs from these studies in that native tone language speakers are compared to speakers of a non-tone language who are not in the process of learning the tone language of the study. Moreover, unlike some of the works cited (e.g., Hombert, 1988; Connell, 2000; Wayland & Guion, 2004), the non-native speakers in the present thesis did not undergo any pre-experiment training. In addition, the present thesis used nonce words and a different methodology from most of the works cited. Perhaps a longer training period prior to conducting the experiments would have resulted in more familiarization with the Yorùbá tones. That, in turn, could have benefitted all participants, more especially the English-Canadian young group who performed poorly in most of the experiments. In addition to more training, it is also possible that a different

methodology could have yielded richer data that give us more insight into the participants' tone perception patterns.

9.5 Language-internal factors

9.5.1 The three basic tones

Among the language-internal factors investigated were the three basic tones, the two complex tones, the number of syllables in the stimuli, and the use of nonce words for the stimuli. As reported earlier (Chapter One, Section 1.1.1), phonological studies of the Yorùbá basic tones (e.g., Bakare, 1995; Hombert, 1976; Hayward et al., 2003; Akinlabi & Liberman, 2000; and Orie, 1997) have described the acoustic features that differentiate the three tones from each other. Perception studies of the Yorùbá tones (e.g., Hombert, 1977) have observed that native speakers typically perceive the High and the Mid tones as steady tones while the Low tone is typically perceived with a falling F0. Meanwhile, studies of the High and the Low tones in other African languages revealed that the two tones share similar characteristics across the languages observed (e.g., Mambila, Dschang, Ibibio, and Kunama, in Connell, 2002; Edo, in Omozuwa, 1991). Based on these findings, the present thesis predicted that both the High and the Low tones will attract relatively more accurate perception than the Mid tone. The results (Identification 1 and 2) agree with the prediction. These results are different from the findings of an earlier perception study by Bakare (1995) which found that the High tone was perceived more accurately than both the Mid and the Low tones. The improved perception of the High and the Low tones relates to the position of the two tones at either extreme ends of the language's tonal space. The High tone occupies the highest frequency zone while the Low tone occupies the lowest frequency zone. Being at the highest or the lowest frequency zone in a language with three-tone system accords the two tones salient acoustic features that make them distinguishable from the Mid tone. The Mid tone assumes the mid-way position and is often considered a neutral tone. Moreover, the creaky phonation typically associated with the production of the Low tone is a tonal feature that differentiates the tone from the other two tones, and especially from the High tone.

Further examination of the results of the first Identification experiment with respect to individual participant group (Tables 4.6 and 4.7), tends to strengthen the findings of the present thesis with respect to the three basic tones. When the remaining five participant groups were compared to the Yorùbá-Nigerian adult group, only the English-Canadian adult group identified the High tones significantly better than the Low tones. Moreover, all the other groups, including the English-Canadian adults, recorded no significant difference in their identification of High versus Mid tones, or Low versus Mid tones. This part of the results possibly suggests that English speakers are sensitive to the high pitch associated with the High tone, which makes it similar to the primary stress system in English. This trend supports the PAM-S model. Further analysis of the Identification results suggests that perhaps due to the closeness of the pitch contour or due to speaker effect, High tones are mostly misidentified as Mid tones and vice versa. Meanwhile, other aspects of the results suggest that both the High and the Low tones are misidentified as Mid tones at similar rates. However, Low tones are rarely misidentified as High just as High tones are rarely misidentified as Low. This misidentification trend suggests that most participants resort to the Mid tone possibly due to its mid-range position between the High and the Low tones. Both High and the Low tones are rarely misidentified as each other due to the distance between the two tones in the language's perceptual space and also due to the presence of salient acoustic cues in the two tones.

Another confirmation of this finding is found in participants' identification of the two tones in the second Identification Experiment, where all the stimuli were either High- or Low-final. The results (Table 5.6) suggest that none of the two tones recorded more significant perceptual accuracy than the other. Thus, while the present thesis does not align with the hypothesis that High-tone features will lead to more accurate perception than the other two tones, it suggests that both the High and Low tones possess different phonological features that may assist listeners in perceiving them better than the 'neutral-sounding' default Mid tone.

Analysis of the tone production results reveals a clear departure from what is observed in the perception results with respect to the three tones. Generally, the Low tone was the best-imitated tone of all with 95% accurate production. It was followed by the Mid tone with 90% and the High tone with 78% accuracy (Table 7.6). A similar trend was observed in the analysis of the production results for the three groups (Yorùbá-Canadian young group, and English-Canadian

adult and young groups). The creakiness feature of the Low tone, which makes it auditorily more noticeable, may have contributed to this successful imitation. In contrast, the confusion of High with Mid tones, possibly due to F0 closeness and similar phonation type, may be cited for the lower rate of imitation accuracy recorded by these two tones. Thus, the present thesis suggests that, regarding the three tones, perception is not symmetric with production.

The observation about Mid-tone's acoustic space is further highlighted by the results of the response analysis vis-à-vis misidentification data (see Chapter 4, Section 4.3; Figure 4.3). Participants tended to misidentify the Low and High tones as Mid more than as any other possible tones. About 67% of total High-tone misidentification and 71% of total Low-tone misidentification were misrepresented as Mid tone. Misidentification of Mid tone itself was almost equally distributed between High (47%) and Low (53%). Moreover, Table 4.8 (Chapter 4, Section, 4.3) reveals that almost all participant groups misidentified the other tones as Mid tones more than as any other tone. Similarly, in the second Identification Experiment which was comprised of only High and Low-final stimuli, about a quarter of each of the other two tones were identified as Mid. Furthermore, the response analysis for this Experiment not only supports the findings on mid-range position of the Mid tone, resulting in less space (cf. Connell, 2000: Mambila Mid tones), it also highlights the issue of High-Mid tone confusion. This is shown in Table 4.8 (see Chapter 4, Section 4.3) where the High tone was mostly misidentified as Mid by most of the participant groups. Regarding production, an analysis of the participants' tonal confusion pattern further supports the medial position ascribed to the Mid tone. When confronted with tone production difficulty, participants from the three groups (the Yorùbá-Canadian young and the two English-Canadian groups), resorted to the Mid tone. They rarely resorted to the High or Low tone due to their salient acoustic cues and their perceived distance.

The tonal confusion and possibly the production difficulty of Mid tone exhibited by the Yorùbá-Canadian young group may be due to incomplete acquisition of the tones or due to tonal attrition. The Yorùbá tonal systems previously acquired are gradually being displaced and subsequently replaced by the systems of the dominant L2 (English). As for the English-Canadian groups, the difficulty may be due to the lack of tonal inventory in their L1 system, or due to narrower pitch range as suggested by Orie (2006b). Based on their tone production pattern, it could be suggested

that perhaps these two non-native groups are producing the tones, especially the High and Mid, too closely in their tonal space.

9.5.2 Complex Tones

Previous studies of Yorùbá complex tones (e.g., Orie, 2012) examined language acquisition by children and adult learners. The present thesis, along with Shittu and Tessier (2014), is among the few to focus on the perception of the Yorùbá complex tones. While Shittu and Tessier (2014) suggest that Yorùbá complex tones are more difficult to perceive than the simple tones, the present thesis goes further by comparing the perception and production of the two Yorùbá complex tones. The analysis (Tables 5.3 and 5.5) suggests that the contour High was identified significantly better than the contour Low. The analysis further suggests that among the groups, the English-Canadian adults identified the contour High significantly better in comparison to the Yorùbá-Nigerian adults. Additionally, the response analysis for the two contour tones reveals that the contour Low was misidentified at higher rates than the contour High, further highlighting the perceptual difference between the two contour tones.

In a departure from the findings of Shittu and Tessier (2014), the present thesis finds that generally, there was no difference in participants' identification of all-contour tones versus all non-contour tones. However, there is a fundamental difference between that study and the present thesis regarding the number of participants and their language context. Shittu and Tessier (2014) had only twenty-eight (28) participants drawn from the Yorùbá-Canadian category, while the present thesis involved one hundred and thirty-nine (139) participants from across the three language context categories identified for the study. Nonetheless, while the performance of the other participant groups showed a general lack of perceptual difference between contour tones and non-contour tones, the English-Canadian adult group was the only group to show significant difference in their identification of contour versus non-contour tones. The results of the mixed-effect models suggest that generally, contour-High tones (LH-rising) were easier to identify than contour-Low (HL-falling). The results further suggest that contour-Low was more difficult to identify than non-contour Low, while the contour-High tone was in no significant way more difficult to identify than non-contour High.

The analysis of the production results, involving the three focus groups, reveals that the contour High (LH, rising) was the best imitated of all the nine tonal patterns with a 92% accurate imitation, while the contour Low (HL, falling) was the worst imitated pattern with 48% (see Chapter 7, Table 7.8). The production experiment further confirms what was observed in the perception results regarding the complex patterns by suggesting that out of the two contour tones, contour Low (HL) was the pattern most likely to be misrepresented (see Chapter 7, Tables 7.9 and 7.10). Thus, as suggested earlier, the improved production and perception accuracy observed for the rising contour (LH) is possibly connected to its tonal pattern that is comparable to the English iambic stress pattern. Additionally, the acoustic similarities between this contour and the falling-rising intonation contour used in the right-most syllable in English interrogative utterances may also explain the improved production as well as identification of the pattern.

Meanwhile, the tone perception results for the English-Canadian adults confirm the suggestion put forward in PAM-S (So & Best, 2008) and also in Connell (2000) that speakers of non-tonal languages can perceive tone adequately. The results further support a form of phonological assimilation between L1 and L2 (So & Best, 2008) through the identification pattern exhibited by the English-Canadian adults. By identifying the High tone more accurately (in Identification 1), these adult participants appeared to assimilate the Yorùbá High tone to the English primary stress. Similarly, when these English-Canadian adults identified and produced the Yorùbá LH pattern more accurately, they appeared to assimilate this pattern to the English iambic meter. Thus, the relatively accurate perception and production of the LH pattern by these participants is further evidence in support of the PAM-S model. However, the present thesis could not account for the reason why these participants' accurate identification and production lean towards the iambic at the expense of the trochaic even though the two patterns are prominent in the prosodic system of their L1. This is another area which merits a future follow-up study.

9.5.3 Syllabicity and Tone pattern

The present thesis does not record broad syllabicity effects on participants' performance as suggested by Perdereau (2013) and Dommergues and Segui (1989). In all the experiments, the results of the mixed-effect models suggest no significant effect for syllable count. However, the

studies cited above differ from the present thesis in the use of real words as opposed to mostly nonce words which were used as stimuli in four of the five experiments for the present thesis.

Thus, the present thesis suggests that the use of nonce words, whether monosyllabic or disyllabic does not affect tone perception or production accuracy. This does not invalidate the finding that nonce-word stimuli affect word identification as suggested by Fox and Unkefer (1985). However, testing the effect of nonce versus real word stimuli on participants' tone perception and production ability was out of the scope of the present thesis.

9.6 External factors

Previous studies have suggested that some external factors may influence participants' perception and production ability (e.g., Yeh & Lu, 2012; Yeh & Tu, 2012). For the two Yorùbá-Nigerian groups, the fact that they are native speakers of the language, coupled with the fact that they are still residing within the native language community play positively into their tone perception and production performance. However, despite being native speakers, there is still the tendency that using nonce words as stimuli for the experiments may affect their accuracy. This may occur mainly due to the expectation that some stimuli might be real words. While this is not measured in this thesis, it is an area which may be studied in the future. In contrast, the English-Canadian groups possess no tone language experience like the native speakers, making them the perfect control participants for the study. However, despite not having any tone language experience, the results suggest that the adult members of this group, due to their older age, coupled with being students taking courses in Linguistics, possess more general language experience than their younger counterparts. The third group, the Yorùbá-Canadian, are also native speakers of the tone language but without continuous residence in a native language community. Some of the younger participants from this group may have undergone incomplete acquisition earlier on in their L1 development. Incidentally, they are also prone to attrition effects due to lack of use of the native tone language. Concerning these younger participants, the effects of demographic factors (Age in months, Months in Canada, and Place of birth) are discussed next.

The length of stay in Canada prior to the test date proved the most crucial of the three. The results from all the experiments suggest that a shorter stay in Canada contributes to better performance in tone perception and production. Consequently, a longer stay leads to worse performance. This is probably the most important factor supporting the suggestion for attrition effects on participants' performance. It indicates that the longer the participants stay in a community where another language is dominant, the faster the onset of attrition of some of the features of the native language.

With respect to the Age at time of testing, the results for four of the five experiments suggest that older age at the time of testing is likely to lead to better performance. The results further suggest that older age at the time of testing significantly contributes to improved performance in the perception tasks while suggesting no such significant effect for the production task. In the perception tasks, the AX Discrimination, the second Identification and the Lexical experiments recorded significant effects for older age at the time of testing while the first Identification and the Production experiments recorded no significant effects.

Participants' place of birth did not significantly affect their performance in three of the five experiments. Significant effects were observed in only the second Identification and the Lexical experiments while no effects were observed in tone Discrimination, the first Identification, and the Production experiments. Meanwhile, in the two experiments where there were significant effects, the non-Nigerian-born participants tended to be significantly more accurate than the Nigerian-born participants. The interaction between Age and Place of birth suggested that even in the two experiments where the non-Nigerian-born participants seemed to have performed better than the Nigerian-born participants, older age during the test coupled with not being born in Nigeria were likely to reduce participants' accuracy significantly in these experiments. At this juncture, one may reasonably question why the Nigerian-born participants, despite being born in the immediate tone language environment, did not perform better than the non-Nigerian-born participants. It could be because some of the Nigerian-born participants were brought to Canada at a very early age (pre-perceptual reorganization stage; pre-Nigerian schooling), or they had had longer stay in Canada before the testing. Longer stay in Canada means more years of Canadian schooling. As such, they may have experienced a rapid displacement of their tone language skills by the L2, or due to an interplay of all these factors.

Thus, out of the three demographic factors considered for the Yorùbá-Canadian young group, longer stay in Canada contributes to lower accuracy in the four perception experiments and the production experiment. Older age at testing leads to more accurate performance in tone discrimination, identification of the contour tones, and lexical recognition. However, the factor does not have any significant effect on tone production, indicating that tone imitation is, perhaps, not as prone to attrition as tone perception. Regarding the Place of birth factor, non-Nigerian born participants are not significantly different from Nigerian-born participants in tone discrimination, identification of the basic tones, and tone production, while they tend to be better than Nigerian-born participants in the identification of the contour tones and in lexical recognition. Thus, as a native speaker of a language who has undergone the perceptual reorganization stage earlier on, it is not necessary to be born in the native language environment to be able to perceive or produce the features of the language. However, some aspects of the results also suggest that with increase in age, these participants born outside Nigeria may lose their tone perception ability over the long term. This is one finding that may need a follow up study.

9.7 Future directions

Some aspects of the results of the present thesis which merit follow-up studies include the effects of the language-internal factors such as same-tone versus different-tone, monosyllabic versus disyllabic and the use of real versus nonce words. While the study suggests no significant effects for same-tone or different-tone pairs on participants' discrimination, the fact that individual groups exhibited significant accuracy in different-tone pairs compared to same-tone pairs calls for a future study on this aspect of the results. Thus, an expanded study with more stimuli will be a fertile avenue to examine possible High-Mid confusion especially in monosyllabic pairs. Such a study may include populations drawn from both native and non-native tone language speakers. Similarly, the thesis suggests no significant effects for Syllable count in all experiments. However, there are hints that monosyllabic stimuli induce more tonal confusion such as the High-Mid confusion observed in the AX Discrimination, the Identification and the Production experiments. Thus, the issue of High-Mid tone confusion as revealed by the present thesis is one

area that merits a follow-up study. Such a future study may include monosyllabic stimuli bearing the three tones, modelled along the present AX Discrimination and the Identification experiments. It may involve native Yorùbá speakers, native speakers of other tone languages, and native speakers of non-tone languages. Moreover, the use of real words in the lexical experiment possibly leads to improved performance by the native tone language speakers. However, this claim also needs follow-up studies which compare nonce-word stimuli with the real words in experiments that involve only the native Yorùbá-speaking population.

One of the areas in which the lexical experiment could be expanded in the future is to investigate the possibility that monosyllabic stimuli may induce tonal or lexico-semantic confusion for listeners. Tonal confusion in monosyllabic pairs of words with High and Mid tones could occur due to the closeness of the pitch height between the two tones. Likewise, semantic confusion between lexical items (e.g., between sē 'to do' and sè 'to cook') is also possible in these monosyllabic sets. This type of confusion is not manifested in disyllabic and trisyllabic stimuli due to the presence of a contrasting tone in the resulting pattern. For instance, in the stimulus '*rēlú*' which has a Mid tone followed by a High tone, the two tones are produced with different pitch heights. The penultimate Mid tone could aid in more accurate perception of the final High tone, thereby eliminating any tonal confusion.

9.8 Conclusion

The five experiments conducted in the present thesis have shown, in various ways, the association between experience and tone perception and production. The experiments have shown that increased experience, i.e., older age, leads to comparable performance at tone discrimination and identification. Further increased experience through language context, i.e., whether you live in a tone language environment or not, generally leads to improved tone perception and production. The perception and production of the three basic tones has revealed that salient acoustic features present in both the High and the Low tones have contributed to comparable perceptual and production accuracy for the two tones. The absence of such features in the Mid tone leads to lower perceptual accuracy for the tone. Throughout the perception and the production experiments, the English-Canadian adults have shown significant accuracy in the

identification of the High tones as well as the rising contour. I suggest in this thesis that the English-Canadian adult participants benefit from their knowledge of the English stress system. Evidence from the demographic data for the Yorùbá-Canadian young participants indicates that the length of stay in Canada correlates with exposure to the tone language, contributing to reduced perceptual accuracy. Thus, the present thesis contributes to the understanding that experience, whether through age or language exposure mediates tone perception and production.

Bibliography

Ajolore, O. (1974). *Learning to use Yorùbá focus sentences in a multilingual setting.* Ph.D. Dissertation, University of Illinois, Urbana-Champaign.

Akinlabi, A. (1985). Tonal Underspecification and Yorùbá Tones, Doctoral Dissertation, University of Ibadan.

Akinlabi, A. & Liberman, M. (2000). The tonal phonology of Yorùbá clitics. In B. Gerlach and J. Grijzenhoul, eds., *Clitics in phonology, morphology and syntax* Pp. 31-62. Amsterdam: John Benjamins.

Arciuli, J., Monaghan, P., & Seva, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational investigations. *Journal of Memory and Language*, 63(2):180–196.

Awobuluyi, 0. (1967). Vowel and Consonant Harmony in Yorùbá. *Journal of African Languages* 6, 1-8.

Awobuluyi, 0. & Bamgbose, A. (1967). Two Views of Vowel Harmony in Yorùbá. *Journal of African Languages* 6, 274-277.

Bakare, C.A. (1995). Discrimination and identification of Yorùbá tones: perception experiments and acoustic analysis. In Owolabi, K. (ed.), *Language in Nigeria.* Ibadan: Group Publishers.

Bamgbose, A. (1967). Vowel Harmony in Yorùbá. *Journal of African Languages* 6, 268-273

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(*1*), 1–48.

Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht, the Netherlands: Foris Publications.

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and non-speech sounds. *Journal of Experimental Psychology-Human Perception and Performance*, *32*(1), 97-103.

Beranek, L.L. (1949). *Acoustic Measurements, New York:* Wiley. Used online at the site: http://www2.ling.su.se/staff/hartmut/umrechnung.htm/.

Boersma, P. & Weenink, D. (2012). *PRAAT: Doing phonetics by computer* [Computer program], Version 5.3.04. Retrieved 06 February 2012 from http://www.praat.org/

Bullock, B. E. & Gerfen, C. (2004). Frenchville French: A case study in phonological attrition. *International Journal of Bilingualism* 8: 303.

Burnham, D. & Mattock, K. (2007). The perception of tones and phones. *Language experience in second language speech learning: In honor of James Emil Flege*, 259-280.

Chen, A. & Kager, R. (2012). Lexical tone perception in early infancy. *Speech Prosody 2012* (pp. 230-233).

Ciocca, V. & Lui, J. (2003). The development of the perception of Cantonese lexical tones. *Journal of Multilingual Communication Disorders*, 1(2), pp.141-147.

Cole, J. (2015). Prosody in context: a review. *Language, Cognition and Neuroscience*. *30*(*1-2*), 1–31.

Connell, B. (2002). Tone languages and the universality of intrinsic F0: evidence from Africa. *Journal of Phonetics* 30, 101–129.

Connell, B. (2000). The perception of lexical tone in Mambila. *Language and Speech* 43, 163–182.

Connell, B. & Ladd, R. (1990). Aspects of pitch realization in Yorùbá. *Phonology* 7, 1-29.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech, 40,* 141–201.

DeCasper, A.J. & Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant behavior and development* 9, 133-150.

Dommergues, J-Y. & Segui, J. (1989). List Structure, Monotony, and Levels of Processing. *Journal of Psycholinguistic Research,* 18 (3), pp.245-253.

Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing 'deafness' in French? *Journal of Memory and Language, 36,* 406–421.

Elliot, C. E. (1991). The relationship between the perception and production of Mandarin tones: An exploratory study. *University of Hawai'i Working Papers in ESL.* 10(2),177-204.

Fowler, C.A., Brown, J.M., Sabadini, L. & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language,* 49, 396–413.

Fox, R. A. & Unkefer, J. (1985). The Effect of Lexical Status on the Perception of Tone. *Journal of Chinese Linguistics*, Vol.13, No.1, 69-90.

Francis, A.L., Ciocca, V., Ma, L. & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*, 268 – 294.

Fromkin, V.A. (Ed.) (1978). *Tone: A linguistic survey.* Academic Press, New York.

Gandour, J.T, & Harshman, R. (1978). Cross-language difference in tone perception: A multidimensional scaling investigation. Language and Speech, 21, 1–33.

Gandour, J.T. (1978). The perception of tone. In *Tone: A linguistic survey*, 41-76.

Gandour, J.T. (1983). Tone perception in far-eastern languages. *Journal of Phonetics*, *11*(2), 149-175.

Gandour, J.T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 235-261.

Gandour J.T., Wong, D., Hsieh, L., Weinzapfel, B., Lancker, D.V. & Hutchins, G.D. (2000). A cross-linguistic PET study of tone perception, *J. Cognitive Neuroscience 12:1*, pp.207-222.

Hallé, P.A., Chang, Y., & Best, C.T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners*, J. of Phonetics, 32(3),* 395-421.

Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40,269-279.

Hao, Y-C. & Jong, K. (2016). Imitation of second language sounds in relation to L2 perception and production. *Journal of Phonetics 54*, 151–168.

Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110. 581-616.

Hayward, K, J. Watkins & A. Oyetade (2003). The phonetic interpretation of register: Evidence from Yorùbá. In J. Local, R Ogden, & R. Temple, *Phonetic Interpretation. Papers In Laboratory Phonology VI*. Cambridge: Cambridge University Press. Pp. 305–321.

Herschensohn, J. R. (2007). *Language development and age*. Cambridge: Cambridge University Press.

Hirsh, K. W., Morrison, C. M., Gaset, S. & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117-128.

Hombert, J-M. (1988). Tonper, un test de perception pour langues tonales: application au Bulu. *Pholia* 3, 169–182.

Hombert, J-M. (1977). Consonant types, vowel height, and tone in Yorùbá. *Studies in African Linguistics* 8, 173–190.

Hombert, J-M. (1976). Perception of tones of bisyllabic nouns in Yorùbá. *Studies in African Linguistics* Supplement 6, 109–121.

Hutz, M. (2004). Is there a natural process of decay? A longitudinal study of language attrition. In M. S. Schmid, B. Köpke, M. Keijzer and L. Weilemar (Eds.*), First Language Attrition: Interdisciplinary perspectives on methodological issues,* 189 – 206. Amsterdam: John Benjamins.

Hyman, L. M. (2001). Tone systems. *Language typology and language universals: An international Handbook*, *2*, 1367-1380.

Jia, G. (2008). Heritage language development, maintenance, and attrition among recent Chinese immigrants in New York City. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language* (pp. 189–203). Honolulu, HI: National Foreign Language Resource Center/University of Hawaii Press.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescent learning English in the United States. *Applied Psycholinguistics, 24,* 131–161.

Jia, G., Strange, W., Wu, Y., Collado, J. & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America,* 119, 1118–1130.

Jusczyk, P.W., Houston, D.M. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* **39***,* 159-207.

Ladefoged, P. (1962). Sub-glottal activity during speech. *Proceedings of the Fourth International Congress of Phonetic Sciences*. Helsinki, 1961, 73 - 91.

Lee, Y., Vakoch, D. A. & Wurm, L. G. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *25*, 527 – 542.

Li, H. (1995). Phonetic errors in Thai students learning Mandarin. *International Chinese Pedagogy.* 32, 66-71.

Liu, L. & Kager, R. (2012). Non-native tone perception from infant to adult: How consistent and flexible is it? http://www.*speechprosody*2012.org/uploadfiles/file/sp2012_submission_107.pdf

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, *76*(2), 190-208.

Mattock, K. & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy,* 10(3):241-265.

Mattock, K., Molnar, M., Polka, L. & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106 (3):1367-1381, 2008.

Montrul, S. (2015). *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor* (Vol. 39). John Benjamins Publishing.

Nazzi, T., Floccia, C. & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development,* 21, 779–784.

Nguyen, T. T. A., & Ingram, J. (2005). Vietnamese acquisition of English word stress. *Tesol Quarterly*, *39*, 309-319.

Omozuwa, V. E. (1991). Acoustic cues for the perception of tones of disyllabic nouns in Edo. *Studies in African Linguistics* 22, 135–156.

Orie, O. (1997). *Benue-Congo Prosodic Phonology and Morphology in Optimality Theory*. Muchen: Lincom Europa.

Orie, O. (2006). L2 Acquisition and Yorùbá Tones: Issues and Challenges. In Arasanyin O. F. & Pemberton, Michael A., (eds,). *Selected Proceedings of the 36th Annual Conference on African Linguistics,* pp.121-128. Somerville, MA: Cascadilla Proceedings Project.

Orie, O. (2012). *Acquisition reversal. The effects of postlingual deafness in Yorùbá. Studies on language acquisition series [SOLA]* 47. De Gruyter Mouton.

Oyebade, F.O. (1990). The Phonology of Yorùbá Child. *Research in African Languages and Linguistics*, Vol., No.1, 17-34.

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English Language Learners: Using Parent Report on First Language Development (PDF). Journal of Communication Disorders, Volume 43, pp. 474-497.

Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism, in B. K̩pke, M. Schmid, M. Keijzer, and S. Dostert [Ed], *Language attrition: Theoretical Perspectives,* 121-133, Amsterdam: John Benjamins.

Perdereau, X. (2013). Decrease of pitch perception ambiguity in tone language processing. *Proceedings of Meetings on Acoustics,* Vol. 19, (1) 060229 (2013). Acoustical Society of America [DOI: 10.1121/1.4800070].

Pulleyblank, D., (2004). A note on tonal markedness in Yorùbá. *Phonology*, *21*(03), 409-425.

Pulleyblank, D. (1986). *Tone in Lexical Phonology*. Dordrecht: D. Reidel Publishing Co.

Pulleyblank D. (1988). Vowel Deletion in Yorùbá. *Journal of African Languages and Linguistics* 10: 117-136.

Professional Psychology Software Tools Inc. (2012). *E-Prime 2.0.* [Computer Software] http://www.pstnet.com.

Quansheng, X., Lan, W. and Gang, P. (2016). Nouns and verbs in Chinese are processed differently: Evidence from an ERP study on monosyllabic and disyllabic word processing. *Journal of Neurolinguistics* 40, pp.66-78.

R Core Team, (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL http://www.R-project.org/.

Seliger, H., & Vago, R. (1991). The study of first language attrition: an averview. *First language attrition.* Cambridge University Press. Cambridge, UK.

Shittu, S., & Tessier, A. M. (2014). Perceptual attrition of lexical tone among L1 Yorùbá-speaking children in Canada. *In Proceedings of the BUCLD 39*. Boston University.

Shockley, K., Sabadini, L. & Fowler, C.A. (2004). Imitation in shadowing words. *Perception & Psychophysics,* 66, 422–429.

Smith, T. S. & Shand, M. (1974). Dichotic perception of initial consonants and lexical tone in Cantonese. *The Journal of the Acoustical Society of America*, *55*(S1), S11-S11.

So, C. K. (2000). *Tonal production and perception patterns of Canadian raised Cantonese speakers.* M.A. dissertation, Simon Fraser University.

So, C. K. (2006). Effect of L1 Prosodic Background and AV Training on Learning Mandarin Tones by Speakers of Cantonese, Japanese, and English, Ph.D. Thesis, Department of Linguistics, Simon Fraser University, 2006.

So, C.K. (2006). Perception of non-native tonal contrasts: Effects of native phonological and phonetic influence. In *Proceedings of the11th Australian international conference on speech science & technology,* pp.438–443.

So, C. K. & Best, C.T. (2008): "Do English speakers assimilate Mandarin tones to English prosodic categories?", In *Interspeech-2008*, 1120.

So, C. K. & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and speech*, *53*(2), 273-293.

So, C. K. & Best, C. T. (2011). Categorizing Mandarin tones into listeners' native prosodic categories: The role of phonetic properties. *Poznań Studies in Contemporary Linguistics PSiCL*, *47*, 133.

So, C. K. & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition, 36,* 195–221. doi:10.1017/S0272263114000047.

Sze, W.-L. (2004). *The development of tone perception in Cantonese-speaking children.* Unpublished B.Sc. dissertation. University of Hong Kong.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of phonetics*. *39*(*4*)*,* 456–466.

Tong, X., Lee, SMK., Lee, MML. & Burnham, D. (2015). A Tale of Two Features: Perception of Cantonese Lexical Tone and English Lexical Stress in Cantonese-English Bilinguals. *PLoS ONE 10(11)*: e0142896. doi:10.1371/journal.pone.0142896.

Valdés, G. (2001). Heritage language students: Profiles and possibilities. In J. K. Peyton, D. A. Ranard, & S. McGinnis (Eds), *Heritage languages in America. Preserving a national resource* (pp. 37–80). McHenry, IL: Center for Applied Linguistics.

Wang, X. (2006). Perception of L2 tones: L1 lexical tone experience may not help. In *Proceedings of Speech Prosody*, Dresden, Germany.

Ward, I. C. (1952). *An Introduction to the Yorùbá language*. Cambridge. Cambridge University Press.

Wayland, R. P. & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning,* 54, 681 – 712.

Whalen, D. H. & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of phonetics*, *23*(3), 349-366.

Whalen, D. H. & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, *49* (1), 25-47.

Xu, Y., Gandour, J.T. & Francis, A.L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *J. Acoust. Soc. Am., 120(2),* 1063-1074.

Yang, B. & Ankenmann, R. (2007). The gap between the perception and production of tones by American learners of Mandarin. Presented at *American Association for Applied Linguistics annual conference.*

Yang, J. & Liu, C. (2012). Categorical Perception of Lexical Tone for Monolingual and Bilingual School-aged Children. *International Journal of Asian Language Processing*, 22, pp.49-62.

Ye, Y. & Connine, C. M. (1999). Processing spoken Chinese: the role of tone information. *Language and Cognitive Processes*, *14*(5/6), 609-630.

Yeh, C-H. (2011). Language Attrition and Tonal Change in Hakka, in Proceedings of the Psycholinguistic Representation of Tone Conference (PLRT), 111-114, Hong Kong: Chinese University of Hong Kong.

Yeh, C-H. & Lin, Y-H. (2013). The attrition of Hai-Lu Hakka's tonal system. *Proceedings of ICPLC 2013.*

Yeh, C-H. & Lu, C-J. (2012). The effect of language attrition on low level tone in Hakka. In Ma, Q., Ding, H. & Hirst, D. (eds.), *Proceedings of the 6th International Conference on Speech prosody* 1, Shanghai, 342-345.

Yeh, C-H. & Tu, J-Y. (2012). The effect of language attrition and tone sandhi on Taiwanese tonal processing. In Ma, Q., Ding, H., Hirst, D. (eds.), *Proceedings of the 6th International Conference on Speech prosody* 1, Shanghai, 87-90.

Yip, M. (2002). *Tone*. Cambridge University Press. Cambridge, UK.

Zhang, Y.H., Nissen, S.L. and Francis A.L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America*, 123(6):4498–4513.

Appendices

Appendix A: Chapter One supplementary information

Appendix A1. Language Environment Questionnaire (Adapted from ALEQ)

A. General Information about the participant:

Name / Participant Code: ----------

Date of birth of the participant: ----------

Gender of the participant: ----------

Date of interview: ----------

Interpreter or broker (*if any*) / Research Assistant:----------

	Age at Test				**Age of Arrival**		
	Year	Month	Day		Year	Month	Day
Date of Testing	-----	-----	-----	Date of Arrival	-----	-----	-----
Date of Birth	-----	-----	-----	Date of Birth	-------	-----	-----
Chronological age	-----	-----	-----	Chronological age	-----	-----	-----

B. Language background:

<u>Questions to the participant's father / mother</u>

1. Where was the participant born? _____
2. How many years have you been in Canada? _____

   Approximate date of arrival (month/year)? _____ / _____
3. Did s/he (the participant) come to Canada at the same time? Yes No
4. If not, when did s/he (the participant) come to Canada? _____

5. What language do you speak with the target child?

   YOR never / ENG always = 0

   YOR seldom / ENG usually = 1

   YOR 50% / ENG 50% = 2

   YOR usually / ENG seldom = 3

   YOR almost always / ENG almost never = 4

6. What language does the child speak with you and his/her mother/father?

YOR never / ENG always = 0

YOR seldom / ENG usually = 1

YOR 50% / ENG 50% = 2

YOR usually / ENG seldom = 3

YOR almost always / ENG almost never = 4

7. What language(s) do you speak with his/her mother/father?

YOR never / ENG always = 0

YOR seldom / ENG usually = 1

YOR 50% / ENG 50% = 2

YOR usually / ENG seldom = 3

YOR almost always / ENG almost never = 4

8. What language do you speak most often with other people in your home?

Mostly English = 0

Mostly Yorùbá = 4

9. Are there other relatives in the home? For example, a grandmother? Yes No

If yes, how many? _____

10. What language does this relative speak with the child?

Mostly English = 0

Mostly Yorùbá = 4

11. What language does the child speak with this relative?

Mostly English = 0

Mostly Yorùbá = 4

12. Does the child have brothers and sisters? Yes No

Sibling	Gender M/F	DOB yy/mm/dd	Age	Lang. spoken with the sibling				
				0	1	2	3	4
1.	------	----------	----	----------				
2.	------	----------	----	----------				
3.	------	----------	----	----------				
4.	------	----------	----	----------				
5.	------	----------	----	----------				

13. What language(s) do the siblings speak with the target child?

Mostly English = 0

Mostly Yorùbá = 4

14. What language does the target child speak with his/her siblings?

Mostly English = 0

Mostly Yorùbá = 4

Language Index: YOR=Yorùbá; ENG=English

Appendix A2.

INFORMATION & CONSENT FORM (ADULT-ENG-CAN)

Research Investigator: Saliu A. Shittu
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3.

Research Supervisor: Dr. Anne-Michelle Tessier
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3.

1. Background

You are being asked to be in this study as a non-Yorùbá person. As a non-Yorùbá speaking adult, you meet the basic requirement of the study. The result of this study will form part of my PhD thesis. The findings of this research will not be used for any commercial purpose.

2. Purpose

The purpose of the research is to test how children, as well as adults are able to notice and identify the three basic Yorùbá tones namely the high, mid and low tones. I will be comparing responses from English speaking adults and children with responses from Yorùbá speaking adults and children. The research will give us some insight about the nature and the extent to which different population groups can identify these tones. I will use the information collected for my research project as part of the PhD requirement.

3. Study Procedures

The research will begin with a questionnaire to record some personal and language backgrounds. This will be followed by conducting the three different types of experiments which are specially designed for this study. In the first experiment, you will listen to two non-Yorùbá words being said with Yorùbá tones and vowels. Then you will decide whether the tones of the two words are the same or different. During the second stage of the experiment, you will listen to other non-Yorùbá words, paying special attention to the last syllable of the word. Then you will identify the tone of the last syllable as either high, mid or low tone. Finally, you will listen to some words and then imitate them as best as you can. I will record your imitations using the recording equipment at the University of Alberta phonology laboratory. The whole meeting will last about 40 minutes.

4. Venue of the study:

All the experiments for the research will take place at the Department of linguistics' Phonology Laboratory.

5. Benefits

There is no personal benefit for you for being in this study. However, there is a reasonable benefit to the Yorùbá community. We hope that the information we get from this study will help us better understand the extent to which Yorùbá children living outside Nigeria could still know the three basic Yorùbá tones, even though they may not be speaking the language regularly. There is no cost for you for being in the research except for the time you will give for taking part. There is no compensation for taking part either.

6. Risks

There is minimal risk involved. If I learn of anything during the research that may affect your willingness to continue being in the study, I will inform you right away.

7. Voluntary Participation

You are under no obligation to participate in this study. Even during the course of the meeting, you are not obliged to answer any specific question you do not wish to answer, or you are not comfortable with. I am also assuring you that even if you agree to be in the study you can change your mind and withdraw at any time. If you withdraw, I will continue to use the data I have collected from you along with other data until the end of the study or at the expiration period when the data will be destroyed.

8. Confidentiality and Anonymity

As I have said above, the present research is part of my PhD programme and you will not be personally identified in the paper. Once collected, your name will be removed (anonymized) and your data put together with others for analysis and reporting, and will not identify you in any way. The information I will collect from you will be kept safe and confidential. Only my supervisor and I will have access to the information. All the information about the study will be saved in a file on a password-protected computer with the file encrypted. We may present the results of the study at academic venues at some point. If it is presented or published in academic

venues, you will not be personally identified in any way**.** The study data including any information about you will be securely stored for five years after the study is over, at which time it will be destroyed. Anonymized data for this study will be retained for research in the future. I will make arrangement for you in case you want to have access to the report of the research findings.

If you have any further questions regarding this study, please do not hesitate to contact me or my supervisor on these numbers:

Me	Supervisor:
Saliu Shittu	Dr. Anne-Michelle Tessier:

The plan for this study has been reviewed for its adherence to ethical guidelines by the Research Ethics Board at the University of Alberta. If you have concerns about this study, you may contact the Research Ethics Office at (780) 492-2615. This office has no direct involvement with this project.

CONSENT

Title of Study: A study of the effects of language attrition on the perception of Yorùbá tones

Principal Investigator(s): Saliu A. Shittu

Study Coordinator: Dr. Anne-Michelle Tessier

	Yes	No
Do you understand that you have been asked to be in a research study?	----	----
Have you read and received a copy of the attached Information Sheet?	----	----
Do you understand the benefits and risks involved in taking part in this study?	----	----
Have you had an opportunity to ask questions and discuss this study?	----	----
Do you understand that you are free to leave the study at any time, without having to give a reason?	----	----
Has the issue of confidentiality been explained to you?	----	----
Do you understand who will have access to your records?	----	----
I agree to take part in this study:	----	----

Signature of Research Participant: ----------

Printed Name: ----------

Date: ----------

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee---------- Date ----------

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

Appendix A3

INFORMATION & PARTICIPANT CONSENT FORM (ADULT- YOR)

Research Investigator: Saliu A. Shittu
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3.

Research Supervisor: Dr. Anne-Michelle Tessier
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3

1. Background

You are being asked to be in this study because you are a Yorùbá person and you meet the basic requirement of the study. The result of this study will be used in support of my thesis. The findings of this research will not be used for any commercial purpose.

2. Purpose

The purpose of the research is to test how children, as well as adults are able to notice and identify the three basic Yorùbá tones namely the high, mid and low tones. The research will give us some insight about the nature and the extent to which different population groups can identify these tones. I will use the information collected for my research project as part of the PhD requirement.

3. Study Procedures

The research will begin with a questionnaire to record some personal and language backgrounds. This will be followed by conducting the three different types of experiments which are specially designed for this study. In the first experiment, you will listen to two non-Yorùbá words being said with Yorùbá tones and vowels. Then you will tell me whether the tones of the two words are similar or different. During the second stage of the experiment, you will listen to other non-Yorùbá words, paying special attention to the last syllable of the word. Then you will identify the tone of the last syllable as either high, mid or low tone. In the third task you will listen to real Yorùbá words. First you will see the word on the screen and then hear the word. Then you will see two pictures on the screen. One of these two pictures matches the word that you hear. You are to choose the one which represent the word you hear. Finally, you will listen to some words and then imitate them as best as you can. I will record your imitations using the recording

equipment I will bring from the University of Alberta Phonology Laboratory. The whole meeting will last about 30-40 minutes.

4. Venue of the study: All the experiments for the research will take place in your house.

5. Benefits

There is no personal benefit for you for being in this study. However, there is a reasonable benefit to the Yorùbá community. We hope that the information we get from this study will help us better understand the extent to which Yorùbá children living outside Nigeria could still know the three basic Yorùbá tones, even though they may not be speaking the language regularly. There is no cost for you for being in the research except for the time you will give for taking part. There is no compensation for taking part either.

6. Risks

I will conduct the meeting in your house and as such there is minimal risk involved. If I learn of anything during the research that may affect your willingness to continue being in the study, I will inform you right away.

7. Voluntary Participation

You are under no obligation to participate in this study. Even during the course of the meeting, you are not obliged to answer any specific question you do not wish to answer, or you are not comfortable with. I am also assuring you that even if you agree to be in the study you can change your mind and withdraw at any time. If you withdraw, I will continue to use the data I have collected from you along with other data until the end of the study or at the expiration period when the data will be destroyed.

8. Confidentiality and Anonymity

As I have said above, the present research is part of my PhD programme and you will not be personally identified in the paper. It is only the results of the experiments which I will use and analyze. I will use them alongside other information which will not identify you in any way. The information I will collect from you will be kept safe and confidential. Only my supervisor and I will have access to the information. All the information about the study will be saved in a file on a password-protected computer with the file encrypted. We may present the results of the study

at academic venues at some point. If it is presented or published in academic venues, you will not be personally identified in any way**.** The study data including any information about you will be securely stored for five years after the study is over, at which time it will be destroyed. I will make arrangement for you in case you want to have access to the report of the research findings.

If you have any further questions regarding this study, please do not hesitate to contact me or my supervisor on these numbers:

Me	Supervisor:
Saliu Shittu	Dr. Anne-Michelle Tessier

The plan for this study has been reviewed for its adherence to ethical guidelines by the Research Ethics Board at the University of Alberta. If you have concerns about this study, you may contact the Research Ethics Office at (780) 492-2615. This office has no direct involvement with this project.

CONSENT

Title of Study: A study of the effects of language attrition on the perception of Yorùbá tones

Principal Investigator(s): Saliu A. Shittu

Study Coordinator: Dr. Anne-Michelle Tessier

	Yes	No
Do you understand that you have been asked to be in a research study?	----	----
Have you read and received a copy of the attached Information Sheet?	----	----
Do you understand the benefits and risks involved in taking part in this study?	----	----
Have you had an opportunity to ask questions and discuss this study?	----	----
Do you understand that you are free to leave the study at any time, without having to give a reason?	----	----
Has the issue of confidentiality been explained to you?	----	----
Do you understand who will have access to your records?	----	----
I agree to take part in this study:	----	----

Signature of Research Participant: ----------

Printed Name: ----------

Date: ----------

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee---------- Date ----------

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

Appendix A4

INFORMATION & CONSENT FORM (PARENT-ENG-CAN)

Research Investigator: Saliu A. Shittu
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3.

Research Supervisor: Dr. Anne-Michelle Tessier
Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2R3

1. Background

This is to request your consent for your child……. to be in a research study. Your child is being asked to be in this study as a non-Yorùbá person. As a child, s/he meets the basic requirement of the study. The results of this study will be used as part of my PhD thesis. The findings of this research will not be used for any commercial purpose.

2. Purpose

The purpose of the research is to test how children as well as adults are able to notice and identify the three basic Yorùbá tones namely the high, mid and low tones. I will be comparing responses from English speaking adults and children with responses from Yorùbá speaking adults and children. The research will give us some insight about the nature and the extent to which different population groups can identify these tones. I will use the information collected for my thesis as part of the PhD programme requirements.

3. Study Procedures

The research will begin with a questionnaire to record some personal and language backgrounds. This will be followed by conducting the three different types of experiments which are specially designed for this study. In the first experiment, s/he will listen to two non-Yorùbá words being said with Yorùbá vowels and tones. Then s/he will decide whether the tones of the two words are the same or different. During the second stage of the experiment, s/he will listen to other non-Yorùbá words, paying special attention to the last syllable of the word. Then s/he will identify the tone of the last syllable as either high, mid or low tone. Finally, s/he will listen to some words and then imitate them as best as s/he can. This will be recorded using recording equipment from the University of Alberta Phonology Laboratory. The whole meeting will last about 40 minutes.

4. Venue of the study:

All the experiments for the research will take place at your child's school or at your home if you wish.

5. Benefits

There is no personal benefit for him/her for being in this study. However, there is a reasonable benefit to the Yorùbá community. We hope that the information we get from this study will help us better understand the extent to which Yorùbá children living outside Nigeria could still know the three basic Yorùbá tones, even though they may not be speaking the language regularly. There is no cost for him/her for being in the research except for the time s/he will give for taking part. There is no compensation for taking part either.

6. Risks

I will conduct the meeting at your home or your child's school and as such, there is minimal risk involved. If I learn of anything during the research that may affect his/her willingness to continue being in the study, I will inform you right away.

7. Voluntary Participation

You are under no obligation to allow your child to participate in this study. Even once you give permission, your child has the right to provide his/her own assent to take part, or to not take part. During the course of the meeting s/he is not obliged to answer any specific question s/he does not wish to answer, or s/he is not comfortable with. I am also assuring you that even if you agree for him/her to be in the study you and/or your child can change your mind and withdraw at any time. If you withdraw your child, you have the right to withdraw his/her data also from the study. If not, I will continue to use the data I have collected from him/her along with other data until the end of the study or at the expiration period when the data will be destroyed.

8. Confidentiality and Anonymity

As I have said above, the present research is part of my PhD programme and s/he will not be personally identified. Once collected, your child's name will be removed (anonymized) and his/her data put together with others for analysis and reporting and will not identify your child in any way. The information I will collect from him/her will be kept safe and confidential. Only my

supervisor and I will have access to the information. All the information about the study will be saved in a file on a password-protected computer with the file encrypted. We may present the results of the study at academic venues at some point. If it is presented or published in academic venues, your child will not be personally identified in any way**.** The study data including any information about him/her will be securely stored for five years after the study is over, at which time it will be destroyed. Anonymized data for this study will be retained for research in the future. I will make arrangement for you in case you want to have access to the report of the research findings.

If you have any further questions regarding this study, please do not hesitate to contact me or my supervisor

Me	Supervisor:
Saliu Shittu	Dr. Anne-Michelle Tessier

The plan for this study has been reviewed for its adherence to ethical guidelines by the Research Ethics Board at the University of Alberta. If you have concerns about this study, you may contact the Research Ethics Office at (780) 492-2615. This office has no direct involvement with this project.

CONSENT

Title of Study: A study of the effects of language attrition on the perception of Yorùbá tones

Principal Investigator(s): Saliu A. Shittu Study Coordinator: Dr. Anne-Michelle Tessier

	Yes	No
Do you understand that your child has been asked to be in a research study?	----	----
Have you read and received a copy of the attached Information Sheet?	----	----
Do you understand the benefits and risks involved in your child taking part in this study?	----	----
Have you had an opportunity to ask questions and discuss this study?	----	----
Do you understand that as a parent, you are free to withdraw your child from the study at any time, without having to give a reason?	----	----
Has the issue of confidentiality been explained to you?	----	----
Do you understand who will have access to your child's records?	----	----
Do you understand that your child has the opportunity to provide individual assent, and that if s/he chooses not to participate, your child's wishes will be respected?	----	----

Who explained this study to you? ----------

I ----------provide parental consent for my child ---------- to take part in this study.

Signature of parent ---------- Parent's name---------- Date:----------

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee---------- Date ----------

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

Appendix A5

INFORMATION & CONSENT FORM (PARENT-NGR)

Research Investigator: Saliu A. Shittu
Department of Linguistics, University of Alberta, Edmonton, AB. T6G 2R3.

Research Supervisor: Dr. Anne-Michelle Tessier
Department of Linguistics, University of Alberta, Edmonton, AB. T6G 2R3.

1. Background

This is to request your consent for your child……. to be in a research study. Your child is being asked to be in this study because s/he is a Yorùbá person and s/he meets the basic requirement of the study. The results of this study will be used as part of my PhD thesis. The findings of this research will not be used for any commercial purpose.

2. Purpose

The purpose of the research is to test how children as well as adults are able to notice and identify the three basic Yorùbá tones namely the high, mid and low tones. I will be comparing responses from English speaking adults and children with responses from Yorùbá speaking adults and children. The research will give us some insight about the nature and the extent to which these groups of children can still identify these tones. I will use the information collected for my thesis as part of the PhD programme requirements.

3. Study Procedures

The research will begin with a questionnaire to record some personal and language backgrounds. This will be followed by conducting the three different types of experiments which are specially designed for this study. In the first experiment, s/he will listen to two non-Yorùbá words being said with Yorùbá vowels and tones. Then s/he will tell me whether the tones of the two words are similar or different. During the second stage of the experiment, s/he will listen to other non-Yorùbá words, paying special attention to the last syllable of the word. Then s/he will identify the tone of the last syllable as either high, mid or low tone. In the third task s/he will listen to real Yorùbá words. First s/he will see the word on the screen, and then hear the word. Then s/he will see two pictures on the screen. One of these two pictures matches the word that s/he hears. S/he

is to choose the one which matches the word s/he hears. Finally, s/he will listen to some words and then imitate them as best as s/he can. This will be recorded using the recording equipment I will bring from the University of Alberta phonology laboratory. The whole meeting will last about 30-40 minutes.

4. Venue of the study: All the experiments for the research will take place in your house or at another safe place.

5. Benefits

There is no personal benefit for him/her for being in this study. However, there is a reasonable benefit to the Yorùbá community. We hope that the information we get from this study will help us better understand the extent to which Yorùbá children living outside Nigeria could still know the three basic Yorùbá tones, even though they may not be speaking the language regularly. There is no cost for him/her for being in the research except for the time s/he will give for taking part. There is no compensation for taking part either.

6. Risks

I will conduct the meeting in your house or at another safe place and as such, there is minimal risk involved. If I learn of anything during the research that may affect his/her willingness to continue being in the study, I will inform you right away.

7. Voluntary Participation

You are under no obligation to allow your child to participate in this study. Even once you give permission, your child has the right to provide his/her own assent to take part, or to not take part. During the course of the meeting s/he is not obliged to answer any specific question s/he does not wish to answer, or s/he is not comfortable with. I am also assuring you that even if you agree for him/her to be in the study you and/or your child can change your mind and withdraw at any time. If you withdraw your child, you have the right to withdraw his/her data also from the study. If not, I will continue to use the data I have collected from him/her along with other data until the end of the study or at the expiration period when the data will be destroyed.

8. Confidentiality and Anonymity

As I have said above, the present research is part of my PhD programme and s/he will not be personally identified. Once collected, your child's name will be removed (anonymized) and his/her data put together with others for analysis and reporting, and will not identify your child in any way. The information I will collect from him/her will be kept safe and confidential. Only my supervisor and I will have access to the information. All the information about the study will be saved in a file on a password-protected computer with the file encrypted. We may present the results of the study at academic venues at some point. If it is presented or published in academic venues, your child will not be personally identified in any way**.** The study data including any information about him/her will be securely stored for five years after the study is over, at which time it will be destroyed. Anonymized data for this study will be retained for research in the future. I will make arrangement for you in case you want to have access to the report of the research findings.

If you have any further questions regarding this study, please do not hesitate to contact me or my supervisor on these numbers:

Me	Supervisor
Saliu Shittu	Dr. Anne-Michelle Tessier

The plan for this study has been reviewed for its adherence to ethical guidelines by the Research Ethics Board at the University of Alberta. If you have concerns about this study, you may contact the Research Ethics Office at (780) 492-2615. This office has no direct involvement with this project.

CONSENT

Title of Study: A study of the effects of language attrition on the perception of Yorùbá tones

Principal Investigator(s): Saliu A. Shittu Study Coordinator: Dr. Anne-Michelle Tessier

	Yes	No
Do you understand that your child has been asked to be in a research study?	----	----
Have you read and received a copy of the attached Information Sheet?	----	----
Do you understand the benefits and risks involved in your child taking part in this study?	----	----
Have you had an opportunity to ask questions and discuss this study?	----	----
Do you understand that as a parent, you are free to withdraw your child from the study at any time, without having to give a reason?	----	----
Has the issue of confidentiality been explained to you?	----	----
Do you understand who will have access to your child's records?	----	----
Do you understand that your child has the opportunity to provide individual assent, and that if s/he chooses not to participate, your child's wishes will be respected?	----	----

Who explained this study to you? ----------

I ----------provide parental consent for my child ---------- to take part in this study.

Signature of parent ---------- Parent's name---------- Date:----------

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee---------- Date ----------

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

Appendix A6.

REMO Approval

https://remo.ualberta.ca/REMO/Doc/0/2DV7QNDK2UJK53S7KRR76D735B/fromString.html

Notification of Approval

Date:	June 11, 2013
Study ID:	Pro00026974
Principal Investigator:	Saliu Shittu
Study Supervisor:	Anne-Michelle Tessier
Study Title:	A study of the effects of language attrition on the perception of Yorùbá tones

Approval Expiry Date:	June 10, 2014	
Approved Consent Form:	Approved Document	Approval Date
	Informed consent_parents	11/06/2013
	Informed Consent Adults	11/06/2013

Thank you for submitting the above study to the Research Ethics Board 1. Your application has been reviewed and approved on behalf of the committee.

A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Research Ethics Board does not encompass authorization to access the staff, students, facilities or resources of local institutions for the purposes of the research.

Sincerely,

Dr. William Dunn

Chair, Research Ethics Board 1

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix A7

Notification of Approval - Amendment

https://remo.ualberta.ca/REMO/Doc/0/KQ67BAPI8S84D12C492NTKUV50/fromString.html

Date: September 21, 2015

Amendment ID: Pro00026974_AME4

Principal Investigator: Saliu Shittu

Study ID: Pro00026974

Study Title: A study of the effects of language attrition on the perception of Yorùbá tones

Supervisor: Anne-Michelle Tessier

Approved Consent Form:

Approval Date	Approved Document
9/21/2015	Info_Consent form_parent_NGR.docx
9/21/2015	Info_Consent form_adult_NGR.docx
9/21/2015	Info_Consent form_parent_CAN.docx
6/11/2013	Informed Consent Adults
6/11/2013	Informed consent_parents
9/21/2015	Info_Consent form_adult_CAN.docx

Approval Expiry Date: Friday, May 13, 2016

Thank you for submitting an amendment request to the Research Ethics Board 1. This amendment to add information about new population groups to be added to the study has been reviewed and approved on behalf of the committee. The following have been approved:

Initial contact_PARENT_CAN dated August 6, 2015

Initial contact_PARENT_NGR dated August 6, 2015

Initial contact_ADLT_CAN dated August 6, 2015

Initial contact_ADLT_NGR dated August 6, 2015

Info_Consent form_parent_CAN dated September 2, 2015

Info_Consent form_parent_NGR dated September 2, 2015

Info_Consent form_adult_CAN dated September 2, 2015

Info_Consent form_adult_NGR dated September 2, 2015

Assent form_older_CAN dated September 2, 2015

Assent form_older_NGR dated September 2, 2015

Assent form_young_NGR_ENGLISH dated September 2, 2015

Assent form_NGR_YORÙBÁ dated September 2, 2015

Assent form_young_CAN dated September 2, 2015

Sincerely,

Anne Malena, PhD

Chair, Research Ethics Board 1

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix B: Chapter Two supplementary information

Appendix B1. *The practice stimuli for the AX Discrimination Experiment: (*Y*=Yes;* N*=No)*

Monosyllabic			**Disyllabic**		
Stimuli pairs		Response	Stimuli pairs		Response
gà	lè	Y	bōlò	rẹ̄dà	Y
gī	tẹ̄	Y	bèlọ̀	fìlọ̀	Y
gbó	dű	Y	kī̃rī	díwọ̀	N
lè	lō	N	pẹ̀tā	ọ̄mì	N
gbō	mì	N			
tḯ	jū	N			

Appendix B2. *Coefficients of the glmer model of the effect of Age, Language Context, Tone and Syllable count on tone discrimination. Random effects: participant and item. N=6672. Col.1= Fixed effects; Col. 2= Coefficient Estimates; Col.3= Standard Error; Col.4= Z value; Col. 5=Pr(>|z|) value***.**

Ref: YorNgr					
	Estimate	Std. Error	Z value	Pr(>|z|)	Sign code
Intercept	1.7698	0.2276	7.775	7.55e-15	***
Sametone	-0.4790	0.1533	-3.125	0.00178	**
Monosyllabic	0.2414	0.1515	1.593	0.11106	
Young	-0.7680	0.1835	-4.185	2.85e-05	***
YorCan	-0.7190	0.2220	-3.239	0.00120	**
EngCan	-0.6281	0.2191	-2.866	0.00415	**
Ref: YorCan					
Intercept	1.05084	0.23092	4.551	5.35e-06	***
YorNgr	0.71893	0.22196	3.239	0.00120	**
EngCan	0.09081	0.21899	0.415	0.67837	
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.'** 0.1 ' ' 1					

Appendix C: Chapter Three supplementary information

Appendix C1. *The practice stimuli for the Identification 1 Experiment.*

Sound	āpá	ìlūkó	jábò	péfọ̄lẹ̀	gbòjẹ̄	mímúdẹ̄
Last Tone	H	H	L	L	M	M

Appendix C2. *Coefficients of the glmer model of the effects of Age, Language Context, Tone, and Syllable count on tone identification. Random effects: participant and item. N=6672.*

Ref: YorNgr; HighTone

	Estimate	Std. Error	Z value	Pr(>|z|)	Sign code
Intercept	1.19890	0.24506	4.892	9.97e-07	***
LowTone	0.8229	0.16517	0.498	0.618362	
MidTone	-0.20429	0.16513	-1.237	0.216025	
Trisyllabic	-0.12782	0.13486	-0.948	0.343226	
Young	-0.89133	0.19334	-4.610	4.02e-06	***
YorCan	-0.88895	0.23633	-3.761	0.000169	***
EngCan	-0.69677	0.23084	-3.018	0.002541	**

Ref: YorCan; HighTone

Intercept	0.30994	0.24757	1.252	0.210595	
YorNgr	0.88897	0.23635	3.761	0.000169	***
EngCan	0.19223	0.23142	0.831	0.406175	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Appendix D: Chapter Four supplementary information

Appendix D1. *Stimuli for the Identification 2 practice session.*

stimulus	múdí	àgá	fálè	wērù	kárē	gbòjẹ̄
Last tone	H	H	L	L	M	M

Appendix E: Chapter Five supplementary information

Appendix E1. *Samples from the Lexical Experiment set up with pictures:*

Monosyllabic Verbs (n=12)				*Disyllabic Nouns (n=10)*			
sound	*pic1*	*pic2*	*?*	*sound*	*pic1*	*pic2*	*?*
[bó] H	 *peel*	 *cover* (L)	1	[ōdó] MH	 *River* (ML)	 *mortar*	2
[rō] M	 *think* (L)	 *plough*	2	[àk͡pá] LH	 *scar*	 *arm* (MH)	1

Appendix E2. *The practice session for the Lexical experiment.*

Picture1		Picture2		Sound1	Response
bẹ́	'jump'	bẹ̀	'beg'	bẹ̀	2
ọ̄kọ̀	'car'	ọ̄kọ́	'hoe'	ọ̄kọ́	2
ọ̄kọ̀	'car'	ọ̀kọ̀	'spear'	ọ̄kọ̀	1
bẹ́	'jump'	bẹ̀	'beg'	bẹ́	1

Appendix E3. *Coefficients of the glmer model of the effect of Age, Language Context, and Syllable count on Lexical recognition accuracy. Random effects: participant and item. N=1892.*

Ref: YorNgr

	Estimate	Std. Error	Z value	Pr(>|z|)	Sign code
Intercept	5.5192	0.5179	10.658	< 2e-16	***
Monosyllabic	-1.1795	0.4627	-2.549	0.0108	*
Young	-2.2238	0.3335	-6.668	2.60e-11	***
YorCan	-2.1025	0.3391	-6.199	5.67e-10	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1