The role of acoustic-phonetic bias in consonant manner perception

Youngja Nam

School of Communication Sciences and Disorders

(Major: Communication Sciences & Disorders)

McGill University, Montreal

December 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

This dissertation is dedicated to my respected and beloved parents.

ACKNOWLEDGMENTS

Finally, I am at the end of my long intellectual journey at McGill thanks to my supervisor, Dr. Linda Polka. I have to say that I am deeply indebted to her for academic supervision and also for her patience with someone who had to start from scratch. She helped me widen my vision (although I am still wearing glasses?) and she helped me grow academically and personally (although I feel like I've gotten shorter?). During my discussions with her, I am quite often amazed at her endless and fiery passion for infant speech perception. Yes, she is literally a speech scientist. If I hadn't been so lucky to have been allowed to board her ship six years ago, I could hardly imagine what I would be now. I have also benefited from her keen insight. She has inspired me in many ways and I owe her a huge debt of gratitude because she made my journey at McGill treasurable and memorable.

I would also like to give my special thanks to Dr. Baum and Dr. Rvachew who reviewed my thesis draft a few weeks ago. I do sincerely appreciate having an opportunity to have comments from you. I also appreciate how lucky I am to have had Dr. Baum on my proposal committee and Dr. Rvachew on my comprehensive examination committee.

Many more people deserve special appreciation from me for their encouragement and emotional support that made my journey more enjoyable. I am indebted to Janet and her boyfriend, Kent, who helped me record my speech stimuli. We had several memorable times together at Second Cup! I am also truly and deeply indebted to Marianne for her help in recruiting French speakers. Raphaëlle, you also deserve my appreciation definitely for your warmth, kindness, encouragement and your great sense of humor.

I also want to express my sincere appreciation to Jeongmin who is in Korea. She recruited twenty Korean volunteers and tested them for my thesis. She is the one whom I can always rely on for any kind of help, at any time, and wherever I am. My appreciation also goes to Yooyoung for her emotional support during my last two years of graduate school. One more person whom I have to give my special thanks to is Director Lee. You are always my mentor in life. I hope that you will enjoy your retirement with your family.

Finally, my parents, who deserve my appreciation from the bottom of my heart; without their unconditional love and care, I couldn't have made it today. Father, I don't know how to express how much I appreciate your strength as you survived several surgeries while I have been away in Montreal. I cannot overestimate how lucky I am to have you two as my parents. Both of you have led me to get to the finish line. I am also indebted to my sisters and brothers, who have always encouraged me with their great senses of humor.

ABSTRACT

The Natural Referent Vowel (NRV) framework developed by Polka and Bohn (2011) proposes the existence of an acoustic-phonetic bias whereby high-salience vowels are preferred over low-salience vowels, resulting in perceptual asymmetries. This dissertation explores the applicability of the NRV framework to consonant manners by conducting two adult and two infant experiments. The first adult study examines the acoustics of stops and fricatives selected for this dissertation and explores English, French and Korean listeners' mapping of English and Persian stops and fricatives to their respective native phonemes. Results of acoustics show that rise time serves as a reliable stop-fricative distinction cue, with stops having significantly shorter (and less variable) rise times than fricatives. Results of identification show that for non-native contrasts, in general, stops are categorized but fricatives are uncategorized and that fricatives are frequently assimilated as stops but stops are rarely assimilated as fricatives. The second study examines whether English, French and Korean listeners exhibit perceptual asymmetries as a function of language experience. Results from a category-based AX same-different task show that for non-native contrasts, in general, the fricative-to-stop manner shift is easier to discriminate than the stop-to-fricative manner shift and the stop-stop same pairs are easier to recognize than the fricative-fricative same pairs; with only one exception. Parallel asymmetries were found for some native contrasts. The first infant experiment demonstrated a directional asymmetry in English- and French-learning 5-6 month old infants' discrimination of the phonemic /bas/-/vas/ contrast; infants detected the /vas/ to /bas/ change but not the reverse in a habituation test paradigm. The second infant experiment revealed a listening preference favoring /bas/ over /vas/ in infants at the same age.

The overall patterns of adult cross-language and infant phonemic perception show that stops with abrupt onsets are more perceptually salient than fricatives with gradient onsets. This supports a view in which acoustics affects perception, such that an acoustic-phonetic high salience bias is present prior to onset of babbling or word recognition. This bias interacts with language experience; accordingly this bias is preserved for most non-native contrasts but is absent or weaker for native contrasts. These findings show that stops can serve as referent consonants in stop-fricative perception, conceptually dovetailing with the NRV framework. The findings are discussed in light of other views concerning directional asymmetries to argue that an NRP (Natural Referent Phone) framework - which extends the NRV framework to include others phones (not just vowels) - will prove to be the best way to account for these findings.

RÉSUMÉ

Le cadre Natural Referent Vowel (NRV) développé par Polka et Bohn (2011) propose l'existence d'un biais acoustique-phonétique dans lequel les voyelles de haute saillance sont préférées aux voyelles de faible saillance, donnant en conséquence des asymétries perceptives. Cette thèse explore l'applicabilité du cadre NRV à des manières de consonnes en effectuant deux expériences adultes et deux expériences chez les nourrissons. La première étude avec les adultes teste l'acoustique des arrêts et fricatives sélectionnés pour cette thèse et explore chez les auditeurs anglais, français et coréens le mappage d'arrêts et fricatives persans et anglais à leurs phonèmes natals respectifs. Les résultats de l'acoustique montrent que le temps de montée sert de repère fiable de distinction entre arrêt et fricative, avec les arrêts ayant un temps de montée nettement plus court (et moins variable) que les fricatives. Les résultats d'identification montrent que pour les contrastes non-natals, en général, les arrêts sont catégorisés alors que les fricatives sont non catégorisées, ainsi que les fricatives sont fréquemment assimilées comme arrêts, mais les arrêts sont rarement assimilés comme fricatives. La deuxième étude examine si les auditeurs anglais, français et coréens présentent des asymétries perceptives en fonction de l'expérience linguistique. Les résultats d'une tâche en fonction de catégorie AX même-différent montrent que pour les contrastes non-natals, en général, le changement de manière fricative-à-arrêt est plus facile à discriminer que le changement de manière arrêt-à-fricative et les pairs même arrêt-arrêt sont plus faciles à reconnaître que les paires même fricative-fricative; à une exception près. Des asymétries parallèles ont été trouvés pour des contrastes natals. La première expérience chez les nourrissons a démontré une asymétrie directionnelle dans la discrimination de la contraste phonétique /bas/-/vas/ chez les nourrissons âgés de 5 à 6 mois qui sont en

train d’apprendre l’anglais ou le français; les nourrissons ont détecté le changement /vas/ à /bas/ mais pas l'inverse dans un paradigme de test d'habituation. La deuxième expérience chez les nourrissons a révélé une préférence d'écoute favorisant /bas/ plus que /vas/ chez les nourrissons au même âge.

Les tendances générales de perception multilingue chez les adultes et phonémique chez les nourrissons montrent que les arrêts débutant brusquement sont perçus plus saillants que les fricatives avec des débuts graduels. Cela appuie le point de vue que l’acoustique affecte la perception, de telle sorte qu'un biais à haute saillance acoustique-phonétique est présent avant le commencement du babillage ou de la reconnaissance des mots. Ce biais interagit avec l'expérience de la langue; en conséquence, ce biais est conservé pour la plupart des contrastes non-natals mais qui est absent ou faible pour les contrastes natals. Ces résultats montrent que les arrêts peuvent servir comme consonnes référents dans la perception des arrêts et fricatives, harmonisant conceptuellement avec le cadre de NRV. Ces résultats sont discutés à la lumière d'autres points de vue concernant les asymétries directionnelles pour argumenter qu'un cadre NRP (Natural Referent Phone) - qui s’étend sur le cadre de NRV pour inclure les autres sons (pas seulement les voyelles) - prouvera être la meilleure façon d’expliquer ces résultats.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii
ABSTRACT v
RÉSUMÉ vii
TABLE OF CONTENTS ix
LIST OF TABLES xiii
LIST OF FIGURES xv

CHAPTER 1. Thesis overview 1

CHAPTER 2. The role of acoustic-phonetic bias in consonant manner perception 4

2.1. Introduction 4
2.2. Acoustic-perceptual salience affects phonetic perception in infants and adults 6
2.3. Asymmetries in vowel perception 8
2.3.1. The natural referent vowel framework 8
2.3.2. Asymmetries in infancy 10
2.3.3. Infants prefer natural referent vowels 12
2.3.4. Asymmetries interact with experience in adults 13
2.4. Asymmetries in consonant manner perception 14
2.4.1. Asymmetries in stop-fricative perception in adults 14
2.4.1.1. The featurally underspecified lexicon (FUL) model 17
2.4.2. Asymmetries in stop-fricative perception in toddlers 18
2.4.2.1. The lexical representation hypothesis 20
2.5. Other models related to directional asymmetries 22
2.6. The privileged status of stops over fricatives 24
2.6.1. The acoustic-phonetic bias hypothesis: An extension of NRV to consonants... 26
2.7. Research aims 27
2.7.1. Stop-fricative perception in adults 28
2.7.2. Stop-fricative perception in infants 28

2.8. Outline of chapters......29

CHAPTER 3. Adult cross-language identification of stop-fricative contrasts 32

3.1. Introduction......32

3.2. The stop and fricative inventories: English, French and Korean 34

3.3. Acoustics of English and Persian stops and fricatives in initial position 35

3.3.1. Speakers 35

3.3.2. Stimulus materials......35

3.3.3. Recording procedure 37

3.3.4. Stimuli selection and verification procedures 37

3.3.5. Stimulus measurements and acoustics 39

3.4. Experiment 1: The identification experiment 42

3.4.1. Methods 42

3.4.1.1. Participants 42

3.4.1.2. Stimulus materials......42

3.4.1.3. Procedure......43

3.5. Results 43

3.6. Discussion 50

CHAPTER 4. Adult cross-language perception of stop-fricative contrasts..53

4.1. Introduction 53

4.2. Experiment 2: The discrimination task......55

4.2.1. Methods 56

4.2.1.1. Participants 56

4.2.1.2. The AX trials......57

4.2.1.3. Set up and Procedure 59

4.3. Statistical results......60

4.3.1. Global performance across all contrasts 60

4.3.1.1. All AX pairs......61

4.3.1.2. Different pairs......62

4.3.1.3. Same pairs 64

4.3.2. Performance on each contrast 66

4.3.2.1. /tas/-/sas/ 67
4.3.2.2. /kas/-/xas/ 69
4.3.2.3. /gas/-/ɣas/ 71
4.3.2.4. /pas/-/fas/ 72
4.3.2.5. /bas/-/vas/ 74
4.3.2.6. /das/-/zas/ 76
4.3.2.7. /tas/-/θas/ 77
4.3.2.8. /das/-/ðas/ 79
4.4. Summary tables 81
4.5. Summary and discussion 83
4.5.1. Phonemic status effect 83
4.5.2. Salience differences play a role in non-phonemic perception 84
4.5.3. Korean listeners' perception of /bas/-/vas/ and /gas/-/ɣas/ 87
4.5.4. Salience differences also play a role in phonemic perception 88
4.6. Beyond the FUL model 89
4.7. Beyond the perceptual assimilations and the developmental period 89
4.8. Conclusion 93

CHAPTER 5. A stop preference in infant stop-fricative perception 95

5.1. Introduction 95
5.2. Experiment 3: Exploring a directional asymmetry in infants 98
5.2.1. Methods 98
5.2.1.1. Participants 99
5.2.1.2. Stimulus materials 99
5.2.1.3. Set up 101
5.2.1.4. Procedure 102
5.2.2. Results 103
5.2.2.1. Task engagement 104
5.2.2.2. Habituation trials 104
5.2.2.3. Test trials 105
5.2.3. Discussion 106
5.3. Experiment 4: Exploring a stop bias in infants 108

5.3.1. Methods 109
5.3.1.1. Participants 109
5.3.1.2. Stimuli 110
5.3.1.3. Procedure 110
5.3.2. Results 111
5.3.2.1. Task engagement 111
5.3.2.2. Test trials 112
5.3.3. Discussion 114
5.4. Lexical specificity? 115
5.5. General discussion 117

CHAPTER 6. Summary of findings and general discussion 119
6.1. Summary of findings 119
6.2. NRV is extended to consonant manners 121
6.3. Directions for future research 124

REFERENCES 126

LIST OF FIGURES

Figure 4-1. Average proportion correct across all the eight stop-fricative contrasts in English, French and Korean listeners: overall performance (A) and performance according to phonemic status (B) (error bars: ±1 SE). P>NP indicates significantly better performance on the phonemic contrasts compared to the non-phonemic contrasts ($p < .05$); 6/2, 4/4 and 1/7 indicate the number of phonemic/non-phonemic contrasts, respectively, for each language group.........61
Figure 4-2. Average proportion correct on different pairs according to phonemic status for English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates that listeners perceived the FS pairs significantly better than the SF pairs ($p < .05$); numbers in the parentheses indicate number of phonemic and non-phonemic contrasts for each language group...63
Figure 4-3. Average proportion correct on same pairs according to phonemic status for English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs ($p < .05$); numbers in the parentheses indicate number of phonemic and non-phonemic contrasts for each language group...65
Figure 4-4. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /tas/-/sas/ contrast in English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs ($p < .05$)..68
Figure 4-5. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /kas/-/xas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates that listeners perceived the FS pairs significantly better than the SF pairs; SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs ($p < .05$)......................................70
Figure 4-6. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /gas/-/ɣas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$)..71
Figure 4-7. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /pas/-/fas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on

the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$)....73

Figure 4-8. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /bas/-/vas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; FF>SS* indicates significantly better performance on the FF pairs than the SS pairs ($p < .05$)....75

Figure 4-9. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /das/-/zas/ contrast in English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$)....76

Figure 4-10. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /tas/-/θas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$)....78

Figure 4-11. Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /das/-/ðas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$)....80

Figure 5-1. Mean listening time as a function of the habituating consonant in each habituation group during the habituation phase (error bars: ±1 SE)....105

Figure 5-2. Mean listening time to old and new trials for the phonemic /bas/-/vas/ contrast in each habituation group during the test phase (error bars: ±1 SE). Arrow indicates the change from old trials to new trials. Asterisk denotes a significant difference ($p < .05$)....106

Figure 5-3. Listening time on the posttest trial according to listening condition (error bars: ±1 SE). Asterisk denotes a significant difference ($p < .05$)....111

Figure 5-4. Listening patterns for each syllable type across listening conditions during the 12 trials (left) and during the 10 trials (right) (error bars: ± 1 SE). Asterisk denotes a significant difference ($p < .05$)....112

LIST OF TABLES

Table 2-1. Deviant detection in consonant manner perception. Adapted from Cornell *et al.* (2013) 18
Table 2-2. Detection of stimulus change in early lexical representations. Adapted from Altvater-MacKensen and Fikkert (2010)..................... 21
Table 3-1. Stop-fricative contrast corpus 36
Table 3-2. Stimulus acoustics. Shaded rows indicate significant distinctions in manner..................... 39
Table 3-3. Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as English consonants by English listeners (n=16) 44
Table 3-4. Assimilation types for each contrast by English listeners..................... 45
Table 3-5. Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as French consonants by French listeners (n=15)..................... 46
Table 3-6. Assimilation types for each contrast by French listeners 47
Table 3-7. Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as Korean consonants by Korean listener (n=20)..................... 48
Table 3-8. Assimilation types for each contrast by Korean listeners..................... 49
Table 4-1. English language background profiles of the French and Korean speakers ... 56
Table 4-2. A set of AX pair matrices for a stop-fricative contrast for a first sub-block ... 58
Table 4-3. A set of AX pair matrices for a stop-fricative contrast for a second sub-block... 58
Table 4-4. English listeners' perception of English and Persian stop-fricative contrasts... 81
Table 4-5. French listeners' perception of English and Persian stop-fricative contrasts... 82
Table 4-6. Korean listeners' perception of English and Persian stop-fricative contrasts..... 82
Table 4-7. Evaluation of the results of listeners' non-phonemic perception in light of PAM..................... 91
Table 5-1. Summary of stimulus acoustics..................... 100
Table 5-2. Syllable corpus employed for habituation and test trials..................... 103

Table 5-3. Syllable tokens used during the test trials according to listening condition........ 110

Chapter 1: Thesis overview

In infant vowel perception, there is robust evidence that the central-to-peripheral vowel shift is easier to discriminate than the other way around and a peripheral vowel is preferred over a central vowel in language-independent manner (see for a review Polka & Bohn, 2011). According to the natural referent vowel (NRV) framework developed by Polka and Bohn, such directional asymmetries are driven by an acoustic-phonetic high salience bias related to acoustic properties; peripheral vowels with formant convergence are acoustically and perceptually more salient than central vowels and thus serve as referents in vowel perception. The NRV framework suggests that the initial acoustic-phonetic bias interacts with language experience.

Perceptual asymmetries have received much less attention in research on consonant perception. Directional asymmetries were reported in toddler phonemic stop-fricative perception (Altvater-MacKensen *et al*., 2010, 2014). In adult consonant manner perception, the English /ba/-/va/ contrast has only been studied (e.g., Tsushima *et al*., 2003, 2005). Infant perception research has not investigated asymmetries in stop-fricative perception. However, it has been suggested that abrupt amplitude change contributes to distinctive auditory enhancement (e.g., Delgutte & Kiang 1984; Stevens and Keyser, 1989) which is relevant given that stops are characterized by abrupt onset acoustics but fricatives are characterized by gradient onset acoustics (e.g., Lisker & Abramson, 1964). The issue motivating this thesis is whether this asymmetry in onset acoustics has perceptual consequences; is there a bias favoring perception of stops over fricatives and if so, do stops serve as natural referents sounds with respect to consonant manner?.

As an attempt to extend the NRV framework to consonant manners, this dissertation explores whether, as with vowels, parallel asymmetries are found in infant phonemic stop-fricative perception and in adult cross-language stop-fricative perception. Toward this end, two adult and two infant perceptual experiments were conducted.

This dissertation is organized as follows. Chapter 2 provides a comprehensive review of the background literature and a detailed description of research objectives and questions. In Chapter 3, the acoustic analyses of test stimuli are presented which confirm that amplitude rise time is a highly reliable cue differentiating stop-fricative consonant manner in syllable initial position. The test stimuli consist of 8 stop-fricative contrasts each produced in a /Cas/ syllable context. Six of contrasts are from English and two are from Persian; multiple instances of each syllable were produced by one male and one female talker (native speaker). Also reported in Chapter 3 are assimilation patterns for the initial consonant of each syllable from all 8 contrasts was assessed in English, French and Korean adults. These data show how the groups differ in their perceptual mapping of each contrast to their native phoneme categories. In Chapter 4, findings from adult discrimination task are presented. Adult native speakers of Korean, English, and French were tested on all 8 stop-fricative contrast in a category-based same-different AX discrimination task. The data were analyzed to determine whether performance is asymmetrical, testing the prediction that on different pairs discrimination of a fricative-to-stop change will be better than stop-to-fricative change, and on same pairs, discrimination of stop-stop pairs will be better than fricative-fricative pairs. In Chapter 5, findings from two infant experiments are presented. In the first experiment English- and French-learning 5-6 month olds were tested in a visual habituation task to examine the presence of a directional asymmetry for one phonemic contrast, /bas-vas/. In the second experiment, English- and French-learning infants at the same ages were presented the same contrast in a listening preference task. Chapter 6 provides a summary of the findings and a discussion of how these data challenge or support different conceptual views of speech perception and phonological development. Limitations of the current research and important directions for future research are also highlighted.

Chapter 2: The role of acoustic-phonetic bias in consonant manner perception

2.1 Introduction

For several decades, research on speech contrast perception in adults and infants has overwhelmingly focused on the listener's ability to discriminate between/among different contrasts. Relatively recently, a new line of research has put the spotlight on perceptual asymmetries in which discrimination of a contrast is depressed when one member of the contrast is presented before the other relative to the reverse order of presentation.[1] There is a growing body of evidence that discrimination asymmetries as a function of stimulus presentation direction are robust and predictable in vowel perception in infants and adults (e.g., Polka & Bohn, 1996, 2011; Pons *et al*., 2012; see also Cowan & Morse, 1986; Repp & Crowder, 1990). The natural referent vowel (NRV) framework put forward by Polka and Bohn (2011) proposes that such directional asymmetries reflect perceptual biases grounded in acoustic property salience. The NRV framework hypothesizes that a shift from a less salient vowel to a more salient vowel is easier to detect compared to the reverse direction.

It has been shown that stimulus direction also affects listeners' performance in consonant perception. Directional asymmetries in consonant perception have been mainly concerned with place contrasts particularly in the neurophysiological literature (e.g., Friedrich *et al*., 2006, 2008). To date, only a few studies have addressed asymmetries in consonant manner perception. All but one study examined asymmetries in the perception of stop-fricative(s) contrasts in adults.

[1] Perceptual asymmetries according to stimulus presentation direction have also been reported in visual perception (Beck, 1974; Cusack & Carlyon, 2003; Treisman & Gormican, 1988; see also Mounts & Tomaselli, 2005).

Most of the findings were basically interpreted as suggesting that not all consonant manners of articulation are equally specified in terms of phonological representations in the mental lexicon (Altvater-MacKensen *et al*., 2010, 2014; Cornell *et al*., 2013). Specifically, stop consonants are assumed to be the default and underspecified manner whereas fricatives are specified as [continuant] in the underlying representation. From these phonological points of view, phonological specifications determine the emergence of perceptual asymmetries for consonant manner contrasts. In this regard, the existing views differ from the NRV view in that the latter is built on the assumption that perceptual asymmetries are driven by acoustic-phonetic biases.

This dissertation is an attempt to explore the applicability of the NRV framework to consonant manner perception. This is approached by examining how infants and adults perceive stop-fricative contrasts. Motivated by the NRV proposition that in general, an acoustic-perceptual bias affects perception in young infancy and comes to interact with language experience in adulthood, this dissertation specifically investigates whether directional asymmetry is evident in phonemic stop-fricative perception in young infants but is revealed as a function of language-specific experience in adults. These investigations have broad theoretical implications for (1) how infants map out their perceptual space as they acquire the phonetic segments of their target language(s), (2) how adult listeners have organized their phonetic perceptual space, and (3) how that space may be re-organized in the context of acquiring a second language. The stop-fricative contrasts are ideal for assessing the potential influence of acoustic properties in consonant manner perception because stops show abrupt onset acoustics and fricatives display gradual onset acoustics. Given the appealing asymmetric nature of onset acoustics between stops and fricatives, this dissertation is particularly concerned with whether more abruptness in onset acoustics enhances perceptual salience or distinctiveness and as shown with vowels, listeners find it easier to discriminate the low-salience fricative to high-salience stop change than the reverse shift.

In the remainder of this chapter, the literature relevant to this dissertation will be reviewed to provide a background overview of the NRV framework and eventually to suggest that as shown in vowel perception, perceptual biases tied to acoustic properties are revealed as perceptual asymmetries in stop-fricative perception. This introductory chapter is concluded by presenting an outline of each of the subsequent chapters.

2.2 Acoustic-perceptual salience affects phonetic perception in infants and adults

Most of the studies on the perception of phonetic contrasts in infancy have demonstrated that the young infant is capable of perceiving many subtle acoustic differences among individual native and non-native phonetic segments (e.g., Aslin *et al*., 1981; Bosch & Sebastián-Gallés, 2003; Polka & Werker, 1994; Werker & Tees, 1984). This broad perceptual sensitivity does not endure into the second half of the first year of life and instead infant perception begins to shift from a language-universal to a language-specific state (e.g., Best & McRoberts, 2003; Cheour *et al*., 1998; Kuhl, 1992, 1994; Kuhl *et al*., 1992; Kuhl *et al*., 2006; Kuhl & Rivera-Gaxiola, 2008; Polka & Werker, 1994; Pons *et al*., 2009; Werker & Lalonde, 1988; Werker & Tees, 1984). Whereas the young infant shows perceptual sensitivity to almost all linguistically-relevant phonetic contrasts regardless of phonemic status, the older infant shows a stable or enhanced perceptual sensitivity to native phonetic contrasts but a diminished perceptual response to many non-native contrasts. This shift in the organization of perceptual space allows the infant to optimize access to language-specific contrasts.

However, there has been emerging evidence that not all non-native contrasts are equally well discriminated in young infancy and not all native contrasts are equally well discriminated in later infancy. It has been found that infants exhibit a different developmental profile for native and non-native phonetic perception as a function of acoustic-perceptual salience. Narayan *et al*. (2010), using the visual habituation paradigm, showed that English-learning infants in three age groups, 4-5 months, 6-8 months and 10-12 months, could discriminate the phonemic Filipino /ma/-/na/ contrast, but could not discriminate the non-phonemic Filipino /na/-/ŋa/ contrast at any age. This was unexpected given that prior studies have consistently shown discrimination of non-phonemic contrasts in young infants (6-8 months or younger). The young English infants' failure to discriminate the non-phonemic /na/-/ŋa/ contrast led to the speculation that perhaps even Filipino-learning infants, who experience this contrast in their ambient language, may succeed only after some period of language experience with the linguistically-relevant contrast. Thus, Filipino infants 6-8 and 10-12 months of age were tested on the native /na/-/ŋa/ contrast. Filipino infants were shown to discriminate /na/-/ŋa/ at 10-12 months but not at 6-8 months. Particularly, given the striking overlap between /na/ and /ŋa/ in F2/F3 space, Narayan *et al*. (2010) attributed the inability of the young infants to discriminate the /na/-/ŋa/ contrast to their acoustic similarity (i.e., low salience of the place cues distinguishing these nasal consonants). These findings show that acoustic distinctiveness and its resulting perceptual salience may play a role in phonetic perception in infancy; a less salient native contrast can require language-specific exposure for successful discrimination.

Similarly, differential discrimination of phonemic contrasts according to acoustic-perceptual salience was evidenced in adult nasal perception. In his earlier study, Narayan (2008) examined Filipino adults' perception of the same nasal contrasts. In the AX discrimination procedure, Filipino adults showed the expected native-level performance on both /ma/-/na/ and /na/-/ŋa/ contrasts. However, a discriminability difference existed between contrasts, with discrimination of /ma/-/na/ significantly better than /na/-/ŋa/. The influence of acoustic-perceptual salience in nasal contrast perception was more evident in adverse listening conditions. Filipino adults' performance stayed at near-ceiling levels on the acoustically-perceptually distinctive /ma/-/na/ contrast regardless of listening conditions but their performance fell as the listening conditions became adverse and performed near chance levels on the acoustically-perceptually less salient /na/-/ŋa/ contrast in the noisiest listening condition.

Apart from the discriminability of nasal place of articulation by infants and adults, acoustic representations and their resulting perceptual impact also appear to affect infants and adults' performance in vowel discrimination tasks when presentation direction is varied. This is discussed in detail below.

2.3 Asymmetries in vowel perception

This section first provides an overview of the rationale for the NRV framework developed by Polka & Bohn (2011) to explain the role of perceptual biases grounded in acoustic properties in phonetic development and then reviews the empirical research explicitly designed to address directional asymmetries in vowel perception. This research brought initial focus on the presence of differences in discrimination performance according to direction and thus is the proper point to begin building the conceptual foundation and motivation for this dissertation.

2.3.1 The natural referent vowel framework

Central to the NRV framework is its concept of acoustic-perceptual salience bias. The NRV framework essentially hypothesizes that the more peripheral a vowel is, the more acoustically-perceptually salient and the more perceptually preferred it is. Peripherality is related to positions in the F1/F2 vowel space; vowels have more focal spectral energy when adjacent formants become closer which naturally occurs as you move from the center to the corners of the vowel space. The NRV framework posits that formant convergence increases acoustic prominence and resultantly perceptual prominence (see also Schwartz *et al*., 2005). It is useful to note that the most peripheral vowels /i/, /a/ and /u/ corresponding to the articulatory extremes are typologically most preferred (Maddieson, 1984) and such typologically preferred vowels may be selected by languages for their perception or production properties (Flemming, 1995). The NRV framework points out that acoustic-perceptual salience differences are revealed as directional asymmetries; a shift from a low-salience central vowel to a high-salience peripheral vowel is easier to detect compared to the reverse direction in which salience overriding occurs. According to NRV, acoustically privileged and resultantly perceptually preferred peripheral vowels serve as natural referent vowels in vowel perception. This is a so-called NRV bias. The NRV model suggests that the infants' selective response to natural referent vowels provides default perceptual structure that infants can use to discover additional vowel categories in their native language. As language experience accumulates, the NRV bias is expected to become weaker for native contrast s as the perceiver becomes highly efficient in perceiving the native categories. The NRV bias (evi dent in directional asymmetries) will be retained for non-native contrasts because (given phoneti c universals) the referent vowel will be more similar to a native vowel category compared to the non-referent vowel. Adults may be able to capitalize on acoustic-perceptual salience differences between vowels in the discrimination of a non-native vowel contrast when learning a second language.

The NRV model emerged to explain a wide range of findings from vowel perception studies. A summary of these findings is presented below, organized to highlight key findings of infant vowel discrimination, infant vowel preference and adult vowel discrimination. The infant studies provide insights into whether the NRV bias remains unaffected in infancy or interacts with language-specific experience at some point.

2.3.2 Asymmetries in infancy

The review starts with the study by Polka and Bohn (1996) because strictly speaking, this was the first study that explicitly aimed to explore how vowel peripherality affects vowel perception in relation to direction. Polka and Bohn (1996) tested English- and German-learning infants at two ages, 6-8 and 10-12 months, with the English /ɛ/-/æ/ contrast (non-phonemic in German) and the German /y/-/u/ contrast (non-phonemic in English), using the conditioned headturn procedure. The results showed that infants at both ages noticed the less peripheral to more peripheral vowel shifts (i.e., /ɛ→æ/ and /y→u/) better than the reverse shifts, regardless of phonemic status.

Some years later, Polka and Bohn (2011) observed that the NRV bias may be modulated by language experience in native vowel perception in later infancy. Polka and Bohn (2011) examined Danish-learning infants' perception of non-phonemic Southern British English /ʌ/-/ɒ/ contrast and Danish /e/-/ɛ/ and /e/-/ø/ contrasts using the same procedure. Danish infants at both ages, 6:15-8:21 and 8:25-11:20 months, performed asymmetrically on the non-phonemic /ʌ/-/ɒ/ contrast, with better discrimination of the less peripheral /ʌ/ to more peripheral /ɒ/ change than the /ɒ/ to /ʌ/ change. Danish infants aged 6:10-9:02 months showed the NRV bias for the phonemic /e/-/ɛ/ contrast by showing depressed discrimination of the shift of the more peripheral /e/ to the less peripheral /ɛ/ relative to their discrimination of the reverse shift. However, Danish infants aged 10:14-11:28 months performed symmetrically on /e/-/ɛ/. Interestingly, Danish-learning infants aged 6:0-7:16 months showed a reversed asymmetry for the native /e/-/ø/ contrast; they detected the change from the more peripheral vowel /e/ to the less peripheral vowel /ø/ better than vice versa. Their older counterparts (7:22-11:9) performed symmetrically o n this native contrast. These results suggest that the NRV bias interacts with language experience toward the end of the first year of life. With respect to the non-conforming /e/-/ø/ contrast, it should be noted that the vowel /e/ is lip unrounded and the vowel /ø/ is lip rounded, a contrast that is not differentiated by F1 or F2 alone but can be distinguished by F3 or F4. Thus, higher formants appear to play a role in some asymmetries, perhaps also through formant convergence, in ways that are not yet fully understood.

A similar developmental profile was found by the Pons *et al*. (2012) study in which the respective role of acoustic salience and frequency distribution in phonemic perception was assessed in the visual habituation procedure. Three age groups of Catalan- and Spanish-learning infants (i.e., 4, 6 and 12 months of age) were tested on the /i/-/e/ contrast which is phonemic in both Catalan and Spanish. Both Catalan and Spanish infants aged 4 months showed an asymmetry conforming to the bias predicted by the NRV model. The NRV bias was retained in Catalan 6- and 12-month-olds. Intriguingly, however, with age the bias shift was underway in the Spanish-learning infants; Spanish 6-month-olds did not reliably discriminate /i/-/e/ in either of the stimulus presentation directions. The bias finally shifted in the Spanish-learning infants; Spanish 12-month-olds detected the change from /i/ to /e/ significantly better than vice versa. Given that the frequency of /i/ is relatively higher than the frequency of /e/ in Catalan but the opposite is the case in Spanish, Pons *et al*. (2012) interpreted the reversed asymmetry in the Spanish 12-month-olds as reflecting the role of language-specific input frequency in infant vowel perception. This observation raises the possibility that a language-universal acoustic-perceptual salience bias may be reversed by a language-specific input-frequency-based bias around the end of the first year of life.

These findings suggest that language experience appears to override the default NRV pattern as needed to optimize access to native vowel categories. It is not yet clear how and when this happens and what factors contribute to the development of native language perception. However, despite the differences between the Spanish and Catalan developmental patterns, it is important to note that an initial asymmetry in discrimination of /i/-/e/ (conforming to the NRV model) was found.

2.3.3 Infants prefer natural referent vowels

Polka *et al*. (2005) more directly assessed the existence of the NRV bias by measuring infant preferences for peripheral vowels over central vowels. Polka *et al*. (2005), using the French-specific /u/-/y/ contrast and the English-specific /i/-/I/ contrast, investigated vowel perception biases at three age groups (i.e., ages 3-5, 6-8, and 10-12 months) in three language groups: English-learning, French-learning and bilingual infants. Bilingual infants were classified as those who were acquiring either English and another language or French and another language. Monolingual infants showed the NRV bias for both contrasts by listening longer to the more peripheral vowels /u/ and /i/ than their respective less peripheral counterparts /y/ and /I/ in all three age groups. Within the bilingual groups, the same preference patterns were observed in 6-8- and 10-12-month-old bilinguals but 3-5-month-old bilinguals did not show a listening preference. It is not yet clear why the bilingual pattern is different. This difference between monolingual and bilingual infants shows that the NRV bias may not be a low level auditory bias, but is a phonetic level bias that is shaped by differences in language experience. Younger infants must be tested to determine whether the NRV bias emerges developmentally (and a bit later in bilinguals) or is present at birth but diminishes temporarily in the bilingual infants, as has been shown in previous research with bilingually exposed infants (Bosch & Sebastián-Gallés, 2003, 2005; Sebastián-Gallés & Bosch, 2009). Importantly, this study provided the first direct evidence that infants prefer more peripheral vowels over less peripheral vowels.

2.3.4 Asymmetries interact with experience in adults

Polka *et al*. (2005) investigated the perception of German vowel contrasts /u/-/y/ and /ʊ/-/ʏ/ in German- and English-speaking adults and two groups of Cantonese-speaking adults (phonet ically naïve vs. trained) at the inter-stimulus interval (ISI) of 2000 ms in a go/no go discriminatio n task. German listeners performed symmetrically on these native contrasts. Within the English li stener group, for both the non-phonemic German contrasts, performance was significantly better in the more central to more peripheral vowel direction (i.e., /y→u/ and /ʏ→ʊ/) than vice versa (s ee Polka & Bohn, 2011, for a similar result in adult non-phonemic perception). Within the Canto nese listener group, phonetically naïve listeners showed an asymmetry for the non-phonemic /ʊ/-/ʏ/ contrast but not for the phonemic /u/-/y/ contrast. However, phonetically trained listeners perf ormed symmetrically on both contrasts. Polka *et al*. (2005) further explored the nature of these asymmetries by testing English listeners' perception of the non-phonemic contrasts at an ISI of 500 ms. It has been known that short ISIs facilitate access to acoustic memory which is believed to rapidly decay relative to phonetic memory and reduce demands on phonetic categorization (Cr owder, 1982; Pisoni, 1973; Werker & Logan, 1985). In the reduced ISI conditions, English listeners performed symmetrically on both the German /u/-/y/ and /ʊ/-/ʏ/ contrasts.

These findings laid the groundwork for the NRV assumption that the initial NRV bias becomes weak or absent with language experience. Prompted by the observations of no asymmetries in listeners knowledgeable of phonetic aspects and in listening conditions placing reduced demands on phonetic encoding, NRV suggests that asymmetries more likely reflect a per ceptual bias relevant to phonetic processing rather than general auditory processing.

2.4 Asymmetries in consonant manner perception

As reviewed above, there is a substantial body of research exploring asymmetries in vowel perception. A number of studies also exist regarding asymmetries in consonant place perception but there is very limited research of this type with respect to consonant manner. This section will critically review findings that reveal the potential role of acoustic-perceptual salience in consonant manner perception with the goal of highlighting current gaps and limitations in this line of investigation. The findings are primarily concerned with the perception of consonant contrasts that differ in stop-fricative manner because this contrast type has received the most attention to date. This section also spends some time addressing the hypotheses that have been proposed to account for the observed asymmetries in consonant manner perception.

2.4.1 Asymmetries in stop-fricative perception in adults

The presence of directional asymmetries in the perception of stop-fricative manner contrast was first documented by Tsushima *et al*. (2003). Tsushima *et al*. (2003) investigated the magnitude of directional asymmetry in the discrimination of a non-phonemic English /ba/-/va/ contrast in Japanese listeners in an AX discrimination experiment before and after discrimination training (see also Tsushima, 2010). Subjects were assigned to one of six listening conditions which differed in the number of talkers and ISI: one talker × 3 ISIs (100, 300, 2000 ms) and two talkers × 3 ISIs (100, 300, 2000 ms). During pre-training, Japanese listeners' discrimination of the /va/ to /ba/ change was significantly better than their discrimination of the /ba/ to /va/ change in all but two conditions (1 talker and 2 talkers at 300 ms) in which there was nonetheless a similar trend. Directional asymmetries also tended to be reflected in response times; Japanese listeners were faster to respond in the /va/→/ba/ direction than the reverse direction except under one condition (2 talkers at 300 ms). The discriminability of the /ba/-/va/ contrast in Japanese listeners was improved during six sessions of discrimination training and directional asymmetries faded under all conditions except one (1 talker at 100 ms) at post-test. Correct response rates after training were greater than 80% but below 100% in all but one condition (2 talkers at 100 ms); asymmetries tended to decrease before subjects reached ceiling performance level.

A point to be noted is that at pre-test, directional asymmetries remained generally unaffected by variations in ISIs, inconsistent with Polka *et al*.'s (2005) findings of the null asymmetries in English listeners' non-phonemic vowel perception at the reduced ISI. It is not clear yet why listeners respond differently to consonants and vowels in relation to ISI conditions.

Subsequently, Tsushima *et al*. (2005)[2] reported that Japanese listeners showed a robust direction effect for the /ba/-/va/ contrast in several versions of the AX discrimination task. For example, Japanese listeners performed asymmetrically in a modified AX task (two talkers at 2000 ms) in which an irrelevant schwa /ə/ was inserted between the 1st and the 2nd stimulus. Japanese listeners also performed asymmetrically across the identification tasks (two talkers at 100 ms, 300 ms and 2000 ms) in which listeners were asked to identify the stimulus order of a given contrast as one of the four directions (i.e., /b → v/, /v → b/, /b → b/ or /v → v/); listeners identified the /v → b/ direction more accurately than the /b → v/ direction. Similarly, the direction effect was evident in an AX repetition task (two talkers at 300 ms) in which the first stimulus was repeatedly presented seven or fifteen times and then the second stimulus was

---

[2]Note that this study was reported in a conference presentation only and thus many of the details are not available.

presented in the same way. Again, the results from these several tasks show that directional asymmetry was robust at very short ISIs.

Tsushima *et al*. (2003, 2005) suggested several possible alternative interpretations but appeared to place emphasis on the language-specific perceptual assimilation, given that the English /b/ is perceptually more similar to the Japanese /b/ than the English /v/. Tsushima *et al*. (2003, 2005) suggested that in the /b/→/v/ direction, /v/ is assimilated to /b/ because for Japanese listeners the more native-like /b/ is easily categorized and held in memory in a stable manner but the less native-like /v/ is difficult to perceive; they assumed that asymmetry is a perceptual pattern specific to non-phonemic contrasts.

However, the language-specific perceptual assimilation suggestion is challenged by the findings of the presence of perceptual asymmetry in the perception of the /ba/-/va/ contrast in native listeners. In a cross-linguistic magnetoencephalography (MEG) study, Zhang *et al*. (2006)[3] examined English and Japanese listeners' perception of synthetic /ba/ and /va/ stimuli using a two-block design in which /ba/s were infrequent deviants and /va/s were frequent standards in one block and then the other way around in the other block. The within-block subtractions of mismatch field (MMF) responses (deviant /va/s − standard /ba/s and deviant /ba/s − standard /va/s) indicated that English listeners' MMF patterns were parallel to those of Japanese listeners, with detection of /ba/ deviants amidst /va/ standards significantly better than /va/ deviants against the background of /ba/ standards. This shows the possibility that perceptual asymmetry may be language independent.

Putting aside the exact underpinning mechanism, the cross-linguistic MEG results of the presence of perceptual asymmetry in native listeners indicate that language experience is not the

[3]Note that this study was presented at conference and no further study results are available.

only determinant contributor to perceptual asymmetry and the nature of acoustic properties may also affect perception. Importantly, this asymmetry in native listeners inspires the possibility that the relative perceptual salience differences tied to acoustic features are attributable to the perceptual asymmetry shown by non-native Japanese listeners in the above studies.

Very recently, asymmetry has also been reported to be evident in stop-nasal manner perception. In an electroencephalographic (EEG) study, Cornell *et al*. (2013) explored German listeners' perception of native nonword stop-nasal /edi/-/eni/ and fricative-nasal /ezi/-/eni/ contrasts. German listeners showed asymmetric mismatch negativity (MMN) amplitude patterns for the /edi/-/eni/ contrast (i.e., larger MMN amplitudes for the /eni/ to /edi/ change than the reverse change) but symmetric MMN patterns for the /ezi/-/eni/ contrast regardless of direction. These results were interpreted in light of the featurally underspecified lexicon (FUL) model (e.g., Lahiri & Reetz, 2002, 2010). Until this study, FUL has been applied to explain asymmetries in consonant place perception. This model is described below because it is also relevant to stop-fricative contrasts.

2.4.1.1 The featurally underspecified lexicon (FUL) model

The FUL model is based on the premise that not all consonant manners of articulation are fully specified in the mental lexicon. That is, a stop manner has sparse or no stored phonological information. Table 2-1 illustrates the FUL model's reasoning about perceptual asymmetry based on the study of Cornell *et al*. (2013). According to FUL, it appears that acoustic features can be equally extracted from all consonant manner types, which is fundamentally at odds with the NRV hypothesis. Specifically, in the case of deviant /eni/ amidst standard /edi/, no conflict occurs between the empty phonological representational information of the stop and the extracted acoustic information of the nasal, resulting in difficulty in detecting the presence of the deviant stimulus. In the reverse case (i.e., deviant /edi/ among standard /eni/), a conflict occurs between the phonological feature [nasal] of the standard stimulus and the acoustic feature [stop] of the deviant stimulus. As for the the /ezi/-/eni/ contrast, conflict occurs in both directions because of their specified manner features.

Table 2-1 Deviant detection in consonant manner perception. Adapted from Cornell *et al*. (2013).

	standard	deviant	lexical representational features activated by the standard	surface representational features extracted by the deviant	conflicting	deviant detection
/edi/-/eni/	/d/	/n/	[empty]	[nasal]	non-conflict	no
	/n/	/d/	[nasal]	[stop]	conflict	yes
/ezi/-/eni/	/z/	/n/	[strident]	[nasal]	conflict	yes
	/n/	/z/	[nasal]	[strident]	conflict	yes

2.4.2 Asymmetries in stop-fricative perception in toddlers

A few studies have examined perception of stop-fricative phones in toddlers. It is important to note that this work did not focus specifically on phonetic perception, but rather on early lexical development and how toddlers use phonetic detail in word recognition or word learning tasks. As well, the researchers pursuing this work are interested in the phonological representations that toddlers are developing and applying as they develop a functional lexicon. As a result, the tasks selected and the focus that is favored differs from the research on vowel perception reviewed earlier.

Altvater-MacKensen and Fikkert (2010) examined the perception of a phonemic /paap/-/faap/ contrast in Dutch-learning toddlers aged 14 months in a spoken label-visual object matching paradigm. Half of the infants were habituated to a stop-initial /paap/object pairing and the other half to a fricative-initial /faap/-object pairing. During the test trials, the /faap/-

habituated toddlers exhibited a significantly longer looking time on same trials in which the habituated label-object pairing remained the same than on switch trials in which the habituated object was labeled as /paap/. In contrast, the /paap/-habituated toddlers showed no looking time difference between same and switch trials. That is, infants noticed the switch when the fricative c onsonant changed to a stop, but not when the stop changed to a fricative. This asymmetrical perf ormance follows the same pattern as observed in Japanese adults, i.e., performance was better wh en a fricative changed to a stop.

In a subsequent experiment, Altvater-MacKensen and Fikkert (2010) examined another group of toddlers' ability to discriminate the 'paap' to 'faap' change in a phonetic discrimination task which does not require establishing a link between a word and an object. The results showed that 14-month-olds detected a stimulus change when the demands of lexical processing were reduced. Altvater-MacKensen and Fikkert (2010) took the result as evidence against an acoustic salience view and ruled out the potential contribution of acoustic factors to asymmetries obtained in the word leaning task. The potential role of perceptual salience related to acoustic features was also disregarded.[4] However, asymmetry was not fully tested. This study did not examine the discrimination of the 'faap' to 'paap' change in 14-month-olds under the assumption that they would detect the 'faap' to 'paap' shift since they had done so in the relatively more demanding word-object matching task. Thus, there still remains the possibility of a group difference in discrimination performance between /paap/- and /faap/-habituated toddlers, which is suggestive of asymmetry according to the habituated consonant.

As a follow-up, Altvater-MacKensen *et al*. (2014) investigated the developmental profile

[4]According to Altvater -MacKensen and Fikkert (2010: 1907), "Obviously, the burst information of a stop is available later than the frication noise of a fricative. This might lead to an asymmetry. However, we did not time-lock infants' looking times to the unfolding acoustic signal but compared looking times over the whole trial." As regards perceptual salience, Altvater -MacKensen and Fikkert (2010) appear to suggest that stops are less perceptually salient than fricatives because of the former being available to the listener later than the latter.

of directional asymmetry by examining 18- and 25-month-old Dutch-learning toddlers' perception of familiar words, using the intermodal preferential looking paradigm. Toddlers were presented a pair of pictures, a target picture and a distracter, while the target object was labeled. The target label was either correctly pronounced or mispronounced by exchanging the manner of articulation in its initial position. Toddlers in the younger age group were assigned to one of the /b/-, /d/-, /v/- and /z/-initial word conditions and toddlers in the older age group were assigned to either /b/- or /v/-initial words. The results showed that 18-months olds who were tested on fricative-initial words looked longer on correct trials than on mispronunciation trials; 18 month olds tested on the stop-initial words did not respond differently to correct and mispronunciation trials. Thus, a stop was not noticed in a mispronounced word but a fricative was noticed in a mispronounced word. In contrast, toddlers at the age of 25 months were equally responsive to a fricative mispronunciation and a stop mispronunciation with a familiar word. Altvater-MacKense n *et al*. (2014) concluded that performance appears to be symmetrical in this task by 25 months of age. However, given that this study used familiar words, this result leaves open the possibility that directional asymmetry may persist into the third year of life if task demands are increased by using unfamiliar words or nonce words.

Taken together, Altvater-MacKensen *et al*. (2010, 2014) take these findings as information about how infants are developing and refining their emerging phonological representations during early lexical development, suggesting the lexical representation hypothesis detailed below.

2.4.2.1 The lexical representation hypothesis

The lexical representation hypothesis assumes that a marked fricative manner is specified by the phonological feature [continuant] but a stop manner functioning as the default manner of articulation is underspecified in children's early lexical representations.

As described in detail in the Table 2-2, according to the lexical representation hypothesis, toddlers habituated to 'paap' have difficulty in detecting mispronunciation because 'paap' has no stored lexical manner feature (i.e., Ø) to compare with the acoustic feature [continuant] extracted from 'faap' and thus accept fricatives as stops. On the contrary, 'faap'-habituated toddlers detect the stimulus change due to the specified representational status of the fricative 'faap'. At first blush, the lexical representation hypothesis appears to have the same view with the FUL hypothesis. However, according to the lexical representation hypothesis, the initially under-refined phonological representation of stops becomes specified in the mature lexicon and thus no asymmetry is assumed to emerge in adult phonemic perception[5], but the FUL hypothesis assumes that the representational status of stops remains underspecified.

Table 2-2 Detection of stimulus change in early lexical representations. Adapted from Altvater-MacKensen & Fikkert (2010).

	'paap' →'faap'	'faap' →'paap'
word learned during habituation	paap	faap
stored phonological feature	[Ø]	[continuant]
mispronounced word during test	faap	paap
extracted acoustic feature	[fricative]	[stop]
matching	no-mismatch	mismatch
detection of stimulus change	no	yes

As noted above, Altvater-MacKensen *et al*. (2010, 2014) point out phonological represent

[5]Altvater -MacKensen and Fikkert (2010: 1908) noted that "children's early lexical representations are not adult-like but unspecified with respect to certain features. We assume that children specify the marked member of a newly learned contrast but that the unmarked member does not get specified in early lexical representations. Representations gradually become more detailed in the course of development."

ations as the primary source of directional asymmetry in toddlers' stop-fricative perception while downplaying other factors such as acoustic properties in the observed asymmetries. This will be discussed in more detail in section 5.4. While the toddler's phonological representation may be an important factor underlying the asymmetries in stop-fricative word processing, the above findings do not effectively rule out an acoustic-phonetic bias as a factor contributing to asymmetrical patterns.

2.5 Other models related to directional asymmetries

Apart from the above-reviewed conceptual models, the perceptual assimilation model (PAM) by Tyler *et al*. (2014) is also, but not exclusively, concerned with directional asymmetries. The PAM developed by Best (1993, 1994, 1995) is one of the most frequently cited theoretical frameworks on non-native segmental contrasts perception in adult listeners. The PAM posits a typology to classify five possible types of perceptual assimilation of non-native contrasts to native phonological categories based on perceived similarity of articulatory properties: a two-category (TC) assimilation, in which both phones of a non-native contrast are assimilated to two different native phonemes; a single-category (SC) assimilation, in which both non-native phones are assimilated to the same native phoneme with the same level of perceived similarity; a category-goodness (CG) assimilation, in which both non-native phones are assimilated to the same native phoneme but with different degrees of perceived similarity; a uncategorized-categorized (UC) assimilation, in which one non-native phone is assimilated to a native phoneme but the other is not assimilated to any native phoneme; and a uncategorized-uncategorized (UU) assimilation, in which both non-native phones are not assimilated as any of the native phoneme categories.

Very recently, the PAM principles were extended to make predictions about directional asymmetries in vowel contrast perception in a study by Tyler *et al*. (2014) in which they examined adult listeners' perception of non-native vowel contrasts. Tyler *et al*. (2014) suggested that directional asymmetries may occur for SC, CG and UU assimilation contrasts in which listeners exploit phonetic information rather than phonological information in discriminating those contrasts but will not arise for TC and UC assimilation contrasts which are discriminated on the basis of native phonological boundaries.

The native language magnet-expanded (NLM-e) model proposed by Kuhl *et al*. (2008) also recognizes the occurrence of directional asymmetries in infancy. In principle, the NLM-e model is about developmental changes in speech perception; that is, infants initially discriminate phonetic units in a language-general mode and, over time, become language-specific perceivers. NLM-e stipulates that directional asymmetries are shown for consonants and vowels during the initial phase of development. Meanwhile, directional asymmetries in infant consonant perception was first evidenced in Kuhl *et al*. (2006)'s study in which English- and Japanese-learning infants at 6-8 and 10-12 months of age were significantly better at discriminating a change from /la/ to /ra/ than the other way around.[6] Thus far, this study is the only published study showing that directional asymmetries are exhibited for consonants in infancy. Another point to be made concerns acoustic salience. Within NLM-e, the role of acoustic salience in speech perception is acknowledged but this is considered to exert only a temporarily influence in early development. Similarly, the PRIMIR framework (Processing Rich Information from Multidimensional Interactive Representations) proposed by Werker and Curtin (2005) also acknowledges that some

[6]It is interesting to note that Tsushima (1999) observed that there was a non-significant tendency for Japanese preschoolers, elementary schoolers and adults to discriminate a stimulus change from /ra/ to /la/ better than the reverse, given that synthetic stimuli were used by Tsushima (1999) as well as Kuhl *et al*. (2006). A similar tendency was also observed by Tsushima (2007).

initial universal phonetic biases can be observed in young infants but, in their view, these phonetic biases are short-lived and will be shaped by specific language exposure very early in development.

It should be noted that PAM, NLM-e and PRMIRI are not explicit about what drives perceptual asymmetries and moreover were not initially motivated to explain perceptual asymmetries. As a result, this dissertation will mainly concentrate on the NRV framework, the FUL model and the lexical representation account. However, PAM, NLM-e and PRIMIR will be also evaluated in light of the results in the following chapters.

2.6 The privileged status of stops over fricatives

The potential role of a perceptual bias grounded in acoustic asymmetry in stop-fricative perception is the focus of this dissertation. In what follows, the inherent acoustic asymmetry between stops and fricatives is first addressed. Then, a working hypothesis regarding perceptual asymmetries in the perception of stop-fricative contrasts is outlined.

To address the presence of a perceptual bias grounded in acoustic property salience in consonant manner contrast perception, this dissertation speculates that not all consonant manners are equally acoustically salient. More particularly, stops are speculated to be privileged over fricatives in terms of acoustic salience. Stops are typically produced by complete oral closure accompanied by a buildup of intra-oral air pressure and a subsequent release accompanied by a brief burst of the maintained air (e.g., Ladefoged, 1993; Shinn & Blumstein, 1984). The burst reveals an abrupt amplitude rise which is marked by an abrupt onset of energy in the spectrogram (e.g., Lisker & Abramson, 1964; Stevens & Keyser, 1989). That is, the little-energy level during the occlusion period suddenly reaches the high-energy noise amplitude immediately after the rapid opening gestures. In contrast, fricatives are produced with incomplete closure that allows air to continuously flow through the mouth and generate audible noise, showing a gradual amplitude rise (e.g., Repp *et al*., 1978). As a result, with specific respect to rise time (i.e., time interval from the onset of consonant noise to the maximum amplitude of the noise duration), stops have remarkably shorter rise times than fricatives (e.g., Shinn, 1985; see also Patil & Rao, 2008).

It is known that the sharper the amplitude change, the greater the response magnitude in the peripheral auditory system (Delgutte & Kiang 1984; Smith, 1979). In this respect, Stevens and Keyser (1989) suggested that the acoustic property of a non-continuant stop is more salient than that of a continuant fricative. Specifically, Stevens and Keyser (1989) suggested that the abruptness in the stop burst and the sharpness in a shift in acoustic landscapes at the release results in an increase in a distinctive auditory response and thus an increase in perceptual salience (see also Wright, 2004). (As a side note, startling alarm sounds have a very rapid rise time.) In a similar vein, Kawasaki (1982) suggested that sharper amplitude or spectral changes in the speech signal increase the acoustic salience and result in better detectability. These suggestions were supported by one MEG study investigating the auditory responses as a function of the onset acoustics of words. Gage *et al*. (1998) reported that stop-initial English words starting with /b, p, d, t, g, k/ produced significantly shorter M100 latencies and higher amplitudes than non-stop-initial English words starting with /f, s, l, r, m/ in native English listeners. Given that the M100 is an index of acoustic information processing and category-level perceptual information processing as well (Roberts *et al*., 2004; Shestakova *et al*., 2004), these results lend support for the crucial role of onset abruptness in stop-fricative perception.

A different line of research provides additional relevant evidence of the possibility that sounds with abrupt onset acoustics is perceptually more distinctive and salient than those with gradual onset acoustics. Easwar *et al*. (2012) examined the effects of rise time of phonetic segments on cortical auditory evoked potential responses in English listeners. Easwar *et al*. (2012) found that the syllable-medial /ʃ/ segments extracted from /aʃɪl/ with shorter rise times (mean: 27.6 ms) elicit larger amplitudes with shorter latencies compared to syllable-initial /ʃ/ segments extracted from /ʃi/ with longer rise times (mean: 113.2 ms). The onset abruptness plays a similar role in tone perception. In an event-related brain potential study, Thomson *et al*. (2009) found that tone stimuli with a short rise time of 15 ms elicited a significantly larger N1 amplitude than those with a long rise time of 185 ms. Thomson *et al*. (2009) also found that the 15 ms rise time tone stimuli elicited a greater N1 amplitude than the 50 ms rise time tone stimuli when their overall intensity was perceptually matched (see also Kodera *et al*., 1979; Onishi & Davis, 1968).

2.6.1 The acoustic-phonetic bias hypothesis: An extension of NRV to consonants

This dissertation begins with the high salience bias hypothesis that the sudden increase of onset energy of stops may enhance acoustic prominence and may well grant a privileged perceptual status to a stop manner over a fricative manner. The more salient unit (stop) grabs attention more effectively than the less salient unit (fricative). This differential salience puts a spotlight on the privileged unit. The spotlighted unit can then serve as a referent for discovering other manner types (fricatives) as needed, i.e., when fricatives are also phonemic in the native language. Such a perceptual bias toward a high-salience unit over a low-salience unit may initially serve as a key perceptual mechanism that optimally facilitates the processing of most pertinent (i.e., most salient) consonant information of the ambient language. In adults, this initial bias would be expected to fade as efficiency in perception of native contrasts develops and overrides or weakens the initial bias. However, the initial bias may re-emerge in adults when perceiving non-native contrasts.

Thus, extending the NRV framework in this context, the high salience bias hypothesis predicts initial biases in consonant manner perception to appear early in development (in infancy) and to show the same basic developmental profile as in vowel perception. In other words, young infants would show a perceptual bias favoring stops over fricatives; however, this bias would diminish (but not disappear entirely) with development of a native contrast and be maintained for a non-native contrast. Applying the NRV framework specifically to discrimination performance as a function of stimulus presentation direction in consonant manner perception, the high salience bias hypothesis proposes that a perceptual bias is revealed as directional asymmetry in which a change from a fricative to a stop is more easily detectable than the reverse change. This is so because in the stop to fricative shift, the high-salience stop overrides the low-salience fricative.

2.7 Research aims

As outlined earlier, prior research suggests the probable existence of a perceptual bias in stop-fricative perception in adults and toddlers. As regards the underlying mechanisms of asymmetry in consonant manner perception, the FUL hypothesis assumes that asymmetry emerges in perception of underspecified stop/specified non-stop contrasts due to asymmetric phonological representational specificity. The toddler data have been explained by the lexical representation hypothesis appealing to a similar assumption. The hypothesis is based on the findings of asymmetries in processing of stop-fricative manner in the context of an associated word learning task in toddlers. On the other hand, the acoustic-phonetic bias hypothesis adopts the concept of the perceptual biases of the NRV model and proposes that as in vowel perception, perceptual biases grounded in acoustic properties underlie perceptual asymmetries in consonant perception.

The purpose of this dissertation is to investigate the presence of perceptual biases in consonant manner perception in an attempt to extend the NRV framework to consonants. This investigation is addressed by examining how adults from different language backgrounds perceive English and Persian contrasts which differ in initial stop-fricative manner and how young infants perceive an English stop-fricative contrast which is phonemic for them.

2.7.1 Stop-fricative perception in adults

Given limited data, it is not clear whether this asymmetry occurs only for the /b/-/v/ contrast or can be observed for other stop-fricative contrasts. Major questions addressed are the following: whether directional asymmetries are observed in the discrimination of other stop-fricative contrasts in adults, and if so, how language experience affects such asymmetries. The acoustic-phonetic bias hypothesis suggests that as shown with vowels, language experience shapes the consonant bias to optimize native language processing. Thus, the hypothesis is viable if a perceptual bias favoring stops over fricatives (i.e., directional asymmetry) is observed in non-native listeners, but this bias is weaker (or absent) in native listeners. To reinforce the hypothesis of the existence of a perceptual bias, a supplementary question is also addressed: whether a perceptual bias is reflected in listeners' identifying stops and fricatives by showing more perceptual consistency for stops than for fricatives.

2.7.2 Stop-fricative perception in infants

Previous studies have shown perceptual asymmetries in processing of phonemic stop-fricative word and non-word contrasts in toddlers during the second year of life. However, there are no data from younger infants. Hence, it is unclear whether the perceptual bias is in place prior to the onset of babbling or word recognition. The question to be addressed here is whether a directional asymmetry is evident in phonemic stop-fricative perception in young infants. This is explored by testing English- and French-learning infants between 5 and 6 months of age. Lastly but probably most importantly, to provide direct evidence of a stop bias, listening preferences are assessed in 5-6-month-old English- and French-learning infants.

2.8 Outline of chapters

The remainder of this dissertation is structured as follows. In Chapter 3, an acoustic analysis of English and Persian stop and fricative stimuli is first presented and then the first of two perceptual tests with adults for this dissertation (Experiment 1) is presented. The acoustic analysis shows that the rise time of stops is significantly shorter than that of fricatives, motivating the question of the role of onset abruptness in stop-fricative perception. An identification task as Experiment 1 is conducted which provides preliminary evidence to suggest the existence of a salience hierarchy of consonant manners. The results show that English, French and Korean listeners perceive stops more reliably than fricatives for non-phonemic contrasts and some phonemic contrasts as well. Listeners regardless of language backgrounds tend to perceive non-native fricatives as stop consonants. These identification patterns suggest the aspect of a stop unit as a referent manner in consonant perception and support the hypothesis of the presence of an acoustic-phonetic bias.

In Chapter 4, the second perceptual test with adults (Experiment 2) is conducted. Experiment 2 examines whether the onset acoustic asymmetry between stops and fricatives is revealed as perceptual asymmetry in adult listeners' stop-fricative perception. In the standard AX discrimination paradigm, English-, French- and Korean-speaking listeners are presented six

English /pas/-/fas/, /bas/-/vas/, /das/-/ðas/, /das/-/zas/, /tas/-/θas/ and /tas/-/sas/ contrasts and two Persian /kas/-/xas/ and /gas/-/ɣas/ contrasts. Listeners' patterns of the perception of different pairs show that for non-phonemic contrasts, listeners in all 3 language groups discriminate fricative-stop pairs (i.e., fricative to stop changes) better than stop-fricative pairs (i.e., stop to fricative changes) with only some exceptions, revealing salience overriding. For phonemic contrasts, listeners in all language groups perform symmetrically regardless of direction with one exception, English listeners' perception of the phonemic /ðas/-/das/ contrast. English listeners show a directional symmetry for this contrast (ðas/→/das/</das/→/ðas/). Asymmetries also exist in the perception of same pairs. For non-phonemic contrasts, all the English, French and Korean listeners perceive the stop-stop pairs better than the fricative-fricative pairs with one exception. These results confirm that pairs of highly salient stops are easier for listeners to perceive than pairs of weakly salient fricatives. This pattern (stop-stop>fricative-fricative) also emerges for certain phonemic contrasts; for example, English listeners' perception of the phonemic /das/-/ðas/ and /tas/-/θas/ contrasts. The overall results suggest that the inherent salience differences generally interact with language-specific experience, consistent with the acoustic-phonetic bias hypothesis, but also suggest that the inherent salience differences may not be fully modulated by language experience. Importantly, the pattern of stop-stop>fricative-fricative challenges the FUL model.

Chapter 5 assesses whether the salience differences affect perception in young infancy. Two experiments (Experiments 3 and 4) are conducted. In Experiment 3, English- and French-learning infants at 5-6 months are tested with the phonemic /bas/-/vas/ contrast in the visual habituation procedure. The results show that whereas /bas/-habituated infants fail to notice the /bas/ to /vas/ change, those infants who are habituated to /vas/ detect the /vas/ to /bas/ change successfully. This evident directional asymmetry in phonemic perception in young infants supports the potential role of acoustic-perceptual salience differences in stop-fricative perception given that the phonetic discrimination task measures infants' ability to resolve acoustic-phonetic details rather than their phonological representational encoding ability. In Experiment 4, English- and French-learning 5-6-month-olds' listening preferences are measured in the sequential preferential looking procedure. The results indicate that infants prefer to listen longer to /bas/ than /vas/. The presence of a preference asymmetry buttresses the high salience bias hypothesis. The stop preference is particularly difficult to be interpreted by the lexical representation hypothesis.

Chapter 6 concludes: first, by summarizing the main findings from the experiments conducted and discussing them in terms of the high salience bias hypothesis. Most importantly, the applicability of the NRV framework to consonant perception is justified; and second, by proposing some future steps addressing the relative role of acoustics and language experience in stop-fricative perception in infancy in an attempt to verify that asymmetries in the perception of consonant manner contrasts are grounded in acoustic properties.

Chapter 3: Adult cross-language identification of stop-fricative contrasts

3.1 Introduction

As mentioned in Chapter 2, the acoustic-phonetic bias hypothesis underpinning this dissertation assumes that stops are acoustically more salient than fricatives. In this study, acoustic salience is related to the abruptness of onset acoustics, particularly a rise time. More specifically, the acoustic-phonetic bias hypothesis assumes that stops with abrupt bursts and shorter rise times are acoustically salient due to the abrupt increase in their onset amplitudes. This assumption takes its inspiration particularly from work by Gage *et al*. (1998) showing that stop-initial words with abrupt onset acoustics elicit larger and faster responses than words starting with other manner types with gradient onset acoustics (described in section 2.6). While a theoretical rationale for differences in acoustic and perceptual salience can be offered that is in line with previous data on infant and adult perception of stop-fricative contrasts (outlined in Chapter 2), additional perceptual tests are needed to confirm the presence of perceptual asymmetries along with relevant acoustic analyses to clarify the specific cues that support such perceptual patterns.

The goal of the present chapter is two-fold: first, to confirm the acoustic properties that are expected to differentiate the stop-fricative contrasts and second, to evaluate manner perception patterns in the identification of native and non-native consonants (Experiment 1). Experiment 1 is the first of two perceptual tests with adults. In this experiment, French- and Korean-speaking adults are presented with eight stop and fricative syllables from English /bas, vas, das, ðas, zas, tas, θas, sas/ and Persian /kas, xas, gas, ɣas/. Experiment 1 is conducted as a preliminary approach to the role of salience effects in consonant manner perception. Salience effects may interact with phonemic status (perceptual assimilation). With respect to non-native perception, instances could arise where assimilation patterns to the native language will reduce or perhaps reverse inherent acoustic-phonetic asymmetries. Asymmetries elicited for contrasting non-native stops and fricatives may also be shaped by language experience. In other words, native language assimilation may lead to mapping of the phones to different manners or to different native phones.

This preliminary assessment investigates whether for the selected contrasts, the contrasting stop and fricative productions are perceived as different manner classes by non-native listeners. Specifically, this assessment investigates whether stop productions are perceived as stops (or non-stops) and fricatives as fricatives (or non-fricatives) in non-native listeners. Asymmetries in consonant manner perception in non-native listeners may depend on whether or not (or how) the listener assimilates contrasting phones to different manner classes. No prior studies have examined non-native stop-fricative contrasts within the language groups in this study. Thus, this preliminary assessment was conducted to provide baseline identification performance for the eight stop-fricative contrasts selected for this study. These initial findings also provide an opportunity to examine broad differences in stop-fricative perception that might emerge if, as hypothesized, stops serve as natural referent consonants. That is, we could expect asymmetrical patterns in non-native consonant identification. Within the acoustic-phonetic bias hypothesis, the specific predictions regarding non-native stop and fricative perception are that (1) stops will show more direct and stable one-to-one (non-native/L2-to-native/L1) assimilations compared to fricatives, which will be less stable, more varied one L2-to-two (or many) L1 mappings and (2) non-native fricatives will often be assimilated as stops but the reverse mapping (non-native stops assimilated to fricatives) will be absent or rare.

The present chapter begins by providing a brief description of the stop and fricative inventories of English, French and Korean to compare inter-language differences in acoustic and perceptual experience with stop-fricative contrasts. Then, the stimulus acoustics of English and Persian stops and fricatives in initial-syllable positions are presented, followed by the preliminary perceptual assessment.

3.2 The stop and fricative inventories: English, French and Korean

English (English refers to North American English) has many fricatives, both voiced and voiceless forms as well as sibilant and non-sibilant fricatives. English has nine fricatives: /f, v, θ, ð, s, z, ʃ, ʒ, h/. However, the glottal fricative /h/ is often considered as a semivowel or the voiceless onset of the following vowel (Ladefoged, 1993; Ladefoged & Maddieson, 1998). English doesn't have a stop-fricative contrast at the velar place of articulation as there is no velar voiced (or voiceless) fricative. In syllable initial positions, English stops are produced with labial, alveolar and velar place; voiced stops are produced either with pre-voicing, [b], [d], [g] or short voicing lag, [p], [t], [k], while voiceless stops are produced as long voicing lag and aspiration, [p^h], [t^h], [k^h] (Lisker & Abramson, 1964).

French (French refers to Quebec French) also has many fricatives, but fewer than English. French has six fricatives /f, v, s, z, ʃ, ʒ/ (Walker, 1984). French stops also come in voiceless-voiced pairs at the labial, alveolar and velar places; voiced stops are produced with pre-voicing, [b], [d], [g], and voiceless stops are produced with short voicing lag, [p], [t], [k] (Laeufer, 1996). As with English, French does not have a stop-fricative contrast at the velar place of articulation due to lack of a velar fricative phoneme.

Korean has only three fricatives /s*, s, h/. In Korean, stops form an unusual three-way laryngeal contrast. The contrast consists of three series of voiceless stops: the tense or fortis series /p*, t*, k*/, the lax or lenis series /p, t, k/, and the aspirated series /p^h, t^h, k^h/ (e.g., Cho & Keating, 2001; Cho *et al*., 2002). The tense stops are laryngealized and unaspirated, the lax stops are breathy and slightly aspirated, and the aspirated stops are strongly aspirated (Cho *et al*., 2002). In standard Korean, the tense stops have shortest voice onset time (VOT) and high F0, the lax stops have intermediate/long VOT and low F0, and the aspirated stops have longest VOT and high F0 (Cho *et al*., 2002). In Korean, homorganic stop-fricative contrast is limited to just two /t-s/ and /t-s*/ due to lack of fricatives.

3.3 Acoustics of English and Persian stops and fricatives in initial position

3.3.1 Speakers

Four native speakers of two American English (1 male) and two Canadian English (1 male) and four native Persian speakers (2 males) were recorded. All of the speakers were in their 20s. The Persian speakers were raised by monolingual Persian-speaking parents and educated from elementary till undergraduate in Iran. Length of residence in Canada ranged from 1 year to 2 years at the time of recording. All English and Persian speakers were affiliated with McGill University except one English male speaker. None of the speakers reported any history of voice or speech disorders. The speakers were phonetically untrained and all were naïve as to the purpose of the experiment. All speakers were compensated for their participation.

3.3.2 Stimulus materials

The stimulus materials consisted of ten English and four Persian consonant-vowel-consonant (CVC) syllables in which the initial consonant was a target stop or fricative consonant.

The target consonant was always followed by /as/. For example: /pas/, /fas/, /bas/, /vas/ …, and so on. Six English and two Persian syllable pairs were constructed; each was a minimal pair contrasting in stop vs. fricative manner of articulation for the initial consonant (see Table 3-1). All the CVC syllables formed nonce words in English, French and Korean. This was done in order to ensure that acoustic-phonetic properties per se of stop and fricative consonants would induce listeners' responses in the subsequent perceptual experiments. Randomized lists of target syllables were read in carrier phrases so that speakers would produce target syllables as naturally as possible. The target syllables were surrounded by vowels for as much consistency as possible across English and Persian – "I will say CVC again." in English and "**اهم بگویم** CVC **من دوباره گاس از اول**." [Man dobare CVC az aval mikhaham beguyam] in Persian.

Table 3-1 Stop-fricative contrast corpus.

Stimulus	Contrast	Phonemic status		
		English group	French group	Korean group
English	/tas/-/sas/	native phonemic	phonemic	phonemic
	/bas/-/vas/	native phonemic	phonemic	**non-phonemic**
	/pas/-/fas/	native phonemic	phonemic	**non-phonemic**
	/das/-/zas/	native phonemic	phonemic	**non-phonemic**
	/das/-/ðas/	native phonemic	**non-phonemic**	**non-phonemic**
	/tas/-/θas/	native phonemic	**non-phonemic**	**non-phonemic**
Persian	/gas/-/ɣas/	**non-phonemic**	**non-phonemic**	**non-phonemic**
	/kas/-/xas/	**non-phonemic**	**non-phonemic**	**non-phonemic**

3.3.3 Recording procedure

All of the speakers were recorded individually in a sound-attenuated booth at McGill Infant Speech Perception Lab in the School of Communication Sciences and Disorders at McGill University. All speech stimuli were recorded via an external pre-amplifier in mono in Praat 5.1.44 (Boersma & Weenink, 2010) at a sampling rate of 44.100 kHz directly onto a LG X-Note E300 laptop computer (16 bit resolution) located in a room adjacent to the booth using a Sennheiser microphone. Two types of lists for target syllable production were created: one written in English for the English speakers and the other written both in Persian and Roman orthographies for the Persian speakers. Before recording, speakers had ample practice time to read the reading lists to ensure familiarity. In the case of the English speakers, practice was focused on /θas/ and /ðas/. The speakers were instructed to read the reading list at a normal speaking rate. The experimenter sat in the adjacent room listening to the speaker producing the target syllables over headphones, giving written instructions (e.g., "Please repeat the sentence again.") through a window between the adjacent room and the recording booth. Each speaker produced 20 tokens or more per target syllable. The entire recording session lasted approximately 1 hour or so for English stimulus recording and 40 minutes for Persian stimuli.

3.3.4 Stimuli selection and verification procedures

The above review of English, French and Korean stops and fricatives suggests that French listeners are likely to assimilate English voiced stops with short-lag VOTs, along with English voiceless long-lag stops, to their voiceless stops. This also leads us to take into account the VOT realizations of Persian stop consonant cognate /g/ and /k/. The Persian /k/ is produced with long voicing lag (aspiration) while the VOT values of the Persian /g/ fall between voicing lead and short lag, but tend to be mostly skewed toward short lag (Bijankhan & Nourbakhsh, 2009). Unlike French counterparts, Korean listeners phonologically identify English voiced unaspirated variants with voicing lead or short lag VOTs as English voiced stops and English aspirated stops as English voiceless stops (e.g., Park & de Jong, 2008; Schmidt, 1996). As a result, only pre-voiced variants were used for the voiced stops in this corpus. This ensures that English and French listeners assimilate the prevoiced [b], [d] and [g] phones to the voiced cognates in their native language and thus the contrasting stop-fricative pairs differ only in the manner of articulation feature; this then allows Korean listeners to assimilate the prevoiced stops to their unaspirated cognates and the voiceless aspirated stops to their aspirated cognates, respectively.

All tokens were edited using the Praat acoustic analysis software. Some acoustic measurements such as duration, F0 and intensity were made on the vocalic portions of the tokens in order to preclude the possible influence of vowel quality in stop-fricative manner perception. Based on the analysis, six tokens of each syllable type were selected from each of those four English and four Persian speakers: 5 tokens × 10 target syllables × 4 speakers (200 English tokens) and 6 tokens × 4 target syllables × 4 speakers (96 Persian tokens). All the English and Persian tokens were presented to 3 independent English-speaking judges and 3 Persian-speaking judges, respectively, for identification and goodness judgments on a five-point scale ranging from extremely unidentifiable and foreign-sounding (1) to highly identifiable and native-like (5). Additionally, 200 tokens except the English interdental /θas/ and /ðas/ syllables (5 tokens × 10 target syllables × 4 speakers) were judged and rated by three French-speaking judges. This was done to ensure that the stimuli selected for the study are highly intelligible items (i.e., good exemplars) to French listeners. Finally, two tokens of each of the target syllables per speaker were selected on the basis of judgmental and rating consistency across native judges. This resulted in the following set of stimuli: 80 English tokens (2 tokens × 10 syllables × 4 speakers) and 32 Persian tokens (2 tokens × 4 syllables × 4 speakers). The native judges were compensated for their participation.

3.3.5 Stimulus measurements and acoustics

Table 3-2 illustrates the results of the acoustic analysis according to manner of articulation for each stop-fricative contrast. As shown in the Table, acoustic measurements were made on rise time (ms), F0, F1, F2 and F3 onset frequencies (Hz), F0 mean frequencies (Hz), F1, F2 and F3 midpoint frequencies (Hz), vowel /a/ duration (ms), final consonant /s/ duration (ms) and mean vowel intensity (dB). The F0, F1, F2 and F3 onsets were measured for the first three pitch periods after vowel onset. The results of the one-way ANOVAs listed in the last column indicate that manner of articulation was reliably differentiated only in terms of rise time, showing that the rise times of stops are significantly shorter than those of fricatives. Additionally, the rise times of stops showed less variation than those of fricatives across all stop-fricative contrasts. None of the other acoustic measurements that were taken approached significance. The present results with the rise time are consistent with the results of Shinn's (1985) cross-linguistic study showing that stops were very reliably differentiated from fricatives along the rise time in Czech, German and Mandarin (see also Patil & Rao, 2008; Weigelt *et al*., 1990).

Table 3-2 Stimulus acoustics. Shaded rows indicate significant distinctions in manner.

Contrast	Manner Variable	Stop Mean	Stop SD	Fricative Mean	Fricative SD	Results
/pas/-/fas/	Rise time	17.15	7.01	119.78	23.48	$F(1, 14) = 151.98, p < .001$
	F0 onset	158.49	68.03	170.56	73.12	$F(1, 14) = 0.002, p = .96$
	F0 mean	153.96	64.25	165.93	70.02	$F(1, 14) = 0.04, p = .95$
	F1 onset	766.64	101.59	699.81	76.28	$F(1, 14) = 0.07, p = .80$
	F2 onset	1227.17	153.67	1237.63	154.66	$F(1, 14) = 3.63, p = .13$
	F3 onset	2605.43	187.00	2489.86	116.24	$F(1, 14) = 0.03, p = .86$

	F1 midpoint	796.66	99.53	800.44	99.65	*F*(1, 14) = 0.09, *p* = .77
	F2 midpoint	1280.33	153.18	1285.73	149.05	*F*(1, 14) = 0.66, *p* = .43
	F3 midpoint	2563.01	144.42	2657.06	173.35	*F*(1, 14) = 0.28, *p* = .60
	/a/ duration	130.76	27.72	141.65	23.28	*F*(1, 14) = 0.20, *p* = .66
	/s/ duration	117.75	26.09	117.66	18.54	*F*(1, 14) = 0.07, *p* = .80
	Intensity	76.01	3.02	75.93	1.55	*F*(1, 14) = 0.2=07, *p* = .80
/bas/-/vas/	Rise time	10.44	3.80	74.33	10.30	*F*(1, 14) = 286.75, *p* < .001
	F0 onset	161.70	70.96	164.59	64.44	*F*(1, 14) = 0.01, *p* = .93
	F0 mean	158.77	65.20	163.31	63.06	*F*(1, 14) = 0.02, *p* = .89
	F1 onset	678.25	95.95	633.45	39.18	*F*(1, 14) = 1.49, *p* = .24
	F2 onset	1233.24	180.89	1170.31	137.52	*F*(1, 14) = 0.61, *p* = .45
	F3 onset	2540.00	171.73	2486.73	176.62	*F*(1, 14) = 0.37, *p* = .55
	F1 midpoint	788.41	123.86	805.68	99.24	*F*(1, 14) = 0.09, *p* = .76
	F2 midpoint	1288.07	168.05	1206.05	114.36	*F*(1, 14) = 1.30, *p* = .27
	F3 midpoint	2659.03	225.48	2601.61	203.48	*F*(1, 14) = 0.29, *p* = .60
	/a/ duration	159.46	26.90	161.97	33.40	*F*(1, 14) = 0.03, *p* = .87
	/s/ duration	116.30	23.37	120.01	21.70	*F*(1, 14) = 0.11, *p* = .75
	Intensity	75.56	3.29	76.29	3.46	*F*(1, 14) = 0.25, *p* = .63
/tas/-/θas/	Rise time	13.53	5.15	125.14	27.56	*F*(1, 14) = 129.06, *p* < .001
	F0 onset	177.48	74.49	507.85	945.65	*F*(1, 14) = 0.97, *p* = .34
	F0 mean	166.89	66.42	166.26	64.33	*F*(1, 14) = 0.0004, *p* = .98
	F1 onset	767.66	135.80	677.56	61.61	*F*(1, 14) = 3.92, *p* = .11
	F2 onset	1267.79	176.36	1383.90	143.28	*F*(1, 14) = 3.06, *p* = .17
	F3 onset	2525.69	273.04	2738.64	219.04	*F*(1, 14) = 3.96, *p* = .11
	F1 midpoint	794.45	106.28	795.65	70.91	*F*(1, 14) = 0.0007, *p* = .98
	F2 midpoint	1233.99	99.60	1231.12	217.57	*F*(1, 14) = 0.001, *p* = .97
	F3 midpoint	2503.34	169.35	2516.60	580.75	*F*(1, 14) = 0.004, *p* = .95
	/a/ duration	134.06	31.11	129.85	23.45	*F*(1, 14) = 0.09, *p* = .76
	/s/ duration	119.46	24.57	109.48	24.39	*F*(1, 14) = 0.67, *p* = .43
	Intensity	73.88	3.46	75.17	3.42	*F*(1, 14) = 1.12, *p* = .31
/das/-/ðas/	Rise time	8.92	3.37	75.00	9.13	*F*(1, 14) = 368.63, *p* < .001
	F0 onset	158.31	67.53	158.06	64.50	*F*(1, 14) = 0.0001, *p* = .99
	F0 mean	155.82	61.95	154.73	64.85	*F*(1, 14) = 0.001, *p* = .97
	F1 onset	621.60	55.39	627.14	78.46	*F*(1, 14) = 0.03, *p* = .87
	F2 onset	1574.54	235.02	1486.48	163.89	*F*(1, 14) = 0.76, *p* = .40
	F3 onset	2573.61	186.64	2691.86	168.11	*F*(1, 14) = 1.77, *p* = .20
	F1 midpoint	1031.01	741.89	804.32	53.04	*F*(1, 14) = 0.74, *p* = .40
	F2 midpoint	1307.37	144.79	1288.09	188.05	*F*(1, 14) = 0.05, *p* = .82
	F3 midpoint	2641.62	277.61	2715.99	161.24	*F*(1, 14) = 0.43, *p* = .52
	/a/ duration	160.53	39.41	163.35	13.11	*F*(1, 14) = 0.04, *p* = .85
	/s/ duration	113.81	25.66	118.57	27.59	*F*(1, 14) = 0.19, *p* = .67
	Intensity	75.23	3.97	76.21	3.67	*F*(1, 14) = 0.47, *p* = .50
/tas/-/sas/	Rise time	13.53	5.15	93.65	23.86	*F*(1, 14) = 93.51, *p* < .001
	F0 onset	177.48	74.49	171.87	70.21	*F*(1, 14) = 0.01, *p* = .88
	F0 mean	166.89	66.42	165.18	66.96	*F*(1, 14) = 0.003, *p* = .96
	F1 onset	767.66	135.80	733.24	119.18	*F*(1, 14) = 0.31, *p* = .59

	F2 onset	1267.79	176.36	1387.50	159.92	$F(1, 14) = 3.02, p = .18$
	F3 onset	2525.69	273.04	2650.70	188.14	$F(1, 14) = 1.14, p = .30$
	F1 midpoint	794.45	106.28	785.98	121.37	$F(1, 14) = 0.02, p = .88$
	F2 midpoint	1233.99	99.60	1323.61	161.42	$F(1, 14) = 1.79, p = .20$
	F3 midpoint	2503.34	169.35	2615.46	150.98	$F(1, 14) = 1.95, p = .18$
	/a/ duration	134.06	31.11	131.50	11.04	$F(1, 14) = 0.05, p = .83$
	/s/ duration	119.46	24.57	120.19	8.28	$F(1, 14) = 0.01, p = .94$
	Intensity	73.88	3.46	74.49	3.48	$F(1, 14) = 0.17, p = .69$
/das/-/zas/	Rise time	8.92	3.37	73.56	11.49	$F(1, 14) = 226.00, p < .001$
	F0 onset	158.31	67.53	161.53	64.23	$F(1, 14) = 0.01, p = .92$
	F0 mean	155.82	61.95	155.81	63.70	$F(1, 14) = 0.000, p = 1.00$
	F1 onset	621.60	55.39	569.70	85.93	$F(1, 14) = 3.06, p = .17$
	F2 onset	1574.54	235.02	1473.56	126.60	$F(1, 14) = 1.17, p = .30$
	F3 onset	2573.61	186.64	2653.69	181.26	$F(1, 14) = 0.76, p = .40$
	F1 midpoint	1031.01	741.89	748.23	109.89	$F(1, 14) = 1.14, p = .30$
	F2 midpoint	1307.37	144.79	1275.10	164.24	$F(1, 14) = 0.17, p = .68$
	F3 midpoint	2641.62	277.61	2694.14	234.32	$F(1, 14) = 0.17, p = .69$
	/a/ duration	160.53	39.41	155.46	28.66	$F(1, 14) = 0.09, p = .77$
	/s/ duration	113.81	25.66	114.07	19.95	$F(1, 14) = 0.01, p = .91$
	Intensity	75.23	3.97	76.10	3.53	$F(1, 14) = 0.28, p = .60$
/kas/-/xas/	Rise time	14.35	4.52	113.86	11.02	$F(1, 14) = 558.61, p < .001$
	F0 onset	193.77	56.62	179.84	55.09	$F(1, 14) = 0.21, p = .65$
	F0 mean	184.97	53.93	178.14	53.07	$F(1, 14) = 0.07, p = .80$
	F1 onset	684.95	121.27	764.77	111.32	$F(1, 14) = 1.88, p = .19$
	F2 onset	1273.39	113.20	1189.92	116.19	$F(1, 14) = 3.07, p = .17$
	F3 onset	2704.49	289.24	2699.04	259.55	$F(1, 14) = 0.002, p = .97$
	F1 midpoint	778.58	103.50	797.07	97.63	$F(1, 14) = 0.14, p = .72$
	F2 midpoint	1260.13	113.81	1250.94	93.92	$F(1, 14) = 0.03, p = .86$
	F3 midpoint	2733.87	298.26	2733.16	241.27	$F(1, 14) = 0.0002, p = .99$
	/a/ duration	195.89	49.06	180.64	17.80	$F(1, 14) = 0.68, p = .42$
	/s/ duration	107.47	17.56	113.66	24.19	$F(1, 14) = 0.34, p = .57$
	Intensity	73.87	3.47	73.13	3.29	$F(1, 14) = 0.39, p = .54$
/gas/-/ɣas/	Rise time	9.66	5.79	77.09	13.58	$F(1, 14) = 166.98, p < .001$
	F0 onset	181.47	55.63	167.20	54.41	$F(1, 14) = 0.27, p = .61$
	F0 mean	178.70	60.04	175.69	60.41	$F(1, 14) = 0.01, p = .92$
	F1 onset	523.94	75.81	606.61	120.62	$F(1, 14) = 3.69, p = .12$
	F2 onset	1361.42	108.88	1281.22	63.85	$F(1, 14) = 3.26\ p = .09$
	F3 onset	2687.97	468.05	2801.08	435.08	$F(1, 14) = 0.25, p = .62$
	F1 midpoint	711.68	134.83	754.39	181.67	$F(1, 14) = 0.29, p = .60$
	F2 midpoint	1255.99	114.97	1189.40	129.42	$F(1, 14) = 1.18, p = .30$
	F3 midpoint	2757.64	311.33	2768.36	423.85	$F(1, 14) = 0.003, p = .95$
	/a/ duration	225.26	44.39	237.12	28.93	$F(1, 14) = 0.40, p = .54$
	/s/ duration	127.42	64.99	117.95	27.64	$F(1, 14) = 0.14, p = .71$
	Intensity	74.38	3.25	74.15	3.85	$F(1, 14) = 0.03, p = .86$

3.4 Experiment 1: The identification experiment

In this experiment, the focus will be placed on how salience effects interact with perceptual assimilation in non-native stop-fricative perception.

3.4.1 Methods

3.4.1.1 Participants

Sixteen North American English-speaking listeners (mean age, 23.3 years; range, 20-30), fifteen French listeners (mean age, 35.2 years; range, 29-45) and twenty Korean listeners (mean age, 28.8 years; range, 19-39) participated. In very much the same way as in the discrimination task, only those functionally monolingual English-speaking participants were included in this experiment. The French-speaking participants had beginning (n=3) to intermediate (n=11) to lower-advanced (n=1) levels of English proficiency. All the French participants started to acquire English after 10 years of age. All of the Korean-speaking participants had basic knowledge of English and had no experience with English native speakers except one with an upper-intermediate level of English proficiency. The nineteen Korean participants had not stayed in an English-speaking environment more than one week. All the English, French and Korean participants had no prior experience with Persian or any Arabic languages. English and French participants were compensated for their participation. Korean participants were all volunteers.

3.4.1.2 Stimulus materials

The stimulus materials are identical to those described in section 3.3.3. Listeners were presented a total of 112 tokens (2 tokens × 4 talkers × 14 syllables). Thus, each syllable was repeated eight times.

3.4.1.3 Procedure

English participants were tested in a quiet room in the School of Communication Sciences and Disorders at McGill University. French participants were tested in their homes. Korean participants were all tested in a quiet room in a speech-language learning center in Jeju Province in Korea. All the English and Korean participants were presented the stimuli at a self-adjusted comfortable listening level over headphones connected to a LG X-Note E300 laptop computer. French participants were also instructed to listen to the stimuli via headphones. Listeners were instructed to identify the initial sound of each syllable in terms of their L1 categories using their respective L1 orthographic symbols (open-response). English listeners were asked to write down “dh” for /ð/ and “th” for /θ/. Once the identification decision was made, then listeners had to rate the goodness-to-fit of the given sound to their selected L1 sound using a 5-point scale. Listeners were told to use “1” if the given sound was totally dissimilar from the selected L1 category and “5” if the given sound was very similar to the selected L1 category. The stimulus presentation order was randomized within each stimulus package for each listener. Listeners were asked to listen to the stimuli once for the identification and once for the rating but they were able to listen to the same stimulus more than two times if needed. No feedback or comments were given during the tests.

3.5 Results

Tables 3-3, 3-5 and 3-7 present mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as English, French and Korean consonants, respectively, by English, French and Korean listeners. The percentages indicate the overall frequency of each L1 consonant which was chosen to identify the given consonant stimuli, summed over all the listeners in each language group. The most frequently chosen identification response is boldfaced. The goodness ratings are based on a 5-point scale (5=very good instances; 1=very poor instances). Stimuli presented are listed in rows and participant responses are listed in columns. Identification responses obtaining less than 1% are not shown. Shaded columns represent phonemes that are lacking in English, French and Korean.

Table 3-4, 3-6 and 3-8 presents a summary of the perceptual assimilation types for the English and Persian consonant stimuli (see section 2.5 for the PAM's classification of assimilation types). Following Antoniou *et al*. (2013), the results were analyzed using a 70% categorization criterion. "C" indicates that the most chosen category for a stimulus was identified with the same native label category in 70% or more of a language group's responses; otherwise, the stimulus was considered as uncategorized ("U"). Meanwhile, it should be noted that there is no clear criterion for deciding whether a stimulus is categorized as an instance of a particular native category (see Bundgaard-Nielsen *et al*., 2011 and Harnsberger, 2001, for discussions of categorization criteria). Again, following Antoniou *et al*. (2013), a *t*-test was performed on the category goodness ratings when two stimuli of a given contrast were categorized as instances of the same native category. When the goodness ratings differed, the contrast was considered as CG assimilation; otherwise, the contrast was considered as SC assimilation. Shaded rows represent non-native contrasts for English, French and Korean listeners, respectively.

Table 3-3 Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as English consonants by English listeners (n=16).

	Stop stimulus						**Fricative stimulus**							
		English			Persian				English				Persian	
	p	b	t	d	k	g	f	v	θ	ð	s	z	x	ɣ
p	**98 (4.2)**	2 (1.7)					1.6 (4.0)							

Stops	b	**91 (4.2)**												
	t		**94 (4.3)**						3 (3.0)	3 (1.8)				
	d			**99 (4.1)**										
	k				**99 (3.9)**								**70 (1.7)**	6 (1.8)
	g					**100 (3.4)**	3.3 (4.7)							20 (3.5)
Fricatives	f						**93 (4.0)**		16 (3.9)					
	v	3 (3.3)						**97 (3.6)**		21 (3.4)				
	θ		6 (4.8)				3 (3.3)		**79 (3.6)**					
	ð									**72 (3.7)**		2 (1.0)		
	s								2 (3.5)		**99 (4.1)**	7 (3.0)		
	z							2 (3.0)		2 (3.0)		**91 (3.9)**		
	h												22 (1.4)	24 (3.1)
Others	r												8 (1.1)	**48 (1.4)**

Table 3-4 Assimilation types for each contrast by English listeners.

Stimulus	Categorized/Uncategorized (≥70%)	Category label	Contrast	Assimilation type
/t/	C	/t/	/t/-/s/	TC
/s/	C	/s/		
/k/	C	/k/	/k/-/x/	CG
/x/	C	/k/		
/g/	C	/g/	/g/-/ɣ/	UC
/ɣ/	U			
/p/	C	/p/	/p/-/f/	TC
/f/	C	/f/		
/b/	C	/b/	/b/-/v/	TC
/v/	C	/v/		
/d/	C	/d/	/d/-/z/	TC
/z/	C	/z/		

/t/ /θ/	C C	/t/ /θ/	/t/-/θ/	TC
/d/ /ð/	C C	/d/ /ð/	/d/-/ð/	TC

As expected, English listeners categorized all English stop and fricative stimuli as intended, showing TC assimilations for the native /tas/-/sas/, /pas/-/fas/, /bas/-/vas/, /das/-/zas/, /tas/-/θas/ and /das/-/ðas/ contrasts. However, English listeners were relatively less successful at identifying /θ/ (79%) and /ð/ (72%) as intended compared to the other native phones (>90%); /θ/ and /ð/ were misidentified as /f/ and /v/ in 16% and 21% of instances, respectively. For the Persian /kas/-/xas/, English listeners categorized both /k/ and /x/ as English /k/ with ratings of 3.9 and 1.7 in 99% and 70% of instances, respectively, exhibiting CG assimilation ($p < .001$). Meanwhile, /x/ was also identified as English /h/ in 22% of instances. For the Persian /gas/-/ɣas/ contrast, English listeners categorized /g/ unanimously as English /g/ but did not meet the categorization criterion for /ɣ/, spreading their identification choices across /r/ (48%), /h/ (24%) and /g/ (20%), exhibiting UC assimilation.

Table 3-5 Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as French consonants by French listeners (n=15).

		Stop stimulus						Fricative stimulus							
		English				Persian		English						Persian	
		p	b	t	d	k	g	f	v	θ	ð	s	z	x	ɣ
Stops	p	**94 (3.6)**	8 (3.3)			2 (3.5)		3 (1.3)		6 (1.9)		2 (3.0)			
	b		**82 (3.6)**						2 (3.5)						3 (3.0)
	t			**98 (3.6)**				2 (3.0)	2 (5.0)	15 (1.2)				8 (3.8)	
	d	2 (4.0)	4 (3.8)		**98 (3.4)**						**39 (3.2)**				
	k					**92 (3.5)**	4 (3.0)							**56 (3.5)**	10 (3.8)

	g						**96 (3.9)**		2 (3.5)		3 (3.0)		7 (4.0)	5 (1.5)	25 (3.7)
Fricatives	f							**91 (3.9)**		**71 (3.1)**				2 (3.0)	
	v		3 (3.7)					4 (3.0)	**93 (3.8)**		38 (3.6)				
	s		2 (4.0)					2 (4.0)		8 (3.7)		**93 (4.2)**	3 (4.7)		
	z		2 (4.0)								8 (1.2)	3 (1.7)	**91 (3.4)**		
	h	2 (3.5)												3 (1.3)	7 (1.6)
Others	r													25 (1.4)	**52 (3.0)**
	l								2 (4.5)		8 (3. 1)				
	a														2 (1.0)

Table 3-6 Assimilation types for each contrast by French listeners.

Stimulus	Categorized/Uncategorized (≥70%)	Category label	Contrast	Assimilation type
/t/	C	/t/	/t/-/s/	TC
/s/	C	/s/		
/k/	C	/k/	/k/-/x/	UC
/x/	U			
/g/	C	/g/	/g/-/ɣ/	UC
/ɣ/	U			
/p/	C	/p/	/p/-/f/	TC
/f/	C	/f/		
/b/	C	/b/	/b/-/v/	TC
/v/	C	/v/		
/d/	C	/d/	/d/-/z/	TC
/z/	C	/z/		
/t/	C	/t/	/t/-/θ/	TC
/θ/	C	/f/		
/d/	C	/d/	/d/-/ð/	UC
/ð/	U			

As expected, French listeners exhibited TC assimilations for the phonemic /tas/-/sas/,

/pas/-/fas/, /bas/-/vas/ and /das/-/zas/ contrasts by identifying these native-like English stimuli consistently as intended in more than 80% of instances. For the Persian /kas/-/xas/ contrast, French listeners categorized /k/ as French /k/ (92%) but did not categorize /x/, dividing their identification choices between French /k/ (56%) and /r/ (25%), resulting in UC assimilation. For the Persian /gas/-/ɣas/ contrast, French listeners categorized /g/ (96%) highly reliably but did not categorize /ɣ/, spreading their identification choices across /r/ (52%), /g/ (25%) and /k/ (10%). For the English /tas/-/θas/ contrast, French listeners categorized /t/ as French /t/ and interestingly categorized /θ/ as /f/ (71%), revealing TC assimilation. For the English /das/-/ðas/ contrast, French listeners categorized /d/ as French /d/ but did not categorize /ð/, dividing their identification choices between /d/ (39%) and /v/ (38%), revealing UC assimilation.

Table 3-7 Mean percentage identification and goodness ratings (in parentheses) of English and Persian initial consonants as Korean consonants by Korean listeners (n=20).

		Stop stimulus						Fricative stimulus							
		English				Persian		English						Persian	
		p	b	t	d	k	g	f	v	θ	ð	s	z	x	ɣ
Stops	p		**48 (3.9)**					13 (3.4)	**83 (3.9)**	3 (3.5)	9 (3.9)				
	p*		38 (3.4)					30 (3.5)	4 (4.3)	15 (3.3)					
	p^h	**87 (3.6)**	4 (3.0)					**44 (3.7)**	2 (4.5)	15 (3.1)	3 (1.7)	2 (4.0)			2 (3.5)
	t		5 (3.8)	7 (3.5)	**83 (3.2)**			7 (3.2)	2 (3.0)	6 (3.0)	**58 (3.6)**				
	t*				7 (3.6)			7 (3.3)		**29 (3.5)**	3 (3.0)				
	t^h	7 (3.1)		**89 (3.6)**	3 (3.7)			2 (3.0)		2 (3.5)				2 (3.5)	
	k		2 (3.5)		3 (4.0)	7 (3.8)	**53 (3.0)**							13 (3.2)	**70 (3.0)**
	k*					3 (3.0)	42 (3.0)			2 (3.5)				8 (3.7)	10 (4.0)
	k^h	3 (3.7)				**90 (3.7)**								**73 (3.9)**	6 (3.0)
Fricatives	s							3 (3.0)		8 (3.7)		40 (3.0)	8 (4.3)		
	s*							2 (1.0)		25 (3.4)		**56 (3.6)**	4 (3.2)		
	h													3 (3.0)	4 (3.0)

Others	ʧ									2 (1.5)		**82 (3.8)**		2 (3.5)
	ʧ*											3 (3.3)		
	$ʧ^h$		3 (3.5)											
	m							3 (1.7)						
	l							4 (3.8)		23 (3.0)				3 (3.0)
	a													4 (3.0)

Table 3-8 Assimilation types for each contrast by Korean listeners.

Stimulus	Categorized/Uncategorized (≥70%)	Category label	Contrast	Assimilation type
/t/	C	/t^h/	/t/-/s/	UC
/s/	U			
/k/	C	/k^h/	/k/-/x/	CG
/x/	C	/k^h/		
/g/	U		/g/-/ɣ/	UC
/ɣ/	C	/k/		
/p/	C	/p^h/	/p/-/f/	UC
/f/	U			
/b/	U		/b/-/v/	UC
/v/	C	/p/		
/d/	C	/t/	/d/-/z/	TC
/z/	C	/ʧ/		
/t/	C	/t^h/	/t/-/θ/	UC
/θ/	U			
/d/	C	/t/	/d/-/ð/	UC
/ð/	U			

For the phonemic English /tas/-/sas/ contrast, Korean listeners categorized /t/ as Korean aspirated /t^h/ but did not categorize /s/, dividing their identification responses between /s*/ (56%) and /s/ (40%), revealing UC assimilation. For the Persian /kas/-/xas/ contrast, Korean listeners categorized both /k/ and /x/ as Korean /k^h/ with ratings of 3.7 and 3.9 in 90% and 73% of instances, respectively, revealing CG assimilation ($p < .001$). For the Persian /gas/-/ɣas/ contrast, Korean listeners did not categorize /g/, dividing their identification responses between Korean lenis /k/ (53%) and fortis /k*/ (42%), but categorized /ɣ/ as Korean lenis /k/ (70%), revealing UC assimilation. For the English /pas/-/fas/ contrast, Korean listeners categorized /p/ as Korean /pʰ/ but did not categorize /f/, spreading their identification responses across Korean /pʰ/ (44%), /p*/ (30%) and /p/ (13%), revealing UC assimilation. For the English /bas/-/vas/ contrast, Korean listeners did not categorize /b/, dividing their identification responses between Korean /p/ (48%) and /p*/ (38%), but did categorize /v/ as Korean /p/ (83%), exhibiting UC assimilation. Korean listeners exhibited TC assimilation for the non-phonemic English /das/-/zas/ contrast; they categorized /d/ and /z/ as Korean lenis stop /t/ (83%) and affricate /ʧ/ (82%), respectively, almost equally reliably. For the English /tas/-/θas/ contrast, Korean listeners reached the categorization criterion for /t/, as mentioned above, but did not categorize /θ/, spreading their identification responses widely across /t*/ (29%), /seig*/ (25%), /p*/ (15%) and /pʰ/ (15%), exhibiting UC assimilation. This shows that Korean listeners treated the English /θ/ as a very foreign sound belonging nowhere in more than 30% of instances. For the English /das/-/ðas/ contrast, Korean listeners reached the categorization criterion for /d/, as mentioned above, but did not categorize /ð/, dividing their identification responses between /d/ (58%) and /l/ (23%), exhibiting UC assimilation.

3.6 Discussion

The results from this preliminary assessment indicate that whereas non-native stops were identified as exemplars of single native categories for the most part, resulting in more direct one L2-to-one L1 assimilations, non-native fricatives were generally identified as exemplars of two or more native categories, resulting in one L2-to-two (or many) L1 assimilations. Non-native fricatives were usually assimilated as stops. English and French listeners perceived the Persian fricative /x/ as stop /k/ for the most part. Such a pattern was most evident in Korean listeners; the non-native fricatives were assimilated as stop categories in almost every case (the one exception was English /z/). Reverse mapping cases in which stops are assimilated to fricatives were absent. These overall results are consistent with the predictions based on the hypothesis that stop consonant manner is more acoustically salient and stable and thus perceptually more favored than fricative consonant manner.

In addition, it appears to be the case that the privileged status of stops over fricatives is found for some native phones. English listeners just reached the categorization criterion for /ð/. As a result, the seemingly TC /das/-/ðas/ contrast yielded a relative categorization asymmetry between /d/ and /ð/. This was also the case with the TC /tas/-/θas/ contrast. That English listeners had some difficulty identifying the native phone /ð/ is consistent with the results of Shafiro *et al*.'s (2012) investigation into English listeners' identification of twenty native consonants.[7] TC assimilations were found for some non-native contrasts. French and Korean listeners assimilated the English /tas/-/θas/ and /das/-/zas/ contrasts as the stop-fricative /tas/-/fas/ and stop-affricate /das/-/ʧas/ contrasts, respectively, although goodness ratings for English /t/ (3.6) and /d/ (3.2) were higher than for English /θ/ (3.1) and /z/ (3.8), respectively. Higher mean goodness ratings for stops than for fricatives were generally found for non-native stop-fricative contrasts. However, there were note-worthy exceptions. For Korean listeners, the voiced Persian fricative /ɣ/ and the voiced English fricative /v/ were more stably perceived than the voiced Persian stop

[7]Shafiro *et al*.'s (2012) reported that English listeners showed the lowest identification accuracy for /ð/ among twenty syllable-medial consonants (/p, t, k, f, ð, s, ʃ, tʃ, b, d, g, v, z, dʒ, r, l, w, j, m, n/) under three vowel contexts (/aCa/, /iCi/ and /uCu/). The voiced /ð/ was least correctly identified across contexts (64%) and was misidentified as /v/ most often (25%). English listeners' identification performance was highly context-dependent; /ð/ was correctly identified in 60%, 44% and 86% of instances in the /aCa/, /iCi/ and /uCu/ contexts, respectively.

/g/ and the voiced English stop /b/, respectively, resulting in the reversed UC assimilations, that is, uncategorized stops and categorized fricatives. Goodness ratings were the same for /g/ and /ɣ/ (3.0) and for /b/ and /v/ (3.9).

Taken together, these findings from the preliminary identification experiment are taken as suggesting that the onset acoustic salience of stop manner leads to more direct and stable assimilation patterns compared to fricative consonant manner. In addition, the listeners' tendency to assimilate non-native fricatives as stop manner of articulation rather than other manners of articulation supports the idea that stops serve as referent consonants in stop-fricative perception. These findings inspire the question of the role of acoustic salience in listeners' discrimination performance. The experiment in the following chapter examines whether onset acoustic salience differences result in perceptual asymmetries for non-native contrasts. Particularly given the Korean listeners' reversed UC assimilations for the /gas/-/ɣas/ and /bas/-/vas/ contrasts, an interesting question arises: whether the same patterns of asymmetries are found for these contrasts. If raw salience differences can be altered by assimilation patterns, then these contrasts might differ from other contrasts.

Chapter 4: Adult cross-language perception of stop-fricative contrasts

4.1 Introduction

In Chapter 3, the acoustic measurement identified the acoustic parameter that straightforwardly differentiates the syllable-initial stop-fricative contrasts. Apart from the acoustics of stops and fricatives, the preliminary perceptual test with adults (Experiment 1) revealed that for non-native stop-fricative contrasts, listeners categorized stops as single stop categories for the most part but were generally unable to categorize fricatives, spreading their responses across two or more native phones, resulting in mostly UC assimilations. These patterns support the concept that stops and fricatives are not equally salient. The potential existence of a salience hierarchy was also supported by the observation that English listeners were less accurate at identifying stop phones relative to fricative phones for the native /das/-/ðas/ and /tas/-/θas/ contrasts. An investigation of listeners' identification responses indicated that listeners, regardless of language backgrounds, often perceptually substitute stops for non-native fricatives but do not show the reverse pattern, i.e. there were no instances where a non-native stop was perceived as a fricative. This supports the hypothesis that stops have a privileged perceptual status over fricatives and suggests the potential role of salient stops as referent consonants in stop-fricative perception.

The main goal of the present chapter is to test the hypothesis that the onset acoustic properties of stops are more salient than that of fricatives and these differences are revealed as perceptual asymmetries in stop-fricative perception. Listeners from the same language groups as in Experiment 1 are presented with the same English and Persian stop-fricative contrasts in an

AX (same-different) discrimination task (Experiment 2). First, Experiment 2 explores whether, in adults, directional asymmetries are observed for stop-fricative contrasts other than the /b/-/v/ contrast reported in earlier research (reviewed in Chapter 3) and, if so, whether directional asymmetries are modulated by language-specific experience. If stops with abrupt onsets are inherently perceptually more salient than fricatives with gradient onsets, listeners' discrimination of stop to fricative changes will be depressed relative to their discrimination of fricative-to-stop changes due to salience overriding. Such directional asymmetries will emerge for non-phonemic contrasts but will be weaker or absent for phonemic contrasts as predicted by the acoustic-phonetic bias hypothesis. Second and also importantly, Experiment 2 also examines whether English, French and Korean listeners will perceive the stop-stop pairs better than the fricative-fricative pairs. If stops are indeed privileged over fricatives in terms of salience, listeners will also find it easier to perceive the sameness of stop phones compared to fricative phones. This outcome would suggest that perceptual stability contributes to the salience of stop perception. As with different pairs such performance differences are expected to emerge for non-phonemic contrasts but to be weaker or absent for phonemic contrasts as would be predicted by the acoustic-phonetic bias hypothesis.

The above-cited questions are essentially concerned with testing the hypothesis that the high salience of stops overrides the low salience of fricatives. For this reason, of particular interest is how Korean listeners perceive the English /bas/-/vas/ contrast and the Persian /gas/-/ɣas/ contrast. Recall that in Experiment 1, Korean listeners showed reversed identification patterns for these contrasts, with stops being uncategorized but fricatives being categorized. If raw salience differences per se result in perceptual asymmetries, Korean listeners will show perceptual asymmetries in the expected direction under the acoustic-phonetic bias hypothesis.

However, if initial salience differences are modulated by language experience, reversed asymmetries might emerge for the contrasts.

Additionally, the present chapter also addresses two questions concerning general patterns of performance. The first is whether listeners perform better on phonemic contrasts than non-phonemic contrasts (i.e., a phonemic status effect). Such a phonemic status effect is expected to be evident in all listener groups, based on the well-documented findings of adult cross-language speech perception performance that adult listeners' non-phonemic contrast perception is affected by their native language experience (e.g., Cho & McQueen, 2006). The second, an extension of the first question, is whether listeners' performance on non-native stop-fricative contrasts also reflects the status of stop-fricative contrasts within the native language. This question is addressed by performance on the Persian contrasts which are non-phonemic for all three listener groups. If the relative status within the native language plays a role, then for these contrasts English and French listeners, who have many stop- fricative contrasts in their native language are expected to outperform Korean whose native language contains very few stop-fricative contrasts.

4.2 Experiment 2: The discrimination task

The AX discrimination task is implemented because with this simple paradigm we can easily and directly compare discrimination performance in each direction. This task was also employed in the previous behavioral studies of directional asymmetry in stop-fricative perception (Tsushima *et al*., 2003, 2005). In this experiment I implemented a category-based AX task which means that the syllables within each AX pair were always different syllable productions. In the present study each AX pair included one syllable produced by a male talker and one produced by a female talker. To succeed in this category-based task, listeners need to attend to category differences (phonetic matches or mismatches) and ignore acoustic differences that are irrelevant to category differentiation.

4.2.1 Methods

4.2.1.1 Participants

Twenty native North American English-speaking listeners (mean age, 22.7; range, 18-30), twenty native French-speaking listeners (mean age, 34.5; range, 24-43) and twenty native Korean-speaking listeners (mean age, 27.1; range, 21-35) participated in Experiment 1. All the English-speaking participants were monolinguals. The inclusion criteria for French participants were those who started to learn English after seven years of age, went to French primary and secondary schools and speak in French at home while not using English often at work or in daily life. The primary inclusion criterion for Korean participants was those who were born in Korea and were educated from elementary to high schools in Korea.

Table 4-1 provides a summary of English language background information for the French and Korean listeners mainly based on their self-appraisal of English proficiency (via questionnaire). None of the English, French and Korean participants had prior exposure to Persian and none had taken any course in phonetics or linguistics. All of them were without hearing problems or histories thereof, as assessed by self-report. All were compensated for their participation.

Table 4-1 English language background profiles of the French and Korean speakers.

Language group	English proficiency	Length of residence in Canada (or other English-speaking countries)
French*	lower-to-middle advanced (n=5)	All were lifetime residents in and around Montreal.

	intermediate-to-upper intermediate (n=5)	
	beginning/lower intermediate (n=10)	
Korean	lower-to-middle advanced (n=1)	3 years
	upper intermediate (n=4)	9 months to 2.4 years
	lower intermediate (n=6)	1 month to 3 years
	beginning (n=9)	1 week to 10 months

*Note that half of the French participants had intermediate-to-middle advanced levels of English proficiency mainly due to exposure to academic English and audio-video mass media in English language.

4.2.1.2 The AX trials

The AX task consisted of a total of 512 trials. The total trials were broken into four blocks of 128 trials and each block contained four sub-blocks of 32 trials each. Each sub-block consisted of 24 English and 8 Persian trials. For each contrast, the AX task consisted of 32 different pair trials including16 stop, fricative (SF) pairs and 16 fricative, stop (FS) pairs, and 32 same pair trials including 16 stop, stop (SS) pairs and 16 fricative, fricative (FF) pairs. Thus, the full task included 16 repetitions of each trial type (FS, FS, SS, FF) for each of the 8 contrasts (8 contrasts × 4 trial types × 16 repetitions = 512 trials). For each trial type, the order of the talker gender was counterbalanced across sub-blocks and blocks. As illustrated in Table 4-2, within a first sub-block, syllable tokens (designated as M_1stop$_1$ and M_1fricative$_1$) produced by English and Persian male talkers (M_1) always preceded syllable tokens (designated as F_1stop$_1$ and F_1fricative$_1$) produced by their respective female counterparts (F_1). As shown in Table 4-3, the order of the talker gender was reversed within a second sub-block; syllable tokens from female talkers were presented before syllable tokens from their male counterparts. In the same way, tokens from the other English and Persian male and female talker pairs (M_2-F_2) constituted a 3rd and a 4th sub-block. Thus, each block had four combinations of talker gender pairs. For each trial type the syllable tokens within the AX pair were produced by a different talker. Block presentations, sub-block presentations within each block and trial presentations within each sub-block were all randomized for each participant. The inter-stimulus interval (ISI) was 1500 ms. The specific ISI was selected because a long ISI is known to facilitate a phonetic mode rather than an acoustic mode in the perception of speech segment pairs.

Table 4-2 A set of AX pair matrices for a stop-fricative contrast for a first sub-block.

Stimulus presentation order (Talker M_1, Talker F_1)		Male talker 1 (M_1)	
		$Stop_1$	$Fricative_1$
Female talker 1 (F_1)	$Stop_1$	M_1Stop_1-F_1Stop_1 Same	$M_1Fricative_1$-F_1Stop_1 **Different**
	$Fricative_1$	M_1Stop_1-$F_1Fricative_1$ **Different**	$M_1Fricative_1$-$F_1Fricative_1$ Same

Table 4-3 A set of AX pair matrices for a stop-fricative contrast for a second sub-block.

Stimulus presentation order (Talker F_1, Talker M_1)		Female talker 1 (F_1)	
		$Stop_2$	$Fricative_2$
Male talker 1 (M_1)	$Stop_2$	F_1Stop_2-M_1Stop_2 Same	$F_1Fricative_2$-M_1Stop_2 **Different**
	$Fricative_2$	F_1Stop_2-$M_1Fricative_2$ **Different**	$F_1Fricative_2$-$M_1Fricative_2$ Same

4.2.1.3 Set up and Procedure

Participants were tested individually in a sound-attenuated booth in the infant speech perception lab at McGill University, using the Paradigm program (Tagliaferri et al., 2010) running on a Windows XP laptop computer. Stimuli were played over Sony noise canceling MDR-NC 7 headphones at a comfortable listening level. Participants were told that they would hear pairs of syllables which share the same middle vowel and the same final consonant. On each trial, participants were asked to press the left arrow key if the syllables started with the same consonants and the right arrow key if the syllables started with two different consonants. The direction of the arrows matched the location of the corresponding choice panel on the computer screen: “Same” in the left panel and “Different” in the right panel. Participants were first given a brief session of task familiarization which consisted of 8 unscored warm-up trials. The practice session was terminated when the participant expressed that he/she felt comfortable with the experimental procedure. The practice sessions for each of the English- and Korean-speaking participants were administered by the experimenter in English and in Korean, respectively, and those for French participants by a native French research assistant. The on-screen instructions were also presented in the native language. After the practice sessions, participants were asked if they fully understood the task and had any further questions about it. Participants were encouraged to respond as quickly as possible, but not so quickly as to impose a sacrifice on performance accuracy. In the experimental session, once the participant made a response, the software program automatically proceeded to the next trial. No feedback was given during practice or during experimental sessions. Periodic self-paced breaks (a maximum of 5 minutes long) were given after every 4 blocks of 128 trials (i.e., after block 1, block 2 and block 3) in order to compensate fatigue. The experiment lasted approximately 70 minutes (on average).

4.3 Statistical results

For each listener an overall percent correct score (across all AX pairs) was computed for each contrast. For each contrast, percent correct was also separately calculated for each type of AX pair, i.e., different pairs and same pairs. Two sets of analyses were then performed; the results are organized in two sections.

The first set of analyses is reported in section 4.3.1. In these analyses, performance was collapsed across all eight contrasts. This section provides a global analysis of performance across listeners with different language backgrounds. In this section overall performance (percent correct across all AX pairs) is analyzed first, and then is followed by separate analyses on different pairs (SF+FS) and same pairs (SS+FF). The latter analyses examine whether stops and fricatives are equally salient percepts and how this interacts with language experience. The analyses on different pairs assess whether perceivers are affected by the direction of change, i.e., stop to fricative (SF) vs. fricative to stop (FS). A direction effect (FS>SF) is expected for non-phonemic contrasts but not for phonemic contrasts. The analyses of same pairs were to ascertain whether category sameness is easier to perceive for one type of phone than the other, i.e., for SS vs. FF pairs. A pair type effect (SS>FF) may be observed if stops are more perceptually stable than fricatives; this effect is also expected for non-phonemic but not for phonemic contrasts.

The second set of analyses (section 4.3.2) provides a detailed performance assessment conducted for each contrast. As in section 4.3.1, an analysis of overall performance (percent correct collapsed across all AX pairs) is followed by separate analyses for performance on different pairs and on same pairs. A summary of findings across the 8 contrasts is provided in Tables 4-4 through 4-6 in section 4.4.

4.3.1 Global performance across all contrasts

4.3.1.1 All AX pairs

Figure 4-1 (Panel A) illustrates overall discrimination performance (percent correct scores summed across all AX pairs) collapsed across all eight contrasts plotted for each language group. Data were analyzed with a one-way ANOVA (Language Group: English, French and Korean). The analysis revealed a significant effect of Language Group [$F(2, 57) = 20.98$, $p < .001$; partial $\eta^2 = .42$]. Tukey post-hoc tests indicated that Korean listeners performed significantly more poorly than both English and French listeners ($p < .001$). The latter two groups did not differ ($p = .50$).

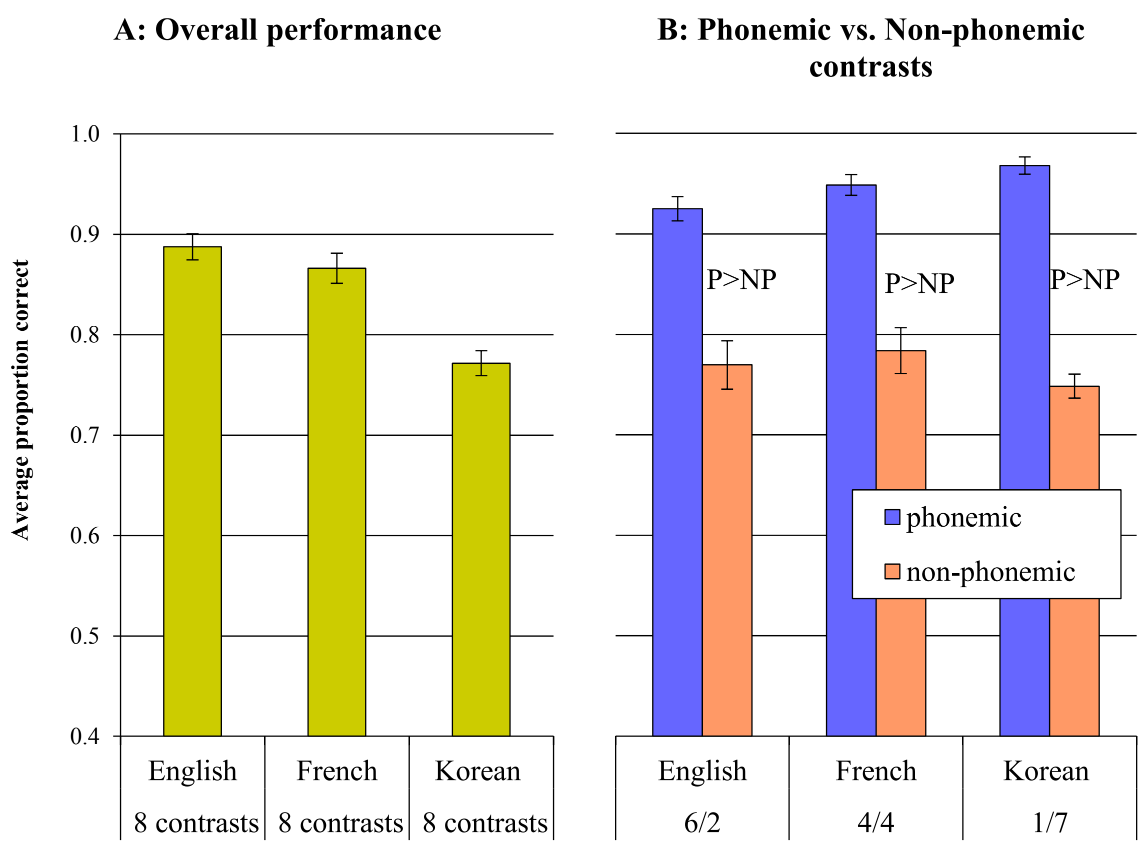

Figure 4-1 Average proportion correct across all the eight stop-fricative contrasts in English, French and Korean listeners: overall performance (A) and performance according to phonemic status (B) (error bars: ±1 SE). P>NP indicates significantly better performance on the phonemic contrasts compared to the non-phonemic contrasts (*p* < .05); 6/2, 4/4 and 1/7 indicate the number of phonemic/non-phonemic contrasts, respectively, for each language group.

Figure 4-1 (Panel B) shows percent correct scores for each language group plotted separately for phonemic and non-phonemic contrasts. The phonemic status profile is identified for each language group; P denotes phonemic contrasts and NP denotes non-phonemic contrasts. Recall that the number of phonemic and non-phonemic contrasts varies across language groups; the number of each contrast type is also indicated for each language group in Panel B. These data were submitted to a 3 (Language Group: English, French and Korean) × 2 (Status: phonemic and non-phonemic) mixed ANOVA. The main effect of Language Group was not significant [$F(2, 57) = 0.53$, $p = .59$]. The main effect of Status was significant [$F(1, 57) = 297.24$, $p = <.001$; partial $\eta^2 = .84$], showing, as expected, that listeners performed significantly better on the phonemic compared to the non-phonemic contrasts. There was an interaction between Language Group and Status [$F(2, 57) = 4.79$, $p = .029$; partial $\eta^2 = .12$]. This interaction was further probed with tests of simple main effects. Simple main effects of Language Group with one-way ANOVAs revealed that language groups performed differently on the phonemic contrasts [$F(2, 57) = 4.34$, $p = .02$] but not on the non-phonemic contrasts [$F(2, 57) = 0.84$, $p = .44$]. With respect to phonemic contrasts, follow-up Tukey tests revealed that Korean listeners were more accurate than English listeners ($p = .01$) but no other group differences were observed. As expected, simple effects analyses confirmed a significant effect of Status (phonemic>non-phonemic) in all three groups ([$t(19) = 7.67$, $p < .001$], [$t(19) = 8.56$, $p < .001$] and [$t(19) = 15.48$, $p < .001$], for English, French and Korean listeners, respectively).

4.3.1.2 Different pairs

Figure 4-2 shows percent correct scores on different pairs for each language group plotted separately for phonemic and non-phonemic contrasts.

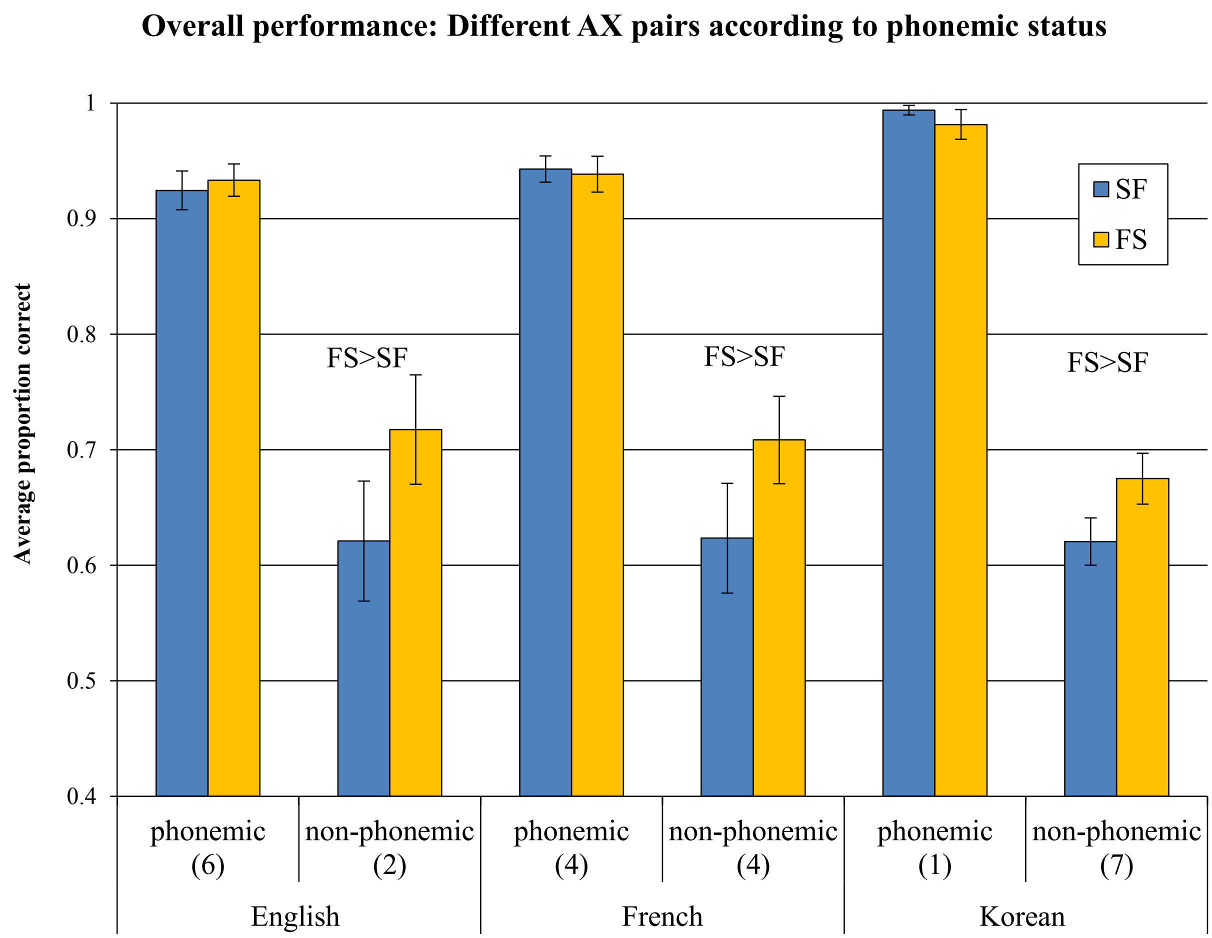

Figure 4-2 Average proportion correct on different pairs according to phonemic status for English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates that listeners perceived the FS pairs significantly better than the SF pairs (*p* < .05); numbers in the parentheses indicate number of phonemic and non-phonemic contrasts for each language group.

Percent correct on different pairs were submitted to a 3 (Language Group: English, French and Korean) × 2 (Status: phonemic and non-phonemic) × 2 (Direction: SF and FS) mixed ANOVA. The main effect of Language Group was not significant [$F(2, 57) = 0.19$, $p = .83$]. The main effect of Status reached significance [$F(1, 57) = 194.60$, $p < .001$; partial $\eta^2 = .77$], showing the predicted effect (phonemic>non-phonemic). The main effect of Direction was significant [$F(1, 57) = 42.25$, $p < .001$; partial $\eta^2 = .43$], with listeners detecting the FS pairs significantly better than the SF pairs. There was also a significant interaction between Status and Direction [$F(1, 57)$

= 39.30, p < .001; partial η^2 = .41]. No interactions with Language Group reached significance: Language Group × Status, [F(2, 57) = 1.37, p = .26]; Language Group × Direction, [F(2, 57) = 2.55, p = .09]; Language Group × Status × Direction, [F(2, 57) = 0.30, p = .74].

Probing the Status × Direction interaction, simple effects of Status revealed that subjects performed better on phonemic compared to non-phonemic contrasts for both SF pairs [t(59) = 14.40, p < .001] and FS pairs [t(59) = 11.98, p < .001]. Simple effects of Direction revealed that the Direction effect differed as a function of Status; as expected, the direction effect (FS>SF) was evident for the non-phonemic contrasts [t(59) = -7.29, p < .001] but not for the phonemic contrasts [t(59) = 0.38, p = .70].

4.3.1.3 Same pairs

Figure 4-3 plots percent correct scores for SS and FF pairs in each language group. The contrast profile for each language group is also indicated on the figure. These scores were submitted to a 3 (Language Group: English, French and Korean) × 2 (Status: phonemic and non-phonemic) × 2 (Pair Type: SS and FF) mixed ANOVA. The main effect of Language Group was not significant [F(2, 57) = 2.02, p = .14]. There were significant main effects of Status [F(1, 57) = 115.24, p < .001; partial η^2 = .67], in the expected direction (phonemic>non-phonemic) and Pair Type [F(1, 57) = 144.86, p < .001; partial η^2 = .72], also in the predicted direction (SS <FF). The interaction between Language Group and Pair Type failed to reach significance [F(2, 57) = 2.84, p = .07]. There were significant interactions between Language Group and Status [F(2, 57) = 5.82, p = .005; partial η^2 = .17] and between Status and Pair Type [F(1, 57) = 40.27, p < .001; partial η^2 = .41]. A marginal three-way interaction of Language Group, Status and Pair Type was also observed [F(2, 57) = 4.13, p = .051; partial η^2 = .10].

To probe the interactions, separate Status × Pair Type ANOVAs were conducted for each language group. Within the English listener group, the analysis revealed significant main effects of Status [$F(1, 19) = 15.69$, $p = .001$; partial $\eta^2 = .45$] and Pair Type [$F(1, 19) = 130.66$, $p < .001$; partial $\eta^2 = .87$] as well as a significant Status × Pair Type interaction [$F(1, 19) = 18.73$, $p < .001$; partial $\eta^2 = .50$]. Simple effects of Status revealed the expected pattern (phonemic > non phonemic) for FF pairs [$t(19) = 4.39$, $p < .001$] but not for SS pairs [$t(19) = -0.51$, $p = .62$]. Simple effects of Pair Type revealed an effect of Pair type (SS> FF) for both phonemic ([$t(19) = 4.30$, $p < .001$] and non-phonemic contrasts [$t(19) = 9.52$, $p < .001$].

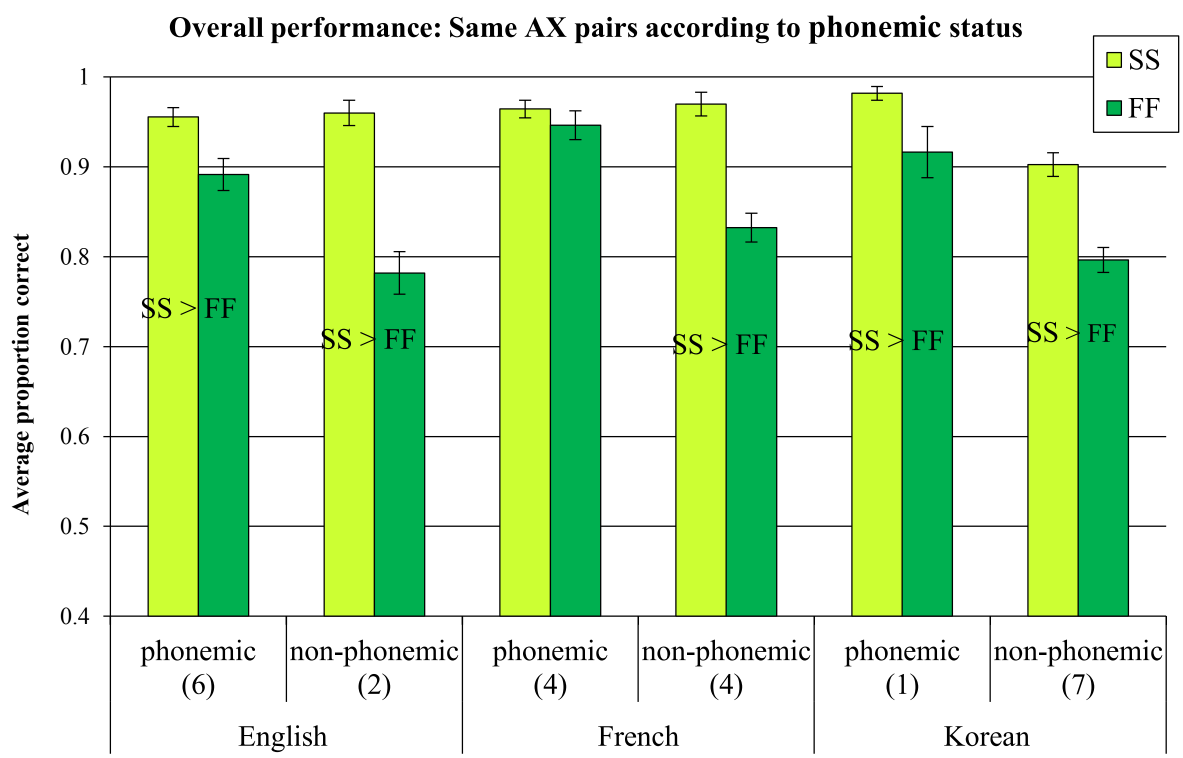

Figure 4-3 Average proportion correct on same pairs according to phonemic status for English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs (*p* < .05); numbers in the parentheses indicate number of phonemic and non-phonemic contrasts for each language group.

Within the French listener group, there were significant main effects of Status [$F(1, 19) = 52.59$, $p < .001$; partial $\eta^2 = .73$] and Pair Type [$F(1, 19) = 58.71$, $p < .001$; partial $\eta^2 = .76$] and also an interaction between Status and Pair Type [$F(1, 19) = 38.19$, $p < .001$; partial $\eta^2 = .67$]. Simple effects of Status revealed that French listeners showed the expected effect (phonemic>non-phonemic) for FF pairs [$t(19) = 7.12$, $p < .001$] but not for SS pairs [$t(19) = -0.85$, $p = .41$]. Simple effects of Pair Type revealed that French listeners showed the expected pattern (SS>FF) for non-phonemic contrasts [$t(19) = 9.12$, $p < .001$] but not for phonemic contrasts [$t(19) = 1.40$, $p = .18$].

Within the Korean listener group, there was a significant main effect of Status [$F(1, 19) = 71.92$, $p < .001$; partial $\eta^2 = .79$], as expected, showing better performance on the phonemic compared to the non-phonemic contrasts. The Pair Type effect reached significance [$F(1, 19) = 21.39$, $p < .001$; partial $\eta^2 = .53$], showing better performance on SS pairs than FF pairs. The interaction between Status and Pair Type was not significant [$F(1, 19) = 2.06$, $p = .17$].

To summarize, as expected, the global analysis indicated a phonemic status effect (phonemic>non-phonemic) across all AX pairs. A direction effect (FS>SF) emerged as a function of phonemic status; listeners in all three groups showed the FS>SF effect for the non-phonemic contrasts but not for the phonemic contrasts. Similarly, there was a pair type effect (SS>FF) for the non-phonemic contrasts in all groups. The SS>FF effect was also evident across phonemic contrasts in English and Korean listeners. The next section will outline the findings within each contrast to assess how well the findings emerging from this global analysis reflect perception of individual contrasts.

4.3.2 Performance on each contrast

In this section, each contrast will be analyzed separately. For each contrast, the results will be presented in a figure with three panels. The first panel (A) on the left shows the total percent correct (for all AX pairs) for the contrast in each language group. The second panel (B) in the middle shows the percent correct on different pairs for each group and the third panel (C) on the right shows the percent correct on same Pairs for each group. See figures 4-4 through 4-11. In each figure, the contrast status is identified for each language group; P denotes a phonemic contrast and NP denotes a non-phonemic contrast.

For each contrast (in sections 4.3.2.1 through 4.3.2.8) three ANOVAs were conducted. First, a one-way between-subjects ANOVA was conducted to analyze the overall percent correct scores across the three language groups (data in Panel A). Second, a Language Group (English vs. French vs. Korean) × Direction (SF vs. FS) mixed ANOVA was conducted to examine performance on different pairs (data in Panel B). Third, a Language Group (English vs. French vs. Korean) × Pair Type (SS vs. FF) mixed ANOVA was conducted to examine performance on same pairs (data in Panel C). In the second and third ANOVAs, simple effects of Direction and of Pair Type were analyzed (using two-tailed paired t-tests) in each language group when either factor was significant, even if the interaction with Language Group was not significant. This was done because Direction and Pair Type are the focal variables of interest in this study. The Direction and Pair type effects noted on the figures represent the outcome of the simple effects analysis.

4.3.2.1 /tas/-/sas/

Figure 4-4 shows performance on the /tas/-/sas/ contrast which is phonemic in all three language groups.

Overall (Panel A): The one-way ANOVA revealed no effect of Language Group [$F(2, 57) = 0.64$, $p = .53$].

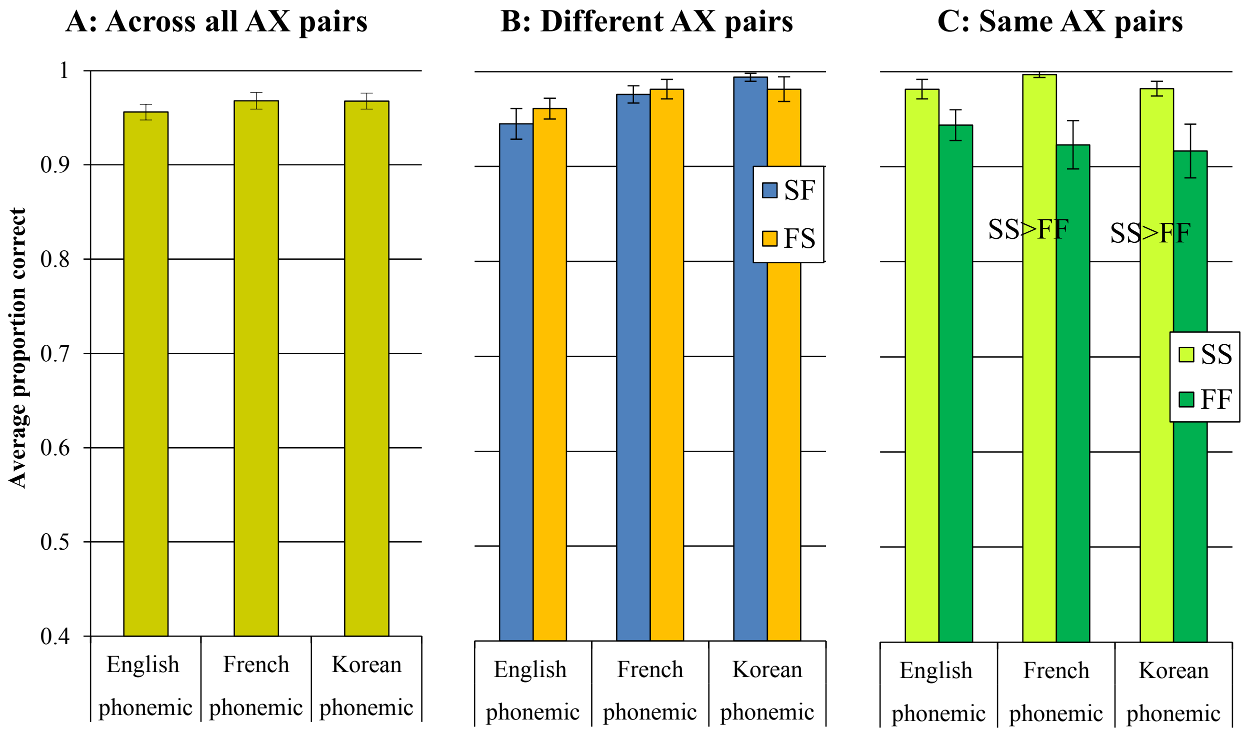

Figure 4-4 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /tas/-/sas/ contrast in English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs (*p* < .05).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 4.29$, $p = .045$; partial $\eta^2 = .10$]. Tukey's post-hoc test indicated that Korean listeners performed significantly better than English listeners ($p = .04$); no other group differences were significant. The main effect of Direction was not significant [$F(1, 57) = 0.26$, $p = .61$]. There was no interaction between Language Group and Direction [$F(2, 57) = 1.97$, $p = .15$]. Simple effect tests failed to show any effects of direction in any of the language groups: English, [$t(19) = -1.77$, $p = .09$]; French, [$t(19) = -0.72$, $p = .48$]; Korean, [$t(19) = 0.94$, $p = .36$].

Same Pairs (Panel C): The Language Group × Pair Type mixed ANOVA revealed a significant main effect of Pair Type [$F(1, 57) = 17.29$, $p < .001$; partial $\eta^2 = .23$], with listeners doing better on SS than FF pairs. The other effects failed to reach significance: Language Group [$F(2, 57) = 0.32$, $p = .72$]; Language × Pair Type [$F(2, 57) = 0.65$, $p = .53$]. Simple effects analyses revealed a marginal effect of Pair Type in the English group [$t(19) = 1.92$, $p = .07$] and a significant Pair Type effect (SS>FF) for French listeners ([$t(19) = 4.12$, $p = .006$] and for Korean listeners [$t(19) = 2.22$, $p = .039$].

4.3.2.2 /kas/-/xas/

Figure 4-5 shows performance on the Persian voiceless /kas/-/xas/ contrast which is non-phonemic in all three language groups.

Overall (Panel A): The one-way ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 7.37$, $p = .001$]. Tukey tests indicated that Korean listeners performed significantly more poorly than both English and French listeners ($p = .043$ and $p < .001$, respectively). The latter two groups did not differ ($p = .40$).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 5.31$, $p = .01$; partial $\eta^2 = .16$]. The Tukey post-hoc analysis revealed that Korean listeners were significantly less accurate than French listeners ($p = .006$). No other group differences were observed. There was a main effect of Direction [$F(1, 57) = 24.90$, $p < .001$; partial $\eta^2 = .30$] in the expected direction (FS>SF). The interaction between Language Group and Direction failed to reach significance [$F(2, 57) = .49$, $p = .62$]. Simple effects analyses revealed a significant effect of Direction (FS>SF) for English listeners, $t(19) = -4.61$, $p = .002$ and for Korean listeners, $t(19) = -4.04$, $p = .001$ and a marginal trend for French listeners $t(19) = -1.87$, $p = .077$.

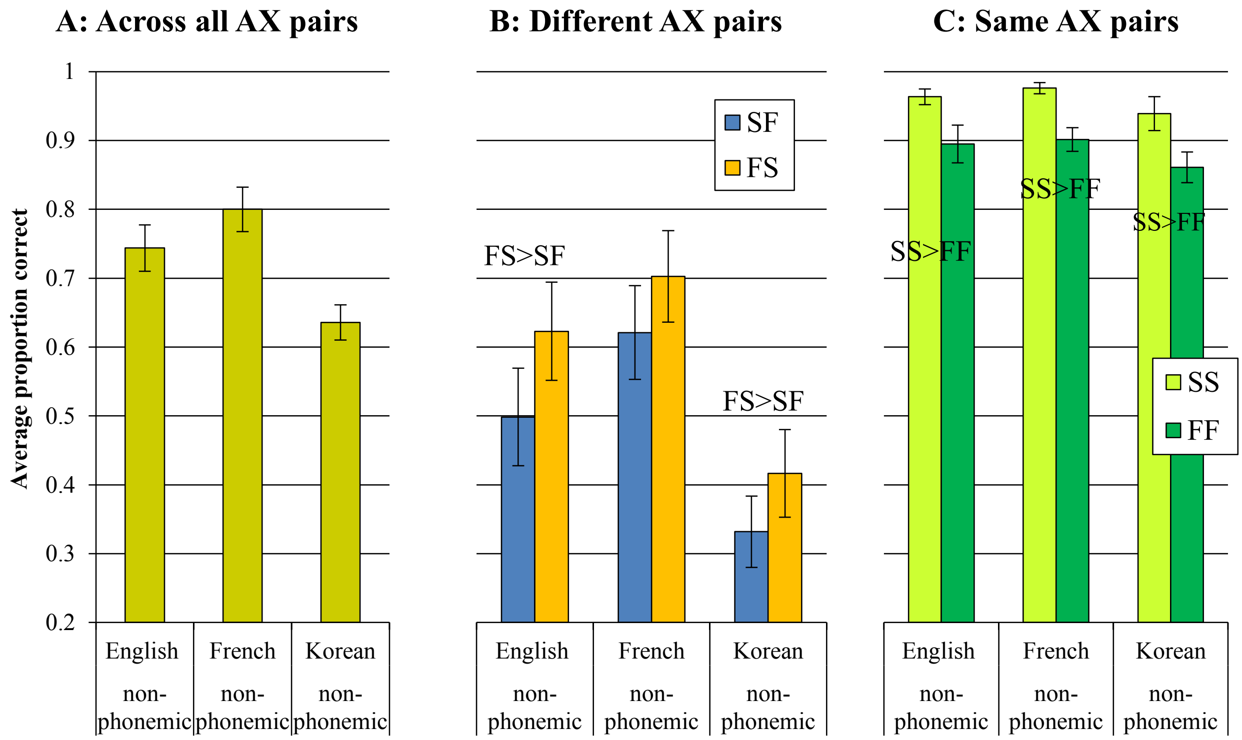

Figure 4-5 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /kas/-/xas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates that listeners perceived the FS pairs significantly better than the SF pairs; SS>FF indicates that listeners perceived the SS pairs significantly better than the FF pairs ($p < .05$).

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Pair Type [$F(1, 57) = 24.37$, $p < .001$; partial $\eta^2 = .30$] in the expected direction (SS>FF). There was no main effect of Language Group [$F(2, 57) = 1.84$, $p = .17$] or Language Group by Pair Type interaction [$F(2, 57) = 0.03$, $p = .97$]. Simple effect tests of Pair Type revealed a significant effect (SS>FF) in all three groups; English [$t(19) = 2.85$, $p = .01$], French [$t(19) = 4.95$, $p = .001$] and Korean listeners [$t(19) = 2.38$, $p = .028$].

4.3.2.3 /gas/-/ɣas/

Figure 4-6 shows performance on the Persian voiced /gas/-/ɣas/ contrast which is not phonemic in any of the three languages.

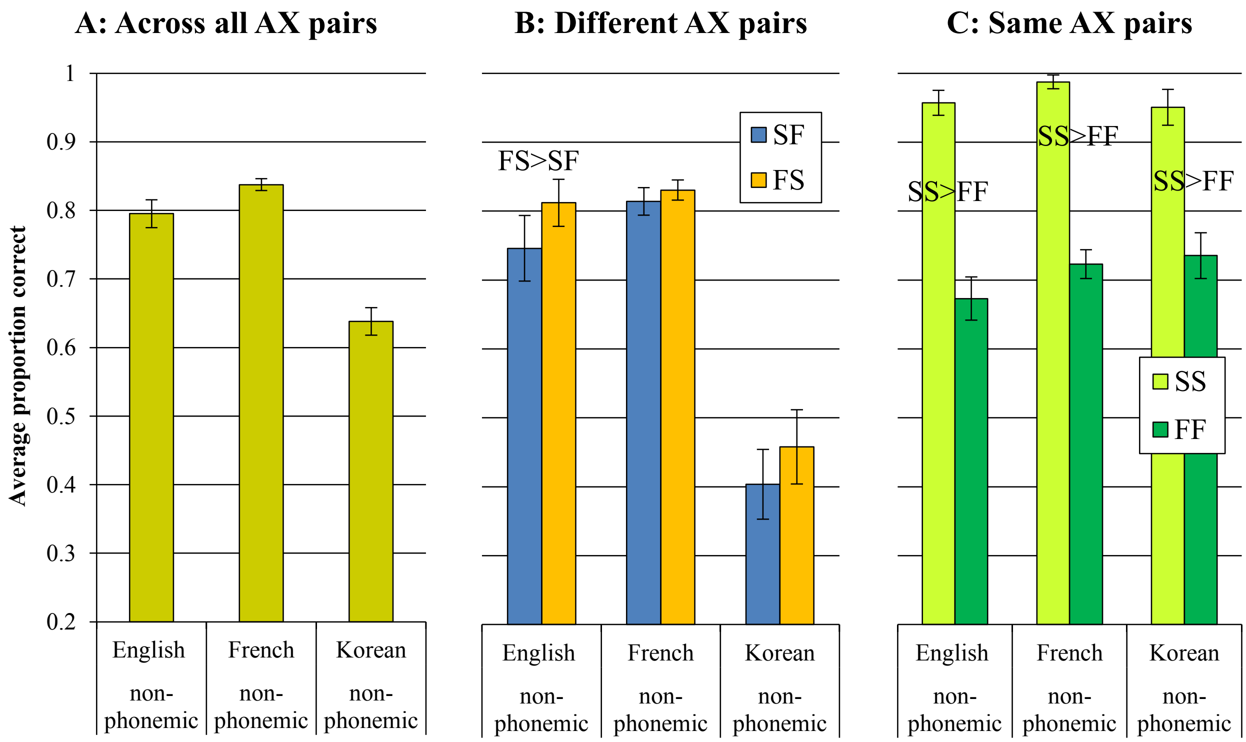

Figure 4-6 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /gas/-/ɣas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs (*p* < .05).

Overall (Panel A): The one-way ANOVA revealed a significant effect of Language Group [$F(2, 57) = 37.91$, $p < .001$]. Tukey tests indicated that Korean listeners performed significantly more poorly than both English and French listeners ($p = .043$ and $p = .001$, respectively). The latter two did not differ ($p = .20$).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 31.96$, $p < .001$; partial $\eta^2 = .53$]. The Tukey's test revealed that Korean listeners performed significantly less accurately than English and French listeners ($p <.001$). The latter two did not differ ($p = .70$). The main effect of Direction reached significance [$F(1, 57) = 10.53$, $p = .002$; partial $\eta^2 = .16$] in the expected direction (FS>SF). The Language Group × Direction interaction was not significant [$F(2, 57) = 1.13$, $p = .33$]. Simple effects analyses revealed a significant effect of Direction (FS>SF) for English listeners [$t(19) = -2.43$, $p = .025$] and a marginal trend for Korean listeners [$t(19) = -1.95$, $p = .066$]. French listeners did not demonstrate a Direction effect [$t(19) = -1.02$, $p = .32$].

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Pair Type [$F(1, 57) = 200.05$, $p < .001$; partial $\eta^2 = .78$] in the expected direction (SS>FF). There was no main effect of Language Group [$F(2, 57) = 1.18$, $p = .32$] or Language Group by Pair Type interaction [$F(2, 57) = 1.30$, $p = .28$]. Simple effect tests of Pair Type revealed a significant effect (SS>FF) in all three groups; English [$t(19) = 10.36$, $p < .001$], French [$t(19) = 10.24$, $p < .001$] and Korean listeners $t(19) = 5.56$, $p < .001$].

4.3.2.4 /pas/-/fas/

Figure 4-7 shows performance on the /pas/-/fas/ contrast which is phonemic in English and French but not in Korean.

Overall (Panel A): The one-way ANOVA revealed a significant effect of Language Group [$F(2, 57) = 18.11$, $p < .001$]. Tukey tests indicated that Korean listeners performed significantly more poorly than both English and French listeners ($p < .001$). The latter two did not differ ($p = .10$).

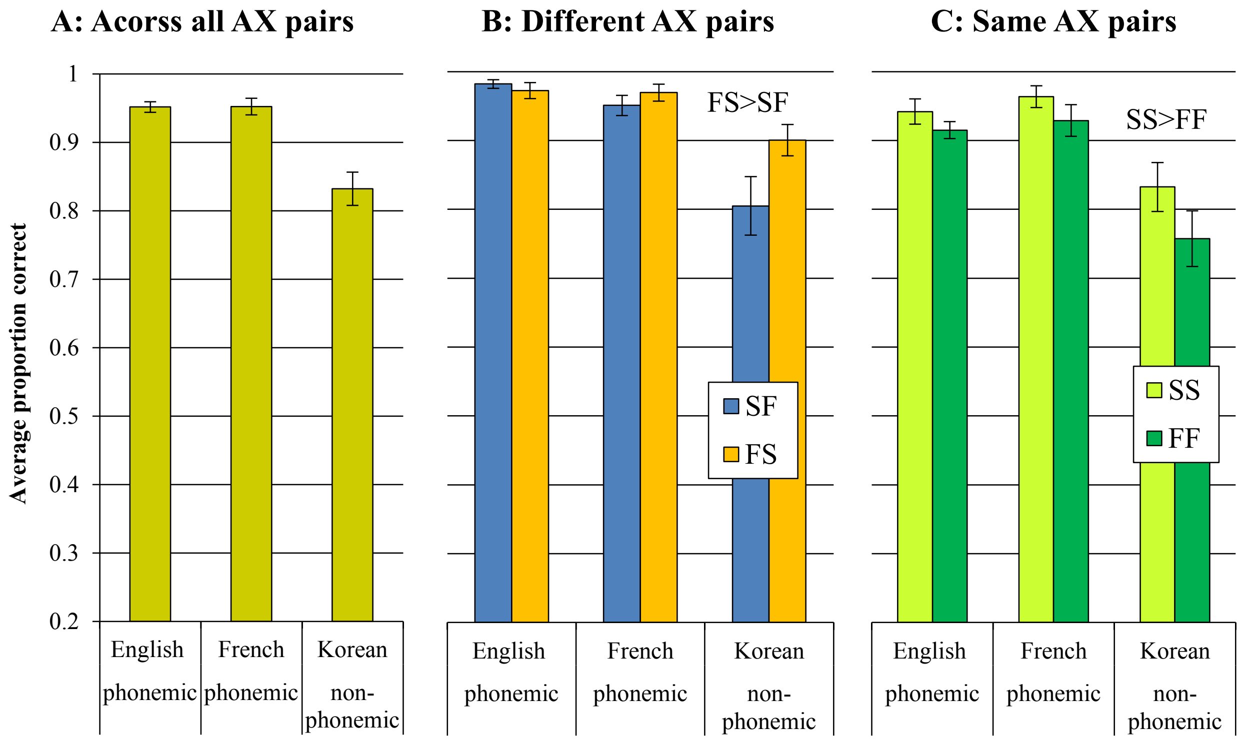

Figure 4-7 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /pas/-/fas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 14.34$, $p < .001$; partial $\eta^2 = .32$]. Tukey tests revealed that Korean listeners performed significantly more poorly than English ($p = .001$) and French listeners ($p = .001$), whereas the latter Groups did not differ ($p = .86$). There was also a significant effect of Direction [$F(1, 57) = 6.80$, $p = .012$; partial $\eta^2 = .11$] in the expected direction (FS>SF) and a Language Group by Direction interaction [$F(2, 57) = 5.53$, $p = .006$; partial $\eta^2 = .16$]. Simple effects of Direction revealed that Korean listeners showed a direction effect (FS>SF) [$t(19) = -2.62$, $p = .017$]. The direction effect was not significant for English or for French listeners ([$t(19) = 0.79$, $p = .44$] and [$t(19) = -1.63$, $p = .12$], respectively).

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 11.97$, $p < .001$; partial $\eta^2 = .30$]. Tukey tests revealed that Korean listeners performed significantly more poorly than English ($p < .001$) and French listeners ($p < .001$). The latter two did not differ ($p = .86$). There was a significant main effect of Pair Type [$F(1, 57) = 12.68$, $p = .001$; partial $\eta^2 = .18$] in the expected direction (SS>FF). There was no interaction between Language Group and Pair Type [$F(2, 57) = 1.36$, $p = .26$]. Simple effects of Pair Type revealed a significant effect (SS>FF) in Korean listeners [$t(19) = 4.15$, $p = .005$] but not in English and French listeners ([$t(19) = 1.27$, $p = .22$] and [$t(19) = 1.62$, $p = .12$], respectively).

4.3.2.5 /bas/-/vas/

Figure 4-8 shows performance on the /bas/-/vas/ contrast which is phonemic in English and French but not in Korean.

Overall (Panel A): The one-way ANOVA revealed a significant effect of Language Group [$F(2, 57) = 30.96$, $p < .001$]. Tukey tests indicated that Korean listeners performed significantly less accurately than both English and French listeners (both ps = <.001). The latter two did not differ ($p = .98$).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 24.22$, $p < .001$; partial $\eta^2 = .46$]. The Tukey's test revealed that Korean listeners performed significantly less accurately than English and French listeners (p <.001). The latter two did not differ ($p = .70$). The main effect of Direction was not significant [$F(1, 57) = 0.55$, $p = .46$]. The Language Group × Direction interaction reached significance [$F(2, 57) = 4.20$, $p = .02$; partial $\eta^2 = .13$]. Simple effects of Direction revealed the expected direction effect (FS>SF) for Korean listeners [$t(19) = -2.26$, p

= .036]. The direction effect was not significant for English or for French listeners ([*t*(19) = 0.85, *p* = .41] and [*t*(19) = 0.92, *p* = .37], respectively).

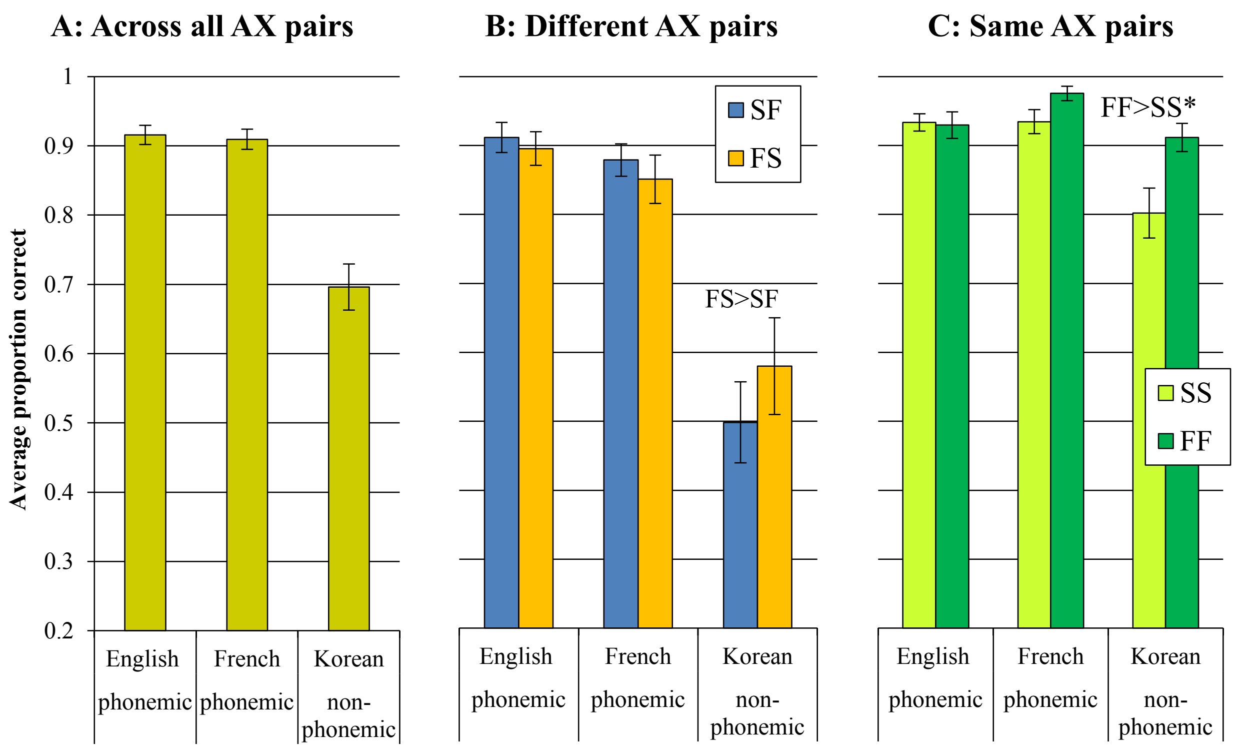

Figure 4-8 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /bas/-/vas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; FF>SS* indicates significantly better performance on the FF pairs than the SS pairs ($p < .05$).

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 9.26$, $p < .001$; partial $\eta^2 = .25$]. Tukeys tests showed that Korean listeners performed significantly more poorly than both English ($p = .008$) and French listeners ($p < .001$). The latter two did not differ ($p = .58$). There was also a significant effect of Pair Type [$F(1, 57) = 11.10$, $p < .001$; partial $\eta^2 = .16$] and a Language Group × Pair Type interaction [$F(2, 57) = 5.15$, $p = .009$; partial $\eta^2 = .15$]. Simple effects of Direction revealed that Korean listeners perceived the FF pairs significantly more accurately than the SS pairs [$t(19) = -4.23$, $p = .004$], which is in the opposite direction (FF>SS) of our prediction. Similarly, French listeners showed a marginal FF>SS trend [$t(19) = -2.01$, $p = .06$]. English listeners did not perform differently according to pair type [$t(19) = 0.22$, $p = .83$].

4.3.2.6 /das/-/zas/

Figure 4-9 shows performance on the /das/-/zas/ contrast which is phonemic in English and French but not phonemic in Korean.

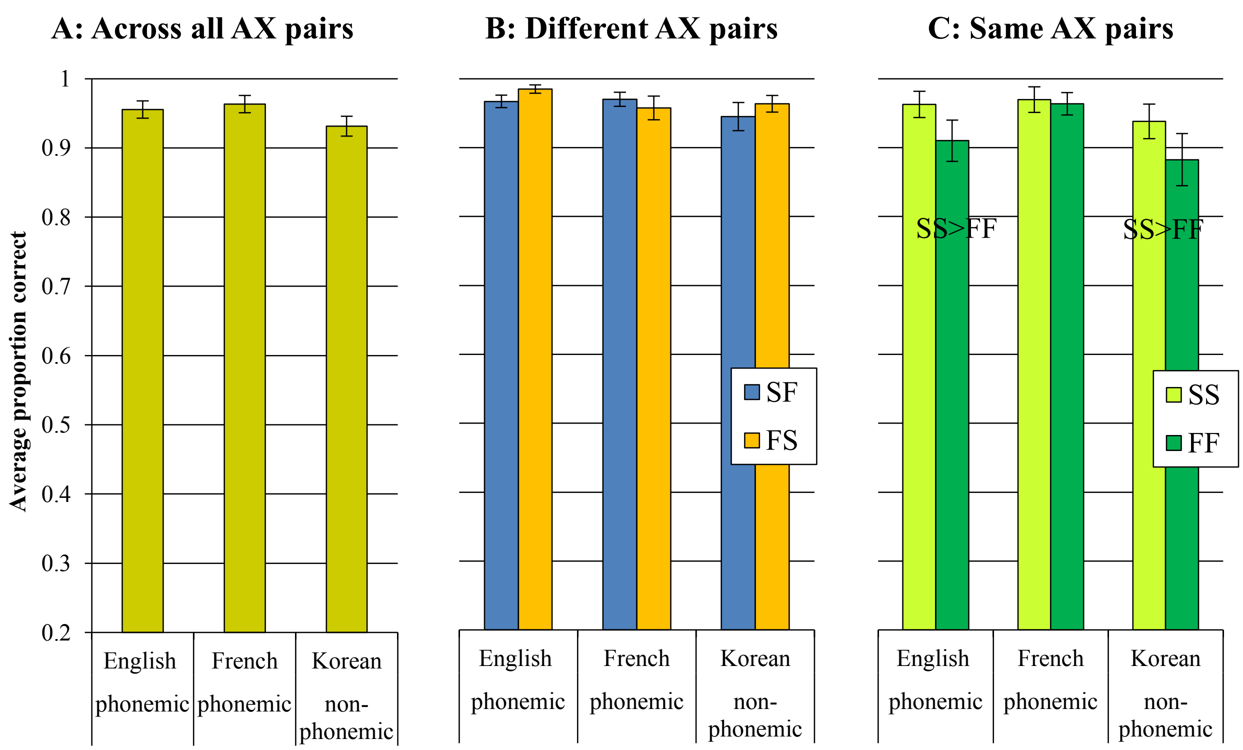

Figure 4-9 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /das/-/zas/ contrast in English, French and Korean listeners (error bars: ±1 SE). SS>FF indicates significantly better performance on the SS pairs than the FF pairs (*p* < .05).

Overall (Panel A): The one-way ANOVA revealed no significant effect of Language Group [$F(2, 57) = 1.53$, $p = .23$].

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed that none of the effects was significant: Language Group [$F(2, 57) = .89$, $p = .42$], Direction [$F(1, 57) = .99$, $p = .32$] and Language Group × Direction [$F(2, 57) = 1.62$, $p = .21$]. Simple effects of Direction revealed a marginal trend (FS>SF) for English listeners [$t(19) = -2.04$, $p = .055$] but no trend for French [$t(19) = 0.83$, $p = .42$] and Korean listeners [$t(19) = -1.11$, $p = .28$].

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Pair Type [$F(1, 57) = 8.64$, $p = .005$; partial $\eta^2 = .13$] in the expected direction (SS>FF). There was no main effect of Language Group [$F(2, 57) = 1.41$, $p = .25$] or Language Group by Pair Type interaction [$F(2, 57) = 1.50$, $p = .23$]. Simple effects tests of Pair Type revealed a significant Pair Type effect (SS>FF) for English and for Korean listeners ([$t(19) = 2.18$, $p = .042$] and [$t(19) = 2.10$, $p = .049$], respectively) but not for French listeners [$t(19) = 0.62$, $p = .54$].

4.3.2.7 /tas/-/θas/

Figure 4-10 shows performance on the /tas/-/θas/ contrast which is phonemic in English but not in French and Korean.

Overall (Panel A): The one-way ANOVA revealed a marginally significant effect of Language Group [$F(2, 57) = 2.87$, $p = .065$]. Tukey tests revealed that Korean listeners performed marginally better than French listeners ($p = .055$). No other group differences were observed.

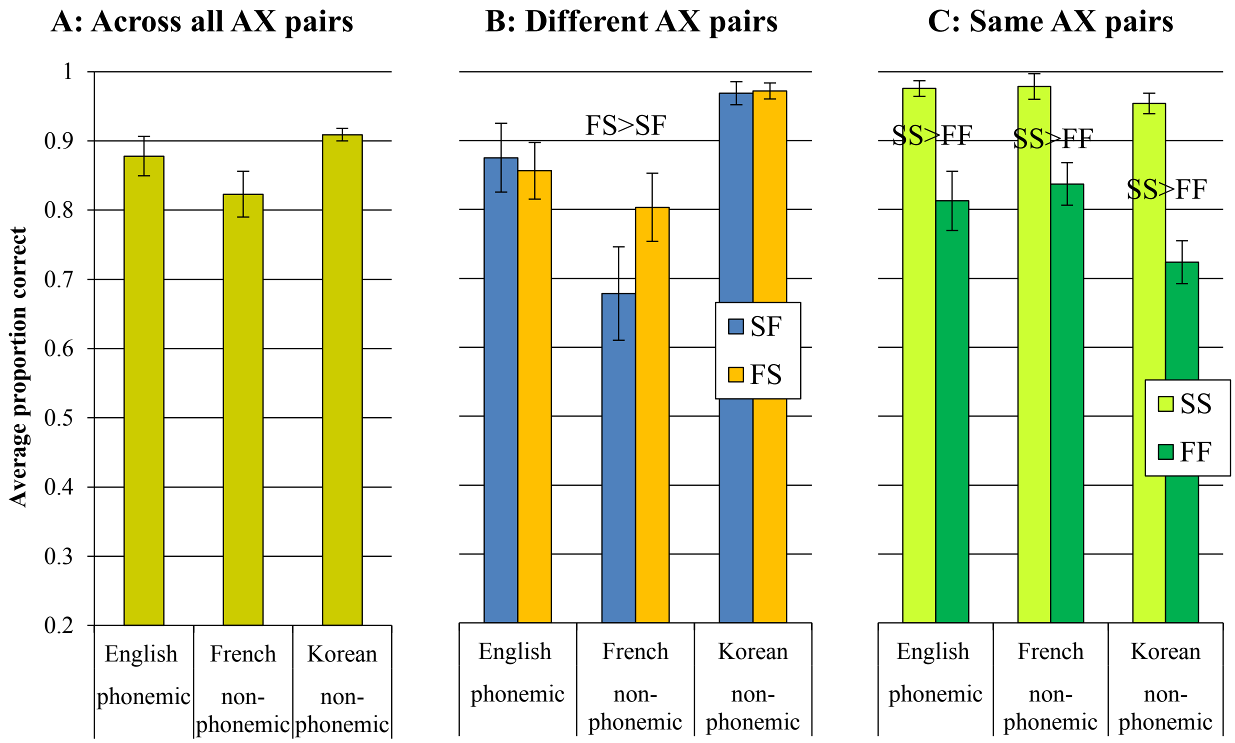

Figure 4-10 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /tas/-/θas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed significant main effects of Language Group [$F(2, 57) = 7.47$, $p = .001$; partial $\eta^2 = .21$]. Tukey tests showed that Korean listeners performed significantly better than French listeners ($p < .001$). No other group differences were observed. There was also a main effect of Direction [$F(1, 57) = 5.69$, $p = .02$; partial $\eta^2 = .09$] and a Language Group × Direction interaction [$F(2, 57) = 8.63$, $p = .001$; partial $\eta^2 = .23$]. Simple effects of Language Group revealed significant group differences on different pairs (SF, [$F(2, 57) = 8.92$, $p < .001$] and FS, [$F(2, 57) = 5.23$, $p = .008$]). Simple effects of Direction revealed a significant direction effect (FS>SF) for French listeners [$t(19) = -$

4.27, p = .004] but not for English and Korean listeners ([t(19) = 0.86, p = .40] and [t(19) = -0.24, p = .81], respectively).

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Pair Type [F(1, 57) = 86.52, p < .001; partial η^2 = .60] in the expected direction (SS>FF). There was no main effect of Language Group [F(2, 57) = 2.85, p = .07] or Language Group by Pair Type interaction [F(2, 57) = 1.96, p = .15]. Simple effects of Pair Type revealed significant Pair Type differences in all the language groups: English, [t(19) = 4.05, p = .001]; French, [t(19) = 5.63, p < .001]; Korean, [t(19) = 7.08, p < .001].

4.3.2.8 /das/-/ðas/

Figure 4-11 plots performance on the /das/-/ðas/ contrast which is phonemic in English but not in French and Korean.

Overall (Panel A): The one-way ANOVA revealed a significant effect of Language Group [F(2, 57) = 24.32, p < .001]. Tukey tests indicated that English listeners performed significantly more accurately than both French and Korean listeners (p <.001). The latter two did not differ (p = .18).

Different Pairs (Panel B): The Language Group × Direction ANOVA revealed a significant main effect of Language Group [F(2, 57) = 26.33, p < .001; partial η^2 = .48]. Tukey's post-hoc analysis revealed that English listeners performed significantly better than both French and Korean listeners (p <.001). The latter two did not differ (p = .79). The main effect of Direction reached significance [F(1, 57) = 25.96, p < .001; partial η^2 = .31] in the predicted direction (FS>SF). The interaction between Language Group and Direction was not significant [F(2, 57) = 0.83, p = .44]. Simple effects of Direction revealed a significant direction effect

(FS>SF) for all language groups: English, [$t(19) = -.30.4$, $p = .007$]; French, [$t(19) = -2.84$, $p = .01$]; Korean, [$t(19) = -4.68$, $p = .002$].

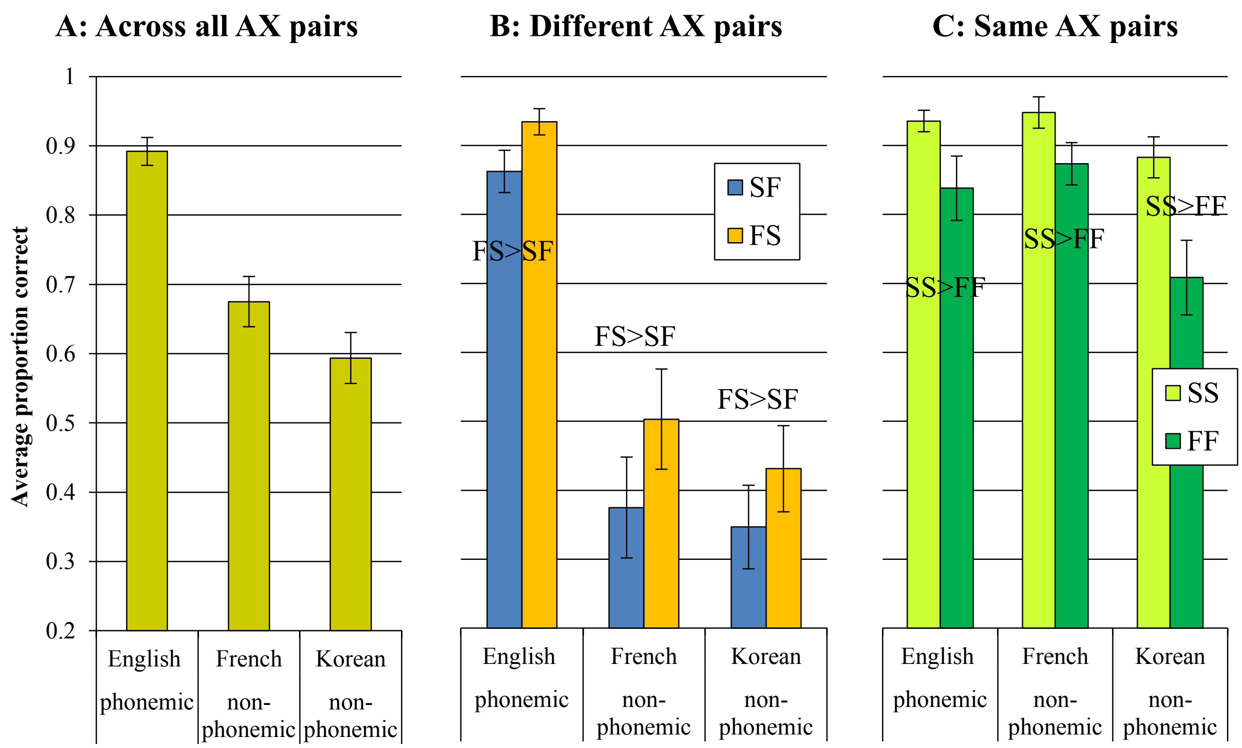

Figure 4-11 Average proportion correct across pairs (A), on different pairs (B) and on same pairs (C) for the /das/-/ðas/ contrast in English, French and Korean listeners (error bars: ±1 SE). FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs ($p < .05$).

Same Pairs (Panel C): The Language Group × Pair Type ANOVA revealed a significant main effect of Language Group [$F(2, 57) = 4.40$, $p = .017$; partial $\eta^2 = .13$]. Tukey's tests indicated that Korean listeners performed significantly less accurately than French listeners ($p = .02$); no other language group differences were significant. There was a significant main effect of Pair Type [$F(1, 57) = 25.11$, $p < .001$; partial $\eta^2 = .31$] in the predicted direction (SS>FF). The Language Group by Pair Type interaction was not significant [$F(2, 57) = 1.60$, $p = .21$]. Simple effects of Pair Type revealed a significant SS>FF effect for all the three groups: English, [$t(19) = 2.33$, $p = .031$]; French [$t(19) = 2.53$, $p = .021$]; Korean, [$t(19) = 4.81$, $p = .001$].

4.4 Summary tables

This section summarizes the discrimination results for each language group in Tables 4-4 through 4.6. In the following tables, % (percentages) indicates discrimination accuracy across all AX pairs; FS>SF indicates significantly better performance on the FS pairs than the SF pairs; SS>FF indicates significantly better performance on the SS pairs than the FF pairs; sig. denotes significant; ns denotes not significant; trend denotes marginally significant; sig.* denotes the reversed effect (FF>SS); trend* denotes the reversed trend (FF>SS).

Table 4-4 English listeners' perception of English and Persian stop-fricative contrasts.

English listeners				
Contrast	Phonemic status	%	FS>SF	SS>FF
/t-s/	Phonemic	96	ns	trend
/p-f/	Phonemic	95	ns	ns
/b-v/	Phonemic	92	ns	ns
/d-z/	Phonemic	96	trend	sig.
/t-θ/	Phonemic	88	ns	sig.
/d-ð/	Phonemic	89	sig.	sig.
/k-x/	Non-phonemic	74	sig.	sig.
/g-ɣ/	Non-phonemic	80	sig.	sig.

Table 4-5 French listeners' perception of English and Persian stop-fricative contrasts.

French listeners				
Contrast	Phonemic status	%	FS>SF	SS>FF
/t-s/	Phonemic	97	ns	sig.
/p-f/	Phonemic	95	ns	ns
/b-v/	Phonemic	91	ns	trend*
/d-z/	Phonemic	96	ns	ns
/t-θ/	Non-phonemic	82	sig.	sig.
/d-ð/	Non-phonemic	67	sig.	sig.
/k-x/	Non-phonemic	80	trend	sig.
/g-ɣ/	Non-phonemic	84	ns	sig.

Table 4-6 Korean listeners' perception of English and Persian stop-fricative contrasts.

Korean listeners				
Contrast	Phonemic status	%	FS>SF	SS>FF
/t-s/	Phonemic	97	ns	sig.
/p-f/	Non-phonemic	83	sig.	sig.
/b-v/	Non-phonemic	70	sig.	sig.*
/d-z/	Non-phonemic	93	ns	sig.
/t-θ/	Non-phonemic	91	ns	sig.
/d-ð/	Non-phonemic	59	sig.	sig.
/k-x/	Non-phonemic	64	sig.	sig.
/g-ɣ/	Non-phonemic	64	trend	sig.

4.5 Summary and discussion

In this section, the results are first reviewed with respect to the phonemic status effect and then are reviewed with respect to perceptual asymmetries. Since the purpose of this dissertation is to test the hypothesis that stops and fricatives are not inherently equally salient, the discussion will focus on findings with respect to performance across different trial types. Specifically, whether listeners discriminate FS pairs better than the SF pairs (FS>SF) and recognize the similarity of SS pairs better than the FF pairs (SS>FF) and whether perceptual asymmetries are modulated by language experience. As pointed out in the introductory section to this chapter, Korean listeners' performance on the English /bas/-/vas/ and Persian /gas/-/ɣas/ contrasts is of particular interest given their reversed identification patterns. Recall that for each of these contrasts Korean adults matched the non-native stop to two native stops categories but they matched the non-fricative to only one native stop while the category goodness ratings were the same. This pattern deviates from the typical assimilation pattern that was observed (in every language group) in which the fricative assimilation was more scattered and assigned lower ratings compared to the stop. For this reason, findings related to Korean perception of these two contrasts will be discussed separately in section 4.5.3.

4.5.1 Phonemic status effect

Looking at overall performance collapsing across all contrasts, the results show, not surprisingly, Korean listeners who had more non-phonemic contrasts performed significantly less accurately overall (77%) than both English (89%) and French listeners (87%) who had less non-phonemic contrasts. Also as expected, the phonemic status effect was straightforward. Listeners were significantly better at discriminating phonemic contrasts than non-phonemic contrasts. English, French and Korean listeners performed similarly on their non-phonemic contrasts. However, language differences emerged for phonemic contrasts, with Korean listeners performing significantly better (97%) than English listeners (92%), although both listener groups performed well. This pattern is likely due to the fact that Korean listeners were tested on one phonemic contrast (i.e., /tas/-/sas/), whereas English listeners were presented six phonemic contrasts and revealed some perceptual variability among contrasts. English listeners discriminated the native /tas/-/sas/, /pas/-/fas/, /bas/-/vas/ and /das/-/zas/ contrasts at accuracy rates of above 90% but their performance was a little less than 90% correct on the native /tas/-/θas/ and /das/-/ðas/ contrasts (88% and 89%, respectively)

Recall that the Persian contrasts, /kas/-/xas/ and /gas/-/ɣas/, were non-phonemic for all three language groups. As expected, Korean listeners who have very few stop-fricative contrasts in their native language were less accurate compared to English and French listeners who have an abundance of stop-fricative contrasts in their native language. These finding shows that the functional load of a feature contrast (in the native language) influences how well adults perceive a novel instance of that feature contrast in another language.

4.5.2 Salience differences play a role in non-phonemic perception

The overall results from Experiment 2 showed that perceptual asymmetries occur for stop-fricative contrasts other than the English /b/-/v/ contrast previously examined by Tsushima *et al*. (2003, 2005) and Zhang *et al*. (2006). Importantly, listeners from different language groups showed parallel asymmetries across a set of contrasts, with only a few exceptions. Recall that for English adults, only the Persian contrasts were non-phonemic. For both Persian /kas/-/xas/ and /gas/-/ɣas/, English listeners discriminated the FS pairs significantly more accurately than the SF pairs. English listeners also recognized the SS pairs significantly better than the FF pairs for both contrasts.

Recall that for French listeners, the English /das/-/ðas/ and /tas/-/θas/ contrasts and the Persian /kas/-/xas/ and /gas/-/ɣas/ contrasts are non-phonemic. French listeners showed the expected perceptual asymmetries (FS>SF and SS>FF) for both English contrasts, /das/-/ðas/ and /tas/-/θas/. For the Persian /kas/-/xas/ and /gas/-/ɣas/ contrasts, French listeners showed the marginal FS>SF trend for /kas/-/xas/ and no asymmetry on different pairs for /gas/-/ɣas/. However, robust asymmetries in the expected direction (SS>FF) were found for same pairs for both Persian contrasts.

These weaker asymmetries in French perception of Persian contrasts appear to reveal some modulation of salience effects by language experience. With respect to /gas/-/ɣas/, it should be noted that Quebec French has a rhotic /r/ which surfaces with various phonetic realizations: apical flap [ɾ], apical trill [r], alveolar approximant [ɹ], velar fricatives [ɣ, x], uvular trill [ʀ], uvular fricatives [ʁ, χ], uvular approximant [ʁ̞] and deleted [Ø] (Milne, 2011; Rose & Wauquier-Gravelines, 2007). The voiced uvular fricative /ʁ/ is now considered as the standard articulation (Little, 2012; Morin, 2013; Walker, 1984). The identification data in Experiment 1 suggest that French listeners treated /ɣ/ as allophones of the rhotic /r/. The allophonic account is supported by the observation that the voiced fricative /ɣ/ which shares manner and voicing with /ʁ/ was remarkably more frequently identified as /r/ than the voiceless fricative /x/ which shares manner but not voicing with /ʁ/ (see Table 3-5). Thus, the null direction effect for the /gas/-/ɣas/ contrast in French listeners is not inconsistent with the hypothesis that salience overriding emerges for non-phonemic contrasts with which listeners have little acoustic-perceptual experience. Rather, the result indicates how acoustic-perceptual experience interacts with the inherent salience differences. Additionally, the asymmetries found for same pairs indicate that the inherently underprivileged salience status of fricatives is readily revealed when two weak acoustic units are matched together, lending support for the notion of a salience hierarchy.

Recall that for Korean listeners, the English /pas/-/fas/, /bas/-/vas/, /das/-/zas/, /tas/-/θas/ and /das/-/ðas/ and the Persian /kas/-/xas/ and /gas/-/ɣas/ contrasts are non-phonemic. (As mentioned earlier, Korean listeners' performance on the /bas/-/vas/ and /gas/-/ɣas/ contrasts is discussed separately below.) For different pairs, Korean listeners showed the expected pattern (FS>SF) for English /pas/-/fas/ and /das/-/ðas/ and also for Persian /kas/-/xas/ but not for English /das/-/zas/ and /tas/-/θas/ where performance was close to ceiling in both directions. A plausible explanation for the latter null results again comes from the identification data (see Table 3-7). For the /das/-/zas/ and /tas/-/θas/ contrasts, the phones within each contrast were consistently mapped to different native language categories. Thus it seems that Korean listeners were able to perceptually isolate one sound from the other for these contrasts given that there is no overlap in their identification patterns.

Nevertheless, for same pairs, Korean listeners showed expected asymmetries (SS>FF) for English /pas/-/fas/ and /das/-/ðas/ and for Persian /kas/-/xas/ as well as English /tas/-/θas/ and /das/-/zas/ where performance on different pairs was not asymmetrical.

Overall, the findings of adult perception of non-phonemic stop-fricative contrasts confirm the hypothesis that salience differences are evident and that they interact to some extent with language experience. The crucial aspect of the results is that listeners regardless of language backgrounds exhibit directional asymmetries in the expected directions. The pattern of these results is reminiscent of that of perception of non-phonemic vowel contrasts in adults who are more accurate at discriminating the central to peripheral vowel changes than the reverse due to salience overriding (Polka & Bohn, 2011). The emergence of the parallel salience overriding effect in stop-fricative perception reinforces the hypothesis that stops with abrupt onsets are more salient than fricatives with gradual onsets. Furthermore, the notion of salience hierarchy is supported further by the observation that listeners across languages are consistently better at matching the sameness of stop phones than fricative phones. This finding suggests that stability contributes to the perceptual robustness of stops relative to fricatives. This stability may also be grounded in the acoustics given that (as reported in Chapter 3), for every contrast, within-category variation in the rise-time cue was significantly lower for stop compared to the contrasting fricative.

4.5.3 Korean listeners' perception of /bas/-/vas/ and /gas/-/ɣas/

Recall that for the English /bas/-/vas/ and Persian /gas/-/ɣas/ contrasts, Korean adults showed uncategorized stop-categorized fricative assimilations. In each case they mapped the non-native stop to two stops in Korean and they mapped the non-native fricative to one Korean phone and assigned the same goodness rating. In all of other non-native assimilations examined in this dissertation (by Koreans as well as French and English), the mapping to native categories was more scattered and/or the goodness ratings were lower for the fricative than the contrasting stop.

Despite this reversed identification pattern, asymmetries were found that follow the predicted patterns with only one exceptions. Specifically, for both contrasts the discrimination of different pairs was asymmetric (FS>SF); this effect was significant for /b-v/ and there was a marginal trend for /gas/-/ɣas/. Discrimination of same pairs was also asymmetric in the expected way (SS>FF) for /gas/-/ɣas/ but it was in the reverse direction (FF>SS) for /b-v/. These data suggest that salience effects for stop-fricative contrasts are preserved across a wide range of assimilation patterns but nevertheless there are ways in which language experience can modulate these effects. It is noteworthy that among UC assimilations, Korean assimilation of English /vas/ is the only instance (in this thesis) where a non-native fricative was mapped to the same native category more than 80% of the time. Moreover, identification differences between /b/ and /v/ were more appreciable than those between /g/ and /ɣ/. The most likely explanation for the reverse asymmetry (FF>SS) for /b-v/ is that Korean listeners were able to perceive /v/ as a highly stable category as evident in the noticeably higher mapping consistency. Here again, however, it should be noted that salience overriding was preserved for different pairs, buttressing the hypothesis of stops' privileged acoustic salience status over fricatives.

4.5.4 Salience differences also play a role in phonemic perception

Under the acoustic-phonetic bias hypothesis, perceptual asymmetries are predicted to be weaker (or absent) for native contrasts. As expected, listeners generally performed symmetrically on native contrasts. However, perceptual asymmetries were found for some phonemic contrasts, typically when performance fell below 100%. For example, asymmetrical patterns on same pairs (SS>FF) were found for /tas/-/sas/ for French and Korean listeners and a trend was noted on /bas/-/vas/ for French listeners. For English listeners, discrimination of different pairs was asymmetrical (FS>SF) for /das/-/ðas/ and there was a marginal effect for /das/-/zas/. For same pairs, English perception was asymmetrical (SS>FF) for /das/-/ðas/, /das/-/zas/ and /tas/-/θas/ and marginal for /tas/-/sas/. The most noteworthy aspect of these results is that asymmetries emerge even for native contrasts when there were only very minor deviations from errorless identification. These findings lend further support to the notion that perception of stop manner of articulation is more salient and stable than fricative manner of articulation.

Overall, the occurrence of perceptual asymmetries in phonemic stop-fricative perception is not readily explained by lack of language experience. The interpretation must take into account the acoustic salience differences of stops and fricatives. The differential salience status is argued to result from the different onset abruptness shown in the acoustic measurements presented in Chapter 3; stops with abrupt rise times are acoustically more salient than fricatives with gradual rise times, resulting in perceptual asymmetries. More importantly, that salience differences affect phonemic stop-fricative perception raises the possibility that salience differences of some stop-fricative contrasts may not be fully modulated by language experience.

4.6 Beyond the FUL model

As discussed in Chapter 2, the FUL model is based on the premise that not all consonant manners are phonologically specified in the mental lexicon and this representational inequality is what underlies perceptual asymmetries. The FUL model does not impose an asymmetry at the acoustic level. However, the results of Experiments 1 and 2 support the view that there are phonetic biases that appear to be grounded in the acoustic signal. The asymmetries observed in this thesis suggest that phonological specification is associated with lower perceptual salience and stability and scattered cross-language assimilations. These findings need to be reconciled with the FUL model. At present the relationship between phonological specification and acoustic –phonetic properties has been addressed in a very limited way that fails to account for the findings reported in this dissertation.

4.7 Beyond the perceptual assimilations and the developmental period

As reviewed in Chapter 3, Tyler *et al*. (2014) applied the PAM principles regarding perceptual assimilations to predict when directional asymmetries will emerge. Tyler *et al*. (2014) proposed that with respect to non-native contrast perception, directional asymmetries are expected to arise for SC, CG and UU assimilation contrasts but not for TC and UC assimilation contrasts. This prediction stems from the assumption that for TC and UC assimilations, performance is quite good (near ceiling) as the perceiver can take advantage of phonological categorization processes, and does not need to tune into phonetic details. For the other assimilation patterns, the perceiver is accessing phonetic level information and asymmetries, which presumably arise from phonetic biases, may then come into play. The PAM view suggests that phonological categorization reduces access to phonetic detail. Let us now turn to the question of whether the results from Experiments 1 and 2 are readily explained by PAM.

Table 4-7 illustrates the results of English, French and Korean listeners' perception of non-phonemic stop-fricative contrasts and the PAM predictions of directional asymmetries according to assimilation types. Shaded rows represent the cases in which the results are inconsistent with PAM. As shown in this Table, the present findings for non-native contrasts do not generally agree with the PAM predictions; directional asymmetries emerged repeatedly for UC and TC assimilations. Moreover, recall that English listeners showed significant perceptual asymmetries for their native /das/-/ðas/ contrast (and a trend /das/-/zas/), which clearly violates PAM predictions. PAM would predict an asymmetry to reverse when the mapping patterns reverses, but this was not the case. Importantly, Korean listeners yielded uncategorized stop-categorized fricative assimilations for the English /bas/-/vas/ and Persian /g/-/ɣ/ contrasts. According to PAM, successful categorization occurs when an L2 phone is very similar to its counterpart L1 phoneme category in terms of articulatory-gestural properties. However, as with other categorized stop-uncategorized fricative contrasts, directional asymmetries emerged in the same directions (FS>SF) despite the reverse in mapping pattern. Overall, PAM does not provide consistent or reasonable explanation for directional asymmetries. Indeed, the pattern of these results would most likely point toward the crucial role of acoustic properties in consonant perception.

Table 4-7 Evaluation of the results of listeners' non-phonemic perception in light of PAM.

Listener group	Contrast	Assimilation type	Directional asymmetry	Predictions by PAM
English	/k/-/x/	CG	Yes	Yes
	/g/-/ɣ/	UC	Yes	No
French	/k/-/x/	UC	trend	No
	/g/-/ɣ/	UC	No	No
	/t/-/θ/	TC	Yes	No
	/d/-/ð/	UC	Yes	No
Korean	/k/-/x/	CG	Yes	Yes
	/g/-/ɣ/	UC	trend	No
	/p/-/f/	UC	Yes	No
	/b/-/v/	UC	Yes	No
	/d/-/z/	TC	No	No
	/t/-/θ/	UC	No	No
	/d/-/ð/	UC	Yes	No

PAM was principally proposed to predict patterns with respect to relative discrimination performance according to assimilation type. For example, PAM predicts excellent discrimination for TC assimilation and very good discrimination for UC assimilation (Best, 1995). However, the present results are rather inconsistent with the PAM prediction about discrimination performance for non-native listeners. This is particularly the case with Korean listeners. Within the Korean listener group, non-phonemic contrasts resulted in UC assimilations for the most part. However, Korean listeners had difficulty discriminating those UC contrasts with two exceptions (/pas/-/fas/ and /tas/-/θas/). Moreover, unlike other native TC contrasts, English listeners' discrimination performance on the /das/-/ðas/ and /tas/-/θas/ contrasts was slightly less than 90%, along with asymmetric identification performance on these contrasts. PAM is principally concerned with naïve adult listeners' performance on non-native speech contrasts via assimilation and their corresponding perceptual variability according to assimilation. However, the results from Experiments 1 and 2 suggest the possibility that native listeners also show perceptual variability for native consonants. As a result, these overall results are difficult to be explained by PAM.

Finally, with respect to the role of salience in speech perception in other conceptual frameworks, as noted in Chapter 3, the NLM-e suggests that acoustic salience plays a temporary role in early speech perception and stipulates that directional asymmetries occur in the initial developmental phase. According to Kuhl *et al*. (2008, p.991), directional asymmetries are observed across languages and ages "at least in infancy". However, Experiment 2 shows that directional asymmetries are found in adult perception of non-native contrasts and even for some native contrasts. Another developmental framework, PRIMIR, shares a similar perspective suggesting that phonetic biases are short-lived. In this respect, Werker and Curtin (2005) notes that as infants establish more robust phoneme categories, the role of salience is reduced in their perception. However, again, phonetic salience affected the perception of native adult listeners who are expected to have robust phoneme categories. As a result, NLM-e and PRIMIR do not seem to offer an adequate explanation for this aspect of speech perception during or beyond the developmental period.

4.8 Conclusion

Taken together, the present experimental results concur with the acoustic-phonetic bias hypothesis that stops with abrupt onset acoustics are perceptually more salient than fricatives with gradient onset properties and thus for non-phonemic contrasts, listeners' discrimination of SF pairs is generally depressed relative to their discrimination of FS pairs due to salience overriding. That listeners generally perform differently according to pair type for non-phonemic contrasts also supports that stop percepts are easier to perceive than fricative percepts. These observations raise the possibility that stops serve as referents in stop-fricative perception. Most importantly, the overall patterns of adult cross-language perception of stop-fricative contrasts conceptually dovetails with the NRV argument and provides empirical support for the applicability of the NRV framework to adult consonant manner perception.

The adult-listener perceptual pattern, specifically the SS>FF pattern, is not reasonably accounted for by the FUL model. The PAM predictions also do not fully agree with the current data. Under the PAM hypothesis, listeners who meet the TC and UC categorization criterion (although there is no standard criterion) for a given contrast will presumably show very successful discrimination of the contrast which in turn precludes a directional asymmetry. A major caveat in this line of reasoning is that it does not acknowledge that not all sounds are equally acoustically and perceptually represented and thus cannot explain the observed perceptual asymmetries for TC and UC contrasts. In addition, the current results show that initial biases go beyond the developmental period, inconsistent with the NLM-e and PRIMIR views.

Importantly, that perceptual asymmetries emerged for some native contrasts suggests that the inherent salience differences may not be fully modulated by years of language experience. This raises an inspiring question of whether infants before babbling show parallel perceptual asymmetries. The evidence of perceptual asymmetries is robust in infant vowel perception, as elaborated in Chapter 3. The following experimental chapter is devoted to exploring whether young infants discriminate a fricative to stop manner change but find it difficult to discriminate the stop to fricative manner shift. Strictly speaking, infants start to be exposed to the ambient language right after birth (or probably before birth). In this regard, the notion of the existence of a default perceptual bias can be further substantiated by examining young infants' phonemic perception. Additionally, the presence of such a bias is directly assessed by examining whether young infants prefer to listen longer to stop stimuli than fricative stimuli. The acoustic-phonetic bias hypothesis can be fully justified by the presence of a stop preference. This will eventually suggest the applicability of the NRV framework to consonant perception.

Chapter 5: A stop preference in infant stop-fricative perception

5.1 Introduction

The adult cross-language discrimination experiment (Chapter 4) revealed that English, French and Korean listeners show evidently better overall discrimination of phonemic contrast(s) compared to non-phonemic contrasts. Aside from the typical phonemic status effect, the results revealed that for non-phonemic contrasts, listeners in all language groups exhibit better discrimination for a fricative to stop change than for a stop to fricative changes (FS>SF). This is consistent with the acoustic-phonetic bias hypothesis that salience differences induce salience overriding. Experiment 2 provides further evidence of the role of salience differences in stop-fricative perception by showing that non-phonemic contrasts, listeners regardless of languages generally are more accurate at perceiving pairs of stop phones than those of fricative phones. The acoustic-phonetic bias hypothesis argues that asymmetries in onset acoustics between stops and fricatives are revealed as perceptual asymmetries in stop-fricative perception. The emergence of perceptual asymmetries even for some phonemic contrasts strengthens the notion that as with vowels, not all consonant manners are equally salient percepts. Importantly, the overall results of Experiments 1 and 2 confirm the NRV argument regarding adult consonant manner perception.

The main purpose of the present chapter is to further assess the validity of the extension of the NRV framework to consonant manner perception in infancy. Motivated by the infant vowel discrimination and preference studies (reviewed in Chapter 2), the validity is tested by first examining whether infants show a directional asymmetry when discriminating a stop-fricative contrast and whether they show a listening preference favoring stops over fricatives.. In a set of experiments (Experiments 3 and 4), English- and French-learning 5-6-month-old infants were presented the stop-fricative /bas/-/vas/ contrast which is phonemic in both English and French; the contrast has been explored in both toddler and adult studies of asymmetric stop-fricative perception. Why *phonemic* stop-fricative perception in *young infants*? The acoustic-phonetic bias hypothesis started from the assumption that stops and fricatives are inherently perceived differently due to different onset acoustic characteristics. The assumption can be neatly verified by ruling out the factor of language experience as a possible source of perceptual asymmetry; specifically, language experience refers to experience with stop manner but lack of experience with fricative manner, given that stops occur more frequently in languages than fricatives. Although the adult discrimination and identification studies (Experiments 1 and 2) provided some support for the assumption, particularly by way of English-speaking listeners' asymmetric performance on some phonemic stop-fricative contrasts, the assumption requires further verification. The emergence of asymmetry in phonemic stop-fricative perception in young infants would probably most convincingly establish the existence of the inherent salience differences between stop and fricative manners of articulation and the role of the salience differences in stop-fricative perception.

In Experiment 3, English- and French-learning 5-6-month-olds are tested in a visual habituation paradigm to examine whether their discrimination performance on the phonemic stop-fricative /bas/-/vas/ contrast differs as a function of stimulus presentation direction. The /b/-/v/ contrast was chosen because the contrast was most studied with respect to directional asymmetries across adult and toddler studies. Under the acoustic-phonetic bias hypothesis, infants are expected to show differential performance according to their habituated consonant; infants who are habituated to the fricative /vas/ will successfully detect the change from the habituated stimulus /vas/ to the new stimulus /bas/ but infants who are habituated to the stop /bas/ will show reduced detectability of the shift from the habituated /bas/ to the new /vas/. The lexical representation hypothesis makes the same prediction about the direction effect FSF<FS) (although it appears somewhat difficult to apply its interpretation to 5-6-month-olds who have a limited receptive vocabulary repertoire). The crucial difference between two hypotheses lies in what drives the direction effect. Specifically, within the acoustic-phonetic bias hypothesis, the stop, with its abrupt increase in onset energy is perceived more saliently than the fricative, with its more gradual rise in onset energy. The salience differences result in a salience overriding effect, whereby discriminability of the /bas/ to /vas/ change is decreased relative to that of the /vas/ to /bas/ change. According to Altvater-MacKensen *et al*.'s (2010, 2014) lexical representation hypothesis, in the /vas/ to /bas/ shift, infants are capable of recognizing the mismatch between the fricative's specified feature and the stop's acoustic feature. However, in the /bas/ to /vas/ shift, infant discrimination performance suffers because the fricative /v/ is considered an acceptable match for the stop /b/ due to its lexically underspecified representational status.

In Experiment 4, the same contrast is presented to 5-6-month-old English- and French-learning infants in a sequential preferential looking procedure to directly measure the presence of a stop bias. The acoustic-phonetic bias hypothesis predicts that infants will prefer to listen longer to the stop /bas/ trials than the fricative /vas/ trials (resulting in a stop preference) because a highly salient acoustic percept would likely impact the infants' perceptual system more effectively than a weakly salient acoustic percept. By showing that one phone is favored over the other by young infants whose ambient consonant inventory exhibits both the stop and fricative consonants, the prediction of the existence of a stop preference can be confirmed that an acoustic entity with abrupt onset acoustics is favored over an acoustic entity with gradual onset acoustics. Altvater-MacKensen *et al*.'s (2010, 2014) lexical representation hypothesis, however, seems difficult (and rather illogical) to make predictions of whether infants show any type of perceptual bias evidence because it does not take into account the role of acoustic attributes and corresponding perceptual biases in stop-fricative perception.

The present set of discrimination and preference experiments are the first to assess the role of salience differences with respect to stimulus presentation direction in consonant perception in infancy. This chapter concludes by showing that the overall findings of the two experiments concur with the acoustic-phonetic bias account that stop and fricative manners are not equally saliently perceived (the inherently privileged salience status of stop manner with abrupt onsets over fricative manner with gradual onsets).

5.2 Experiment 3: Exploring a directional asymmetry in infants

Experiment 3 (a discrimination experiment) examines whether English- and French-learning infants between 5-6 months of age show a directional asymmetry for the phonemic /bas/-/vas/ contrast in the visual habituation paradigm. If stop manner is inherently more salient than fricative manner, the inherent salience differences would affect young infants' discrimination performance. The acoustic-phonetic bias hypothesis predicts that the /vas/ to /bas/ change will be easier to discriminate compared to the /bas/ to /vas/ change in which salience overriding takes place.

5.2.1 Methods

The visual habituation paradigm has been used extensively over the past few decades to examine infants' ability to discriminate speech contrasts (e.g., Best *et al*., 1988; Polka *et al*., 2008; Narayan *et al*., 2010). This may be due to the fact that this paradigm has yielded reliable and interpretable results while encompassing infants over a relatively wide age range (Polka *et al*., 1995) and being executed with relative ease (Werker *et al*., 1998). The crucial premise of the paradigm is that infants will prefer a novel stimulus over a familiar stimulus that has been presented repeatedly; this preference (i.e., a novelty effect) indicates that the infant can differentiate the familiar and novel stimuli (Aslin, 2007, for a recent review).

5.2.1.1 Participants

Forty eight normally-developing full-term 5-6-month-old infants (27 boys; M_{age} = 151 days; range = 130−180 days) were included in this study: 21 monolingual English-learning infants (M_{age} = 150 days; range = 130−170 days) and 27 monolingual French-learning infants (M_{age} = 152 days; range = 130~180 days). All the infants were estimated to have at least 90% exposure to either English or French according to parental reports of language experience. An additional 11 infants were excluded from analysis due to fussiness and crying (n=5), failure to habituate within the maximum of 12 trials (n=1), parental interference (n=1), equipment or program errors (n=3), experimenter error (n=1).

5.2.1.2 Stimulus materials

One native female speaker of American English produced multiple tokens of each target syllable. The recording procedure and the token selection procedure were identical to those used in the adult experiments. Six tokens of each syllable type were selected, showing minimal variation within each manner class. Table 5-1 summarizes the results of the acoustic analysis of the stimuli. As shown with the stimuli employed for the adult experiments (section 2.3.5), the acoustic analysis revealed that the rise times are significantly shorter for the stop /bas/ than for the fricative /vas/ [$F(1, 10) = 625.08$, $p = <.001$]. No other variables were significantly different between /bas/ and /vas/.

Table 5-1 Summary of stimulus acoustics.

	Variable	Token 1	2	3	4	5	6	Mean	SD
Stops	Rise time	10.98	8.95	8.47	12.52	9.57	13.09	10.6	1.92
	F0 onset	189.14	189.84	188.9	194.76	198.05	191.52	192.04	3.66
	F0 mean	192.05	192.87	191.39	193.59	195.63	189.47	192.5	2.09
	F1 onset	811.95	791.26	735.25	763.31	778.86	764.25	774.15	26.35
	F2 onset	1241.72	1201.99	1281.63	1215.88	1198.68	1232.75	1228.78	30.88
	F3 onset	2695.51	2712.93	2662.35	2575.31	2613.88	2603.35	2643.89	54.81
	F1 midpoint	901	882.77	884.57	888.63	916.94	868.28	890.37	16.74
	F2 midpoint	1362.9	1319.48	1367.24	1331.71	1306.97	1377.55	1344.31	28.8
	F3 midpoint	2741.55	2774.06	2741.94	2632.07	2704.88	2633.84	2704.72	59.75
	/a/ duration	195.83	180.45	197.89	191.67	198.41	190.08	192.39	6.73
	/s/ duration	127	133.82	136.1	136.48	133.34	131.74	133.08	3.47
	Intensity	79.19	79.5	78.24	79.05	79.02	77.47	78.75	0.75
Fricatives	Rise time	91.24	89.35	78.95	75.49	76.05	81.34	82.07	6.74
	F0 onset	191.53	193.78	191.01	200.28	192.94	190.82	193.39	3.57
	F0 mean	190.62	190.3	192.12	189.19	192.93	191.45	191.1	1.34
	F1 onset	659.59	775.8	735.45	779.47	763.04	675.85	731.53	52.04
	F2 onset	1287.77	1215.06	1233.42	1362.29	1326.56	1220.57	1274.28	61.19
	F3 onset	2690.39	2652.39	2642.14	2703.77	2494.9	2485.93	2611.59	96.66
	F1 midpoint	909.35	935.37	868.64	922.09	915.06	895.72	907.71	23.23
	F2 midpoint	1328.64	1394.4	1334.62	1336.75	1423.34	1437.36	1375.85	48.66
	F3 midpoint	2705.96	2639.38	2570.5	2646.34	2641.65	2690.03	2648.98	47.42
	/a/ duration	192.7	195.26	197.53	192.37	191.65	193.58	193.85	2.19
	/s/ duration	132.08	135.44	133.85	122.65	134.94	125.14	130.68	5.44
	Intensity	77.84	78.99	78.16	77.49	78.75	77.26	78.08	0.69

5.2.1.3 Set up

The experiment was conducted in a dimly lit sound-attenuated booth. The parent was seated in the pre-positioned chair facing a 21" Sony TV monitor holding the infant on her lap securely. The TV monitor was approximately 150 cm from the infant. The parent wore Sony noise canceling MDR-NC 7 headphones over which vocal music was delivered from a Sony CFD-S38 portable stereo to mask the experimental auditory stimuli. This was done to ensure that the parent did not influence the infant's looking response. The testing booth was surrounded by a thick black curtain to reduce visual distraction to the fullest extent possible such that all the stimulus presentation equipment, except the TV monitor, was draped in the black curtain. Audio TRAK BSI-90 loudspeakers were below the TV monitor; the auditory stimuli were played at 65 dBA. A SONY digital video camera mounted on a movable tripod was located directly below the TV monitor. The lens of the camera protruded through an aperture in the black curtain. The video camera broadcast of the infant's face was outputted to a Sharp video monitor in the adjacent observation room. In this way, the experimenter observing the infant via the video monitor outside the testing booth could code the infant's gaze onset and offset to looks to the TV monitor by pressing designated keys on the laptop computer keyboard. As a protection against experimenter bias, the experimenter was blind to the experimental syllable category presented to the infant. The entire experiment was implemented using the Habit 2000 software program (Cohen *et al*., 2000) from which both auditory and visual stimuli were presented. The software also recorded the infant's looking time online and calculated the total looking time per listening trial.

5.2.1.4 Procedure

The testing session consisted of pretest, habituation, test, and posttest. On the pretest and posttest trials, the infant was presented a music audio file as an auditory stimulus. The trials served as indicators of the infants' overall task engagement. Each trial was initiated when the infant fixated on a flashing 4-centimeter-red disk (i.e., an attention-getter) on the TV monitor. Upon the infant's fixation, the visual attention getter disappeared. Immediately afterwards, a static black-and-white checkerboard pattern appeared on the monitor and an auditory stimulus began to play. The auditory stimulus continued as long as the infant fixated on the checkerboard. Each trial was terminated when the infant looked away from the checkerboard for more than 2 consecutive seconds or when the maximum trial length of 25 seconds was reached. When the infant looked back at the flashing red light, the next trial began; thus, every trial began with the flashing red light. On each trial, the computer program calculated the total amount of time that the infant fixated on the checkerboard throughout the presentation of an auditory stimulus. The infant's fixation time served as an index of his or her listening time. If the infant's looking time failed to reach 1 second on a trial, the program did not count the infant's fixation as an actual look. The computer automatically reset the trial and replayed the sound file.

Table 5-2 shows syllable corpus employed for each trial during the habituation and test trials. Each trial lasted approximately 25 seconds long and contained fourteen tokens separated by an ISI of 1500 ms. As shown in the Table, four tokens per syllable category were used for habituation trials. On each habituation trial, two of the four tokens were presented three times and two of them were repeated four times. The tokens were arranged in a predetermined pseudorandom order in which the same token was never played twice in a row. The syllables presented on old trials were acoustically different but phonetically identical to those presented during the habituation phase; during new trials the syllables belonging to new phonetic category were presented.

Table 5-2 Syllable corpus employed for habituation and test trials.

Phases		Habituation group	
		/bas/-habituated group	/vas/-habituated group
Habituation		$bas_{/1,2,3,4/}$	$vas_{/1,2,3,4/}$
1st set of test trials	new trial	$vas_{/5/}$	$bas_{/5/}$
	old trial	$bas_{/5/}$	$vas_{/5/}$
2nd set of test trials	new trial	$vas_{/6/}$	$bas_{/6/}$
	old trial	$bas_{/6/}$	$vas_{/6/}$

During habituation, the computer calculated the average looking time of the infant over a sliding three-trial window and tracked the maximum average (3-trial) looking time. The habituation criterion was set to a decrease of 65% (in the running 3-trial average) from the maximum 3-trial average (as in Polka *et al*., 2008). When the infant's looking time decreased to the level of habituation criterion or when the infant was presented with a maximum of 12 trials, the habituation phase ended and the test phase began. The test phase consisted of two sets of one new and one old trial. The presentation order of the test trials was fixed (as in Polka *et al*., 2008). The tokens presented were counterbalanced across infants within each habituation condition. Infants were alternately assigned to either /bas/-habituated listening condition or /vas/-habituated listening condition.

5.2.2 Results

5.2.2.1 Task engagement

The first step in the analysis was to assess whether /bas/- and /vas/-habituated infants maintained their interest in the task in a comparable manner. Toward this end, infants' listening time on the posttest trial was compared using an independent samples t-test. There was no significant difference between habituation groups [$t(46) = .89$, p (two-tailed) = .38], suggesting that infants' task engagement was comparable across the two habituation groups.

5.2.2.2 Habituation trials

Since this dissertation focuses on a potential stop bias in infant stop-fricative perception, infants' habituation data were analyzed to assess whether there were any differences in total habituation time according to the habituating consonant. Recall that infant listening time is infant-controlled in this experiment which allows for variability in absolute levels of listening across individual infants. To compare overall listening time during habituation across groups total habituation time for each infant was calculated by summing the listening time across all of the trials presented during the habituation phase of the task.

Figure 5-1 plots infants' listening time according to habituation group. A one-way ANOVA performed for a habituation group difference revealed a marginal effect of habituation group [$F(1, 44) = 3.75$, $p = .059$]. As depicted in Figure 5.1 (left), during the habituation phase, infants in the /bas/-habituated group (M = 82.52 s, SD = 50.42) tended to listen longer than infants in the /vas/-habituated group (M = 60.47 s, SD = 23.86).

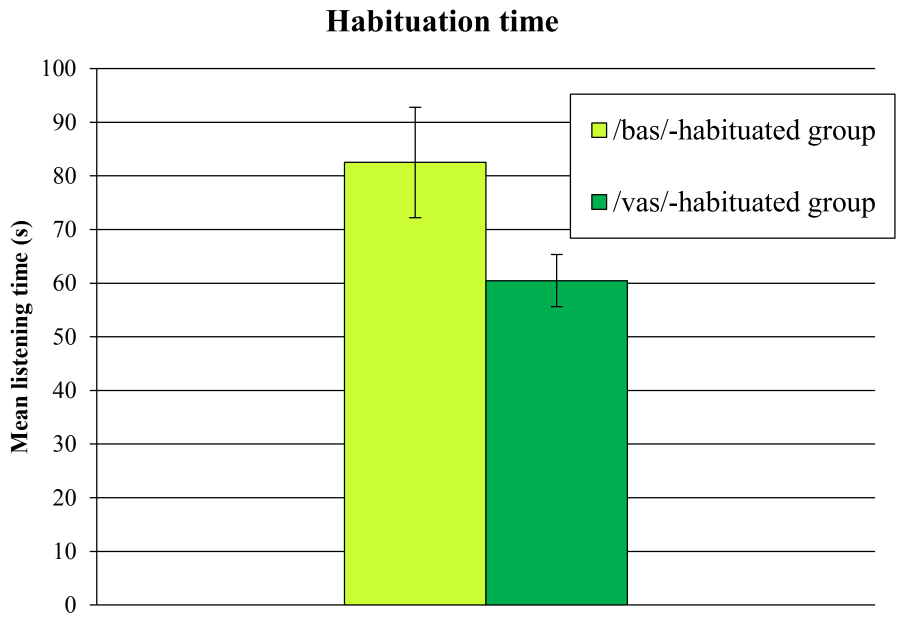

Figure 5-1 Mean listening time as a function of the habituating consonant in each habituation group during the habituation phase (error bars: ±1 SE).

5.2.2.3 Test trials

The next set of analyses addressed whether infants showed differential listening time patterns on test trials as a function of the habituated consonant.

Figure 5-2 plots infants' listening time over the two pairs of test trials as a function of the habituated consonant. Data were submitted to a 2 (Habituation Group: /bas/-habituated vs. /vas/-habituated) × 2 (Trial Type: old trials vs. new trials) mixed ANOVA, with the latter as a within-subjects factor. The main effect of Trial Type failed to reach significance [$F(1, 46) = 1.04$, $p = .31$; $M_{/old\ trials/} = 5.95$ s vs. $M_{/new\ trials/} = 6.49$ s]. The main effect of Habituation Group reached significance [$F(1, 46) = 5.49$, $p = .02$, partial η2 = .11], showing that /bas/-habituated infants ($M = 7.13$ s) listened longer to test trials than /vas/-habituated infants ($M = 5.31$ s). The interaction between Habituation Group and Trial Type was significant [$F(1, 46) = 4.48$, $p = .04$, partial η2 = .09], which was further probed with simple effect tests.

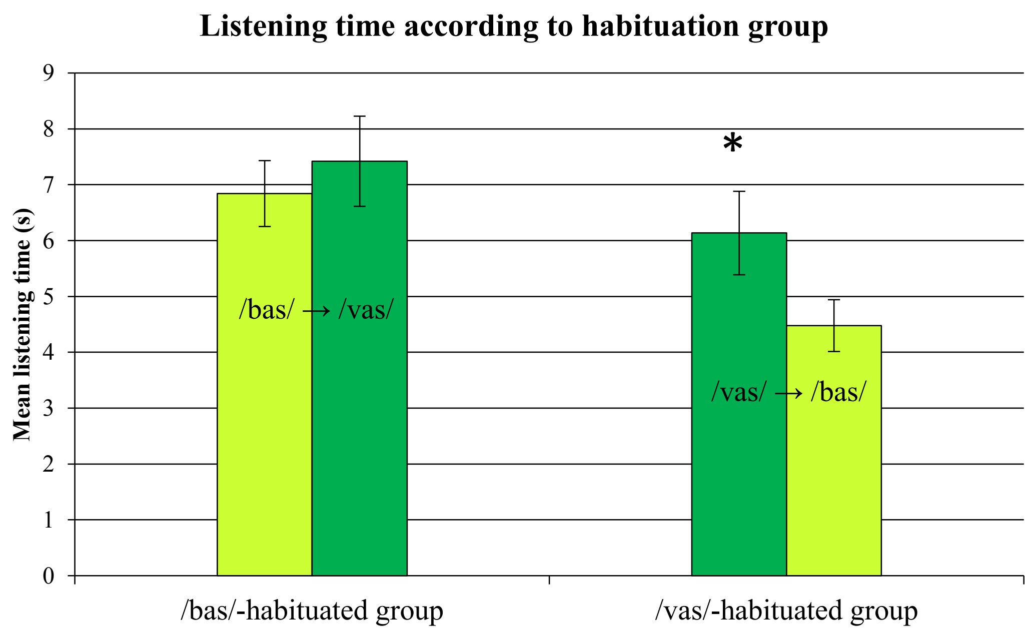

Figure 5-2 Mean listening time to old and new trials for the phonemic /bas/-/vas/ contrast in each habituation group during the test phase (error bars: ±1 SE). Arrow indicates the change from old trials to new trials. Asterisk denotes a significant difference (*p* < .05).

Results of simple effects of Habituation Group using one-way ANOVAs indicated that there was a significant group difference on old trials [$F(46) = 11.23$, $p = .002$] but not on new trials [$F(1, 46) = 0.50$, $p = .48$]. Simple effects of Trial Type using paired two-tailed t-tests revealed that there was a significant novelty preference in /vas/-habituated infants [$t(23) = 2.43$, $p = .023$], with infants listening significantly longer to new /bas/ trials than old /vas/ trials [$M_{/old\ trials/} = 4.48$ s vs. $M_{/new\ trials/} = 6.14$ s]. Eighteen of the 24 infants tested listened longer to new /bas/ trials than old /vas/ trials. There were no significant differences between new and old trials in /bas/-habituated infants [$t(23) = -0.72$, $p = .48$]. Within the /bas/-habituated group, infants listened longer to old /bas/ trials than new /vas/ trials [$M_{/old\ trials/} = 7.42$ s vs. $M_{/new\ trials/} = 6.84$ s].

5.2.3 Discussion

The pattern of discrimination of the phonemic /bas/-/vas/ contrast shown by English- and French-learning infants at 5-6 months of age in Experiment 3 corroborates the predictions of the acoustic-phonetic bias hypothesis that infant discrimination performance would differ as a function of the salience of their habituated consonant, suggesting that a stop manner is inherently more salient than a fricative manner. Some evidence of a salience effect emerged early in the task during the habituation phase; infants in the /bas/-habituation group listened longer before habituating compared to infants in the /vas/-habituation group. During the test phase, infants who were habituated to the fricative /vas/ syllables showed a novelty preference; they listened significantly longer to new trials than habituated trials which shows that they were able to discriminate the fricative to stop manner change. In striking contrast, infants who were habituated to the stop /bas/ syllables did not show a novelty preference; they listened with equal interest to both syllables during the test phase. Moreover, during the test phase, /bas/-habituated infants listened significantly more overall, than /vas/-habituated infants. Specifically, the habituation group difference was evident on old trials but not on new trials. The asymmetry arises because the /bas/-habituated infants continued to perceptually engage with their habituated consonant, while the vas-habituated had lost interest in theirs. Infants' perceptual clinginess toward the stop /bas/ syllables over the fricative /vas/ syllables during both the habituation and test phases further suggests the existence of a perceptual bias toward a stop over a fricative.

Importantly, the expected directional asymmetry (FS>SF) was observed in phonemic stop-fricative perception in young infants who are sensitive to many acoustic-phonetic subtleties (e.g., Kuhl *et al*., 1997; Polka & Werker, 1994; Werker & Tees, 1984), as noted in Chapter 2, is consistent with the acoustic-phonetic bias view that stops and fricatives are not inherently equally salient percepts. The present findings casts doubt on the lexical representation interpretation because it seems difficult to attribute the direction effect shown by young infants in this kind of phonetic discrimination task to lexical representational specificity. Rather, the emergence of a directional asymmetry in the task involving no lexical processing explicitly highlights the role of acoustic properties and their perceptual consequences, suggesting the extension of the NRV framework to consonants. From the NRV perspective, directional asymmetries reflect perceptual high salience biases grounded in acoustic properties. As described in Chapter 2, infants exhibit perceptual biases toward high-salience peripheral vowels over low-salience central vowels. The application of the NRV framework to consonants begs the question of whether such high salience biases are evident in consonant perception as well.

Experiment 4 explores whether a parallel high salience bias emerges in phonemic stop-fricative perception in infancy. As in Experiment 3, English- and French-learning infants between the ages of 5 and 6 months are tested on the phonemic /bas/-/vas/ contrast.

5.3 Experiment 4: Exploring a stop bias in infants

Findings from Experiment 3 reveal young infants' success in detecting the shift from /vas/ to /bas/ and their failure to detect the shift from /bas/ to /vas/, in line with the acoustic-phonetic bias hypothesis. There is another important aspect of the infant discrimination results that deserves to be mentioned. Infants in the /bas/-habituation group tended to listen longer to their habituating consonant compared to infants in the /vas/-habituation group during the habituation phase. Similarly, during the test phase, /bas/-habituated infants listened longer to their already habituated consonant than the new test consonant, while /vas/-habituated infants did not show a relative perceptual proclivity for their habituated consonant. The infant listening patterns in relation to consonant manner salience throughout the experiment raise the possibility of a high-salience stop bias in stop-fricative perception. In fact, the acoustic-phonetic bias hypothesis suggests that a high-salience stop bias underlies stop-fricative perception. The direction-dependent performance difference and the relative proclivity toward a stop consonant observed in Experiment 3 are rather indirect evidence rendering it plausible that young infants might have been more involved in processing the stop /bas/ than the fricative /vas/. The existence of a stop bias in stop-fricative perception needs to be established empirically.

As an attempt at further validation of the acoustic-phonetic bias hypothesis and also as a further step of the extension of the NRV framework to consonants, Experiment 4 examines the existence of a stop bias in stop-fricative perception. The listening preference of 5-6-month-old English- and French-learning infants is assessed in the sequential preferential looking procedure where strings of /bas/ and /vas/ syllables are presented in alternating succession in order to directly measure whether young infants respond selectively to stop and fricative manners of articulation. The acoustic-phonetic bias hypothesis predicts that infants will listen longer to the high-salience stop /bas/-syllable strings than the low-salience fricative /vas/-syllable strings.

5.3.1 Methods

5.3.1.1 Participants

Twenty 5-6-month-old monolingual English-learning infants (n=8) and monolingual French-learning infants (n=12) participated in the study (9 boys; M_{age} = 154 days; range = 133~182 days). An additional 8 infants were tested, but had to be excluded from data analysis due to excessive crying (n=1), general fussiness (n=3), equipment failure (n=2) or experimenter error (n=2). All the infants were exposed to either English or French at least 90% of the time.

5.3.1.2 Stimuli

The stimuli used in Experiment 4 were identical to those employed in Experiment 3. Table 5-3 shows syllable tokens used for each test trial according to listening condition. As illustrated in the Table, six tokens from each syllable type were used for each 25-second trial. Each trial contained fourteen tokens. For each trial, tokens were presented in a predetermined pseudorandom order. For each listening condition, the /bas/ and /vas/ syllables were alternated every other trial. The order of syllable types was counterbalanced across infants. All the tokens were presented at a 1500-ms ISI.

Table 5-3 Syllable tokens used during the test trials according to listening condition.

Listening condition	Test1	Test2	Test3	….	Test11	Test12
/bas/-first	$\text{bas}_{/1,2,3,4,5,6/}$	$\text{vas}_{/1,2,3,4,5,6/}$	$\text{bas}_{/1,2,3,4,5,6/}$	…..	$\text{bas}_{/1,2,3,4,5,6/}$	$\text{vas}_{/1,2,3,4,5,6/}$
/vas/-first	$\text{vas}_{/1,2,3,4,5,6/}$	$\text{bas}_{/1,2,3,4,5,6/}$	$\text{vas}_{/1,2,3,4,5,6/}$	…..	$\text{vas}_{/1,2,3,4,5,6/}$	$\text{bas}_{/1,2,3,4,5,6/}$

5.3.1.3 Procedure

The test setup was the same as in Experiment 3. In this experiment, however, the sequential preferential looking procedure was employed to directly assess whether infants during the first half year of life have an inherent preference for a stop manner over a fricative manner. Research has shown that this relatively simple procedure is successful in ascertaining the presence or absence of preference for one type of stimulus over another in infants (e.g., Cooper & Aslin, 1994; Shi, Werker & Cutler, 2006; Vouloumanos & Gelfand, 2012; Vouloumanos & Werker, 2004).

Half of the infants were assigned to the /bas/-first listening condition, alternating between /bas/ and /vas/ syllable trials and the other half to the /vas/-first listening condition, alternating between /vas/ and /bas/ syllable trials. The experiment session comprised one pre-test trial, twelve test trials and one post-test trial. Once the infant was judged to be fixating on the flashing red light, the auditory stimuli were presented with the checkerboard pattern. For each listening condition, the computer calculated the amount of time the infant fixated on the checkerboard pattern while listening to each of the /bas/ or /vas/ syllable trial. Infants' mean listening time for the total six trials for each syllable type was the primary dependent measure.

5.3.2 Results

5.3.2.1 Task engagement

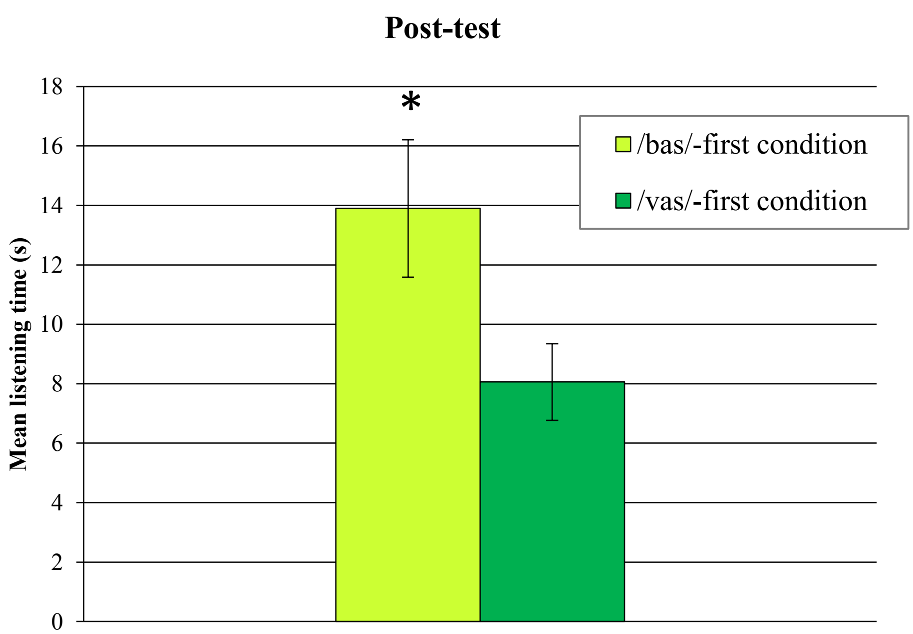

Figure 5-3 Listening time on the posttest trial according to listening condition (error bars: ±1 SE). Asterisk denotes a significant difference (*p* < .05).

Figure 5-3 shows infants' listening time on the posttest trial for each listening condition. As in Experiment 3, as a first step, an independent samples t-test was performed on infants' mean listening time on the posttest trial to assess infants' overall task engagement. The results showed a significant difference between listening conditions [$t(18) = 2.21$, p (two-tailed) = .04][8], with infants in the /bas/-first group listening significantly longer than infants in the /vas/-first group. Interestingly, the /bas/-first infants renewed their interest in the task to the level of the first few test trials during this final trial.

5.3.2.2 Test trials

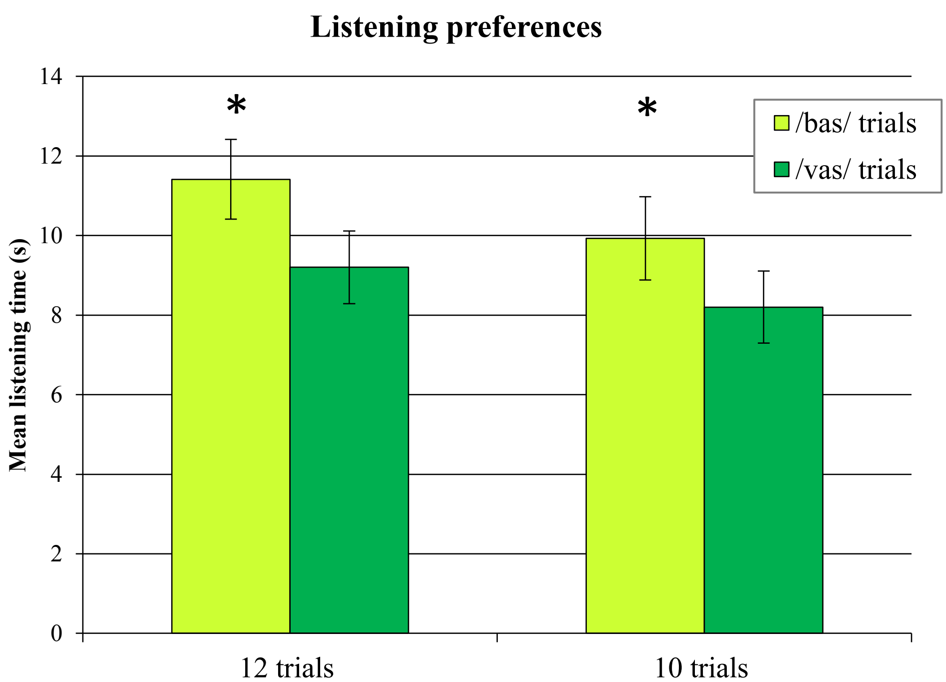

Figure 5-4 Listening patterns for each syllable type across listening conditions during the 12 trials (left) and during the 10 trials (right) (error bars: ±1 SE). Asterisk denotes a significant difference ($p < .05$).

Figure 5-4 plots infants' listening time for each syllable type both with the first two starting trials included (left) and with them removed (right). The analyses were conducted to

[8]However, no such difference was found between listening conditions on the pretest trial [$t(18) = 0.16$, p (two-tailed) = .88].

explore the presence of a stop bias. Given the potential influence of starting syllable types on infants' listening time, all analyses were performed twice, once with the first two trials included and again with them removed. First, infants' listening time for each syllable type during the total 12 trials was entered into a 2 × 2 mixed ANOVA with Listening Condition (/bas/-first vs. /vas/-first) as a between-subjects factor and Syllable Type (/bas/ vs. /vas/) as a within-subjects factor. As expected, there was a significant main effect of Syllable Type [$F(1, 18) = 18.19$, $p = <.001$, partial $\eta^2 = .50$]. Infants across listening conditions preferred to listen significantly longer to /bas/ trials than /vas/ trials [$M_{/bas/} = 11.41$ s vs. $M_{/vas/} = 9.20$ s]. Sixteen out of the 20 infants tested listened longer to /bas/ syllables than /vas/ syllables. The main effect of Listening Condition was not significant [$F(1, 18) = .11$, $p = .75$], showing that infants did not perform differently according to their assigned listening condition group. There was no significant interaction between Listening Condition and Syllable Type [$F(1, 18) = 0.79$, $p = .50$]. Additional analysis using paired t-tests (two-tailed) showed that infants showed a significantly longer listening time to /bas/ syllables than /vas/ syllables in both the /bas/- and /vas/-first conditions: [$M_{/bas/} = 10.88$ s vs. $M_{/vas/} = 9.13$ s; $t(9) = 3.04$, $p = .014$] and [$M_{/bas/} = 11.94$ s vs. $M_{/vas/} = 9.27$ s; $t(9) = 3.10$, $p = .013$], respectively.

Second, with the first two trials removed, infants' listening times for each syllable type over the remaining 10 trials served as the dependent variable. The syllable type difference remained significant [$M_{/bas/} = 9.93$ s vs. $M_{/vas/} = 8.20$ s; $F(1, 18) = 10.76$, $p = .004$, partial $\eta^2 = .37$], showing that infant preference was not affected by the starting syllables. Sixteen out of the 20 infants listened longer to /bas/ syllable trials than /vas/ syllable trials. No other effects were significant: Listening Condition [$F(1, 18) = .28$, $p = .60$] and Listening Condition × Syllable Type [$F(1, 18) = 0.91$, $p = .35$]. Additional analysis also showed that infants tended to listen longer to /bas/ syllables than /vas/ syllables in both the /bas/- and /vas/-first listening conditions: [$M_{/bas/}$ = 9.18 s vs. $M_{/vas/}$ = 7.95 s; $t(9) = 1.63$, $p = .14$] and [$M_{/bas/}$ = 10.69 s vs. $M_{/vas/}$ = 8.45 s; $t(9) = 3.02$, $p = .014$], respectively.

5.3.3 Discussion

The results of the current preference task (Experiment 4) are in accordance with the hypothesis of an acoustic-phonetic bias. As expected, infants at 5-6 months of age show a clear perceptual bias favoring stop /bas/ syllables over fricative /vas/ syllables regardless of their assigned listening condition and also regardless of the exclusion/inclusion of the first two starting trials. These findings provide the first evidence for the presence of a perceptual preference toward a stop over a fricative in consonant manner perception in infancy, providing new insights into the possible role of abruptness of onset acoustics in consonant manner perception in infancy. The presence of a stop preference in phonemic stop-fricative perception naturally raises the question of what essentially makes a stop manner a more preferred acoustic percept. Given the inherent asymmetry in onset acoustics between stop and fricative manners of articulation, the stop preference exhibited by young infants suggests that a more abruptness in onset acoustics has an effective perceptual impact and this grants a stop percept a perceptual privilege over a fricative percept. The privileged stop percept is arguably very perceptible and perceptually unchallenging. In contrast, the acoustically underprivileged fricative percept may be relatively difficult and perceptually challenging. Thus, a highly salient stop may be more likely to ‘entertain’ infants due to a small perceptual burden compared to a weakly salient fricative. That infants perceive a stop-fricative contrast in a manner which is similar to the well-known pattern of perception of peripheral-central vowel contrasts in infants (Polka *et al*., 2005) lends further support for the role of acoustic attributes and their perceptual consequences in infant consonant perception. Such a similarity suggests that extension of the NRV framework to consonant manners may be valid. However, the stop preference cannot be explained by the lexical representation hypothesis as will be addressed further in the following section.

5.4 Lexical specificity?

As reviewed in Chapter 2, Altvater-MacKensen *et al*. (2010, 2014) downplay the role of acoustic and perceptual factors underlying asymmetries in toddlers' stop-fricative word processing. With respect to acoustic factors, Altvater-Mackensen and Fikkert (2010) conclude that their 14-month-olds' successful detection of the 'paap' to 'faap' shift in the phonetic discrimination task allows them to rule out the potential role of acoustic factors in asymmetric discrimination. However, it should be noted that their discrimination task did not test discrimination performance in both directions; no toddlers were habituated to 'faap' and then presented 'paap' (described in section 2.4.2). Given that perceptual performance is often graded (not an all-or-none phenomenon), these findings do not effectively rule out an asymmetry in discrimination performance between 'paap'- and 'faap'-habituation groups. As such, the potential role of acoustic salience in the object labeling task remains unclear.

The infant discrimination results reported here suggest that acoustic-phonetic biases play a potential role in infant stop-fricative perception. This is also supported by the observation that young infants show the clear stop preference in the preference task involving acoustic-phonetic processing. Although Altvater-MacKensen *et al*. (2010, 2014) downplay the role of acoustics in directional asymmetries found in their word recognition studies, they do not dismiss the possibility altogether. Moreover, in their view the specified phone (fricative) is predicted to be more salient than the underspecified phone (stop). Following Altvater-MacKensen *et al*.'s (2010,

2014) reasoning, infants should show a preference toward a lexically specified fricative over a lexically underspecified stop. It is problematic for their view that infants show the opposite pattern in our discrimination and preference tasks.

One further point to note is that under the lexical representation hypothesis, asymmetric discrimination is temporary and will disappear once both stop and fricative units are represented in the mental lexicon. However, Experiment 2 revealed that English adults show asymmetries for some native stop-fricative contrasts. These results are inconsistent with the lexical representation hypothesis which predicts that native adults, should not show asymmetrical patterns because according to Altvater-MacKensen *et al*.'s (2010, 2014), both stop and fricative units are specified in their mental lexicon. The more crucial finding is that English adults' perception performance suffers on some pairs of fricative phones but their perception performance is highly accurate for all stop pairs. In addition, Experiment 1 revealed that adult English listeners do have some difficulty identifying some native fricative phones but do not show such a perceptual variability for stop phones. Such relative identification difficulty only for some fricatives seems rather difficult to be explained by the lexical representation hypothesis in which fricatives are considered more salient than stops.

Although the overall results presented here are difficult to link to lexical development in young infants with little or no lexical knowledge, this does not mean that lexical development does not contribute to stop-fricative perception in infancy but rather suggests that perceptual biases are in place *before* infants begin to recognize words. It seems more likely that these biases may help infants form phonological representations and that lexical-phonological development builds on these initial perceptual biases. In other words, the effect of phonological representation and acoustic-phonetic biases do not have to be mutually exclusive. Both could operate at different points in development or even simultaneously as development advances. Thus, exploring how these facets interact can help us understand how infants develop and use phonetic detail as they develop functional speech processing skills.

5.5 General discussion

Taken together, the results of the current set of infant discrimination and preference studies are consistent with a view that perceptual asymmetry in stop-fricative perception in young infants reflects a perceptual bias grounded in acoustic properties, consistent with the acoustic-phonetic bias hypothesis. As a result, the present findings are interpreted as supporting the hypothesis that the abruptness of energy increase in a stop onset enhances perceptual focalization, which is akin to the perceptual consequence of convergence between consecutive formants of peripheral vowels in vowel perception. That is, stop consonants displaying abrupt onset acoustics are acoustically salient and correspondingly perceptually unchallenging whereas fricative consonants displaying gradual onset acoustics are acoustically weakly salient and resultantly perceptually demanding. Most importantly, the results of Experiments 3 and 4 suggest that the NRV framework may be extended to infant consonant perception by showing parallel salience overriding effect and high salience bias as with infant vowel perception. Meanwhile, the pattern of discrimination of the /bas/-/vas/ contrast in young infants in the phonetic discrimination task highlights problems associated with interpreting asymmetries in early word learning in terms of lexical representational features alone. Crucially, the presence of a stop preference (an acoustic-phonetic bias) challenges the lexical representation hypothesis even though the hypothesis has considerable logical appeal for explaining directional asymmetries in early stop-fricative word processing.

Given the main purpose of this dissertation, now comes the most essential question: whether the NRV framework is applicable to consonant manner perception. The concluding chapter brings together the main findings from the adult and infant experiments to assess whether the NRV framework can be extended to perceptual asymmetries in stop-fricative perception. However, the most important aspect of the present findings is that stops and fricatives are not equally salient acoustic-phonetic units. This finding raises several important questions: 1) what is the developmental profile of the salience effect in stop-fricative perception, specifically are perceptual asymmetries evident in newborns and in infant perception of non-native stop-fricative contrast, and 2) are there the salience effect in the perception of other consonant manner features? and 3) what steps are needed to more directly test the claim that asymmetries in consonant manner are grounded in specific acoustic-phonetic properties? The final chapter is dedicated to discussing these pertinent questions with regards to the perception of consonant manner contrasts. Directions for future research will be highlighted.

Chapter 6: Summary of findings and general discussion

The current dissertation was aimed at exploring the applicability of the NRV framework developed by Polka and Bohn (2011) to stop-fricative perception. Central to the NRV framework is the notion that not all vowels are equally acoustically and perceptually salient. The NRV framework hypothesizes: (1) salience differences result in perceptual asymmetries in which the high-salience peripheral to low-salience central vowel change is more difficult to discriminate than the reverse due to salience overriding, (2) high-salience vowel percepts are preferred over their low-salience counterparts and (3) salience differences interact with language experience but do not entirely disappear. Two sets of perceptual experiments were conducted to assess whether this vowel-specific framework can be extended to consonant perception. The first set of experiments (Experiments 1 and 2) examined adult cross-language perception of 8 stop-fricative contrasts, six were English contrasts (/pas/-/fas/, /bas/-/vas/, /das/-/ðas/, /das/-/zas/, /tas/-/θas/ and /tas/-/sas/) and two were Persian contrasts(/kas/-/xas/ and /gas/-/ɣas/). The second set of experiments (Experiments 3 and 4) examined perception of the one phonemic stop-fricative contrast, /bas/-/vas/, in 5- to 6-month olds who are exposed to either English or French.

6.1 Summary of findings

In Chapter 3, acoustic analyses of initial consonant of the syllables employed for Experiments 1 and 2 revealed that rise time is a highly reliable cue distinguishing stop-fricative manner distinction in syllable initial position, with stops having remarkably shorter and less variable rise times than fricatives. This onset acoustics asymmetry, in light of prior research (Gage *et al*., 1998), motivated the hypothesis that stops with abrupt onsets are acoustically and perceptually more salient than fricatives with gradient onsets.

In Experiment 1 (Chapter 3), English, French and Korean listeners completed a native language assimilation task, i.e. they were presented the syllables from the 8 contrasts and were asked to identify which native categories best match the initial consonant in syllable and rate the goodness of fit. This assessment revealed two findings. First, listeners across languages assimilated non-native stops as a single native category for the most part but they typically assimilated fricatives to two or more native categories. Thus, non-native stop-fricative contrasts resulted in generally categorized stop-uncategorized fricative (UC) assimilations. As expected, listeners showed mostly strong TC assimilations for native contrasts, showing highly stable categorization for both phones. Intriguingly, English listeners showed weaker TC assimilations for the native /t/-/θ/ and /d/-/ð/ contrasts (i.e., stop categorization was highly stable but fricative categorization was somewhat less stable). More intriguingly, Korean listeners showed a reversed UC assimilation for the English /b/-/v/ and Persian /g/-/ɣ/ contrasts in which the fricatives were categorized but the stops were uncategorized. Second, listeners across languages usually preferred to choose stops to identify non-native fricatives over other manners of articulation. Overall, these assimilation patterns point to a privileged salience status of stops over fricatives. It was predicted that this bias would be associated with asymmetries in other perceptual tasks.

Overall, findings of Experiment 2 (Chapter 4) confirmed this prediction. In experiment 2, English, French and Korean adults were tested on perception of the English and Persian stop-fricative contrasts in a same-different AX discrimination task to assess whether, as with vowels, perceptual asymmetries are present in adult consonant perception and if so, whether asymmetries emerge as a function of language experience. When presented non-native contrasts, listeners across languages showed asymmetrical performance on different pairs (FS>SF) and on same pairs (SS>FF) that were predicted on the basis of acoustic salience. These same asymmetries were even found for several native contrasts.

These patterns reveal a robust acoustic phonetic bias that interacts with language experience but doesn't disappear entirely, consistent with the NRV framework. Even when UC assimilation patterns reversed (Korean perception of /gas/-/ɣas/ and /bas/-/vas/) showing that fricative categorization was more stable, asymmetries in discrimination performance still point to a more salient stop percept.

In Experiment 3 (Chapter 5), 5-6 month old English- and French-learning infants were tested on the phonemic /bas/-/vas/ contrast in a visual habituation task to assess the role of salience differences in infancy. As expected, infants discriminated the /bas/ to /vas/ change but not the reverse. Other patterns in experiment 3 suggested that infants have a perceptual bias favoring /bas/ over /vas/. This was confirmed in Experiment 4 (Chapter 5) which directly examined whether 5-6 month old English- and French-learning infants are indeed perceptually biased toward the high salience /bas/ over the low salience /vas/ using a preferential looking task. Consistent with the NRV framework, infants showed a stop bias by listening significantly longer to /bas/ than /vas/. This finding confirms that a stop bias is in place before infants can produce these stops or fricative sounds or recognize words, consistent with the notion that the bias is related to inherent salience differences between stops and fricatives.

6.2 NRV is extended to consonant manners

Taken together, the results of two sets of experiments in this dissertation support the view that acoustic properties play a role in stop-fricative perception, consistent with the NRV notion. With respect to non-native stop-fricative contrasts, adult cross-language identification results demonstrate that stop percepts are more perceptible and more stable than fricatives. Furthermore, the infant findings show that this acoustic-phonetic high salience bias is present in early development. The overall patterns of adult and infant stop-fricative perception indicate that the initial acoustic-phonetic bias interacts with language experience; accordingly this bias is maintained for most non-native contrasts but is absent or weaker for native contrasts. The overall results are consistent with the NRV framework.

As mentioned in Chapter 2, several other models are also concerned with directional asymmetries. To begin with, the FUL model (e.g., Lahiri & Reetz, 2002) stipulates that stops are phonologically underspecified but fricatives are specified and this asymmetry in phonological specification drives asymmetric perception. The FUL model does not acknowledge the possibility that acoustics affects perception. However, adult perceptual experiments demonstrate robust perceptual biases that can be attributed to acoustic-phonetic properties. Similarly, the lexical representation account proposed by Altvater-MacKensen *et al*. (2010, 2014) shares the notion of specification asymmetry. Altvater-MacKensen *et al*. (2010, 2014) interpreted the observed FS>SF effect in toddler listeners' word recognition to be due to the specification differences between stops and fricative. However, the infant discrimination and preference results in this thesis reveal acoustic-phonetic biases in infant stop-fricative perception, before lexical representations are formed. It might be argued that acoustics play a lesser role in toddler perception compared to infant perception. However, given that Experiments 1 and 2 revealed that acoustics affect even native adult perception, it is still difficult to rule out the potential role of acoustic-phonetic biases in toddler perception. Meanwhile, the presence of asymmetries in adult native listeners is also inconsistent with the lexical representation view proposed by Altvater-MacKensen and colleagues (2010, 2014) in which perceptual asymmetries exist in early acquisition but not in adults, for whom both stop and fricative units are assumed to be specified in the mental lexicon. Importantly, the SS>FF pattern robustly observed in this dissertation seems difficult to be explained under the specification asymmetry perspectives in which fricatives appear to be more perceptible than stops.

PAM also has some difficulties explaining the current adult perception results. According to Tyler *et al*. (2014), PAM proposes that directional asymmetries are expected for SC, CG and UU assimilations requiring phonetic processing but not for TC and UC assimilations in which listeners benefit from phonological processing. However, the current results showed directional asymmetries repeatedly for UC assimilations and in native perception as well. The PAM perspective, which emphasizes the role of phonological knowledge in cross-language phonetic perception, cannot be dismissed. However, the present findings suggest that universal acoustic-phonetic biases also shape perception across the life span, suggesting that a model that integrates PAM and NRV principle may be needed.

Finally, the NLM-e model suggests that directional asymmetries occur in early development during which acoustic salience affects perception. Similarly, the PRIMIR stipulates that phonetic biases exist in early acquisition but these are considered transient effects that can be over-ridden by language experience in the first year of life. Thus, neither by the NLM-e and PRIMIR models provides an adequate account for the robust biases displayed by both infants and adults in this thesis.

In summary, the results of the two sets of experiments are best explained within a NRV-based framework. The overall results shows that stops with abrupt onset acoustics are more salient and preferred than fricatives with gradient onset acoustics and such salience differences are revealed as perceptual asymmetries. Most importantly, these results suggest that stops can serve as referent consonants in stop-fricative perception. At present, other conceptual views fail to reconcile these findings in a coherent way.

6.3 Directions for future research

This dissertation has advanced our understanding of the role of acoustic properties and their corresponding perceptual salience in adult and infant stop-fricative contrast perception, particularly in relation to perceptual patterns as a function of stimulus presentation direction. Research in developmental changes in perception of vowel contrasts shows that the salience effect consistently affects infant perception of non-phonemic vowel contrasts during the first year of life but the salience effect is modulated by language experience toward the end of the first year (e.g., Polka & Bohn, 2011; Pons *et al*., 2012). These research findings motivate the question of whether, as with vowels, the salience effect in stop-fricative contrast perception in infancy provides a similar developmental profile with respect to phonemic status. Moreover, the exploration of whether perceptual asymmetries are evident in newborns would provide insight into the nature of perceptual asymmetries.

There are two important limitations of this research. First, the interpretation of the infant results is limited since infants were tested with the /b/-/v/ contrast which is phonemic in English and French. Thus, we cannot determine whether the present asymmetries reveal a default, universal pattern or one that is shaped by language experience already at this age. Second in both English and French stops occur more frequently than fricatives, thus distributional patterns in the input may be critical in formation of this bias or they may reinforce or maintain a default asymmetrical pattern. To make a claim that the bias measured here is acoustically driven will require more direct evidence the bias depends on differences in the rise time cue and does so independent of frequency effects.

One insightful direction for future studies would be to test newborns. If newborns favor a stop over a fricative, the interpretation should acknowledge the role of acoustic properties in stop-fricative perception. Another related research would be to explore infant perception of non-native stop-fricative contrasts. A stop-fricative contrast that merits examination in future experiments is the English /bas/-/vas/ contrast in that the present study reveals that Korean-speaking adults show the expected FS>SF effect for different pairs but the reversed FF>SS effect for same pairs. An interesting question arises: whether the salience effect affects Korean infant perception in the same way it does Korean adults. If so, Korean-learning infants are expected to exhibit the expected direction effect (FS>SF) in a discrimination task but a fricative bias rather than a stop bias in a preference task.

Another insightful direction for future studies would be to explore whether salience effects are present in the perception of other consonant manner features. I would like to approach this topic by exploring infant perception of homorganic affricate-fricative contrasts such as /ʤ/-/ʒ/. This will be informative because exploration is motivated by the fact that affricates are less common than fricatives in languages but affricates have shorter rise times than fricatives. The investigation of whether infants successfully discriminate the fricative to affricate change but not the other way around will decisively uncover the role of onset acoustics in the perception of consonant manner contrasts, combined with the role of lexical-phonological specifications. In this way, I can also provide more direct evidence for the claim that asymmetries in consonant manner perception are grounded specific acoustic-phonetic properties. Most importantly, the presence of such perceptual asymmetries will serve to strongly reinforce the NRV hypothesis that perceptual asymmetries are induced by a perceptual high salience bias related to acoustics.

References

Altvater-Mackensen, N., & Fikkert, P. (2010). The acquisition of the stop-fricative contrast in perception and production, *Lingua, 120***,** 1898–1909.

Altvater-Mackensen, N., van der Feest, S. V., & Fikkert, P. (2014). Asymmetries in Early Word Recognition: The Case of Stops and Fricatives. *Language Learning and Development*, *10*(2), 149-178.

Antoniou, M., Best, C. T., & Tyler, M. D. (2013). Focusing the lens of language experience: Perception of Ma'di stops by Greek and English bilinguals and monolinguals. *The Journal of the Acoustical Society of America*, *133*(4), 2397-2411.

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*, 48–53.

Aslin, R.N., Pisoni, D. B., Hennessy, B.L., & Perey, A. V. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development*, *52*, 1135–1145.

Beck, J. (1974). Relation between similarity grouping and peripheral discriminability. *Journal of Experimental Psychology, 102*(6), 1145.

Best, C. T. (1993). Emergence of language-specific constraints in perception of nonnative speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, pp. 289–304. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman and H. C. Nusbaum (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken*

Words, pp. 167–224. Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 171–204. Baltimore, MD: York Press.

Best, C. T., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and speech*, *46*(2-3), 183-216.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human perception and performance*, *14*(3), 345.

Bijankhan, M., & Nourbakhsh, M. (2009). Voice onset time in Persian initial and intervocalic stop production. *Journal of the International Phonetic Association*, *39*(03), 335-364.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computers [Computer program]. Version 5.1.44. Cho, T., Jun, S., & Ladefoged, P.

Bosch, L. & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech* 46, 217-243.

Bosch, L. & Sebastián-Gallés, N. (2005). Developmental changes in the discrimination of vowel contrasts in bilingual infants. Paper presented at the *4th International Symposium on Bilingualism*, Arizona State University, USA

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in*

Second Language Acquisition, *33*(03), 433-461.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience, 1*(5), 351-353.

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155-190.

Cho, T., Jun, S., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*, 193–228.

Cho, T., & McQueen, J. M. (2006). Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants. *The Journal of the Acoustical Society of America*, *119*(5), 3085-3096.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2000). *Habit 2000: A new program for testing infant perception and cognition* (Version 2.2.5c). [Computer software]. Austin, the University of Texas.

Cooper, R. P., & Aslin, R. N. (1994). Developmental Differences in Infant Attention to the Spectral Properties of Infant-directed Speech. *Child Development*, *65*(6), 1663-1677.

Cornell, S. A., Lahiri, A., & Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 757.

Cowan, N., & Morse, P. A. (1986). The use of auditory and phonetic memory in vowel discrimination. *The Journal of the Acoustical Society of America, 79*(2), 500-507.

Crowder, R. G. (1982). Decay of auditory memory in vowel discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 153-162.

Cusack, R., & Carlyon, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 713.

Delgutte, B., & Kiang, N. Y. (1984). Speech coding in the auditory nerve: III. Voiceless fricative consonants. *The Journal of the Acoustical Society of America*, *75*(3), 887-896.

Easwar, V., Glista, D., Purcell, D. W., & Scollie, S. D. (2012). The effect of stimulus choice on cortical auditory evoked potentials (CAEP): consideration of speech segment positioning within naturally produced speech. *International journal of audiology*, *51*(12), 926-931.

Flemming, E. (1995). *Auditory representations in phonology.* Unpublished doctoral dissertation, UCLA.

Friedrich, C. K., Eulitz, C., & Lahiri, A. (2006). Not every pseudoword disrupts word recognition: an ERP study. *Behavioral and Brain Functions*, 2(10).

Friedrich, C. K., Lahiri, A., & Eulitz, C. (2008). Neurophysiological evidence for underspecified lexical representations: asymmetries with word initial variations. *Journal of Experimental Psychology*: *Human Perception and Performance*, *34*(6), 1545.

Gage, N., Poeppel, D., Roberts, T. P., & Hickok, G. (1998). Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain research*, *814*(1), 236-239.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *The Journal of the Acoustical Society of America*, *110*(1), 489-503.

Kawasaki, H. (1982). *An acoustical basis for universal constraints on sound sequences*. Doctoral dissertation, University of California, Berkeley.

Kodera, K., Hink, R. F., Yamada, O., & Suzuki, J. I. (1979). Effects of rise time on simultaneously recorded auditory-evoked potentials from the early, middle and late ranges. *International Journal of Audiology*, *18*(5), 395-402.

Kuhl, P. K. (1992). Infants' perception and representation of speech: Development of a new theory. In J. J. Ohala, T. M. Nearey, B. L. Derwing, M. M. Hodge, & G. E. Wiebe (Eds.), *Proceedings of the International Conference on Spoken Language Processing* (pp. 449-456).

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, *4*, 812-822.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684-686.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B*, *Biological Sciences*, *363*, 979–1000.

Kuhl, P., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience, 31*, 511–534.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, 13–21.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606-608.

Ladefoged, P. (1993). *A Course in Phonetics (Third Edition)*. New York: Harcourt Brace Jovanovich.

Ladefoged, P., & Maddieson, I. (1998). The sounds of the world's languages. *Language*, *74*(2), 374-376.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. *Laboratory phonology*, *7*, 637–676.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics, 38*(1), 44-59.

Laeufer, C. (1996). The acquisition of a complex phonological contrast: Voice timing patterns of English initial stops by native French speakers. *Phonetica*, 53, 86–110.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422.

Little, S. (2012). A Sociophonetic Study of the Metropolitan French [R]: Linguistic Factors Determining Rhotic Variation.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Milne, P. M. (2012). The Effects of Syllable Position on Allophonic Variation in Québec French/R. *University of Pennsylvania Working Papers in Linguistics*, *18*(2), 9.

Morin, Y. C. From apical [r] to uvular [ʀ]: what the apico-dorsal r in Montreal French reveals about abrupt sound changes. In Studies in phonetics, phonology and sound change in Romance (ed.), *Fernando Sánchez Miret and Daniel Recasens*, 65–93. München: Lincom Europa.

Mounts, J. R., & Tomaselli, R. G. (2005). Competition for representation is mediated by relative attentional salience. *Acta Psychologica, 118*(3), 261-275.

Narayan, C. R. (2008). The acoustic–perceptual salience of nasal place contrasts. *Journal of Phonetics*, *36*(1), 191-217.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, *13*, 407-420

Onishi, S., & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked V potentials. *The Journal of the Acoustical Society of America*, *44*(2), 582-591.

Park, H., & de Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics*, *36*(4), 704-723.

Patil, V., & Rao, P. (2008). Acoustic cues to manner of articulation of obstruents in Marathi. *Proceedings of Frontiers of research on Speech and Music (FRSM)*. Kolkata, India.

Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics, 13*, 253–260.

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*, 577-592.

Polka, L., & Bohn, O.,-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*, 467-478.

Polka, L., Bohn, O.-S., & Molnar, M. (2005). Natural referent vowels guide the development of vowel perception. *Journal of the Acoustical Society of America, 117*, 2398.

Polka, L., Jusczyk, P. W., & Rvachew, S. (1995). Methods for studying speech perception in infants and children. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 49–89). Timonium, MD: York Press.

Polka, L., Rvachew, S., & Molnar, M. (2008). The effect of a nonspeech distractor on infant speech perception. *Cognition*, *13*, 421–439.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 202,* 321–435.

Pons, F., Albareda-Castellot, B., & Sebastián-Gallés, N. (2012). The interplay between input and initial biases: Asymmetries in vowel perception during the first year of life. *Child development*, *83*(3), 965-976.

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences, 106*(26), 10598-10602.

Repp, B. H., & Crowder, R. G. (1990). Stimulus order effects in vowel discrimination. *The Journal of the Acoustical Society of America, 88*(5), 2080-2090.

Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D. (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 621–637.

Roberts, T. P., Flagg, E. J., & Gage, N. M. (2004). Vowel categorization induces departure of M100 latency from acoustic prediction. *Neuroreport*, *15*(10), 1679-1682.

Rose, Y., & Wauquier-Gravelines, S. (2007). French speech acquisition. *The international* *guide to speech acquisition*, 364-384.

Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *The Journal of the Acoustical Society of America*, *99*(5), 3201-3211.

Schwartz, J. L., Abry, C., Boë, L. J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. *Speech Communication*, *45*, 425-434.

Sebastián-Gallés, N. & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Developmental Science* 12, 874-887.

Shafiro, V., Levy, E. S., Khamis-Dakwar, R., & Kharkhurin, A. (2013). Perceptual Confusions of American-English Vowels and Consonants by Native Arabic Bilinguals. *Language and speech*, *56*(2), 145-161.

Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V., & Huotilainen, M. (2004). Orderly cortical representation of vowel categories presented by multiple exemplars. *Cognitive Brain Research*, *21*(3), 342-350.

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*(2), 187-198.

Shinn, P., & Blumstein, S. E. (1984). On the role of the amplitude envelope for the perception of [b] and [w]. *The Journal of the Acoustical Society of America*, *75*(4), 1243-1252.

Shinn, P. C. (1985). *A cross-language investigation of the stop, affricate and fricative manners of articulation*. Unpublished doctoral dissertation, Brown University, Providence.

Stevens, K. N., & Keyser, S. J. (1989). Primary Features and their Enhancement in Consonants. *Language*, *65*, 81-106.

Smith, R. L. (1979). Adaptation, saturation, and physiological masking in single auditory-nerve fibers. *The Journal of the Acoustical Society of America*, *65*(1), 166-178.

Tagliaferri, B., Turner, C., & James, T. (2010). Paradigm (Version 1.0.2.479) [Computer software]. Lawrence, Kansas: Perception Research Systems.

Thomson, J. M., Goswami, U., & Baldeweg, T. (2009). The ERP signature of sound rise time changes. *Brain research*, *1254*, 74-83.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review, 95*(1), 15.

Tsushima, T. (1999). *Developmental changes in nonnative speech segment discrimination.* Unpublished doctoral dissertation, Boston University, Boston, MA.

Tsushima, T. (2007). Asymmetries in perception of an American English /r-l/ by adult Japanese learners of English. *Journal of the Japan Society for Speech Sciences, 8*, 45-62.

Tsushima, T. (2010). A preliminary study on stimulus order effects in discrimination of the English /b/-/v/ contrast during categorical AX discrimination training. *The Journal of Communication Studies, 31*, 153-170.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. *Acoustical Science and Technology***,** *24*(**6**), 410-412.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2005). Stimulus order effects in discrimination of a nonnative consonant contrast, English /b-v/, by Japanese listeners in the AX discrimination procedure. *Paper presented at the First Acoustical Society of America Workshop on L2 Speech Learning*, Vancouver, Canada.

Tyler, M. D., NIL, M., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4-21.

Walker, D. C. (1984). *The pronunciation of Canadian French*. Ottawa, ON: University of Ottawa Press.

Weigelt, L. F., Sadoff, S. J., & Miller, J. D. (1990). Plosive/fricative distinction: The voiceless case. *The Journal of the Acoustical Society of America*, *87*(6), 2729-2737.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech. *Language Learning and Development, 1*, 197–234.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental psychology*, *24*(5), 672.

Werker, J. F., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psycholinguistics, 37,* 35-44.

Werker, J. F., Shi, R., Desjardins, R., Pegg, J. E., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. *Perceptual development: Visual, auditory, and speech perception in infancy*, 389-420.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kircher,

& D. Steriade (Eds.), *Phonetically based phonology* (pp. 34–57). Cambridge: Cambridge University Press.

Zhang, Y., Imada, T., Kawakatsu, M., & Kuhl, P.K., (2006). Neural basis of perceptual asymmetry for the stimulus order effect: A cross-language MEG study. *Poster presented at the Biomag 2006*, Vancouver, Canada.