ACQUISITION OF NOVEL PERCEPTUAL CATEGORIES IN A THIRD LANGUAGE: THE ROLE OF METALINGUISTIC AWARENESS AND FEATURE GENERALIZATION

By

DIVYA VERMA GOGOI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010
© 2010 Divya Verma Gogoi
To my parents
my husband, Ankur
and our daughter Kaavya
ACKNOWLEDGMENTS

Working on this dissertation has been a challenging and enriching experience for me. Without the support and encouragement from several people, it wouldn’t have been possible for me to undertake this research. I would first like to thank my academic advisor and committee chair, Dr. Caroline Wiltshire and my Co-Chair, Dr. James Harnsberger, for their guidance, constant encouragement and willingness to be available for discussions, their incise comments and patient feedback, even if I went off track several times have proved invaluable to this work. I really appreciate Dr. Harnsberger’s efforts to pull me out of the statistical ‘twister’. I thank my dissertation committee members Dr. Wayland, Dr. Lord and Dr. Shrivastav for lending their invaluable guidance as phonetic experts and language specialists towards my dissertation work. To Dr. Wayland for making international students feel welcomed during our initial stay at UF and her constant willingness to give feedback whenever sought. I would like to take this opportunity to thank the Language Learning dissertation Grant Committee for approving the partial funding for this project. I express my gratitude to all the subjects who participated in this research. Without their co-operation, I wouldn’t have been able to complete the project in time.

To my parents who patiently waited for my visit home, the ‘next month’, which never came. They have always supported me to do what I feel is right and then be ready to face the consequences, good or bad. I thank Bisha, for her immense support in every way during the data collection process in India. I thank all the wonderful people with whom I had a chance to be friends with at UF. Especially, Monica, Ashwini, Srikant and my fellow graduate students: Hanye, Weihua Zhu, Zoe Ziliak, Sharon Barkley and Quinn Hansen, who have been a big support to me and most importantly I would like to thank my husband, Ankur for being a constant source of energy that fuelled me to take on a lot of things unbridled.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>8</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>9</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>13</td>
</tr>
<tr>
<td>Overview</td>
<td>13</td>
</tr>
<tr>
<td>Goals of the study</td>
<td>14</td>
</tr>
<tr>
<td>Literature Review</td>
<td>15</td>
</tr>
<tr>
<td>Third Language Acquisition</td>
<td>15</td>
</tr>
<tr>
<td>Factors Influencing Acquisition of L2/L3 Speech Sounds</td>
<td>17</td>
</tr>
<tr>
<td>Proposed Factors</td>
<td>19</td>
</tr>
<tr>
<td>Metalinguistic awareness</td>
<td>19</td>
</tr>
<tr>
<td>Feature generalization</td>
<td>21</td>
</tr>
<tr>
<td>Perceptual assimilation</td>
<td>23</td>
</tr>
<tr>
<td>Cross Language Speech Perception Models</td>
<td>25</td>
</tr>
<tr>
<td>Speech Learning Model (SLM)</td>
<td>25</td>
</tr>
<tr>
<td>Perceptual Assimilation Model (PAM)</td>
<td>27</td>
</tr>
<tr>
<td>The Present Study</td>
<td>31</td>
</tr>
<tr>
<td>Organization</td>
<td>34</td>
</tr>
<tr>
<td>2 METHODOLOGY</td>
<td>35</td>
</tr>
<tr>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td>Description of languages</td>
<td>35</td>
</tr>
<tr>
<td>Malayalam</td>
<td>36</td>
</tr>
<tr>
<td>Nasals</td>
<td>37</td>
</tr>
<tr>
<td>Fricatives</td>
<td>39</td>
</tr>
<tr>
<td>Laterals</td>
<td>37</td>
</tr>
<tr>
<td>Rhotics</td>
<td>40</td>
</tr>
<tr>
<td>Bengali</td>
<td>41</td>
</tr>
<tr>
<td>Spanish</td>
<td>43</td>
</tr>
<tr>
<td>American English</td>
<td>45</td>
</tr>
<tr>
<td>Stimulus Materials</td>
<td>46</td>
</tr>
<tr>
<td>Stimulus Material Evaluation</td>
<td>49</td>
</tr>
<tr>
<td>Stimuli Evaluation: Preliminary Experiment</td>
<td>50</td>
</tr>
<tr>
<td>Stimulus Preparation</td>
<td>51</td>
</tr>
<tr>
<td>Subjects</td>
<td>54</td>
</tr>
</tbody>
</table>
D  PERCEPTUAL ASSIMILATION RESPONSE SHEET ................................................................. 136
E  RESULTS IN TABULAR FORMAT ....................................................................................... 137
LIST OF REFERENCES ............................................................................................................. 139
BIOGRAPHICAL SKETCH ........................................................................................................ 147
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Phonemic consonant inventory of Malayalam</td>
<td>37</td>
</tr>
<tr>
<td>2-2</td>
<td>Phonemic consonant inventory of Bengali</td>
<td>42</td>
</tr>
<tr>
<td>2-3</td>
<td>Phonemic consonant inventory of Spanish</td>
<td>44</td>
</tr>
<tr>
<td>2-4</td>
<td>Phonemic consonant inventory of American English</td>
<td>45</td>
</tr>
<tr>
<td>2-5</td>
<td>Talker information; all talkers are native speakers of Malayalam and born in the state of Kerala, India;</td>
<td>48</td>
</tr>
<tr>
<td>2-6</td>
<td>Result of the preliminary AX discrimination task across three Malayalam speakers. Mean D prime values averaged across contrasts as well as across vowels.</td>
<td>51</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2-1</td>
<td>A map of the state of Kerala with cities of origin for the Malayalam speakers. A smaller map of India on the left hand bottom corner shows the location of the state of Kerala.</td>
<td>49</td>
</tr>
<tr>
<td>2-2</td>
<td>An outline map of Mexico, Central America and South America showing the regional distribution of the subjects.</td>
<td>58</td>
</tr>
<tr>
<td>2-3</td>
<td>Different trials possible in an AX discrimination task using two tokens of one speech sounds.</td>
<td>64</td>
</tr>
<tr>
<td>3-1</td>
<td>Mean percent correct identification scores averaged over all four contrasts at the pretest and the posttest level.</td>
<td>73</td>
</tr>
<tr>
<td>3-2</td>
<td>Mean d’ scores of AX discrimination test over all four contrasts at the pretest and posttest level (Maximum value of a d’ score = 4).</td>
<td>74</td>
</tr>
<tr>
<td>3-3</td>
<td>Mean percent correct identification of contrasts spoken by trained-on talker(posttest) and new talker (generalization test).</td>
<td>76</td>
</tr>
<tr>
<td>3-4</td>
<td>Individual comparison of the monolingual group with BE bilingual group (top chart) and SE bilingual group (bottom chart).</td>
<td>77</td>
</tr>
<tr>
<td>3-5</td>
<td>Mean percent correct response in identifying the lateral non-native contrast [ l-l ] in the identification test at the pretest and posttest level.</td>
<td>79</td>
</tr>
<tr>
<td>3-6</td>
<td>Mean d’ scores of discrimination of the lateral non-native contrast [ l-l ] at the pretest and posttest levels by the subjects in the three language groups.</td>
<td>81</td>
</tr>
<tr>
<td>3-7</td>
<td>Mean percent identification scores for the nasal non-native contrast [n-n] by the subject groups at the pretest and the posttest level.</td>
<td>82</td>
</tr>
<tr>
<td>3-8</td>
<td>Mean percent identification scores for the fricative non-native contrast [j-s] by the three language groups at the pretest and the posttest level.</td>
<td>83</td>
</tr>
<tr>
<td>3-9</td>
<td>Mean percent scores of identification test across all four contrasts by the bilingual groups (BE and SE) and the monolingual group (AE) at the pretest and the posttest levels.</td>
<td>86</td>
</tr>
<tr>
<td>3-10</td>
<td>Mean d’ scores of AX discrimination test across all four contrasts, by the bilingual groups (BE and SE) and the monolingual group (AE) at the pretest and the posttest levels</td>
<td>87</td>
</tr>
</tbody>
</table>
3-11 Mean percent correct response for identification of the three place contrast (laterals, nasals, fricatives) by bilingual (BE and SE) and monolingual (AE) groups at the pretest and posttest levels.................................................................90

3-12 Mean d’ scores from the AX discrimination tests for the three place contrasts (Laterals, nasals, fricatives) by the language groups at the pretest and posttest levels. ....91

3-13 Mean percentage of assimilation elicited from all three language groups, pretest vs posttest. ..........................................................................................................................94

3-14 Percentage of difference scores (pretest to posttest) of assimilation types for the individual contrasts elicited from all three language groups. ........................................95

3-15 Percentages of assimilation types elicited from individual language groups at the pretest and posttest levels are provided in three different charts. .................................97

3-16 Percentage of difference scores (posttest-pretest) of assimilation types elicited from BE, SE bilingual groups and AE monolingual group over all four contrasts. .................98

4-1 Hypothetical representation of the learning continuum displaying the ambiguity of ranking the UU, UC and CG assimilation types in order to assess the direction of learning. .............................................................................................................118
Many factors that contribute towards the perception of second language (L2) sounds and the subsequent establishment of L2 phonetic categories have been the focus of past research. However, learning a third language, unlike SLA, may be influenced by additional factors attributed to the presence of two language systems in a bilingual instead of one language system in a monolingual. Previous third language acquisition research, mainly lexical processing or syntax acquisition, have suggested that being a bilingual provides a positive influence to learners of a third language in attaining general proficiency. Whether bilingualism or a having multilingual benefit as a factor influences the acquisition of novel non-native speech contrasts by adult bilinguals, remains to be explored. Another factor that may facilitate in the phonetic acquisition of a third language is the role that phonetic features may play in the development of new phonetic categories. Features utilized in native contrasts may generalize in the learning of novel non-native contrasts, even if they play a limited or no role in the initial perception of these non-native contrasts. Such factors influencing third language acquisition within specific areas of proficiency like phonological acquisition have not been previously investigated.

This study involves a perceptual training experiment examining the influence of factors of multilingual benefit and/or feature generalization. The training incorporates the acquisition of
four speech contrasts (dental/alveolar – retroflex distinction) of a third language by bilingual (Bengali-English and Spanish-English speakers) and monolingual (American English speakers) groups. The perceptual data is gathered through measures like consonant identification, AX discrimination and perceptual assimilation. The perceptual assimilation data is analyzed using the PAM model to determine the extent to which the initial assimilation patterns change towards direction of learning following the training.

The results across the testing measures show subtle effects of the multilingual benefit factor playing a role in facilitating the acquisition of the non-native speech contrasts when comparing individual language groups. This effect is observed as robust when both bilingual groups are combined to increase sample size during reanalysis in all tests, except for AX discrimination. The factor of feature generalization does not show an influence on the acquisition of non-native speech contrasts by Bengali-English bilinguals in any of the testing measures. Therefore, the results are suggestive of only the multilingual benefit factor playing a role in enhancing the acquisition process of a third language by adult bilinguals. The perceptual assimilation analysis across language groups show a change of assimilation types after training revealing a shift towards the direction of learning. The bilinguals show a greater shift from SC assimilation type to TC assimilation types than monolinguals. The study thus suggests that multilingualism as a factor facilitates the acquisition of non-native speech contrasts.
CHAPTER 1
INTRODUCTION

Overview

The globalization of the world has led to many scenarios where people with different native languages come face to face with each other and the need to communicate arises. The field of second language acquisition at the cognitive level has a long history of exploration of the issue of whether bilinguals have a single system of memory representation or separate systems for each language that they use. Previous studies on this issue have suggested presence of a single or shared system for more than one language at the lexicon level but acknowledge the probability of the presence of more language-specific levels as well (see Flege, 1995; Kirsner et al., 1984; Grosjean, 1992; Krashen, 1981; Kroll and Tokowicz, 2005; De Bot, 1992). SLA studies have looked into various factors that influence the process of acquiring a new language by adults who have only one language system. However, the scenario is not that simple. Acquisition of more than two languages has become more a norm than exception. Therefore, research in the field of third language acquisition or in broader perspective, multilingualism has made strides in the last few years.

Since the last decade researchers have turned their attention to investigating the factors that influence the acquisition of a third language (Cenoz, Hufeisen & Jessner, 2001; Baker, 2001; Bialystok, 2001; Munoz, 2000; Sanz, 2000). While there has been growing interest in L2 acquisition and bilingualism, many studies have looked into the factors such as age of acquisition, amount of use of language, context, status of the L1 etc. and their effect on learning L2 either simultaneous or sequential. Fewer studies have looked into third language acquisition within specific areas of proficiency like phonological acquisition (Enomoto, 1994; Werker, 1986). Previous third language acquisition research have suggested that being bilingual provides
a positive influence to learners of third language in attaining general proficiency (Cenoz & Valencia, 1994; Ardeo, 2000; Munoz, 2000; Bild & Swain, 1989). These studies were mostly done on children in classroom settings. Whether or not factors such as metalinguistic awareness and/or feature generalization enhance the acquisition of the phonetics of a third language for the bilinguals has not been previously investigated. In the present study, an experiment is set up in this study to examine such factors after providing training with the new contrast in a laboratory setting.

**Goals of the study**

The goal of the present study concerns the exploration of certain factors that may positively influence the acquisition of a third language by adult bilinguals. Specifically, it examines the effects of bilingualism in terms of two factors; firstly, benefits of possessing two linguistic systems often termed as metalinguistic awareness and secondly, transfer of phonetic feature knowledge from L1/L2 to L3. The study investigates the effects of these two factors on the acquisition of a third language, Malayalam by two bilingual groups Bengali-English and Spanish-English speakers and one monolingual group of American English speakers. By presenting the subjects with novel speech contrasts (retr oflex sounds in manners that are lacking in the subjects’ L1/L2 sounds inventory) over a limited period of training, the study aims to examine if any additive effects of being a bilingual are seen in the course of learning. The specific sound contrasts from Malayalam that are used as stimuli are lateral [l - l], nasal [n-ŋ], fricative [ʃ-ʃ] and rhotic [r-ɻ]. The experiment set up for this study includes various tests of identification, discrimination and perceptual assimilation to observe the effects of the factors in different levels of perception.
The main hypothesis being tested in the present study is that the bilingual groups are expected to perform better than the monolingual group in perceiving and identifying the speech contrasts from the target language. Such results would argue for positive influence of bilingualism on the third language acquisition in the specific area of phonetics. Within the bilingual groups, the Bengali-English learner group, who have phonemic experience with retroflex feature but only in manners other than what is presented in the experiment, is expected to perform better than the Spanish-English learner group, who have no exposure to retroflex feature, phonemic or phonetic. By incorporating previous language experience as a factor within the bilingual groups, the results will provide information for an additive effect of feature generalization in enhancing the acquisition of speech contrasts from third language over the effect of bilingualism. The results will also be analyzed in terms of whether they fit the predictions of the Perceptual Assimilation Model (PAM) proposed by Best & Tyler 2007. The study aims to examine the changes in assimilation patterns of nonnative speech contrasts by the listener groups at both the initial perceptual stage and after training stage by adopting the assimilations types described in PAM.

**Literature Review**

**Third Language Acquisition**

A vast corpus of literature exists concerning second language acquisition (SLA) in adulthood. Research on third language acquisition as a separate entity is a relatively recent endeavor. Therefore, research in this field is not as extensive. Third language acquisition refers to the learning of a non-native language by an individual who has already acquired two languages either simultaneously (before puberty) or sequentially. The scenario of third language acquisition although quite similar to SLA still presents differences since, unlike second language learners, third language learners have more experience as language learners and have access to
two linguistic systems when acquiring a third language (see Cenoz, 2003). Most of the previous studies on third language acquisition have looked into the question of whether bilingualism facilitate the acquisition of third language? (Cenoz, Hufeisen & Jessner, 2001; Baker, 2001; Bialystok, 2001; Munoz, 2000; Sanz, 2000) Most of these studies have been conducted on school children in a classroom setting. The focus of these studies have been primarily to examine the level of general proficiency in the target language being learnt through various linguistic tasks like word awareness, grammar, speaking, reading and writing. Their results have suggested that being a bilingual provides positive influence on the acquisition of the third language (Cenoz & Valencia, 1994; Ardeo, 2000; Munoz, 2000; Bild & Swain, 1989). In addition to the effect of bilingualism, third language acquisition also presents a diverse range of other factors that may influence the learning of the target language. These factors are quite similar to the typical factors that affect acquisition of the second language, such as age of acquisition or the amount of use of the language (see e.g. Flege, 1998; Flege et al., 2002) or phonetic, phonemic or acoustic factors (Polka, 1991), coupled with other influential factors that are traditionally explored in the field of bilingualism such as L1/L2 status, socioeconomic factors, degree of bilingualism etc (see Hammarberg, 2001). More specific to the current study, the effect of bilingualism, also termed as metalinguistic awareness, along with feature generalization, here referred as the transfer of phonetic/phonemic feature knowledge from L1/L2 sound inventory to facilitate in the acquisition of L3 sounds, is examined in the area of phonological acquisition of third language. The following section provides a background review of certain factors that may influence acquisition of sounds in both L2 and L3 along with the detailed review of the proposed factors - metalinguistic awareness and feature generalization.
Factors Influencing Acquisition of L2/L3 Speech Sounds

Through extensive research many factors that contribute towards the perception of second language (L2) sounds and subsequent establishment of L2 phonetic categories have been examined. Based on such factors, theories on cross language speech perception by naïve as well as experienced listeners have been developed (e.g., perceptual assimilation model, or PAM (Best, 1994, 2007; Flege, 1995). The term naïve listener here refers to a listener who has no prior exposure to the non-native sounds of the target language and an experienced learner is acquainted with the non-native sounds although he/she has not formed any phonetic categories corresponding to these sounds. Phonetic, phonemic and acoustic factors, among others, have been determined to influence the learning of L2 sounds (Polka, 1991). Previous studies that have addressed such factors are briefly discussed below.

The factor of phonetic experience defined as, “experience with specific phonetic segments as allophones or free variants of native phoneme categories” (Polka, 1991: 2962), has been explored in cross-language speech perception research. It has been hypothesized (Werker et al, 1981; Tees and Werker, 1984; Strange and Dittmann, 1984; Polka, 1992) that a non-native contrast that contains a phonetic segment identical to a segment frequently experienced only in phonetic contexts by a subject in his language(s), is more easily differentiated by the learner as opposed to non-native contrasts which may not employ allophonic experience. The findings from Werker et al (1981) and Tees & Werker (1984) showed naïve American English adults’ increased difficulty in differentiating (after training) the [t - t] place contrast compared to the [d-th-t] voicing contrast from Hindi. They credited allophonic experience as an explanation for these results suggesting that experience with VOT contrasts in the subjects’ native language (AE) facilitated the differentiation of the non-native voicing contrast. However, a limited number
of subjects who had experience with the Hindi language in their early childhood were able to perceive the place contrast after training, suggesting early linguistic experience facilitates better perception of the place contrast than the AE subjects with no early life experience with Hindi (Tees & Werker, 1984). Polka’s (1992) study tested the above given hypothesis but found no conclusive results due to presentation order effects. However, the allophonic experience hypothesis was contested in subsequent studies (Best et al., 1988; Pruitt et al., 2006), suggesting that phonetic experience alone could not explain the variability found in perceptual performance on different non-native contrasts and could even have a detrimental effect in the perception of non-native contrasts. Phonemic and phonetic experience together (presence of linguistic experience), are able to better explain the facilitation effects that are seen in the perception of non-native contrasts (Best et al., 1988; Sundara et al., 2006).

Another factor of acoustic salience that may influence the acquisition of L2/L3 sounds was proposed by Burnham (1986) to account for variable perception levels of non-native contrasts. Burnham claimed that the distinction between the easily discriminable and the poorly discriminable non-native speech contrasts can be explained in terms of their inherent psychoacoustic basis. He used the terms “robust” and “fragile” to maintain the psychoacoustic distinction between non-native contrasts. According to him, a contrast was robust if the reliable acoustic cues in that speech contrast are too salient to be lost during memory demanding tasks. However, psychoacoustic salience individually as a factor can explain perceptual performance only when some independent acoustic cue measures can be determined. Further research needs to be conducted in this respect.

The above discussed factors have been determined to substantially influence the perception of non-native speech sounds of a target language. However, without explicit definition, a target
language can be assumed to be a second language (L2), a third language (L3) or a novel language (language unknown to the listener). The area of third language acquisition (or multilingualism) until recently was considered equivalent to second language acquisition (SLA) (as noted in Hammarberg, 2001). However, learning a third language, unlike SLA, may be influenced by additional factors like effects of bilingualism (Hoffman, 2001; Sanz, 2000; Chenoz, 2001; Bialystok, 1992, 2001). The section that follows provides detailed information on the factors explored and presents them in the context of the current study.

Proposed Factors

Metalinguistic awareness

Metalinguistic awareness allows the individual to step back from the comprehension of an utterance; to attend to the structural features of the sounds in order to abstract the knowledge of its distinguishing cues in a speech contrast, a task which demands certain cognitive and linguistic skills (Malakoff, 1992: 518). This awareness helps a bilingual to acquire non-native contrasts more easily than monolinguals. The assumptions in the formulation above are based on the results of studies on lexical processing and word learning by bilingual and monolingual children showing that bilingual children are more adept at doing metalinguistic tasks and have cognitive benefits from knowing two or more languages (Peal and Lambert 1962; Munoz, 2000; Klein, 1995; Sanz, 2000 etc.). The study by Klein (1995) showed that during lexical acquisition, multilinguals learned the same number of lexical items as monolinguals but at a quicker rate of acquisition. Klein based the explanation for this phenomenon on enhanced cognitive skills in multilinguals which helped them tease out the potential relevant data for resetting the particular parameter, preposition stranding in this case, for the new language. Numerous studies looking at
cognitive processing\(^1\) at a linguistic level have explored the concept of metalinguistic awareness in bilingual and multilingual children as well as adults (Bialystok, 2004). The studies emphasize the fact that bilinguals with near native proficiency in both languages would show evidence of metalinguistic awareness (Bialystok 1992, 2001; see also review in Jessner, 2006). In her study, Sanz (2000) suggests that metalinguistic awareness may be an enhancing factor in language processing. She promotes bi-literacy as the cause of heightened metalinguistic awareness.

“…metalinguistic awareness, which results from exposure to literacy in two languages, gives bilinguals the capacity to focus on form and pay attention to the relevant features in the input” (Sanz, 2000: 36).

This study extends the concept of metalinguistic ability from the field of lexical processing to the field of phonetics. This core concept of metalinguistic awareness has not yet been explored in the field of cross-language speech perception and in the area of phonological acquisition of a third language. It is yet to be seen if implicit knowledge of two languages, in the case of phonetic features in the sound systems, can result in heightened capacity to acquire non-native speech sounds. However, the concept of metalinguistic awareness is a very broad usage for various metalinguistic skills. This study specifically assesses metalinguistic awareness/metalinguistic ability from the perspective of the effects of bilingualism where having two or more languages, well established within the language system of an individual benefit him for linguistic processing over a monolingual individual. Therefore, rather than using the term metalinguistic awareness loosely, the more operational term of multilingual benefit factor will be used in this study.

According to previous studies in lexical processing, the bilinguals may function at a faster rate and at a higher level for reorganizing information in order to deal with the high-demand task of

---

\(^1\)For this study the term cognitive processing is operationally defined as the act of abstraction of relevant information from a speech utterance, required for a linguistic skill, as a result of higher level of concept formation.
acquiring new non-native speech contrasts when compared with monolinguals. The assumption that bilinguals develop meta-linguistic awareness which benefits them over monolinguals in acquiring a third language, can be examined through the experimental design set in which bilinguals are compared to monolinguals in the acquisition of speech sounds in the third language, as is done in this study. For instance, in the current study scenario, the retroflex feature, which is non-existent in the Spanish-English bilingual’s phonological space, may still be acquired by them at a relatively faster rate of acquisition than English monolinguals. This may reflect that bilinguals employ their metalingual knowledge to pay attention to distinctions in auditory-acoustic stimuli and to tease apart relevant acoustic cues in order to accurately distinguish between the newly acquired speech contrasts. The results of perceptual performance after training would be compared to results of the monolingual group. Therefore, in case of successful and faster acquisition, results would suggest the presence of meta-linguistic awareness among bilinguals (Spanish-English) which facilitates the acquisition of a novel contrast without the limitation of feature generalization of an existing feature.

**Feature generalization**

Features, an important set of information contained in the established phonetic category, play a crucial role in the development of new phonetic categories (Nosofsky, 1986, 1987; Kruschke, 1992; Lively et al., 1993; Francis and Nusbaum, 2002). Features may refer to phonemic attributes of the sound system of a language, phonetic features or acoustic cues attached to the specific segments of the system. According to Francis and Nusbaum (2002) features can be stretched or shrunk according to relative weights given to them. For this study, features refer to the phonetic features within a phonemic segment of a language. The factor of feature generalization or feature productivity as defined by Polka (1992: 40) is “whether experience with a particular place distinction facilitates perception of the same phonemic place
distinction in a different manner context…” The factor of feature generalization, operationally defined for this study, refers to the concept where the listener being sensitive to individual features in his/her native phonetic categories is able to identify the same feature in a novel non-native contrast and thus develop new phonetic categories. Few studies have looked at the feature generalization phenomenon in the field of cross-language speech perception. Polka (1992) and Harnsberger (1998) both found results contrary to the expected results in this study that feature generalization will reveal better perceptual performance. However, one factor that could explain their results is that these previous studies involved only the perceptual presentation of new non-native contrast to naïve listeners, rather than conducting a perceptual training experiment and subsequently observing any effects of feature generalization during a perceptual testing phase.

The present study is different from previous works in that it takes the next logical step and examines the role of feature generalization in the perceptual performance data (posttest data) of TRAINED listeners as opposed to naïve listeners (see Polka, 1992; Harnsberger, 1998). Recently, a study by McAllister, Flege, & Piske (2002) found that experience with the duration feature did facilitate native American English speakers’ perception of Swedish vowels (phonemic lengthening) more than that of native Spanish speakers (both subject groups had lived in Sweden for at least 10 years). The difference between McAllister et al.’s study and others was that the subjects were experienced learners of Swedish and used the target language often. This provides support for the hypothesis of feature generalization. To test for feature generalization in the current study, a group of Bengali-English bilinguals (BE) undergo perceptual training to acquire the speech contrasts of retroflex nasal as opposed to alveolar nasal. The Bengali language has extensive use of the retroflex feature for stops in its phonemic inventory but lacks this feature in manners like nasal stop [ɾ], lateral approximant [ɿ], fricative [ʃ] and central
approximant \([ \approx ] \). An observation of faster acquisition of these novel sounds by the BE bilingual group when compared with either the bilingual Spanish-English group or the monolingual American English group, will indicate a successful generalization of the feature retroflex from stops to novel manners of articulation like nasal stops. The results of training on the identical stimuli from the group of English monolinguals would be observed as a control group.

One criticism on the choice of subject group could be that the phenomenon of feature generalization could be evident in monolingual Bengali speakers as well. The current study deals with feature generalization as a factor of enhancing learnability of non-native speech contrasts, regardless of whether the learner is a bilingual or monolingual. However, taking bilingual learners of L3 for this study provides us with the opportunity to assess both feature generalization and multilingual benefit within a single group of subjects. Moreover, if the results do not provide evidence for multilingual benefit (in the Spanish-English (SE) group) but show enhanced learnability in the Bengali-English group, it would suggest that it is not the factor of multilingual benefit that is facilitating learnability in Bengali-English but the factor of feature generalization. On the contrary, if the multilingual benefit factor is observed in the SE group, then the feature generalization factor is assumed to have an additive effect that can be observed in BE group as significantly higher performance level than SE group.

**Perceptual assimilation**

The third focus of the research project concerns the role of the bilinguals’ particular perceptual category inventory in the acquisition of non-native contrasts from a third language. Looking at the perceptual assimilation patterns elicited from the learners at the post-training stage will provide an opportunity to examine the formation of new perceptual categories, if any, developed after a limited time of training. Moreover, perceptual assimilation analysis would be
another dependent measure to test for performance level of acquisition among the subject groups. This measure will show greater ecological validity as opposed to just perception tests conducted in previous studies that looked at the acquisition of L2/L3 through perceptual assimilation. Unlike a perception test, training involves feedback on performance, an essential aspect of learning. Since a training experiment where an individual is perceptually trained on non-native speech contrasts would simulate more closely the natural learning process, eliciting and examining perceptual assimilation (PA) types at the post-training stage and comparing them with pre-training PA types will prove to be more ecologically valid than simply eliciting the PA types in a perception experiment without any controlled training.

Previously, several assimilation analytical models have been developed to account for the influence of native perceptual categories on the perception, production, and acquisition of non-native speech sounds, including the Speech Learning Model (SLM) and, of greater interest for this project, the Perceptual Assimilation Model (PAM) since it predicts the assimilation patterns not only at the initial exposure stage but also predicts learnability beyond the first stage. Likewise, the current study also looks at the assimilation patterns at the post-training stage. PAM concerns the discriminability and learnability of non-native contrasts based on the relationship between the non-native and native category inventories. These relationships have been encapsulated in several assimilation types: Two-category (a non-native contrast that is highly similar to a native contrast), Single category (a non-native contrast that is consistently identified as a good exemplar of a single native speech sound), Category goodness (a non-native contrast consistently of two speech sounds that are consistently identified with a single native sound but differ in goodness of fit), uncategorizable, either UC or UU (involving one/both non-native speech sounds that are not consistently identified with a particular native category) and non-
assimilable (a contrast that is perceived in a nonlinguistic manner). These assimilation types have been discussed in detail in the next section. The present study tests these predictions of the PAM model in two ways: 1) Determining the extent to which initial patterns of assimilation change, if at all, following extensive training experience and 2) measuring the extent of assimilation pattern change in individual subject groups in order to assess any effects of the influencing factors explored in this study.

The following section discusses the above mentioned cross language speech perception models that have been developed predicting perceptual assimilation patterns of naïve listeners as well as subsequent changes in assimilation patterns in the learning of a second language by experienced learners.

**Cross Language Speech Perception Models**

Since the advent of speech perception studies on non-native contrasts during the 1980-90s, various speech perception and acquisition models that address bilingualism have been proposed. This section of the chapter describes the pertinent models and examines the core claims of each model.

**Speech Learning Model (SLM)**

The speech learning model focuses on the L2 phonological acquisition in production and perception (Flege 1995). It is claimed in this model that many second language production errors have an underlying perceptual basis. The model considers the factor of perception as the underlying principle behind the acquisition of L2 speech sounds and effective learning. The model propagates the notion that if learners of second language sounds are not able to accurately perceive particular L2 sounds, the consequent production of those L2 sounds will be inaccurate.

One of the preliminary assumptions of the model is that the production and perception of phonetic systems “remain adaptive over the life span” and they are susceptible to reorganization
to accommodate L2 sounds, when encountered during the life span. Another postulate of this model is that “bilinguals strive to maintain contrast between L1 and L2 phonetic categories (long term memory representations) which exist in a common phonological space” (Flege, 1995: 239), bringing forth the relevance of the phonetic /phonemic contrasts between the L1 and L2 sounds, in the process of acquiring the second language. Hypotheses of this model reflect the need to acknowledge the phonetic differences between the L1 and L2 sounds that may be phonetically closer or more distant, which may influence the perception of L2 sounds by the learners.

The factor of mental representation or categorization of L1 and L2 sounds has been directly addressed in this model. According to SLM, if the bilinguals learn to distinguish some phonetic differences between the L2 sound and the closest L1 sound to it, a new phonetic category is likely to be established for that L2 sound. Nevertheless, the category formation for an L2 sound can be blocked if the mechanism of “equivalence classification” comes into play. In other words, for the bilingual the L1 and L2 sounds can be perceptually linked and therefore, one phonetic category is used for both sounds, which may eventually resemble each other in production. The term implies that ‘equivalent’ or similar sounds are difficult to acquire due to the perception of these specific L1 and L2 sounds by the listener as being quite similar (non-perceivable phonetic distance by listener), whereas ‘new’ or dissimilar (terms used interchangeably) sounds are easier because the listener easily distinguishes the phonetic differences.

SLM also predicts that if a new phonetic category for an L2 sound is in fact established, the phonetic system would try to maintain contrast between the existing L1 phonetic category and the newly formed L2 category since both L1 and L2 categories are assumed to ‘exist in a common phonological space’ (Flege, 1995: 239)). Thus, the new L2 category established may
not be identical to the native speaker category, that is, the target sound. Therefore, the predictions for mental representations of the L1 as well as L2 sounds are well documented in this model.

The model differs from an older contrastive analysis approach as it proposes a less abstract phonetic level of analysis than phonemic level of analysis. Contrastive analysis of two languages involves the examination of phonemic differences between the inventories of those languages. According to one SLM hypothesis, learners perceptually relate positional allophones in the L2 to the closest positionally defined allophone in the L1. On the other hand, L1 for L2 substitution described in contrastive analysis is seen as either failure to discern the phonetic differences between the L1 and L2 sounds or inability to articulate the correctly perceived sound.

According to SLM, during the course of second language acquisition, the phonetic category that has already been established for an L1 sound during childhood evolves gradually if it is perceptually linked to an L2 sound. Such developments are influenced by age of learning and perceived phonetic distance between an L1 sound and the closest L2 sound. The model claims that the greater the perceived cross-language phonetic distance, the more probability for the development of a separate phonetic category for the L2 sound. Also, age of learning (AOL) determines how much perceived phonetic distance is required to stimulate the process of separate category formation for the L2 sound (early AOL + smaller perceived phonetic distance needed = category formation). In summary, the model claims that the most L2 production errors observed have an underlying perceptual problem.

**Perceptual Assimilation Model (PAM)**

The Perceptual Assimilation Model (Best, 1995, 2007) lays its foundation on the phonetic/phonological nature of the L2 speech sounds versus L1 sounds, similarly to the other speech perception models. The preliminary assumption of the model is that the nature of the
discrepancies and similarities that a non-native listener picks up in the auditory signal may be perceived in terms of articulatory or acoustic properties. The model takes the ‘ecological perspective’ and relates to the Direct Realist Theory (e.g., Gibson, 1991; Gibson, 1966, 1979) that the listener’s perception is stimulated mostly through the evidence about the articulatory simple gestural events in the speech signal (Best, 1995). The model extends the direct realist theory in the way that the adult listener perceives information about gestural similarities in the L2 speech with the native phonemes, when receiving the non-native signal. According to the model, there is no two-extremes phenomenon (full assimilation or none) in perceiving non-native speech. Also, all non-native contrasts do not pose same level of difficulty to the listener. The perceptual difficulty should be predictable from the differences in the gestural similarity/discrepancy between non-native and native sound distinctions. There are five main possibilities during perception that the model predicts as outcome of listeners’ mental representation of non-native speech contrasts. The predictions of the discrimination task are also given alongside the category type speech contrasts.

Firstly, non-native speech contrasts may be treated as two category type (TC). In this case, the speech contrasts may be gesturally similar to two different native phoneme categories making the non-native speech contrast easily discriminable by the L1 listener. An adult listener is expected to show near native-like discrimination of this type of non-native speech contrast.

Secondly, non-native speech contrasts may be treated as single category type (SC) if listener’s perception of the non-native speech contrast is equally good or equally poor and both sounds are assimilated to a single category.

Thirdly, non-native speech sounds may show differences in category goodness type (CG). When the non-native contrast is assimilated to a single category but one member of the contrast
is seen as a better token of the native phoneme and the other as a marginal token of that category, this makes the discrimination easier than the SC type but not equivalent to TC type of assimilation. Again, because category goodness type of pairs may be perceived with varying degree of discrepancy between the two phones, the level of discrimination may vary.

Finally, Uncategorizable type (UU) are formed where even though none of the phones from the non-native pairs is assimilated with any native phoneme category, yet the phones are considered as speech sounds of a foreign language. For Uncategorized versus categorized type (UC), where only one phone of the non-native pair is assimilated to a native phoneme category, the adult listener is expected to show a very good discrimination performance for this type of non-native speech contrast.

Therefore, concluding from the above given possible outcomes of discrimination performance, the pattern for an adult listener would be highest discrimination for Two category types, moderate to good discrimination for both category goodness type, non assimilable type (for different reasons) as well as for the uncategorized versus categorized type, and lowest for the single category type. The uncategorizable type varies from good to poor depending upon the acoustic proximity of the two phones in the pair. Also, PAM assumes a strong L1 phonological system influence on the perception ability. The PAM model, in conclusion, emphasizes that the perception of non-native segments is performed according to their similarities or discrepancies from the native segmental system, being in close proximity to the native phonological space as well as with each other (within the contrast) (Best, 1995).

Although the model provides intricate details of perceptual abilities of adults for non-native speech contrasts with highly probable outcomes for the categorization of the non-native phones by the listener, the model does not really deal with the learning process of these non-
native phones by the L1 speaker, unlike SLM. The earlier PAM model provides the predictions of perception of non-native speech contrasts by listeners who have not had any experience with L2. These discrimination outcomes provided in the model are based on the first exposure of the listener to these sounds. The learning process that takes place after the learner is exposed repeatedly to the L2 sounds and the reorganization of the L1 phonetic categories that takes place due to the influence of the L2 phones is not addressed in this model.

These differences are addressed in the revised PAM model (Best & Tyler, 2007). The Perceptual Assimilation Model extends its predictions from the earlier model focusing on the naïve non-native listener to a revised version of the model where it accounts for the experienced L2 learner apart from the initial encounter with L2 as a naïve listener. Best refers to the revised version as PAM-L2. Four possible cases of L2 minimal contrasts are discussed starting with PAM’s initial predictions.

- Only one L2 phonological category is perceived as equivalent (perceptually assimilated) to a given L1 phonological category (TC). The discrimination of the speech contrasts would be excellent as the L1 learner would always perceive the fully assimilated L2 phone separate from other less assimilated L2 phones. The L2 learning for this L2 phonological category will not take place as it has been entirely assimilated and recognized with an L1 phonological category.

- Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but one is perceived as being more deviant that the other (CG). The discrimination level for this type would range from moderate to good. During the learning process, the learner is likely to gradually form a new phonological category for the perceived poor/deviant phone of the L1 phonological category. The good exemplar L2 phone may remain assimilated to the L1 phonological category.

- Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but as equally good or poor instances of that category (SC). The listener will poorly discriminate the contrast since the two phones would be perceived as having very little phonetic difference and be considered as exemplars (either equally good or bad) of single L1 category. For the learners to distinguish these two phones apart, relevant phonetic differences between the phones will have to be learnt, forming a separate phonological category over time.
No L1-L2 phonological assimilation: There could be two situations with these types of minimal contrasts. The listener may not perceptually assimilate either of the phones with any L1 phonological category. If the two L2 phones are in existence with relatively larger distance in the learner’s L1 phonological space, the learning process will be easier for the listener. However, if the L2 phones are relatively close to each other in the L1 phonological space, the listener would be not able to discriminate them easily.

The PAM-L2 model makes the assumption that the learner begins the acquisition of L2 as a naïve but phonologically sophisticated L1 listener. With the initial introduction of the phones of L2 and new L2 phones perceptually assimilated with L1 phonological/phonetic categories, the learning process for the L2 learner starts. The PAM-L2 model makes predictions for the learning process of this, now experienced, L2 learner. The target phonological system, that the learner may achieve over a life span, which the model hypothesizes, is that of a common L1 and L2 system whereby phonetic categories from both L1 and L2 share a common phonological space.

All these models of cross-language speech perception look at the second language learning by adults. These models can be extended to look at the factors facilitating third language learning. The current study proposes factors like metalingual benefit and feature generalization as some of those factors that may facilitate third and/or second language acquisition.

The Present Study

This study is an attempt to provide a window to phenomena not explored well in the field of cross-language speech perception and phonetics. Firstly, the multilingual benefit concept has been attested only in psycholinguistics, lexical and cognitive processing or word learning fields. Moreover, until recently, multilingual benefit was examined only in bilingual children. The current study will test this factor’s influence on third language acquisition in the area of cross-language speech perception and third language learning. Secondly, the phenomenon of feature generalization has been evaluated in studies only dealing with naïve listener’s perceptual performance. The main experiment of the study deals with training the subjects and testing their
perceptual performance in order to seek evidence for feature generalization. Observing the perceptual performance of a trained listener rather than a naïve listener may be the missing factor in finding the role of feature generalization in facilitating the learnability of non-native phonetic categories. Apart from the pioneering steps in exploring metalinguistic ability and feature generalization, the study adopts the assimilation types established by the revised PAM model and tests the predictions laid out by addressing the issue of the relation of initial perception modifying into different assimilation patterns during the course of training. The relevance of chosen stimulus set and the appropriate subject groups is discussed as follows.

A stimulus set which is relatively difficult to discriminate based on acoustic differences is archetypal for this kind of training study so as to bring out subtle distinctions in the perceptual performance of the various subject groups. Dental/alveolar-retroflex place distinction as a speech contrast is considered to have a reasonable level of difficulty in discrimination due to differences in spectral features as opposed to temporal features in speech contrasts (Pruitt et. al., 2006; Polka, 1991; Strange and Dittmann, 1984). Therefore, dental/alveolar-retroflex place distinction was chosen as the place contrast for the non-native speech contrast as the stimulus set. The speech contrasts were taken from Malayalam, recorded as natural stimuli from eight speakers. Malayalam was chosen as the target language as it has a rich inventory of consonants employing the retroflex feature extensively in various manners. This allows a flexible choice of speech contrasts as per the requirements of the study. An appropriate and diverse range of challenging non-native speech contrasts were selected from Malayalam for the study.

1) [ l - ] (alveolar versus retroflex lateral approximant)
2) [η - æ ] (dental/alveolar versus retroflex nasal)
3) [ʃ - s ] (palato-alveolar versus retroflex fricative)
4) [ t - ] (alveolar tap versus retroflex approximant)
These speech contrasts differ primarily in place of articulation, although the \([r - ɹ]\) contrast also differs in manner. The dental/alveolar-retroflex place distinction, common to all of the chosen speech contrasts, allowed examination of factors (a) feature generalization and (b) multilingual benefit in three populations of subjects. Furthermore, selection of four different manners with same place contrast allows for the examination of the learnability level of these speech contrasts. They can also provide new data testing the predictions of PAM in perceptual assimilation.

Three subject groups were chosen for this study. Firstly, the reason for choosing these particular speech contrasts, when looking at Bengali-English (BE) language group, is primarily for the purpose of evaluating feature generalization. The speech sounds with retroflex feature \([ɹ, ɹ̃, ɾ̃, ɹ̪]\) are not found in Bengali. However, Bengali employs the retroflex feature quite productively but only in the oral stop and flap manners of articulation. Therefore, if the retroflex feature gets generalized from existing phonetic categories to newly formed phonetic categories, the evidence of any variability in generalization of the feature in different manners will be reflected during the acquisition of these four new speech contrasts.

Secondly, Spanish speakers altogether lack any experience with the dental/alveolar-retroflex place distinction in phonemic or phonetic environments. However, as the participants of this study are bilingual (Spanish-English), they are assumed to have phonetic experience similar to American English speakers, who have alveolar place distinction phonemically and sometimes have allophonic experience with retroflex oral stops (See Polka, 1991). No prior native language(s) experience with the retroflex feature makes the Spanish-English bilingual group ideal to examine multilingual benefit factor. This bilingual group, having significant
improvement in identification of stimulus sets after training, if any, can be safely attributed to multilingual benefit.

The third population of subjects chosen for this study is a control group. Monolingual American English (AE) speakers were chosen in order to control for the bilingualism factor. Also, American English speakers lack dental/alveolar-retroflex place distinction at the phonemic level. However, retroflexion of alveolar stops in particular environmental contexts (e.g. [dɪta]) ‘dry’ (Polka, 1991) is sometimes found in the speech of American English speakers.

**Organization**

This chapter provided an introduction to the current study along with a review of previous literature regarding cross-language speech perception as well as the posited factors. A detailed discussion of the research questions evaluated by the study along with the relevance of the stimuli and subject populations were presented. Chapter 2 gives a comprehensive description of the phonemic and phonetic inventory of relevant segments in the four languages utilized in the experiment. It also provides the experimental design of the perceptual training and testing paradigm used for the main experiment of the study. Chapter 3 presents experimental findings of the training experiment undertaken in this study. Lastly, Chapter 4 summarizes the findings and discusses the implications and limitations of the current study.
CHAPTER 2
METHODOLOGY

Introduction

The experimental design of this study is directed towards evaluating the hypotheses that (a) feature generalization and/or multilingual benefit may facilitate the learnability of non-native speech contrasts by bilinguals and (b) assimilation patterns in the initial perception of non-native speech contrasts can be used to predict subsequent learnability in perception and accuracy in production. The study involves the perception training experiment which consists of learning of novel non-native speech contrasts. The next section describes the phonemic segments and their phonetic realizations in the four languages employed in the main experiment. The following sections describe the preparation and preliminary evaluation of the stimuli for the main experiment and discuss the recruitment of various language groups as subjects for the current study. It also highlights the appropriateness of the choice of the three language groups in order to directly examine the research questions put forth in the study. Relevant literature review on the method of training and their underlying theoretical assumptions is also provided in subsequent section. Towards the end of this chapter, the adapted experimental design from the perception training paradigm is described in detail.

Description of languages

The experiment outlined in this chapter entails the involvement of four languages. The perception training stimuli was obtained from Malayalam (target language). The subjects who underwent training were from three different language groups, bilingual Bengali-English and Spanish-English speakers and monolingual American English speakers. With the description of each language, its distinctive segments are discussed with their phonetic allophones. For the proposed experiment, description of phonemic inventories alone would be insufficient to capture
the initial perceptual exposure as well as the developmental acquisition of the novel non-native speech contrasts. The comparison of all four languages in question regarding the retroflex feature is relevant to the assessment of factors such as feature generalization and/or multilingual benefit since exposure to phonemic or allophonic experience with the retroflex feature may be either beneficial or detrimental to the perception of a novel speech contrast (Polka, 1991; Strange, 1995; Pruitt et al., 2006). The study looks at the acquisition of Malayalam speech contrasts by Spanish-English bilinguals, Bengali-English bilinguals and English monolinguals.

Malayalam

A member of the Dravidian language family, Malayalam uses the retroflex feature in most of the manners it employs. This language was chosen to be the target language for the experiment since the retroflex place of articulation is highly productive in this language.

Generally, retroflex sounds are taken as a specific place of articulation whereby a voiced retroflex stop in one language would have similar articulatory characteristics in another language. Contrary to this opinion, Ladefoged and Bhaskararao (1983) revealed that retroflex speech sounds found in three languages, one Indo-Aryan and two Dravidian, displayed a varying degree of retroflexion. Hey claimed that the degree of retroflexion and which place of articulation the tongue blade touches could be compared with vowel heights in the way that it is possible to make a continuously changing set of sounds from an apical dental to retroflex place of articulation. In their study, Tamil, a Dravidian language closely related to Malayalam, showed a high degree of retroflexion whereas Hindi, an Indo-Aryan language related to Bengali, showed least degree of retroflexion. Malayalam, like Tamil, is assumed to have a high degree of retroflexion.

For the purpose of the experiment, four speech contrasts were chosen from this language to serve as training stimuli. The common attribute in all of the four contrasts is the presence of the
retroflex place feature in one of the segments in the contrast. The segments containing the retroflex feature are novel speech sounds for the listener groups. Table 2.1 shows the phonemic and phonetic segments of Malayalam (adapted from Asher, R.E. and Kumari T.C. (1997)).

Table 2-1. Phonemic consonant inventory of Malayalam (adapted from Asher, R.E. and Kumari T.C. (1997).

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td>η</td>
<td>η</td>
<td>η</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>v</td>
<td>ι</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill/ Tap</td>
<td></td>
<td>r</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td>s</td>
<td>§</td>
<td>f</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
</tr>
</tbody>
</table>

**Nasals**

Malayalam exhibits a rich inventory of nasal stops employing a six-way place contrast starting with bilabial, dental, alveolar, retroflex, palatal up till the velar. Researchers agree on most of the phonemic distinctions among the different place contrast except for the dental – alveolar place contrast. The dental – alveolar place contrast has been discussed at length describing its phonological as well as phonetic distinction in the language (Kumari, 1972; Sreedhar, 1972; Kalackel, 1985; Raja, 1960; Mohanan and Mohanan, 1984; Asher and Kumari, 1997).

On articulatory grounds, this particular place contrast is marked universally in the world’s languages, since dental and alveolar place contrasts have very subtle distinctions articulatorily and acoustically. It has been described in various ways on the basis of its articulation and provided with phonemic pairs; terms such as ‘spread-contact vs. point contact (Aiyar, 1972),

---

1 There is a disagreement among researchers about the phonetic properties of this sound. Asher and Kumari (1997) refers to it as a sublamino-palatal approximant (discussed further in this chapter).
apico-laminal dental vs. apico-alveolar (Asher and Kumari, 1997) or simply dental vs. alveolar (Subramoniam, 1973) have been used for this place contrast.

On phonological grounds, this place contrast’s productivity in Malayalam is limited to only one kind of distribution – as geminates occurring medially. Due to this functionality, dental nasal and alveolar nasal are considered phonemic in Malayalam (Kumari, 1972; Kalackel, 1985; Asher and Kumari, 1997). However, Mohanan & Mohanan (1984) noted that dental and alveolar nasals were perceived as separate in word-initial as well as intervocalic positions in nonsense words by native Malayalam speakers, which perhaps implies that native speakers of Malayalam can distinctly perceive the dental and alveolar nasals even in distributions not found in Malayalam.

On the other hand, apart from the geminate-occurring-medial position, the rest of the distributions in Malayalam for this contrast appear to be complementary to each other. The distribution for this contrast is described as [ŋ] occurring non-finally and [n] occurring non-initially (Kumari, 1972; Sreedhar, 1972). In addition, the orthographic script of the language does not distinguish between the dental and alveolar place contrast (Asher and Kumari, 1997).

For the purpose of this experiment, dental and alveolar nasals in Malayalam are assumed to maintain phonemic distinctions. Of particular interest for this experiment is the distribution of this place contrast in intervocalic position. Unlike the dental nasal, the alveolar nasal occurs in intervocalic position (Raja, 1960; Kumari, 1972; Sreedhar, 1972; Kalackel, 1985; Asher and Kumari, 1997). The training stimuli employ VCV as the only syllabic position. Therefore, the speech contrast [n] – [ŋ] results in being produced as alveolar nasal vs. retroflex nasal since intervocalically the nasal is realized as alveolar. This assumption does not conflict with Mohanan & Mohanan’s (1984) observation that Malayalam speakers can hear and reproduce the dental and alveolar nasals distinctly in an intervocalic position. The perception of these segments in new
phonetic contexts cannot be interpreted to mean that speakers can produce dental nasal intervocalically without any auditory prompt. Also, the reproduction of segments correctly in Mohanan & Mohanan’s observations may have been due to short-term memory effect. Therefore, for the current experiment the contrast alveolar – retroflex nasal was used as one of the training stimuli.

**Fricatives**

The voiceless fricatives in Malayalam do not have voiced counterparts. The fricatives in question, namely the lamino palato-alveolar [ʃ] and retroflex fricative [ʂ] were introduced in Malayalam through Sanskrit loan words (Asher and Kumari, 1997). However, these fricatives in Modern Malayalam have been so well integrated in the language that they’re considered native to the language (Kalackel, 1985; Sreedhar, 1972 etc). The place contrast of palato-alveolar – retroflex fricative ([ʃ] - [ʂ]) was employed as one of the training stimuli from Malayalam.

**Laterals**

Laterals were chosen as one of the stimuli speech contrasts due to the fact that the lateral approximants in Malayalam exhibit a two-way place contrast, alveolar and retroflex. This place contrast most appropriately measures up to nasal or fricative place contrasts in terms of homogeneity of place of articulation as compared to the rhotic place-manner contrast in the present experiment’s stimuli. Moreover, the retroflex lateral approximant is uniquely characteristic of Dravidian languages. This provides an opportunity to investigate the acquisition of this place contrast by Indo-Aryan languages including Bengali which do not maintain alveolar-retroflex lateral approximant place contrast.
Rhotics

Most researchers talk of another lateral/continuant sound which is considered unique to the Dravidian language family (orthographic symbol: Ɪ). However, there is disagreement to a large extent on its phonetic properties. It has been described in Malayalam as

Voiced retroflexed palatal fricitionised lateral [ɭ] (Kumari, 1972)
Retroflex lateral fricative (no contact at the post-alveolar region) [ɭ] (Sreedhar, 1972)
Voiced sublamino-palatal approximant [ɭ] (Asher & Kumari, 1997)
Voiced apico-prepalatal approximant [ɭ] (Kalackel, 1985)
Velar retroflex frictionless continuant [ɭ] (Raja, 1960)
Retroflex approximant [ɭ] or [z] (Bright, 1998; Krishnamurti, 2003)

In articulatory terms, Kalackel (1985) describes it as

The apex of the tongue is grooved as if for pronouncing /s/, but it is bent back so that the tip of the tongue is pointed at the root of the mouth where it is highest. The voiced sound which is thus produced is not actually a lateral (since the air escapes not over the sides of the tongue, but over the tip).

Clearly, a detailed phonetic description is required for this peculiar segment. For convenience, the term retroflex approximant along with the IPA symbol [ɭ] is used as reference to this particular speech sound in this study. This speech sound is contrasted with alveolar tap [ɾ] to create a paired contrast with rhotic properties.

The case of alveolar tap - retroflex approximant ([ɾ] – [ɭ]) speech contrast is more complex than other alveolar/palato-alveolar – retroflex speech contrasts, in that the speech sounds contrast in manner of articulation as well as place of articulation. This speech contrast is not homogenous with other speech contrasts chosen in terms of contrasting attributes between speech sounds.

In all, four speech contrasts were chosen from Malayalam to examine the acquisition of novel non-native speech contrasts by three listener groups. The novel speech sound in all of the contrasts is the one containing the retroflex place feature.
The following subsections deal with the three listener languages, Bengali, Spanish and American English. These languages will be discussed focusing on the presence or absence of retroflex features in their phonemic or phonetic segments.

**Bengali**

Bengali, an Indo-Aryan language, displays productiveness of the retroflex feature in manners like plosives and flap. It was chosen as one of the listener languages. Unlike Malayalam, Bengali lacks the alveolar – retroflex place contrast in manners like nasals, laterals and fricatives (Bhattacharya, 1988; Ferguson and Choudhary, 1960; Kostic and Das, 1972; Tunga, 1995; Ray et. al., 1966). These gaps in the consonant inventory make Bengali an appropriate choice of language to investigate the acquisition of alveolar-retroflex place contrasts in these manners. In addition, comparing this language group (Bengali-English bilinguals) with the rest of the listener language groups allows us to investigate feature generalization of the retroflex feature.

Bengali has a three-way place contrast in nasal stops, namely bilabial, alveolar and velar. Unlike north-western Indo-Aryan languages (E.g Hindi), Bengali has lost the retroflex place feature in the nasal manner of articulation, although the Bengali script retains the orthographic symbol for the retroflex nasal. Therefore, one phoneme /n/ is used for two graphemes (ফ, ঙ) (Bhattacharya, 1988). However, Bengali speakers have phonetic experience of the retroflex nasal as an allophone. When /n/ precedes a voiced retroflex stop in a consonant cluster, it undergoes assimilation and is realized as a retroflex nasal (Ferguson & Choudhary, 1960; Kostic & Das, 1972; Tunga, 1995). Although Polka (1991) and Strange (1995) noted that phonetic experience

---

2 Tunga (1995) noted that Midnapore dialect of Bengali (bordering the Oriya language area) uses the retroflex nasal [ɾ] more often than the dental nasal [n]. None of the subjects spoke this dialect.
of a sound can influence the perception of a non-native speech contrast, it remains to be seen if the retroflex feature with varying degrees of retroflexion can have similar effect due to exposure in allophonic experience.

The degree of retroflexion in Indo-Aryan languages has been observed to be much less than that of Dravidian languages (Ladefoged and Bhaskararao, 1983). The retroflexion in Bengali is claimed to have a slighter degree of retroflexion than Malayalam (Chatterji, 1921; Kostic and Das, 1972; Haldar, 1986; Tunga, 1995).

<table>
<thead>
<tr>
<th>Table 2-2. Phonemic consonant inventory of Bengali</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labial</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Stops</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
</tr>
<tr>
<td>Nasals</td>
</tr>
<tr>
<td>Tap/flap</td>
</tr>
<tr>
<td>Lateral approximant</td>
</tr>
</tbody>
</table>

* Adapted from Bhattacharya, K. (1988)

The fricatives in Bengali have similar story. Table 2-2 displays the phonemic inventory of Bengali. Phonemically, Bengali displays palato-alveolar [ʃ] and dental [s] fricative. The palatal fricative occurs in all syllabic positions as well as intervocally as geminates. The dental fricative however, can only be realized in consonant clusters positioned word-initially or occurs in loan words (Bhattacharya, 1988; Ferguson and Choudhary, 1960; Kostic & Das, 1972; Ray et. al., 1966; Haldar, 1986). Thus, the palatal place contrast is the most productive in fricative manner of articulation. Bengali has lost the retroflex [ʂ] fricative originally derived from Sanskrit. However, as with the nasals, the Bengali script retains the orthographic symbol for the retroflex fricative. There are three graphemes (শ,ষ,স) for a single phoneme /ʃ/.
Bengali maintains a single place in the case of lateral approximants, alveolar /l/ (orthographic symbol: ল). It lacks the retroflex place of contrast in laterals as a distinctive segment. However, Bengali speakers do experience retroflex lateral [ɭ] as an allophone of /l/ in certain environments. Alveolar laterals preceding retroflex consonants undergo assimilation and are realized as retroflex lateral approximants, especially in consonant sequences like /ɭɾ/ or /ɭd/ (Bhattacharya, 1988; Ferguson and Choudhary, 19603; Tunga, 1995).

Rhotics in Bengali have a two-way place contrast, namely alveolar /ɾ/ and retroflex /ɿ/ flap. The alveolar flap occurs in medial and word-final positions. Alveolar trill [ɾ] is realized as an allophone of this phoneme when it occurs word-initially. Ferguson & Choudhary (1960) observed that sometimes an alveolar flap occurring after a labial or retroflex stop gets realized as an approximant (like English [ɿ]). Similarly, the retroflex flap when occurring in final position or followed by a consonant sometimes is realized as a prolonged retroflex continuant. Therefore, Bengali speakers may have allophonic experience of retroflex approximant as well. Nevertheless, none of the Malayalam retroflex speech sounds being tested for acquisition occur phonemically in Bengali.

Spanish

Spanish, one of the Romance languages, uses the palatal place of articulation extensively. Most Spanish dialects allow manners of articulation like nasal, fricatives, approximants and laterals at the palatal place of articulation (Hammond, 2001). This language was chosen as another listener language involving the Spanish-English bilingual group, based on the

3 Ferguson & Choudhary (1960) also claims that a sequence like/ɾɭ/ (where alveolar lateral follows a retroflex consonant, both consonants result in a long retroflex lateral [ɭ]).
nonexistence of the retroflex feature in its phonemic as well as phonetic segments. The lack of experience with retroflexes makes the Spanish speaking subjects ideal to test the effects of meta-linguistic awareness since any kind of linguistic experience with retroflex is controlled in this language group. Table 2.3 shows phonemic consonant inventory of Spanish for the relevant manners of articulation, namely rhotics, nasals, approximants and lateral approximants.

Table 2-3. Phonemic consonant inventory of Spanish (Only manners that are relevant for the study are displayed here.)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labio-dental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td></td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>[β]</td>
<td>f</td>
<td>θ [ð]</td>
<td>s</td>
<td>x [γ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td>w (labio)</td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td>k</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Adopted from: Hammond (2001)

Rhotics in Spanish are considered another unique attribute of Romance languages. The alveolar trill is phonemically contrasted with the alveolar tap (Lindau, 1985). The alveolar trill is not found as a phonemic sound in either English or Bengali. In the case of the alveolar tap, while being a phonemic consonant in Bengali, it occurs only in phonetic relationship with alveolar /t/ and /d/ stops in American English (Roach, 1983; Olive et. al., 1993; Ladefoged and Maddieson 1996).

Nasals as well as lateral approximants in Spanish have a two-way place contrast described as alveolar and palatal. This contrast is prevalent in most of the South American dialects of Spanish. In the case of fricatives, most of the Spanish dialects maintain the alveolar place contrast in sibilants and do not have an equivalent of palatal or palato-alveolar fricative; the exception being Argentinean Spanish where in the fricative manner a palatal allophone also occurs along with the alveolar one.
There appears to be complete lack of retroflex features in the sound inventory of Spanish whether phonemic or phonetic. Therefore, the acquisition of retroflex sounds or specifically the alveolar-retroflex place contrasts from Malayalam may pose a degree of difficulty, attributed to acquisition of new non-native speech sounds, for the speakers of Spanish.

**American English**

American English was included as the third listener language for the experiment. This listener group (monolingual American English group) acted as the control group for the experiment. English employs the alveolar place of articulation extensively to create phonemes in different manners of articulation. Table 2.4 shows the phonemic consonant inventory of American English of the relevant manners of articulation.

Table 2-4. Phonemic consonant inventory of American English (Only manners that are relevant for the study are displayed here.)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labio-dental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>m</td>
<td>v</td>
<td>s</td>
<td>z</td>
<td>ζ</td>
<td>η</td>
<td>n</td>
<td>h</td>
</tr>
<tr>
<td>Fricative</td>
<td>θ</td>
<td>δ</td>
<td>ѵ</td>
<td>ζ</td>
<td>j</td>
<td>ι</td>
<td>j</td>
<td>w(labio)</td>
</tr>
<tr>
<td>Approximant</td>
<td>l</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
*Adapted from: Ladefoged (1999)*

The lateral approximant in American English employs only the alveolar place. The phoneme /l/ has a velarized lateral approximant allophone, syllable-finally. Nasals in American English have a three-way place contrast namely, bilabial, alveolar and velar. The post-alveolar fricative, similar to palato-alveolar fricative in Malayalam, does exist in American English. However, the language lacks any existence of retroflex fricative.

The alveolar tap, while being a phonemic consonant in Malayalam, occurs only in a phonetic relationship with alveolar /t/ and /d/ stops in American English (Ladefoged and Maddieson, 1996). In the case of retroflex feature exposure, unlike Spanish, English speakers
may have some phonetic experience with the retroflex feature in incidental contexts. Polka (1991) noted that English speakers may have some experience of retroflex stops at the phonetic level. The alveolar stops /t/ and /d/ may be produced as retroflex stops /tᵢ/ and /dᵢ/ when the alveolar stops precede /ʌ/ in words like “try” and “dry”. However, as noted by Polka (1991), this is not a predictable phonological process and may vary on a dialectal or individual level. Also, it should be noted that the consonants in question (/l, n, r, sh/) do not precede the rhotic approximant in syllable-initial position in American English. Therefore, AE speakers do not have the experience of their becoming retroflexed as is the case with /t, d/.

**Stimulus Materials**

The novel non-native speech contrasts for the study were taken from Malayalam. It was chosen as the target language as it has a rich inventory of consonants employing the retroflex feature extensively in various manners. This allows a flexible choice of speech contrasts as per the requirements of the study (as discussed below).

The speech contrasts chosen for the study are:

1. [l - l] (alveolar versus retroflex lateral approximant)
2. [n - n] (dental/alveolar versus retroflex nasal)
3. [f - s] (palato-alveolar versus retroflex fricative)
4. [r - ɾ] (alveolar tap versus retroflex approximant)

These speech contrasts differ primarily in place of articulation, although the [r - ɾ] contrast also differs in manner. The reason for choosing these particular speech contrasts, when looking at the Bengali-English (BE) language group, is primarily for the purpose of evaluating feature generalization. One of the speech sounds in each speech contrast from Malayalam is marked with the retroflex feature [l, n, s, ɾ]. These sounds are not found in Bengali which employs the retroflex feature but only in the oral stop and flap manners of articulation.
Therefore, if the retroflex feature gets generalized from existing phonetic categories to newly formed phonetic categories, the evidence of any variability in generalization of the feature in different manners will be reflected during the acquisition of these four new speech contrasts.

In case of the monolingual language group, American English (AE) speakers do not have experience with retroflex sounds phonemically. However, retroflexion of alveolar stops in particular environmental contexts (e.g. [dɪˈʌl] ‘dry’ (Polka 1991)) is found in the speech of American English speakers. Spanish-English (SE) bilinguals have no experience with the retroflex feature in phonemic or phonetic environments when using Spanish. However, being bilingual, they are assumed to have phonetic experience similar to American English speakers. Having no phonemic exposure to retroflexion makes the Spanish-English bilingual group ideal to examine the multilingual benefit.

The materials were recorded as natural stimuli by Malayalam talkers producing each of the eight speech sounds from the four speech contrasts discussed above. Each speech sound was embedded between two identical vowels (vowel contexts: [a, i, u]) resulting in a set of non-words. Six repetitions of each consonant with each vowel context were recorded in order to use the best possible tokens for the stimuli.

Recording for the stimuli was done in India (Chennai, Chandigarh, Bangalore, Cochin, and Ahmadabad). Currently, all speakers that were recorded have been residing out of the state of Kerala. Ten native speakers of Malayalam were recorded (See figure 2-1 for their region distribution). However, due to errors in elicitation, two speakers (M03 and M05) were not considered for this study. Detailed information of the Malayalam speakers is provided in table 2-5.
The data were recorded and digitized at 44.1 KHz on a Marantz Digital recorder and transferred directly onto a personal computer. The recordings took place in a quiet room with no significant reverberation or background noise. The subjects were provided with the non-words (in Malayalam script) on the computer with an inter trial interval (ITI) of 2.5 seconds. The talkers were asked to read through the list before the recording in order to be familiar with the non-word materials. The data were segmented using PRAAT speech analysis software.

Table 2-5. Talker information; all talkers are native speakers of Malayalam and born in the state of Kerala, India; Years= “years not lived in their home state”.

<table>
<thead>
<tr>
<th>Talker Code</th>
<th>Gender</th>
<th>Age</th>
<th>Years</th>
<th>Other Languages known</th>
</tr>
</thead>
<tbody>
<tr>
<td>M01</td>
<td>F</td>
<td>28</td>
<td>4</td>
<td>Kannada, English</td>
</tr>
<tr>
<td>M02</td>
<td>F</td>
<td>37</td>
<td>1</td>
<td>Hindi, English</td>
</tr>
<tr>
<td>M04</td>
<td>M</td>
<td>45</td>
<td>15</td>
<td>Hindi, Punjabi, English</td>
</tr>
<tr>
<td>M06</td>
<td>F</td>
<td>36</td>
<td>9</td>
<td>Hindi</td>
</tr>
<tr>
<td>M07</td>
<td>M</td>
<td>28</td>
<td>1</td>
<td>English, Tamil</td>
</tr>
<tr>
<td>M08</td>
<td>M</td>
<td>29</td>
<td>5</td>
<td>Hindi, English</td>
</tr>
<tr>
<td>M09</td>
<td>M</td>
<td>32</td>
<td>8</td>
<td>Hindi, English</td>
</tr>
<tr>
<td>M10</td>
<td>M</td>
<td>29</td>
<td>5</td>
<td>Hindi, English</td>
</tr>
</tbody>
</table>
Stimulus Material Evaluation

Before developing the stimulus material for the main study, the recorded data from the Malayalam talkers was evaluated. Native Malayalam listener ratings were elicited for tokens from eight talkers. Four native speakers of Malayalam evaluated the intelligibility of the recorded stimuli through a consonant identification task. Tokens that were identified consistently as the target sound 83% - 100% were selected for the stimuli. Three out of the selected tokens, closest in duration and mean fundamental frequency, were selected for each consonant with all three vowel contexts, making a total of 72 stimuli (8 consonants x 3 vowel contexts x 3 tokens) per talker.
Stimuli Evaluation: Preliminary Experiment

Conducting the main experiment would have been unproductive for the purpose of evaluating the role of assimilation if ceiling or floor effects were observed for any of the speech contrasts in the monolingual language group (AE). Therefore, to avoid this possibility, a preliminary experiment was conducted to examine the discrimination level of these contrasts by monolingual American English speakers since the AE monolinguals acted as the control group in the main experiment, against which the results of the bilingual groups are to be compared. Thus, it is considered reasonable to test the difficulty level of the stimuli with the control group.

The goal of the preliminary experiment was to examine the level of discriminability of the novel non-native speech contrasts by American English speakers. Eleven monolingual American English subjects were recruited for this experiment. All were undergraduates at the University of Florida. They were compensated with extra credit or paid cash ($5/40 minutes). None of them reported average proficiency in any language other than American English.

For stimuli, two tokens of each non-word (8 consonants x 3 vowel contexts) per talker were used. Data from three talkers was used. In all, there were 1152 stimuli pairing up to make 576 trials (16 trials per speech contrast – different trials, identical trials and physically different trials).

Subjects were given a categorical AX discrimination task. (Please refer to the section later in this chapter for a description of an AX discrimination task.) Because physically different tokens of the stimuli are used, the discrimination task has to be administered as a categorical task. Naturally produced stimuli may have subtle irrelevant acoustic differences which may influence subjects’ decision away from category-level processing. Therefore, the subjects are asked to decide whether both stimuli sound like they are from same category or different
category. For this task, the inter stimulus interval (ISI) was provided as 1 sec and inter trial interval (ITI) was 4 sec.

The results of the preliminary experiment were as follows. Each subject’s performance in the AX test was converted to a d’ score (scale of 0 to 4). Table 2.6 shows the results of the discrimination task for the American English speakers.

Table 2-6. Result of the preliminary AX discrimination task across three Malayalam speakers. Mean D prime values averaged across contrasts as well as across vowels.

<table>
<thead>
<tr>
<th></th>
<th>[l-l]</th>
<th>[n-n]</th>
<th>[r-r]</th>
<th>[f-f]</th>
<th>Average across contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a]</td>
<td>1.09</td>
<td>2.55</td>
<td>2.99</td>
<td>2.70</td>
<td><strong>2.33</strong></td>
</tr>
<tr>
<td>[i]</td>
<td>0.94</td>
<td>1.11</td>
<td>1.69</td>
<td>2.28</td>
<td><strong>1.51</strong></td>
</tr>
<tr>
<td>[u]</td>
<td>1.37</td>
<td>1.25</td>
<td>2.89</td>
<td>1.09</td>
<td><strong>1.65</strong></td>
</tr>
<tr>
<td>Average across vowels</td>
<td>1.13</td>
<td><strong>1.64</strong></td>
<td>2.53</td>
<td>2.02</td>
<td></td>
</tr>
</tbody>
</table>

The [r-r] contrast averaging over all three vowel contexts was the most discriminable of the four speech contrasts (d’ value 2.5). Overall, the [l-l] speech contrast was least discriminable (1.1) followed by the nasal contrast [n-n] (1.6) and fricative contrast [f-f] with d’ value at (2.02). Among the vowel contexts, the [i] context was least discriminable (1.5) closely followed by [u] (1.6) whereas [a] was the most discriminable (2.3). Across the four speech contrasts, the [i] vowel context was the least discriminable (with the exception of fricative [f-f] contrast where [u] was the least discriminable) and [a] was the most discriminable. No ceiling or floor effects were found for any of the non-native speech contrasts.

**Stimulus Preparation**

Based on the results of the preliminary experiment as well as native listener ratings, the stimulus materials for the perception training experiment were selected. For each talker, three physically different tokens per consonant and vowel context (8 consonants x 3 vowel contexts x
3 tokens) were chosen. The tokens of one talker (M08) were used solely for the familiarization task. Talker M04 was used for pretest and posttest as well as discrimination test and perceptual assimilation test, since all of these tests have to be identical in stimuli and procedure in order to be directly compared. M04 was also included during training since the testing (pretest and posttest) is on a trained talker. M07 was used for the generalization test only. This would be the novel talker, whose voice was not exposed to the subjects anytime during the pretest or training phase. Data of six talkers (M01, M02, M04, M06, M09, M10) was used for training sessions.

- Familiarization phase: one talker (8 consonants x 3 vowels x 2 tokens x 2 repetitions) 96 trials in all.

- Pretest/posttest phase
  - Pre-test/post-test: one talker (8 consonants x 3 vowels x 2 tokens x 6 repetitions) 288 trials in all.
  - Discrimination test: one talker (8 consonants x 3 vowels x 2 tokens x 4 orders) 192 trials in all.
  - Perceptual assimilation test: one talker (8 consonants x 1 vowel x 2 tokens x 5 repetitions) 80 trials in all.
  - Production pretest/posttest: one talker (8 consonants x 3 vowels x 1 token x 6 repetitions) 144 trials in all.
  - Generalization test: one talker (8 consonants x 3 vowels x 2 tokens x 6 repetitions) 288 trials in all.

- Training phase: two talkers per session (8 consonants x 3 vowels x 3 tokens x 2 repetitions) 288 trials per session. Over a total of six training sessions, data from six talkers is used.

**Sound-symbol association issue:** The issue of sound-symbol association was foreseen in the design of the training experiment. Since the subjects (all three language groups) are not familiar with Malayalam orthography or vocabulary, it would be difficult to train them without sound-symbol association. This kind of issue has not arisen in the previous studies (Logan et al., 1991 – learning of /r/-/l/ distinction by Japanese learners of English) as the subjects were familiar
with the vocabulary and alphabets of the target language. In the current study, the question of orthography associated with the 8 consonants was resolved by using arbitrary symbols for each consonant. The arbitrariness of the symbols being associated with the sounds is inevitable since these are novel non-native sounds for the listeners.

A number of alternate methods of representing the sounds were considered before appropriate symbols were finalized. Firstly, using the original Malayalam script or IPA symbols with the sounds would have taken the subjects much more time and effort for the sound and symbol association resulting in decreased focus on the perception of sounds. Secondly, purely arbitrary geometric shapes as symbols were dismissed as potential candidates for the same reason as above. In addition to that, the confounding possibility of learners perceiving the sounds correctly but clicking on the wrong geometric shape could be detrimental to the validity of the results. The third alternative was using the English alphabets as representative symbols. The option was considered as the subjects in all the language groups were familiar with English alphabets and their corresponding English consonants. The small and upper case of the alphabet ‘n’ could be used for the \([n] – [\eta]\) sound distinction so that whenever subjects hear a nasal, there would not be the basic confusion of finding the correct symbol out of the 8 symbols, as would be the case with purely arbitrary representations. One caveat with this type of representation could be that the subjects may not easily learn the distinction between the small and upper case letters representing the alveolar-retroflex distinction. Therefore, to avoid this possibility, all dental/alveolar sounds in the four sets of contrasts that were presumed to be assimilated to their L1/L2 category were assigned the small case letters. The retroflex sounds in all the four sets of contrasts were assigned the corresponding upper case letters. In addition to this, during familiarization process, the sounds in each set were auditorily prompted (along with their
corresponding symbol) next to each other. The subjects were told that the sounds they will hear are not English sounds but from a language not known to them.

After careful consideration, the orthographic representation of the sounds was finalized as small and upper case English alphabets. Hence, the study uses the following symbols in training:
l= dental lateral approximant; L=retroflex lateral approximant; n=dental nasal; N=retroflex nasal; sh=palato-alveolar fricative; S=retroflex fricative; r=alveolar tap and R=retroflex approximant. These arbitrary symbols were familiar to all three language groups and their efficacy was verified in pilot testing.

Subjects

The study employed a diverse range of challenging non-native consonant contrasts from Malayalam with three populations of listeners,

- Bengali-English (BE) bilingual group
- Spanish-English (SE) bilingual group
- American English (AE) monolingual group

Twenty (20) participants were recruited for each group from the University of Florida campus. At the time of the experiment, the participants, ages ranging from 18 to 35, were enrolled in undergraduate or graduate courses at the university.

For all subjects, the level of proficiency was self-reported through a language background questionnaire. Self-reported language proficiency may not be an accurate measure for screening of potential subjects for the study. Therefore, certain precautionary measures were taken while recruiting the subjects for the three language groups. In the case of American English monolinguals, the experimenter asked for specific information regarding the knowledge of other languages or extent of exposure to other languages apart from English (L1). The preliminary (introductory) email sent out to the interested individuals has been provided in Appendix A. If
the candidate had taken Spanish language classes in middle/ high school, which was mostly the case, a brief test of spoken Spanish was conducted to assess the level of L1 transfer from English. The procedure and content of the spoken Spanish test has been explained in Appendix A. In the case of Bilingual Spanish-English subject recruitment, in addition to conducting the language background questionnaire, the experimenter assessed the level of proficiency in spoken American English. These individuals were also asked to take the spoken Spanish test. The screening for bilingual Bengali-English subjects involved assessing their level of spoken proficiency in English by observing any noticeable signs of L1 transfer in consonantal segments during conversation. More specifically, the Bengali-English participants were screened for any exposure to Dravidian languages and if present then to what extent. Candidates with even minimal awareness of retroflex sounds were not recruited for this study.

During the screening process, bilinguals, preferably early bilinguals with an age of acquisition (AOA) for L2 of fewer than 12 yrs, were considered as potential subjects. If age of acquisition (AOA) for L2 was more than 12 yrs, then candidates with self-reported proficiency level of L2 as HIGH were considered as potential subjects. Specifically for Bengali-English subjects, the English refers to Indian English since most of the student population is graduate students staying in U.S no more than 5 years. Subjects’ proficiency in English was self-reported and also assessed by the experimenter.

For the monolingual subjects, self reporting of any language exposure was required including conversational ability in another language or passive exposure to another language. Subjects with very minimal knowledge of another language (mostly Spanish from High School language classes< 2 years) were considered in the monolingual American English language group. It was almost impossible to find pure monolingual subjects with no exposure to another
language especially in the age group of 18-35 years due to the educational curriculum in North American Schools. None of the participants had any prior exposure to Malayalam.

In addition to this, subjects were asked in the screening questionnaire (Appendix A) of any hearing or speech problem that they may have had in the past. All subjects reported normal hearing with no hearing or speech impairment history. No preliminary clinical screening for normal hearing was conducted. Therefore, the information collected on normal hearing was self-reported.

**Bengali-English Bilinguals**

The Bengali-English language group was chosen for their phonemic experience with the relevant place feature (retroflex in oral stop and flap manners of articulation), though Bengali and English lack any direct correspondents to the Malayalam contrasts under study. The results from the BE language group when compared with Spanish-English bilinguals were expected to provide evidence for the feature generalization factor in the learning of a third language.

Twenty (20) speakers of Bengali with English as L2 participated in the study. The mean age of participants in this group was 28 years. All subjects were from Kolkata, the capital of West Bengal state with the exception of two subjects. BE16 and BE17 are from Barddhaman, a southern district in West Bengal, close to Kolkata. They speak the standard colloquial Bengali, similar to what is spoken in Kolkata. Most of the subjects have acquired a working knowledge of Hindi through entertainment media, but only four subjects, BE03, BE05, BE11 and BE12, have had formal education in Hindi. Some subjects have been exposed to languages like Assamese, Oriya, Urdu, Nepali and Telugu through the surrounding environmental context from a few months up to two years. However, even low proficiency in comprehension or speaking in these languages was not reported. Therefore, any influence from the above mentioned languages can be ruled out. The details of all the subjects have been provided in the appendix A.
Spanish-English Bilinguals

The Spanish-English (SE) bilingual group was chosen for their lack of linguistic experience with retroflex place feature. The results from SE language group when compared with other groups were predicted to provide evidence for effects of multilingual benefit in bilinguals.

Twenty (20) speakers of Spanish with English as their L2 or another L1, participated in this study. The mean age of subjects within this language group was 20 years. All subjects have lived in North America all or most of their lives. In cases where subjects were born outside of North America, they had shifted to North America within the first 6 years of their lives. Therefore, these bilinguals acquired Spanish at home as their first language and picked up American English either simultaneously with Spanish or during preschool years (age 3 to 6). The dialects of Spanish spoken by these subjects are of Central American origin with the exception of one subject who spoke Argentinean Spanish. Figure 2-2 shows the origins of the Spanish dialects of the SE bilingual subjects. The table providing detailed information on individual subjects is given in appendix A.
Figure 2-2. An outline map of Mexico, Central America and South America showing the regional distribution of the subjects.

**American English Monolinguals**

American English speakers were recruited for the monolingual language group acting as a control group for the experiment. The subjects (20.2 years as the average age within this language group) acquired American English as their first language. None of the subjects reported more than minimal proficiency in any language other than American English. Since most of the
students had lived most of their lives in Florida, many of them had to take Spanish as a required language class. This factor could not be controlled for therefore, subjects were asked to read out two brief tongue twister phrases in Spanish. This was done to determine if the subject had actually acquired any perceptual categories corresponding to Spanish language. Only those participants were recruited as subjects whose speech did not provide any evidence of Spanish categories but in fact displayed transfer from American English. The Spanish phrases as well as detailed information on individual subjects are provided in Appendix A.

**Experimental Design**

**Perception Training Method**

The present experiment used a perception training method with identification paradigm containing high variability stimulus (High-variability perception training). This training emerged as an effective training method in learning non-native speech contrasts through the academic disagreement between two schools of thought regarding the mental representation of phonetic sounds in the minds of L2 learners: the prototype model (Strange and Dittmann, 1984; Jamieson and Morosan, 1986, 1989; Kuhl, 1991, 1992) and the exemplar-based model (Jamieson and Morosan, 1986; Logan et al., 1991; Lively et al., 1993, 1994; Bradlow et al., 1997). In order to comprehend the reason for adopting High-variability perception training method for this study, it is imperative to discuss the theoretical assumptions underlying these two models of long term memory representation.

The earlier school of thought, the prototype model, claimed that the phonetic sounds acquired by a listener were represented as context-invariant units of sounds, the content of which was the ideal example of that phonetic sound, otherwise named as the prototype, or in a phonologically-driven view, phoneme. Every representation of a phonetic category in the mind of an L2 learner was assumed to have foci, the most representative member of the category
The foci was expected to have more stable representations in long-term memory, since only linguistically relevant information was stored in that prototype and could overlap concrete natural members of that category when perceived. However, empirical testing of this training type failed to prove effective in developing long term memory representation since the subjects in these studies failed to generalize to new phonetic environments. A limitation in the notion of prototypical representation was that it was assumed that only linguistically relevant information will be extracted to form the representation of the phonetic category. However, by postulating this, the model assumed that other detailed information in those stimuli, such as talker’s voice, speaking rate or the phonetic environment, is filtered out. This detailed information was established to be of relevance in the representation of phonetic categories by subsequent empirical results (Logan et al., 1991; Lively et al., 1993).

The results from Logan et al. (1991, 1993), Lively et al. (1994) and Bradlow et al. (1997) clearly indicated that for generalization to occur, detailed information like talker variability and phonetic environments are as relevant as other linguistically significant information. The perceptual learning and development of stable representations of phonetic categories result from extensive stimulus variability in natural stimuli. This led to the formulation of the exemplar model of phonetic representation, extending the exemplar-based model of categorization (Nosofsky, 1986, 1987; Kruschke, 1992) based on the selective attention weighing mechanism. Multidimensional representations are stored in the memory during training while selective attention weighs the importance of various stimulus dimensions. Changes in selective attention ‘stretch’ or ‘shrink’ the perceptual space for these dimensions and alter the internal category structure resulting in phonetic categories being less similar to other members as dimensions are
stretched and more similar to other members when dimensions are shrunk (as cited in Lively et al., 1993).

The choice of training method for the current research is clear from the above discussion. The present experiment used a perception training method with identification paradigm containing high variability stimulus (High-variability phonetic training). The choice of perception over production or pronunciation training is also based on the debate of mental representation of phonetic categories in long-term memory. Since production/pronunciation training is based on providing the learners with the ideal or best exemplar of the phonetic category, the input lacks any variability at all (Dalby and Kewley-Port, 1999; Akahane-Yamada et al., 1996; Neri et al., 2002 etc).

Procedure

A pretest-training-posttest design was employed, adapting the procedure used by Lively et al. (1993, 1994). The effect of training was measured by comparing the performance in pretest and posttest tasks. A consonant identification procedure was used throughout the training phase of the experiment. Apart from the consonant identification task, an AX discrimination task and perceptual assimilation task were also employed during the pretest and posttest phase of the experiment. A minimum of seven-hour gap was used between sessions to avoid any confounding effects of working memory influencing the perceptual performance before and after training and also to prevent any adverse effects of fatigue during consecutive training sessions. The term working memory, in this case, refers to immediate retrieval of information stored in the short term memory, as a result of preceding training. On the contrary the results would be more valid if the subject’s response is based on their long term memory representation of sounds. This arbitrary period of seven hours as the minimum gap between two consecutive train sessions was considered assuming that the routine task of the individuals would negate any working memory
effects. All training and testing took place in a sound-attenuated room equipped with individual computer stations (contains a keyboard, a CRT monitor and headphones) for subjects. Stimulus materials were presented to subjects over a set of calibrated head phones. The software collected individual responses during all phases of the experiment. All subjects were tested and trained individually.

**Familiarization**

Subjects were familiarized with the symbols associated with each consonant. The talker used for this task hyper-articulated the sounds during recording and, thus, was considered ideal for the familiarization task. The stimuli consisted of one talker (8 consonants x 3 vowel contexts x 2 tokens x 2 repetitions). In all, there were 96 trials and the ITI was 2 seconds. The sound file was played while the symbol was displayed on the computer screen. The subjects were asked to pay attention to both the speech stimulus and its corresponding symbol. They were told that the words are from a language not known to them and the arbitrary symbols represent different sounds. The non-words contrasting in place (eg. [ili] vs [iLi] ) were presented within same vowel context alternating between the two words for maximum impact of familiarization task.

**Pretest**

During the pretest phase, the perceptual performance of the non-native stimuli was determined at the naïve listener level through various tasks. Pre-testing of perception (consonant identification task) was incorporated to test the level of accuracy in identification of the target sounds. The stimuli for the pretest (identical task also given as posttest) consisted of one talker (8 consonants x 3 vowel contexts x 2 tokens x 6 repetitions), whose data is also used for the training (Trained-on talker). In all there were 288 trials. The pretest consisted of a consonant identification task presenting the randomized stimuli over the headphones. The sound file was played and the response options became available for the subject to click on. The response
options provided were all of the eight symbols corresponding to the eight non-native sounds in question, thus technically making 12.5% as chance level though no confusion between the manners is expected since all L1s have dental/alveolar nasals, laterals and rhotics. The subject had the option to change her/his response before going on to the next trial. Once the next trial button was pressed, the answer was recorded by the software and no changes could be made (same for all tasks). The task was self-paced and therefore no ITI was assigned. No correct answer feedback was provided during pretest.

A categorical AX discrimination task was used to measure listener sensitivity to non-native speech contrasts at the naïve exposure level (pretest). The same task was administered during the posttest phase so as to measure the perceptual performance on the same stimuli after training.

An AX discrimination task tests the subjects’ level of perceptual performance on a given stimulus set. Each trial consists of a set of two stimuli (either category-different or category-same). Subjects are required to decide if the stimulus X (the second stimulus they hear) is the same or different than stimulus A (the first of the stimuli set). The decision of similarity or difference is required to be made on the basis of perceptual categories (phonemic sound, word etc).

Two types of trials can be used in such discrimination task: physical identity trials and categorical trials. The physical identity trials only pair up the same token for same trials so that the trial represents the same exemplar of the speech sound. This provides insight into the perception of purely physical properties of the sounds. On the other hand, categorical trials, along with using physical identity trials, also employ the physically different tokens of the same speech sounds as pairs for the same trials. This provides insight into the perceptual processing at the phonemic level. For the discrimination task in this study, categorical trials were employed.
Twelve possible trials per speech contrast are usually employed. An example of these possible trials is given in figure 2-3. In this test, the two same trials are repeated in order to maintain equal number of same and different trials and therefore the number of same trials become four making a whole set of 16 trials per speech contrast.

Inter stimulus interval (ISI), the time gap between the presentations of the stimuli in a single trial, is also important in defining the difficulty level of the task. The longer the ISI, the more difficult it is for the subject to rely the response only on working memory. Thus, phonemic level processing comes into play and the longer ISI tests the perception at a categorical level.

<table>
<thead>
<tr>
<th>Same trials</th>
<th>A1</th>
<th>A2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A2</td>
<td>A1</td>
<td>X2</td>
<td>X1</td>
</tr>
<tr>
<td>Different trials</td>
<td>A1</td>
<td>X1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1</td>
<td>X2</td>
<td>X1</td>
<td>A1</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>X1</td>
<td>X1</td>
<td>A2</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>X2</td>
<td>X2</td>
<td>A2</td>
</tr>
</tbody>
</table>

Figure 2-3. Different trials possible in an AX discrimination task using two tokens of each speech sounds.

Another task known as the perceptual assimilation task was administered during the pretest phase to examine the similarity of the speech sounds presented as stimuli when compared with the listener’s native speech sounds. This test provides an insight into listener intuitions of classifying the non-native speech sounds according to the already established phonetic categories of the native language(s). In this study, the listeners were asked to listen to each stimulus sound (provided with only one vowel context) and write down the perceptually closest speech sound (in the native language orthography) in her native language(s). In case of bilingual groups, the subjects were allowed to choose the response (the closest sound) from both L1 and L2. They were asked to write down the alphabet(s) from their native language corresponding to the sound heard and specify the language (especially in the case of Spanish-English bilinguals since the
writing system of both English and Spanish overlaps substantially). This assimilation task was an open-set test, where the listener was not bound by a few response options to choose from but was allowed to choose from her past native linguistic experience.

Along with that, for each trial, listeners were asked to provide a category goodness rating of how similar the native speech sound is to the newly heard stimulus sound. The goodness rating scale consisted of 1 to 7 levels where 1 was to be chosen when the target sound was ‘very different’ from the sound closest to it (according to the listener). Option 7 was to be chosen when the target sound was ‘exactly same’ or in other words highly similar to the sound closest to it (modal response for the token, according to the listener). Listeners were able to specify the degree of difference between the sounds by circling a number between 1 and 7. The same task was administered at the both the pretest as well as the posttest phase of the experiment in order to evaluate any changes in assimilation patterns as an indication of developing new phonetic categories. In all, the pretest phase consisted of a consonant identification test, an AX discrimination test and a perceptual assimilation task.

**Training**

Subjects began the 6-session training phase after the pretest phase. The training procedure used the same consonant identification task. The difference between the pretest task and the training task is that feedback is given on every trial during the training sessions. If the response was correct, the message ‘you are correct!’ was displayed. After 500 ms interval, the sound file was replayed along with the correct symbol for reinforcement of the association. If the response was incorrect, the message “Incorrect!!! Please listen to it again” was displayed. After 500 ms interval the sound file was replayed and the correct symbol was displayed for reinforcement. The subjects were not required to respond during repetition. The task was self-paced with no ITI between trials. Data gathered from six talkers was used as stimuli for the training portion of the
experiment. Subjects heard only two talkers per training session. The sequence of presenting talkers remained the same throughout training. Therefore, subjects heard each talker two times during the 6-day training phase. The stimuli contained (8 consonants x 3 vowel contexts x 3 tokens x 2 repetitions) per talker. In all, there were 288 trials (2 talkers per session). The task was paused for 60 seconds after one block of trials (stimuli of one talker). The subjects were allowed to take breaks for longer time if required.

Posttest

A final posttest was given at the end of the training phase. The posttest consisted of the consonant identification task, discrimination task and perceptual assimilation task (all three identical to pretest). Another similar identification task was administered but with untrained stimuli and one talker (novel talker) to measure the generalization (generalization test) of the training to novel stimuli with new talker voice information. The design of the generalization test was identical to the consonant identification test except that the stimuli were taken from a novel untrained talker.

Task Order Presentation

The order of presentation for the various tasks during the pretest phase was kept same for all the subjects across language groups. Once the familiarization task was over, the pretest phase began with the consonant identification pre-training test as the initial task. It was followed by the AX discrimination test. After a brief rest for five minutes, the subjects were given the perceptual assimilation task. The presentation of the stimuli in each of the tests was identical for all the subjects.

The consonant identification test, AX discrimination test and the perceptual assimilation task that were conducted at the pretest level were repeated at the posttest level of the experiment as well. In addition to that, a generalization test was also conducted during the posttest phase.
The presentation order of these tasks was as follows: consonant identification post-training test followed by the AX discrimination test and the perceptual assimilation task. Towards the end, the generalization test, where perceptual performance on stimuli from novel talker was assessed, was conducted.

**Data Analysis**

The results of the pretest and posttest phase were analyzed across contrast and vocalic context as well as across language groups. The mean percentage scores for the consonant identification tasks and d prime scores (discussed further in detail) for the discrimination tasks were obtained. Results from the generalization test were analyzed across groups and individually when compared with posttest results. ANOVAs were conducted separately for each contrast to test the hypotheses under consideration.

The calculation of d prime scores and perceptual assimilation types are discussed herewith. For the analysis of discrimination test results, the testing software captured ‘1’ if the subject clicked on the response item ‘same’ and ‘2’ if the response item was ‘different’. These responses were then converted to a binary code, termed for this study as signal present. Hit rate, which means that the subject correctly identifies the distinction within a contrast, was calculated by averaging the signal present for all the ‘different’ trials. False alarm rate, which means that the subject has marked all responses the same regardless of the trial type (either same or different), brings out such discrepancies in the result and helps in accurately obtaining the d prime score. The false alarm rate was calculated by averaging the signal present for all the ‘same’ trials. Based on these, the d prime score was then calculated by subtracting the NORMINV value of false alarm rate from that of the hit rate. NORMINV formula returns the inverse of the normal cumulative distribution for the specified mean and standard deviation. Once the d prime values were obtained for the pretest discrimination test as well as the posttest discrimination test, a
difference score (posttest value-pretest value) was obtained and later averaged vowel context per contrast for each subject. These difference scores were submitted to the statistical tests for the discrimination results.

The calculation of perceptual assimilation types from the data collected during the PA task was as follows. The responses written by the subjects were recorded in excel files. To arrive at the modal response for each sound, a matrix was developed which calculated the percentage of a particular response for that speech sound. Once the modal response for each of the sounds was reached at, the general assimilation types were obtained. The resultant general assimilation types for each contrast (pre as well posttest values) were either within category (WC) or between categories (BC). A contrast was BC if the modal responses for the sounds in that contrast were different. On the other hand, a contrast was WC if the modal responses were same and the overlap between them was less than or equal to 0.1 (maximum). For the between category assimilation types, if the overlap was more than or equal to 0.9, it was termed as two category assimilation (TC) since the overlap of modal responses was minimum. If the percentage count for modal responses for both speech sounds within a contrast was less than 0.9, then it was termed as both uncategorizable assimilation types (UU) otherwise it was termed as uncategorizable-categorizable assimilation type (UC). In the case of within category assimilation types, the ratings for the speech sounds in that contrast were submitted to TTest (statistical measure) in order to examine any significance difference between the goodness ratings given for the modal responses corresponding to each speech sound. If significance was found between the ratings given, the contrast was termed as category goodness assimilation type (CG), otherwise single category (SC). (Please refer back to Chapter 1 for detailed description of the various assimilation types under Perceptual Assimilation Model.)
Summary

In this chapter, the methodology of the main experiment was provided. Starting with the description of the target language (Malayalam) and three listener languages (Bengali, Spanish and American English), the chapter provides details of stimuli preparation and evaluation as well as issues related to stimuli presentation. Information on subject groups was given, including the explanation to the appropriateness of the choice of the subject groups. A review of the perception training paradigm involving discussion of theoretical issues in mental representation of perceptual categories is considered essential for the understanding of the perception training method chosen. Finally, a detailed procedural description of the experiment was provided.
CHAPTER 3
RESULTS

Introduction

For this study, three goals were put forth in chapter one: to investigate the factor of multilingual benefit in cross-language speech perception; to examine the role of feature generalization in facilitating the acquisition of non-native speech contrasts by trained / experienced listeners; and to investigate changes in assimilation patterns of these speech contrasts, if any, before and after training in individual language groups as well as across groups. Of these, the third goal is descriptive and is discussed in the last section of this chapter.

The design of the experiment, mainly the choice of the subject groups, has its basis on the hypothesis formed for this study. It is expected that a positive influence of multilingual benefit on the acquisition of speech contrast will be seen in both bilingual groups through their better performance in Identification and discrimination tasks when compared to monolingual group. The factor of feature generalization is hypothesized to show an additive effect over multilingual benefit in the acquisition of the non-native contrasts presented. Also, a shift in perceptual assimilation types reflecting a trend towards learning is hypothesized in relation to the third goal of the study. The appropriateness of the choice of language groups in order to examine the first two goals has been discussed in previous chapters. The a priori possibilities from the three language groups’ comparisons can be made as follows. One or more of these states may emerge as results of this study depending upon different hypotheses stated below.

Possibility (1) results if the presence of multilingual benefit emerges as dominant through the perceptual performance of both bilingual groups, that is, Bengali-English (BE) and Spanish-English (SE), will be significantly higher than that of the monolingual group, American English subjects (AE). Possibility (2) reflects the scenario where only BE group may show higher
perceptual performance than either the SE or AE group, evidencing that only feature
generalization is facilitating the BE group to learn to perceive the non-native speech contrast
faster than the rest. There will be no evidence for multilingual benefit effect if the SE group fares
as equal to AE, the monolingual group in this case. Possibility (3) is concerned with the
assumption that feature generalization may be seen as an additive effect over the multilingual
benefit factor. In this case then, the BE group will perform significantly higher than the SE group
which in turn will perform higher than the AE group. In other words, the BE group will have
added advantage of enhancing the learning process since both multilingual benefit as well as
feature generalization factors will facilitate this process. Possibility (4) represents the fourth
logical alternative, that of null hypothesis. In this case, there will not be any significant changes
in the perceptual performances of all the three groups displaying no evidence for either
multilingual benefit or feature generalization.

Three different tests have been employed to investigate the hypotheses posited above. Data
of a total of sixty subjects, twenty from each language group, have been analyzed. The
perceptual performance of subjects in various tests like the identification test, AX discrimination
test and perceptual assimilation test has been measured by comparing the performance at the
pretest level to that at the posttest level. In order to normalize any performance differences at the
pretest level, difference in the scores of the posttest from the pretest have been submitted as data
to assess significance level.

The results presented in this chapter were tested for statistical significance using a number
of mixed-model ANOVA (Analysis of Variance) on various tests employed during pretest and
posttest phase of the experiment. The independent variables for the AX discrimination test,
identification tests (Pre-post ID task and Generalization task) as well as the perceptual
assimilation test remained same, such as native language group and contrast type. ANOVAs reported the main effects of the independent variable, language group, as well as the interaction between these variables, language group x contrast. In the discrimination test ANOVAs, the dependent variable was mean percent correct score of a subject on sixteen trials of a particular contrast type. In the varied identification test ANOVAs, the dependent variable was mean percent correct score of a subject on twelve trials of a particular contrast. The percent change in category assimilation type was used as the dependent variable for the perceptual assimilation ANOVAs. If a particular fixed effect or their interaction was found to be significant at p≤0.05 or marginally significant such as p<0.1, then Tukey post hoc tests were conducted to examine the differences of the least-squares means of the fixed effects in order to determine the statistical significance of performance by individual language groups.

Results

This section gives a detailed account of the results in the areas of testing, such as Identification task, AX discrimination task, generalization test and the perceptual assimilation task. Since the results of perceptual assimilation require a more descriptive approach than a statistical analysis, they are discussed in a separate sub-section below. Once the overall results containing the main effect of independent variables are stated for the test of identification, discrimination and generalization (provided immediately below), specific post hoc results are discussed in detail under the headings for the relevant factors examined in this study.

Main Effects and Interaction of Independent Variables

In the case of the identification test on a trained-on talker, comparing the results of pretest and posttest, no main effect for language groups was observed (F(2,57)=2.42, p>0.05). Fig 3-1 shows the mean percent correct response in pretest and posttest level of identification test by subjects from the three language groups. All three listener groups improved from pretest to
posttest, with the largest increases in performance observed with the Bengali-English group, followed by the Spanish-English and American English trainees. The main effect for contrast type was observed to be significant, as expected ($F(3,171)=5.46, p<0.05$). The interaction between language groups and contrast was observed to be statistically significant ($F(6,171)=5.74, p<0.05$).

Since the language group * contrast interaction was significant, post hoc tests using Tukey’s HSD procedure were conducted for only lateral, nasal and fricative contrast types. The rhotics contrast showed no group differences. These results of individual contrasts are discussed in post-hoc analysis under the section of individual factors.

The results of perceptual performance from the comparison of pretest and posttest level of AX discrimination test show no main effect for language group ($F(2,57)=1.55, p>0.05$, n.s). Figure 3-2 represents the mean d’ scores of AX discrimination test at the pretest and posttest
level for the three language groups. All three listener groups improved from pretest to posttest, with the largest increases in performance observed with the Bengali-English group, followed by the Spanish-English and American English trainees. The interaction between language group and contrast was observed to be significant ($F(6,171)=4.04, p<0.05$). Of the post-hoc analysis using Tukey test, only the lateral contrast showed statistical significance for language group effect.

Figure 3-2. Mean d’ scores of AX discrimination test over all four contrasts at the pretest and posttest level (Maximum value of a d’ score = 4).

In the case of the generalization test (an identification task), the scores were compared to scores of the identification test (posttest) test in order to examine any evidence for the development towards forming new phonetic categories for the non-native sounds by subjects in any of the language groups. Posttest scores represented the performance level of subjects’ perception of the trained-on talker while the generalization scores displayed the accuracy in identifying the stimuli from the novel talker (a voice not heard before during the experiment). The difference of the posttest scores and generalization scores were submitted as data for the ANOVA test. The results for the test of generalization, where similar stimuli from a new talker were tested for accurate identification, showed no main effect for language groups ($F(2,57)=1.10, p>0.05, n.s$). The
interaction between language group and contrast was observed to be statistically insignificant (F(6,171)=1.21, p>0.05, n.s). None of the contrasts showed any significance concerning the language groups. Therefore, no post hoc testing was conducted for the test of generalization.

Mean identification scores of the generalization test when compared with the posttest scores, displayed unexpected results. Previous studies have provided evidence of development of robust phonetic categories by L2 learners through generalization tests (Lively et al. 1994). Based on their results, it was expected that subjects will perform either equally well or worse in identifying the stimuli from a novel voice than the trained-on voice, as was the case in posttest ID test. If any effects of multilingual benefit and/or feature generalization were to be observed, then the expected results would be that either the BE (for feature generalization) or both the BE and the SE groups (for multilingual benefit) would perform equally well, providing evidence for the development of separate phonetic categories for the non-native contrasts. On the other hand, the AE group, acting as the control group for both factors, would perform significantly worse in the generalization test, displaying that AE group may require more training or exposure to these contrasts in order to develop separate phonetic categories. The results from the current experiment in terms of mean identification scores for the generalization test are provided in Figure 3-3. For all three listener groups, the identification of the non-native contrasts was more accurate in tokens from the new talker when compared with the trained-on talker, with the slightly larger increase in performance observed with the Bengali-English group, followed by the Spanish-English and American English trainees.
Figure 3-3. Mean percent correct identification of contrasts spoken by trained-on talker (posttest) and new talker (generalization test).

Therefore, looking at the overall results of the tests used to factor out any presence of facilitating effects of multilingual benefit and/or feature generalization, it is difficult to conclude that any of the hypotheses have been proven true. However, some of the post hoc analyses of the ANOVA results, where marginal significance is observed in either main effect for language groups or interaction between language groups and contrasts, are discussed in the following section. These analyses may shed some light on the not-so-robust (perhaps subtle) effects of the factors explored in this study.

**Examining Evidence for Individual Factors**

**Multilingual benefit**

As we have seen in the above section, the main effect for language in the pre-post identification test was on the borderline of significance level (F (2,57)=2.42, p>0.05) (p=0.09). The averaged scores over all four contrasts at the pretest and posttest level reflect this trend. Looking at figure 3-4, we can see that both bilingual groups show a noticeable increase in the
accuracy scores from pretest to posttest level when compared with AE, the monolingual subjects’ acting as the control group.

Figure 3-4. Individual comparison of the monolingual group with BE bilingual group (top chart) and SE bilingual group (bottom chart).

The mean percentage of improvement in the bilingual BE group is calculated as 19% whereas the improvement percentage in monolingual AE group is observed as 12%. The mean percent improvement in accuracy scores of the bilingual SE group is also seen as 19%, identical to that of the bilingual BE group.
These differences of the posttest result from that of pretest results reflect the actual perceptual performance by normalizing any initial performance differences at the pretest level. Looking at the averaged improvement for all three groups, the prediction that both bilingual groups individually have higher performance level than monolingual group appears to be acceptable.

In addition to the above results, the interaction between language groups and contrasts was significant \( (F_{(6,171)}=5.74, p<0.05) \) \( (p=0.0001) \). Therefore, post hoc analysis was conducted using Tukey procedure. Apart from the rhotic non-native contrast, all other contrasts (laterals, nasals, fricatives) showed significance for the language group*contrast interaction. The lateral non-native contrast was observed to be significant \( (F_{(2,171)}=7.62, p<0.05) \). Similar significance results were observed for the nasal non-native contrast \( (F_{(2,171)}=4.54, p<0.05) \) and the fricative non-native contrast \( (F_{(2,171)}=3.75, p<0.05) \).

Of these, the only lateral non-native contrast was observed to show any evidence for the facilitative effect of multilingual benefit in bilingual groups (BE and SE) since the performance level of identification of the lateral contrast was significantly more among bilinguals than the monolingual group. The nasal non-native contrast reflects a trend towards feature generalization. Therefore the nasal along with the fricative contrast is discussed in detail under the next section.

Next, in order to determine the significant differences between each of the language group pairs, the differences of least squares means were obtained using the post hoc Tukey test. When analyzing the post hoc results for the lateral non-native contrast, it was observed that the perceptual performance of accurately identifying the sounds in this non-native contrast by the Bengali-English bilinguals was significantly better than the American English monolinguals \( (p<0.01) \). The Spanish-English bilingual group’s accurate perception of the lateral contrast was
marginally different than that of the American English language group (p = 0.09). Figure 3-5 shows the mean accuracy scores of identification of the lateral contrast at the pretest and posttest levels by the three language groups. The mean percent improvement in performance by BE group was the largest (25%), followed by SE group (18%) and AE group (8%). On the other hand, when the differences of least squares means of the bilingual groups, BE and SE, were compared using the Tukey post hoc test, the pair showed no statistical significance in their performance from pretest to posttest level (p = n.s).

![Bar chart showing mean percent correct response in identifying the lateral non-native contrast [ l- ] in the identification test at the pretest and posttest level.]

Figure 3-5. Mean percent correct response in identifying the lateral non-native contrast [ l- ] in the identification test at the pretest and posttest level.

These results for the lateral non-native contrast provide evidence for the facilitative effect of multilingual benefit since Spanish speaking subjects, without any linguistic experience with retroflex feature, were able to perform nearly as well as Bengali speaking subjects. The monolingual language group (AE) was significantly (marginal with SE group) different from both the bilingual groups.
Few of the post hoc testing for the AX discrimination test is suggestive of multilingual benefit factor. As stated before, the main effect for language group in the improvement of performance in case of discrimination test was not significant. However, the interaction between language groups and contrasts was observed to be statistically significant. Only the lateral non-native contrast was observed to have significant interaction (F (2,171)=9.86, p<0.05). The non-native contrasts with manners as nasal (F (2,171)=0.46, p=n.s); fricative (F (2,171)=0.62, p=n.s); as well as the rhotics (alveolar tap versus retroflex approximant) non-native contrast (F (2,171)=0.04, p=n.s) were observed to have no significant interaction of language groups with contrasts.

Analysis of the differences of least squares means in the post hoc Tukey test for the lateral non-native contrast, showed that the improvement in performance by both bilingual groups (BE and SE) individually was statistically significant over the performance by the monolingual group (AE). Improvement in the Bengali speaking subjects’ discrimination of the lateral non-native contrast was more than the monolingual American English speakers (p < 0.001). The Spanish speakers discriminated the lateral non-native contrast better than the monolingual American English speakers (p < 0.05). However, the discriminability of the lateral contrast when compared between the bilingual groups showed no significant difference (p = n.s). These results show that the bilinguals were performing significantly better than the monolinguals in the case of lateral contrast perhaps implying the presence of multilingual benefit effect. Figure 3-6 represents the averaged d’ scores of the discrimination test during pretest and posttest for the three subject groups.
Figure 3-6. Mean d’ scores of discrimination of the lateral non-native contrast [l-ʃ] at the pretest and posttest levels by the subjects in the three language groups.

Looking at the post hoc analysis of the lateral non-native contrast in the identification test, which was the same task as used during training, as well as the AX discrimination task provided evidence for a trend that may direct towards demonstrating that among bilinguals, multilingual benefit may perhaps be at work in enhancing the acquisition of the non-native contrasts more than is observed in monolinguals.

Feature generalization

Feature generalization/feature productivity refers to the previous linguistic experience with a native phonetic feature that may aid in the acquisition of a non-native contrast that exploits this same feature but differs from any native contrast along other dimensions.

One of the hypothesis postulated as outcome for this study was that the Bengali-English group may perform significantly better than the Spanish-English group due to its experience with a retroflex feature in native contrasts. Additionally, both bilingual groups may perform significantly better than the monolingual American English group. In this case, feature generalization is expected to have an additive effect over multilingual benefit factor in this case.
It implies that when compared, within the bilingual groups, the Bengali-English group is expected to perform significantly better than the Spanish-English group.

However, the results showed that feature generalization was not a significant factor. In most of the ANOVAs and the Tukey post hoc analysis, no significant differences were observed between BE and SE bilingual groups’ improvement in perceptual performance. In the case of Tukey post hoc analysis of the nasal non-native contrast during the identification test at the pretest and the posttest level, the Bengali speaking subjects’ improvement in performance was significantly more than the American English speakers (p < 0.05) but not significant from those of the Spanish speaking subjects (p = n.s). The statistical significance was also not observed for the Spanish-English bilinguals and the American English monolinguals (p = n.s).

![Figure 3-7. Mean percent identification scores for the nasal non-native contrast [ñ-ñ] by the subject groups at the pretest and the posttest level.](image)

Figure 3-7 shows the averaged identification scores for the nasal non-native contrast by the language groups at the pretest and posttest level. All three listener groups improved from pretest to posttest, with the largest increases in performance observed with the Bengali-English group,
followed by the Spanish-English and American English trainees. These results may suggest the possibility that perhaps feature generalization may have been at work facilitating the Bengali speakers to generalize the retroflex feature to the nasal non-native contrast.

One area where statistical significance is seen between Bengali-English and the Spanish-English language group (p < 0.05) is the post hoc analysis of the fricative (palatal vs. retroflex) non-native contrast within the pretest-posttest identification test ANOVA. Figure 3-8 provides the averaged scores of identification of the fricative non-native contrast during the pretest and posttest levels. All three listener groups improved from pretest to posttest, with the largest increases in performance observed with the Spanish-English group, followed by the Bengali-English and American English trainees.

![Figure 3-8](image)

Figure 3-8. Mean percent identification scores for the fricative non-native contrast [ʃ-ʒ] by the three language groups at the pretest and the posttest level.

However, surprisingly in this analysis, both bilingual groups’ performance in identifying this particular contrast is similar to that of the monolingual group (p = n.s). This result does not support the feature generalization hypothesis since the statistical significance between the BE
and the SE groups is showing inverse results (SE>BE). This result does not support the multilingual benefit hypothesis also since BE bilinguals did not perform any better than the AE monolinguals.

The interpretation of these results is discussed in chapter four. These results are suggestive of the effect of multilingual benefit but not of the effect of feature generalization factor proposed in the hypothesis. The feature generalization effect, expected to be seen in BE bilinguals because of their linguistic experience with retroflex feature, appears to have no additive effect over the general factor of multilingual benefit, expected in both the bilingual groups. Therefore, any significant differences seen between the BE and the AE language groups can be suggestive of multilingual benefit more so than of feature generalization.

The effect of multilingual benefit factor in enhancing the acquisition of non-native contrast is observed exclusively in the non-native contrast with lateralization as the manner. The lateral non-native contrast was considered the most difficult contrast for this study since all the three language groups have only a single phonetic category for a lateral sound. Therefore, perhaps the subtle presence of multilingual benefit and its effects seen in the bilingual groups’ performance surfaced in the acquisition of relatively difficult contrasts. The results seen with lateral non-native contrast may reflect the positive influence that is not prominent enough to be observed with easily discriminable contrasts. Perhaps by increasing the sample size of the language groups may bring forth the factor of multilingual benefit as a robust effect. Therefore, to inspect for more robust effects regarding the multilingual benefit factor, a reanalysis of the experimental data was conducted. The reanalysis is discussed in the following section.

Re-Analysis

As seen in the results above, none of the tests or individual contrasts reflect any robust effects of feature generalization. However, the presence of multilingual benefit is marginally
evident in the perceptual performance (both identification and discrimination tests) of the lateral non-native contrast and as an overall result suggesting that the scenario of the bilingual groups (BE and SE) performing better at identifying the non-native contrasts than the monolingual group (AE) is accurate. These results, although not robust enough to be conclusive, provide us with reasonable doubt and reflect towards a trend of the subtle presence of multilingual benefit playing a role in enhancing the acquisition of non-native contrasts, particularly the difficult place distinctions. Therefore, it is appropriate to further explore this area in a different light.

Since the results indicated improved perceptual performance by bilingual groups over the monolingual group, it was thought suitable to examine the factor of bi/multilingualism as a whole rather than two separate individual groups. Therefore, a re-analysis of the present data (from the two tests provided at pretest and posttest levels: Pre-Post ID test and AX discrimination test) was conducted.

Re-analysis: Multilingual Factor

In this re-analysis, data from both the bilingual groups, Bengali-English and Spanish-English were put collectively under the bilingual group (N=40). The American English subject groups was reiterated as the monolingual group (N=20). The pooled data was submitted to a mixed-model ANOVA test where the independent variable, this time, was bilingualism vs monolingualism, instead of individual language groups. The results of all the tests mentioned above were re-examined with this ANOVA design.

Figure 3-9 represents the averaged accuracy scores for identifying all four non-native contrasts at the pretest and the posttest level by the bilingual groups (BE and SE) and the monolingual group (AE). The improved performance in percentage seen in subject groups after training was 19% for the bilingual group and 12% for the monolingual group across all four contrasts.
Interestingly, when pooled as a group, the bilingual subjects’ perceptual performance in the identification test was significantly better than the monolingual subjects’ performance. The main effect for multilingualism showed statistical significance ($F (1,58)=4.75$, $p <0.05$). The interaction between the multilingualism and contrast was also significant ($F (3,174)=4.45$, $p <0.05$). On examining the tests of effect slices, it was observed that of the four contrasts, the lateral non-native contrast ($F (1,174)=11.72$, $p<0.05$) and the nasal non-native contrast ($F (1,174)=5.03$, $p<0.05$) showed statistical significance between the independent variable, bilingualism-monolingualism.

In case of the AX discrimination test results, the pooled data set of bilinguals’ performance revealed that $d'$ scores of bilinguals were marginally different from that of monolingual subjects. The averaged $d'$ scores of discrimination at the pretest and the posttest levels by the bilinguals collectively compared to those of monolingual subjects has been represented in figure 3-10. The
main effect for multilingualism was borderline \( (F(1,58)=2.95, \ p=0.09) \) unlike the insignificant results \( (p = \text{n.s}) \) of discrimination test when examining individual language group effects.

![Figure 3-10](image_url)

Figure 3-10. Mean \( d' \) scores of AX discrimination test across all four contrasts, by the bilingual groups (BE and SE) and the monolingual group (AE) at the pretest and the posttest levels

The interaction between group and contrasts was observed to be significant \( (F(3,174)=6.59, \ p<0.05) \). Therefore, post hoc analysis was conducted on individual contrasts using the Tukey procedure. Of the four contrasts, only lateral non-native contrast showed statistical significance between the bilingual and the monolingual groups \( (F(1,174)=18.50, \ p<0.05) \). The post hoc results from the re-analysis were quite similar to the post hoc results from the main analysis in terms of which contrasts showed significant differences among the independent variables.

The results from the pooled dataset exhibit the effects of multilingual benefit being employed by the bilinguals in order to enhance the acquisition of the non-native contrasts. However, this re-analysis reveals the relevance of certain methodological issues. The effects of certain factors like multilingual benefit perhaps cannot be captured at the standard sample size of
twenty subjects used in this study. These effects could be subtle and perhaps obscured by individual subject variations. This limitation is further discussed in chapter four.

**An Alternative Re-analysis: Only Place Contrasts**

Alternative re-analysis of the data was conducted by examining individual contrasts used for this study. The results from the main analysis and subsequently the results from the pooling analysis reflected the need to investigate individual contrasts and to look for inconsistencies among the four contrasts chosen for this study. The results that looked at individual language groups showed enormous amount of variability in the means procedure of ANOVAs conducted. In addition to that, the perceptual performances varied from one contrast to the other. Based on the three reasons stated below, a re-examination was conducted of the four non-native contrasts as hindsight.

Looking at the preliminary experiment, where data of initial AX discrimination of these contrasts were gathered from native American English speakers, the highest d’ score was attained by the rhotics non-native contrast (d’=2.53) (refer to Chapter 2, pg 18, for details). Though not at ceiling (Max d’=4), this particular contrast stood out as it showed ease of discriminability at above chance (63%) whereas the other three contrasts showed discriminability at chance level (50%). Therefore, ease of discriminibility of this contrast may have suppressed the subtle effects that were seen in more difficult contrasts like laterals (prelim d’ score = 1.13) since the statistical analyses were conducted on the difference scores averaged over all contrasts. This possibility gets support from the pretest scores for the rhotics contrast which were well above chance level for all three groups (AE=59%, BE=67% and SE=67%). Thus, the probability to observe subtle effects of multilingual benefit and/or feature generalization would have been minimal.
Looking more closely at the phonetic features of this particular contrast, it appears to evade the overall homogeneity of the stimuli for this study. The non-native contrasts with manners lateral approximant, nasal and fricative only differ in terms of place distinctions. The lateral non-native contrast has an alveolar and a retroflex [l-l]. The nasal non-native contrast has an alveolar and a retroflex [n-n] (see chapter Two for detailed discussion on dental/alveolar distributions in Malayalam). The fricative non-native contrast has a palato-alveolar and a retroflex [ʃ-ʃ] distinction. However, the rhotic non-native contrast [r-t] differs in not only the place distinction but also the manner. This contrast involves an alveolar tap and a retroflex approximant. Therefore, the rhotic contrast stands apart from the rest of the three only-place contrasts. Perhaps this could have provided the contrast with inherent extra salience of acoustic features which may have provided the subjects with ease of discriminability. This interpretation might raise questions on the reasons for choosing rhotics contrast as one of the stimuli. In hindsight, initially it was assumed that feature generalization might be a robust factor and will thus emerge even if there were additional secondary cues to the contrast involving manner or other features that may provide easier discrimination of the contrast. However, since this is not the case as is evident from the main results, a reanalysis exclusively of the place contrasts is called for. One more point of contention within this particular contrast is that the phonetic definition of retroflex approximant sound in Malayalam has been a debatable issue among the phoneticians researching this language (see chapter Two, pg 6, for details). In hindsight, no particular phonetic description has been agreed upon making this particular contrast of concern regarding as stimuli for this study.

Therefore, a second re-analysis of the data set, where only three non-native contrasts were analyzed, was conducted. In this case, the scores on the non-native contrasts which differed only
in place distinctions (lateral, nasal, fricative) were used in a mixed-model ANOVAs for three datasets: Identification test, AX discrimination test and the generalization test.

In the case of the identification tests at the pretest and posttest levels, the main effect for language groups was observed to be significant ($F (2,57)=5.12, p<0.05$). Figure 3-11 shows the averaged accuracy scores of identification test over the three place contrasts by all language groups at the pretest and posttest levels. All three listener groups improved from pretest to posttest, with the largest increases in performance observed with the Bengali-English group, followed by the Spanish-English and American English trainees. The improvement percentage for the bilingual groups was significantly higher. For the BE group it was 19% and for the SE group it was 17%. The percentage of improvement in perceptual performance seen in AE group was 9%.

![Figure 3-11. Mean percent correct response for identification of the three place contrast (laterals, nasals, fricatives) by bilingual (BE and SE) and monolingual (AE) groups at the pretest and posttest levels.](image)

The interaction of language groups with the three contrasts was also observed to be significant ($F (4,114)=5.81, p<0.05$). The main effect of contrast showed no significant
difference, unlike the main analysis results, providing evidence for homogeneity within the three place contrasts (F (2,114)=1.81, p=n.s.) (p=0.16).

The post hoc analysis using Tukey test was conducted. All three non-native contrasts showed significant differences between the language groups. The post hoc results were the same as the results from the main analysis since there was no change in the data set within the individual contrasts.

In the case of AX discrimination test, the ANOVA results indicated marginal difference between language groups. Figure 3-12 represents the averaged d’ discriminability scores over the three non-native contrasts at the pretest and posttest levels. All three listener groups improved from pretest to posttest, with the larger increases in performance observed with the Bengali-English group, followed by the Spanish-English and American English trainees. The main effect for language groups was seen at the borderline significance (F (2,57)=2.81, p=0.06). The interaction between the language groups and the three place contrasts was observed to be significant (F (4,114)=5.45, p<0.05).

Figure 3-12. Mean d’ scores from the AX discrimination tests for the three place contrasts (Laterals, nasals, fricatives) by the language groups at the pretest and posttest levels.
The results for the generalization test scores when compared with the posttest scores showed no statistical significance, similar to the main analysis results. The area of testing for robust phonetic category development did not reveal any significant differences among the language groups in any of the analyses. One explanation could be the within group variability. Individual subjects’ performance may have dampened any differences between language groups. Another explanation could be the issue of baseline talker intelligibility. These various explanations will be discussed in detail in chapter four.

In summary, conducting re-analyses of the data demonstrated a different pattern of results when 1) both bilingual groups were pooled as one group and 2) problematic contrast was excluded. The positive influence of the multilingual benefit factor was supported with the dependent measures when the issue of sample size was accounted for by pooling the BE and SE groups together. The results for the identification tests showed significant group differences between bilinguals and monolinguals and those of AX discrimination displayed marginal significance in the performance of bilinguals over monolinguals. The reanalysis with only place contrasts also revealed a significant main effect for language groups. However, effects of the feature generalization factor did not surface prominently even in the reanalysis results. The next section presents the results of the perceptual assimilation tests conducted at the pretest and posttest levels during the training experiment.

**Perceptual Assimilation Results**

The perceptual assimilation task conducted at the pretest and posttest levels requires a more descriptive interpretation of its results. Therefore, this section presents results solely from these tests. Because of its descriptive nature, no prior predictions could be made on the results of perceptual assimilation patterns. However, based on PAM-L2 predictions (Best & Tyler 2007), we could expect to see a direction towards learning. For this study, learning was defined in terms
of a shift from non-categorical assimilation types like single category (SC) or UU (both uncategorizable phonetic sounds) to assimilation types such as category goodness (CG) or uncategorizable and categorizable type (UC) which reflect a trend towards higher sensitivity of discrimination or even up to two-category assimilation which represents excellent discrimination of the non-native contrasts.

The following sections describe results from the perceptual assimilation test in terms of changes in assimilation types after training across all language groups. In addition, the question of whether the effects of factors like multilingual benefit are seen in terms of perceptual assimilation patterns will be explored by comparing the language groups’ learnability trend through the difference scores of assimilation types elicited from the pretest and posttest results.

**Changes in Assimilation Types after Training**

For this study, the assimilation types (for details, refer to chapter 1) described in PAM model have been applied for interpreting the results. Five major assimilation types which reflect various levels of assimilation, that is, two category (TC), single category (SC), category goodness (CG), uncategorizable-categorizable (UC) and both uncategorizable (UU) have been used. The results from the perceptual assimilation test reveal a consistent shift of assimilation types from none / minimal separate categorization of non-native contrasts before training to assimilation types like TC, CG and UC, which reveal improvement taking place in discrimination of the non-native contrasts during training. This implies a direction towards learning of the non-native contrasts as separate phonetic categories.

Figure 3-13 shows the mean percentage of number of assimilation types drawn from subjects over all three language groups collectively. The results showed a decrease of 18% in SC assimilation type and 5% in UU assimilation type from pre training to post training levels. The TC and CG assimilation types displayed an increase of 11% and 10% respectively. The UC
assimilation type showed a modest increase of 2% from pretest to posttest. The overall change of assimilation types from showing less categorization to a higher sensitivity towards categorization provides evidence of learning taking place during the limited training period. The changes in these assimilation types were not similar for all the four contrasts.

![Perceptual Assimilation Types](chart.png)

Figure 3-13. Mean percentage of assimilation elicited from all three language groups, pretest vs posttest.

Looking at the results for the four non-native contrasts (figure 3-14) across all language groups, decrease in SC assimilation type was observed for all contrasts. For the lateral and the nasal non-native contrasts, there was found to be a steep decrease (-23% and -29% respectively) when compared to the rhotic or the fricative contrasts (-12% and -10% respectively). This steep increase is manifested as a large increase in CG (22% and 17% respectively) and TC (11% and 14% respectively) assimilation types for the nasal and lateral contrasts. This provides evidence of possible learning of the non-native contrasts taking place. The rhotic contrast showed a modest increase in CG (2%) but a comparable increase in TC (13%) assimilation types. In case of UU (2%), it showed an increase unlike other non-native contrasts. The fricative contrast, on the other hand, showed very modest increase in CG (2%) and TC (7%) types which were expected to
show learnability. Nevertheless, it displayed learnability through UC assimilation type where the positive difference (improvement) is 12%.

![Figure 3-14. Percentage of difference scores (pretest to posttest) of assimilation types for the individual contrasts elicited from all three language groups.](image)

On the whole, the SC and UU assimilation types displayed a decrease from pretest to posttest whereas the CG and TC types showed an interesting increase. Overall, the subjects showed more learning in lateral and nasal non-native contrasts than the other two contrasts.

**Changes in Assimilation Types: Multilingual Benefit Factor**

The perceptual assimilation test was one of the areas to examine the effects of factors like multilingual benefit and feature generalization on the acquisition of non-native contrasts. These chosen contrasts were phonemically assumed to be assimilated as the single category type at the initial exposure (pretest) level. The assimilation results from the previous section provide evidence of a definite shift towards learning these novel non-native contrasts in a very limited training period. However, the question arises whether any language group differences were seen pointing towards the effects of factors explored in this study.
The results from the identification tests and AX discrimination tests revealed no evidence supporting the presence of feature generalization as a facilitating factor in the acquisition process by the bilingual group, BE. Marginal effects were observed for multilingual benefit in case of the lateral non-native contrast. On two alternate reanalysis of the data set, the presence of this factor emerged as significant for the bilingual group when compared with monolingual group. Therefore, it is reasonable to examine the bilingual groups as a whole comparable to monolingual group rather than individual language groups.

Figure 3-15 displays percentage of assimilations elicited from individual language groups at the pretest and posttest levels. The bilingual groups, BE and SE, showed similar shifts in assimilation types from pretest to posttest, unlike the monolingual group. The SC type of assimilation pattern was observed to have similar percentage at the pretest level for both BE (38%) and SE (41%) groups, whereas the same assimilation type had a very high percentage of 60% at the pretest level for the monolingual AE group.

Interestingly, a sharp decline was seen in the SC type at the posttest level for both bilingual groups, unlike the monolingual group. An inverse trend was observed for the TC and in some cases for the CG assimilation type from pretest to posttest. Both bilingual groups showed a high rise in the TC assimilation type at the posttest level contrary to the monolingual group. The SE group was deviant from the BE group in case of UC assimilation type. The SE group displayed a modest decrease of (-)5% whereas the other two groups showed similar percentage but increase in UC type.

Taking a comparable view of the three groups, figure 3-16 shows the difference-scores of the percentage of perceptual assimilation types comparing the two bilingual groups with the monolingual group over all four contrasts. A positive percentage denotes an increase in the
Figure 3-15. Percentages of assimilation types elicited from individual language groups at the pretest and posttest levels are provided in three different charts.
percentage of elicited assimilation type from pretest to posttest and a negative percentage
denotes a decrease in the assimilation types gathered by the language groups.

A steep decline was observed in the SC type along with a nearly equal steep increase in the
TC assimilation type, and CG type for SE group, during the posttest for bilinguals. On the other
hand, the monolingual group showed a decline in SC type lower in percentage than that of the
bilinguals. Moreover, the SC type remained high in percentage even at the posttest level for the
monolingual group. Only a modest increase was seen in the TC, UC and CG assimilation types
which reflected limited learning taking place within that group. Thus, the results described from
the perceptual assimilation tests provide evidence for the effects of multilingual benefit
facilitating the acquisition of novel non-native contrasts even with such limited training.

Although this section seemed to have required a more descriptive analysis of the
perceptual assimilation results, unlike the statistical results from other tests such as the
identification test and the discrimination test, a statistical analysis (one-way ANOVA) on proportion change in assimilation types was conducted to examine significance based on the individual language groups as well as the multilingual factor. The main effect for language was observed to be significant ($F(2,57)=4.16, p<0.05$). In the post hoc Tukey test, BE group showed proportionately more change in assimilation types than the AE group ($p<0.05$). The BE and SE group showed no significant difference in proportion change ($p=\text{n.s}$). The SE group showed no significant difference from AE group for proportion change ($p=\text{n.s}$).

These results lead to the reanalysis of the dataset by pooling the bilingual groups together and comparing with the monolingual group. To look at the multilingual factor, a one-way ANOVA was conducted which showed significant differences between the bilingual group and the monolingual group ($F(1,58)=7.78, p<0.05$). These results, again, are suggestive of the effect of multilingual benefit on acquiring the non-native contrasts by the bilinguals as a group, regardless of their L1 languages.

**Summary**

The results of the current study are suggestive of the effects of multilingual benefit. The main language effects in reanalysis and post hoc tests of the lateral non-native contrast provided support for the hypothesis stated for the multilingual benefit factor. The evidence for this factor was mainly seen in the lateral non-native contrast in the dependent measures like the identification test and the AX discrimination test. Both the reanalysis provided evidence for multilingual benefit factor by showing significant group differences. However, neither the effect of feature generalization nor its additive effect surfaced in any of the results. The only marginal effect of feature generalization was seen in the identification test results for the nasal non-native contrast, which does not show strong support for this factor. The test of generalization results were observed as non-significant in terms of group differences as the mean percentage of
performance revealed that on an average subjects performed better with the novel voice (generalization test) than with trained-on voice (posttest). The results from the perceptual assimilation tests revealed a consistent shift towards learning of the non-native contrasts. Also, they showed trends of effects of multilingual benefit among bilingual groups as opposed to the monolingual group.

These results of this very limited training are only suggestive, of the hypothesis that bilinguals can acquire non-native contrasts faster than monolinguals, given the facilitating effects of multilingual benefit. However, no study is without its caveats. There were some unexpected results such as subjects did better with novel voice than a trained-on voice. In addition, a lot of within group variability was found for the generalization test. All the results, their interpretation and caveats are discussed in the next chapter.
CHAPTER 4
DISCUSSION AND CONCLUSION

Introduction

The experiment was designed to address a number of important questions concerning third language acquisition and cross-language speech perception: Would multilingual benefit and/or feature generalization facilitate the acquisition of novel non-native contrasts among bilinguals? Can perceptual assimilation patterns be modified through a limited training period? Would these changes in assimilation patterns, if any, reflect differences between the bilingual and monolingual subjects in terms of acquisition of non-native contrasts? These broad questions are addressed through an overview of the results. In addition to that, a discussion regarding the possible explanations for the unexpected results seen in this study will be put forth discussing some unforeseen caveats in the methodology as well as issues with current assimilation models. The results of this study will also be discussed in tandem with previous literature in order to embed the results in a larger perspective, to further future potential hypotheses in the field of language acquisition.

Overview of the Results

The design of the current study had three main aims. The first aim concerned the assessment of the effects of metalinguistic ability developed as a result of having two languages already in place. The second aim pertained to the assessment of the effects of previous linguistic experience with certain phonetic features present in the target stimuli. The third aim was to describe changes in assimilation patterns and examine any trends towards learnability. These aims were narrowed down to more specific hypotheses. The choice of subject groups was determined based on the formulation of the hypotheses. Briefly stated, Bengali-English (BE) bilinguals were chosen because of their linguistic experience with the retroflex phonetic feature...
and more so because of the fact that they were balanced bilinguals. Spanish-English (SE) bilinguals were appropriate as the counteracting group solely because of their lack of any linguistic experience with retroflex feature and also for having acquired two languages relatively simultaneous. American English (AE) language speakers were chosen as the control group due to lack of any substantial knowledge of another language, making them a monolingual group.

Based on these language groups, more strictly defined hypotheses concerning the main aims were constructed. The effects of multilingual benefit would be evident if both bilingual groups performed significantly better at identifying the non-native contrasts than the monolingual group. The facilitating effects of feature generalization would be seen if the BE group performed significantly better than the SE and the AE group. Additive effects of feature generalization over multilingual benefit would be observed if the BE group identified and discriminated the non-native contrasts more accurately than the SE group, which in turn performed better than the AE monolingual group. Apart from these hypotheses, the results for the generalization test were expected to be fairly equal to or lesser than the posttest identification test results, since the successful generalization to a new voice would indicate the formation of new phonetic categories corresponding to the non-native contrast. This prediction was based on the results from the previous training studies examining generalization of newly acquired contrasts to new voices (Logan et al 1991, 1993, Lively et al. 1994, Bradlow et al. 1997).

The results of the current study do not show robust empirical evidence to support the hypotheses. However, this may be due to the complex nature of the stimuli (4 pair of novel non-native contrasts to acquire) as well as very limited time period for implicit perceptual training (6 sessions- 35 minutes each). The subjects were required to perceive distinctions among eight target sounds, four of which were novel for the subjects [l-ɾ, n-ŋ, r-ɻ, ʃ-ʒ]. Acquiring all these
four contrasts within the training period of only 3.5 hours (6 sessions of 35 minutes each) would have been a very daunting task. Perhaps, over a longer period of time, the robustness of the results will be seen. Therefore, to observe any robust effects in such limited training, the influence of the factors being explored would have had to be really strong.

The hypothesis for multilingual benefit was supported in one particular contrast (lateral non-native contrast) out of the four contrasts used as stimuli. The evidence for multilingual benefit was consistent in the lateral contrast in the discrimination results as well as in all the subsequent reanalysis conducted. The results of the perceptual assimilation test revealed effects of multilingual benefit prevalent in the bilingual groups. The statistical analysis of the proportion of change in assimilation patterns among the language groups was observed to be significant, with the BE group having a larger change in assimilation patterns after training as compared to the monolingual AE group but not significantly different from the bilingual SE group. The SE group revealed a marginal significant difference in the proportion change from that of AE group. Therefore, although the evidence supporting the effect of multilingual benefit in facilitating the acquisition of non-native contrasts in bilinguals is not robust enough, it is suggestive of presence of this factor.

The hypothesis of feature generalization/feature productivity did not surface as a facilitating effect where multilingual benefit was not seen, such as a scenario where the BE group would perform better than the SE and AE groups with no group differences between SE and AE. The only instance which could be interpreted as perhaps an effect of feature generalization was the results of the nasal non-native contrast in the case of the identification test. However, even in this case, the SE group was at borderline significance with the AE group, resulting in a very weak interpretation of the feature generalization hypothesis. In addition to
that, an additive effect that was hypothesized to result from having previous linguistic experience with the retroflex feature in addition to multilingual benefit was not observed in any area of testing.

The perceptual assimilation test results along with requiring a descriptive approach allow us to make a priori prediction concerning the multilingual factor. It was expected that the changes in assimilation types before and after training, if any, would reveal effects of multilingual benefit; that is, both bilingual groups would show a greater shift towards the assimilation types that clearly reflect learning of the non-native contrasts than the monolingual group. The results of the study show support for this prediction. Group differences were observed between the bilingual groups when compared with the monolingual group, where the bilingual groups show a considerable decrease in SC (single category) and UU (uncategoriable) assimilation types and an increase in the TC (two-category) and CG (category goodness) assimilation types. This indicates that learning took place even during the very limited time period of training. The learning was seen more in bilinguals than monolinguals.

An overall shift in assimilation types from within category non-discriminable types to more discriminable categorical assimilation types were observed across all language groups and all four contrasts. The PAM-L2 model (Best & Tyler 2007) provides us with an approachable system of assimilation types that can be used to examine the perceptual assimilation performance before and after training. However, some caveats in the model restrict us from further exploring the direction of learnability through these specific assimilation types. These are discussed in the next section in detail along with other unexpected results.

**Alternate Explanations**

If the results of the current study show support for certain hypotheses, there are other results and cross-language group differences that are not so easily explained. This section
provides either possible explanations or interpretations of the various unexpected results found in the study.

**Test of Generalization Results**

The results in the area of test of generalization during the posttest phase of the training were contrary to the prediction. It was expected that the posttest scores of identification would be higher than or fairly equal to the generalization test scores, as seen in previous literature. Moreover, if the explored factors were in fact effective in enhancing the acquisition of the non-native contrasts, it was expected that the bilingual groups would show near equal test scores for posttest identification (trained-on voice) and generalization (novel voice) whereas the monolingual group would show a lower score in generalizing to a novel talker than scores attained with the trained-on talker. However, the results of this study supported no such claims. On the contrary, there were no group differences between all three language groups. Moreover, all three language groups displayed generalization scores (novel talker) that were higher than the posttest scores (trained-on talker). On analyzing the individual contrasts in order to tease out perhaps one particular contrast that may have been the easiest to discriminate, it was found that no one contrast was consistently being identified accurately by all speakers. There was a considerable variation in identification scores within groups and across contrasts. Therefore, one particular contrast could not be held responsible for these unexpected results.

Another explanation for these surprising results could be the issue of baseline talker intelligibility issue. This inherent/inbuilt intelligibility of an individual talker is perhaps the factor affecting the results of the generalization test. Perhaps, the hyper-articulated sounds produced by the novel talker made it easy for the all the subjects, across the language groups, to identify the stimuli more accurately than was the case with trained-on talker. The term talker intelligibility can be explained in regards to the articulation of the sounds by individual talker. It
could be seen as a continuum where on one extreme a talker produces “clear” hyper-articulated speech which enhances the intelligibility of the speech for the listener of the same language. On the other end of the continuum, a talker produces “plain” speech which may be unintelligible to the listener, with same phonetic category inventory, in the context of various kinds of noises. The terms “clear” refer to speech that is distinct and a part of speech production where the acoustic/articulatory features are indicated clearly so as to be intelligible in adverse listening conditions. It may perhaps allow the talker to enhance the distinctions between contrasts in a phonological space (Uchanski, 2000; Smiljanic and Bradlow, 2005, 2007). The “plain” speech refers to the normal rate of speech that a talker will produce in ideal listening conditions, that is, not hyper-articulate the sounds in order to make one’s speech intelligible.

Therefore, based on this interpretation, the results of the generalization test scores can be explained in the following way. Probably, the novel talker used in the test of generalization provided the stimuli, of which certain sounds were clearly hyper-articulated among the four speech contrasts, unlike the stimuli from the trained-on talker used in the posttest. The hyper-articulated sounds may have provided all the subjects with easy identification of those sounds resulting in better accuracy than with the trained-on voice. With introduction to multiple talkers during training, the formation of new phonetic categories, which was the desired effect, to accommodate the distinctions between the new non-native contrasts may have started among the subjects of the bilingual group. Therefore, with the baseline talker intelligibility varying considerably for the trained-on talker and the novel talker, the subjects were able to perform slightly better with the new voice than the trained-on voice. Although, no significant outcomes were observed for this test, the results are still suggestive of the fact that a direction towards development of robust phonetic categories corresponding to the non-native contrasts could be
seen since all groups fared better in generalization scores. The monolingual groups scored equally well as they did with the trained-on voice, like the bilingual groups. Contrary to hypothesis, it could be interpreted as an effect of the enhanced intelligibility of the novel talker in increasing the performance level for the monolingual group. This interpretation of the results also reflects the robustness of the training program already observed in many high variability training studies (Logan et al., 1991; Lively et al., 1993, 1994; Bradlow et al., 1997).

Many perceptual training studies where multiple talkers are introduced during training acknowledge the variations found in individual talker’s intelligibility by the listeners as a confounding factor for the extreme variations in the results (Lively et. al., 1993, 1994; Iverson et al., 2005; Semiljanic and Bradlow, 2007 etc.). However, this area of inherent talker intelligibility requires a deeper probe than a mere mention in the training studies as a confounding variable. This aspect of perceptual training method can be looked at exclusively in future research work which can contribute substantially in enhancing the efficacy and accuracy of the high variability perceptual training method.

**Necessity for Reanalysis**

The design of the experiment was constructed with the assumption that the factors being explored would show robust effects within the individual bilingual groups. Therefore, a standard sample size of twenty subjects in each group seemed quite appropriate for the design. Additionally, having a number of contrasts in the stimuli was expected to provide the study with a range of perceptual data that could potentially contribute to the existing pool of explored contrasts for the cross-language speech perception field and could be used for future work.

On the contrary, the outcome of the existing design showed no effects of the feature generalization factor. Additionally, the effects of multilingual benefit showed only trends towards its presence, and only in certain contrasts. No strongly suggestive support emerged
through any of the testing measures for these factors that may influence the acquisition of the non-native contrasts among bilinguals when compared with monolinguals. A further inspection seemed appropriate in order to examine the reasons for these varied unexpected results. Therefore, two reanalysis were conducted looking at two issues: ineffectiveness of the sample size and individual examination of the non-native contrasts.

Since the presence of the multilingual benefit factor was considered as merely suggestive within each bilingual group in the main results of the study, looking exclusively at the multilingual factor was considered as the next step. Combination of the two bilingual groups and comparison with the monolingual group revealed significant group differences implying that bilinguals had performed better at acquiring the non-native contrasts than the monolingual groups. The pooling of bilinguals into one group provided evidence for multilingual benefit as the bilingual group showed significantly better identification results than the monolingual group. The reanalysis by pooling the two groups together displayed a significant effect of multilingual benefit that gave the individuals with two or more language systems an ability to extract relevant acoustic cues for the place distinctions of the non-native contrasts and be able to learn to identify the sounds in these contrasts separately. However, this reanalysis revealed some caveats in the original design of the experiment, namely, the issue of sampling size. It appears that by increasing the sample size of the language groups, the effects of multilingual benefit become robust. Additionally, a large degree of individual variation within the language groups pointed towards the need for a larger sample size. The individual variability could have resulted from the uncontrolled factors such as the individual aptitude for learning. The power issue in this study brought forth the fact that multilingual benefit effects are not dominant enough to be seen as a robust effect within a small sample of the population. In order to see significant group
differences brought forth with multilingual benefit as the determining factor, a larger sampling size may be required, especially with very limited laboratory training period as in the current study.

The other basis for an alternate reanalysis was examining all the non-native contrasts individually. The reason behind this examination was to tease out any particular contrast whose results may be suppressing an effect of multilingual benefit. The large variation in the perceptual performance on these contrasts pre- and post-training was another motivating factor to conduct a reanalysis.

Upon a closer inspection of the four contrasts in question, the non-native contrast that differed in both place and manner appeared to be most different from the rest of the stimuli. There were several reasons that emerged to exclude the rhotic non-native contrast from the dataset and reanalyze it. The reasons have been discussed in detail in chapter three and will be briefly restated here. Firstly, the results from the preliminary experiment showed that of the four non-native contrasts, rhotics was the most easily discriminable by the monolingual American English group. Therefore, we see high perceptual performance scores for the rhotics contrast. Therefore, the results of this contrast may have suppressed the group difference significance noticeably. Secondly, the sounds \([r-t]\) of the rhotics contrast are distinguished in place as well as manner unlike the other three contrasts the sounds of which differ only in place distinction. This criterion perhaps provided the listeners with more of the salient acoustic cues resulting in easy identification and consequently easier learning of this non-native contrast, which may have again been the reason for not showing significant group differences in the main result. Moreover, the phonetic definition of the Malayalam retroflex approximant has not reached consensus among the phoneticians who have examined the Malayalam consonant inventory. Lastly, the fact that
Spanish-English bilinguals will already have separate established phonetic categories for the alveolar tap, acquired through Spanish, and the alveolar approximant, acquired through American English, making this contrast the most easily discriminable for the Spanish-English bilinguals. This would not have been the case with Bengali-English bilinguals since they used the Indian English dialect which substitutes the alveolar tap with the American English approximant leaving the BE bilinguals as the only language group with no added advantage from this contrast.

Together, these reasons motivated the exclusion of the rhotic contrast from the analysis. Therefore, a reanalysis of the data set with only the place contrasts was conducted. The results of the only-place contrast reanalysis, excluding the rhotic contrast results, for the identification test revealed significant group differences among the three language groups, as expected. On post hoc testing, the results of the lateral non-native contrast were suggestive of the effects of multilingual benefit. The bilingual groups performed equally well at identifying the contrast. The BE group performed significantly better than the AE group with SE group showing marginal significance with the AE group. The results of AX discrimination test which had originally showed no significant group differences, revealed a marginal main effect for language groups. The results from this reanalysis were suggestive of the assumption that perhaps the results of the rhotic non-native contrast were preventing the effects of multilingual benefit to surface as a robust effect.

The results also shed light on the phenomenon that factors like multilingual benefit, which enhance the perceptual performance or facilitate in the learnability of non-native contrasts, may not be evident in a laboratory research with average difficulty contrasts, such as rhotics in this study. In terms of inherent acoustic difficulty in perceiving the relevant cues of the contrast,
studies have shown that spectral properties are less easily discriminable than the temporal properties when they are the distinguishing acoustic cue within the contrast. Alternatively, in terms of previous linguistic experience/ native perceptual sound inventory, the PAM model of cross-language speech perception (Best & Tyler, 2007) details that non-native contrasts which are assimilated to a single phonetic category as equally good exemplars or equally bad exemplars (SC type) or UU (no assimilation with less phonetic distance between them) type, will have a higher level of difficulty in discriminability. According to PAM, during the learning process as well, the learner will exert more time in overcoming the difficulty level of this contrast and forming separate phonetic categories corresponding to the contrast. Therefore, only when the target new non-native contrast is inherently at a higher level of difficulty will the effects of multilingual benefit emerge as robust. Thus, conducting reanalysis of the dataset gathered from this experiment revealed different perspectives of the factors explored in this study, which could be potential candidate hypothesis for future work.

**Caveats in Perceptual Assimilation Analysis**

The results of the perceptual assimilation test were descriptive in nature. Reiterating the results, an overall shift in assimilation patterns was observed across the three language groups and all four non-native contrasts. A shift from assimilation types like SC (single category) or UU (uncategorizable) towards more categorical/ higher sensitivity for discrimination assimilation types like TC (two-category), CG (category goodness) or UC (uncategorizable-categorizable) were observed. Noticeably, these results reflect the growth of the learning curve for all the three language groups since a decrease in SC assimilation types was consistently observed from pretest to posttest. Similarly, an increase was observed in assimilation types like TC, CG and in some cases UC from pretest to posttest.
Although these learning shifts were seen for all of the language groups, a difference in the percentage of shift was noticeable between the monolingual group versus the bilingual groups. The bilingual groups, however, fared equally well, and their shifts in assimilation types were comparable. Therefore, in order to examine group differences statistically, an analysis of the proportion of change in the assimilation patterns was conducted which revealed significant main effect for language groups as well as significant effect for the multilingual factor (during reanalysis). The post hoc tests provided suggestive evidence for the multilingual benefit factor.

However, this analysis of proportion change concerns only the comparison of proportion of change in the assimilation type which in no way describes the direction of learning. In order to examine learnability among the groups in terms of perceptual assimilation, a definitive ranking of the assimilation types, needs to be established. For this study, five assimilation patterns (SC, UU, UC, CG and TC) were used to assess the perceptual performance of the learner groups. These assimilation types were adopted from the revised PAM model (Best & Tyler 2007). Since the current model of perceptual assimilation does not delve into the prediction of learning stages based on the assimilation patterns revealed in the perception of an experienced learner before and after training, the hypothesized learning continuum is proposed as an extension of the PAM model.

To assess the direction of learning in the perceptual assimilation mode a continuum of learning is postulated where SC (single category) and TC (Two-category) assimilation type form the two extremes. According to PAM, SC assimilation type at the initial exposure reflects poor discrimination of the given contrast since both sounds fall under one phonological category of the existing language of the learner. The model predicts that the learner may not form a new phonological category at all where SC assimilation takes place. In case of TC (two-category)
assimilation at the initial stage of perception, the type reflects the learner’s good discrimination of the contrast. The learner already distinguishes the sounds of the contrast into two different phonological categories. Therefore, after becoming an experienced learner (in a classroom setting or natural environment) or at the post-training stage (in a controlled setting), as is the case in this study, if the learner maintains (pretest-to-posttest) the SC type of assimilation for a particular contrast, no learning has taken place. In case the learner shows TC assimilation before training and maintains the same level of assimilation post training, again no learning has taken place since the learner is already maintaining a good distinction within the contrast. However, if the learner progresses from SC type at the initial exposure stage to TC type at the post-training stage, evidence of successful learning can be inferred. This shift in assimilation pattern would provide evidence that the learner has been able to place the sounds into two different phonetic categories at the end of the training. In general, any shift in assimilation patterns towards TC during the course of training would evidence learning. Consequently taking these assimilation types as the two extreme ends, a learning continuum is proposed (Figure 4-1).

Other assimilation types that are placed in this continuum are CG (Category-goodness), UC (Uncategorizable-categorizable) and UU (both uncategorizable). The ranking of these assimilation types can be speculated as follows. According to PAM, CG is a within-category assimilation, like SC type, and allows poor to good discrimination. UU and UC are both between-category assimilation patterns, similar to TC type, and may show moderate to excellent discrimination. Therefore, to determine direction of learning, a ranking similar to this can be assumed SC<CG<UU<UC<TC. This placement of assimilation types imply that if a learner shows shift in assimilation pattern from SC to CG or for that matter from any lower rank to a higher rank (SC to UC, CG to TC or CG to UC), learning is taking place.
However, this ranking sequence is questionable. The ranking of SC and TC assimilation patterns appear to be decisive on the continuum but this is not the case with CG, UU or UC. Looking back at the PAM model for a more detailed description of these assimilation types, a definitive distinction between the CG, UU and UC assimilation patterns seems missing.

According to PAM, CG assimilation type reflects the perception of a particular contrast as being within the same L1 phonological category but one phone being considered as the good exemplar of the category and the other as being the deviant exemplar of the same category. During the learning process, the learner may form a new phonetic category for the perceived poor/deviant phone of the L1 phonological category but only gradually. The good exemplar L2 phone may remain assimilated to the L1 phonological category. This definition ranks the CG type above SC type and lower to the TC type in terms of learning. Also, as a general assimilation pattern, CG is a within-category type which restricts the placement of CG on the continuum closer to SC as related to TC.

However, the assimilation type CG representing a within category assimilation pattern, which may equate to non-learning type, may be seen with a different perspective. Based on the results of the current study, it is observed that if the learner provides variable goodness ratings (e.g. ratings 2 and 5 on the scale of 1 to 7) for the target contrast implying that he perceives the sounds as good and bad exemplar of the same phonetic category at the pre-training level, and over the course of training shifts to perceiving the same target contrast consistently as two extremes of the ratings scale (at post-training ratings 1 and 7 on the scale of 1 to 7), then perhaps successful learning does take place. However, according to PAM, the general assimilation pattern of CG type is still within-category, that is, the learner still labels the modal response for each sound of the contrast as the same. Yet the learner consistently perceives these sounds as
different (one sound at the goodness rating of 1 (highly dissimilar to the modal response chosen from their native sound inventory) and the other sound at the goodness rating of 7 (highly similar to the modal response)) at the post-training level. The problem here lies with the orthography of the language. Since this study was a controlled laboratory training experiment and learners were asked to use an open set from their language systems for perceptual assimilation task, the learners did not have new orthographic representations for the newly formed phonetic categories to substitute and thus represent a TC type rather than a consistent CG type. However, in real life situation, this dilemma may not arise as a learner may use the orthographic representations corresponding to the phonetic categories of the target language while acquiring the non-native contrast. In other scenarios, where orthography of the target language is not available or not being used, the learner will have lexical items through which the phonemic distinction will be clear and thus, the learner can eventually form separate phonetic categories for these sounds and assimilate the contrast to a TC type. This may not be the case in controlled laboratory training settings with a limited time period which may result in subjects still using the same modal response but consistently discriminating the non-native contrast. Therefore, in this light, CG types do show a trend towards learning.

Another case of perceptual assimilation is that an assimilation pattern is termed UC if only one of the L2 phones is perceptually assimilated with an L1 phonological category and considered a good exemplar of that category. According to PAM, the discrimination of the speech contrasts would be excellent as the learner would always perceive the fully assimilated L2 phone separate from other less assimilated L2 phones. This places the UC type closer to the TC type on the continuum which means that having UC type at post-training stage would show a trend towards learning. However, with PAM specified definitions, the ranking of UC type to CG
type cannot be ascertained since CG type can allow for a wide range from poor to good
discrimination of the L2 contrast. Moreover, the UC assimilation type is not limited to just one
scenario as explained in the PAM model. Based on the results from this study it has been
observed that the learner provides various responses for categorization of the other L2 phone (the
uncategorizable one), one of which may overlap with the categorization of the fully assimilated
L2 phone. The discrimination, then, may not be as clear with the minimally contrasting words in
the target language. For example, the learner may assimilate one L2 phone consistently with L1
phonological category ‘x’ but assimilate the more deviant L2 phone to various other L1
phonological categories ‘y’, ‘z’ including ‘x’ category. In this scenario, UC assimilation type
cannot be interpreted to successful learning since at least ten to fifteen percent of the times the
L1 phonological category with which one L2 phone is fully assimilated, overlaps with the other
uncategorizable L2 phone. Therefore, even though the learner consistently identifies one L2
phone with one L1 category, the identification of the minimal contrast of the sounds is not learnt.
Perhaps, in the real life situation, when the uncategorizable L2 phone is placed in minimal
contrast with other sounds of L2, the learner may gradually form a new phonological category
for it.

Similarly, the case of UU assimilation type is also quite ambiguous when it comes to the
placement on the learning curve. The UU type represents assimilation where the learner does not
assimilate the L2 phones to any one of his L1 phonological categories. PAM claims that the
discrimination of this kind of speech contrast may depend on not just the comparative features of
the given sound to its closest L1 counterpart but also its relationship to other phones in the same
phonological space. If the L2 phones are assimilated to similar set of L1 sounds then
discrimination of this contrast is difficult for the learner since the acoustic proximity of the L2
phones to each other and to the set of assimilated L1 phones will be less. On the other hand, if the L2 phones are assimilated to different set of L1 sounds the acoustic distance between the L2 sounds and the set of assimilated L1 sounds will be much more and therefore would result in easier discrimination by the learner. Thus in the case of UU type, the level of discrimination depends on the inherent acoustic distance between those two sounds as well as the L1 categories to which they’ve been assimilated. Again, the possibility of overlapping of L1 categories among the sets of sounds that are assimilated to the L2 contrast is imminent in the UU type. Moreover, the listener may be providing more than one modal response for each category and still not be consistent with the goodness ratings. All these scenarios make it difficult to rank the UU assimilation type against the UC and CG types.

Following definitions specified in the PAM model, assimilation types like category goodness (CG) and the categorical (UC and UU) patterns cannot be allocated a ranking on the learning continuum decisively. If the general assimilation pattern is considered then CG (within-category) has to be placed lower than UU and UC (between-category). On the other hand, predictions of PAM allocate a very broad range of discrimination levels from poor to good for both CG and UU. Also, both UU and UC can possibly have lot of scenarios with varying assimilation results as discussed above. Therefore, ranking these highly sensitive assimilation types on the learning continuum solely based on PAM definitions of the assimilation patterns is problematic.

Based on the perceptual assimilation results in this study, an attempt is made to speculate a learning continuum which shows the assimilation types representing various stages of learning L2/L3 by adults. Figure 4-1 shows the hypothetical learning continuum with five different assimilation types representing different types of learning.
Figure 4-1. Hypothetical representation of the learning continuum displaying the ambiguity of ranking the UU, UC and CG assimilation types in order to assess the direction of learning.

The SC and TC assimilation types form the two extreme ends of the continuum where SC type at the post-training stage equates to no learning taken place and the TC type represents successful learning. The ranking of the CG, UU and UC assimilation types towards learning cannot be decided without more empirical evidence supporting the ranking. Therefore, based on the results of this study, the ranking of these three assimilation types can only be speculated. Between CG and UU types, even though CG assimilation is a within-category pattern its goodness ratings of the deviant L2 phone exemplar reveal that the learner perceives the L2 phones apart. Also, the perceptual assimilation results show a drastic increase in CG assimilation types with consistently large goodness rating differences when pre-training and post-training level assimilations are compared. UU assimilation type is a between-category assimilation pattern which implies that the L2 phones are assimilated to different set of L1 phones. Other between-category assimilation patterns like TC denote successful learning but same cannot be applied for UU type. As discussed above, UU assimilation type can result from many scenarios which may not confirm as evidence of learning. Therefore, CG can be ranked above UU assimilation type.

UC assimilation type can also result from many scenarios as discussed above. However, in this case, one L2 phone is fully assimilated to one of the L1 phonetic category which results in better discrimination of the L2 contrast. Moreover, a shift towards UC assimilation type at the
post-training stage signifies that the learner has perhaps progressed from not distinguishing the L2 phones at all to being consistently assimilating one L2 phone to one L1 phonetic category. This indicates progression towards learning and thus places the UC type above UU type and more towards the right extreme of the learning continuum, that is, the TC extreme. The overall perceptual assimilation results also validate this ranking as they show a decrease in the UU assimilation type and an increase in UC assimilation type at the post-training stage. Between UC and CG assimilation types, no definitive ranking can be stated between these two assimilation types since both the types show an increase at the post-training stage. Therefore, both these assimilation types can be positioned anywhere between UU and TC.

Thus, the ranking of assimilation types representing the learning continuum can be hypothesized as SC<UU<UC, CG<TC. However, this ranking is based on only the results of the current study. There is a need to examine these assimilation types over a learning period and determine whether they represent a direction towards learning. Further theoretical and empirical exploration of these assimilation types will contribute substantially to confirm the learning continuum and extend the current theoretical models of cross-language speech perception to models of speech perception in the due course of learning. It is essential to assess the efficacy of assimilation types like UU and UC in determining the perceptual assimilation patterns especially regarding those that are elicited after training or in real life situations from early learners where a shift in assimilation patterns can be determined in order to evaluate the progress towards learning.

Results in Relation To Previous Studies

This study was an attempt to answer the following questions - Does being a bilingual facilitate in the acquisition of a third language? If so, is this facilitation influenced by the effects of multilingual benefit which embody a developed general ability in balanced bilinguals to be
more sensitive to distinctions in auditory-acoustic stimuli and being able to abstract the linguistically relevant information to contrast the speech sounds and develop a new phonetic category? Secondly, would effects of feature generalization be observed in performance of learners who may already have linguistic experience with a phonetic feature which is present in the new non-native contrast as well? Lastly, what kinds of changes are seen in the assimilation patterns of non-native contrasts post training? The results of this study contributed information as well as a new range of data set to the existing literature in cross-language speech perception.

Previous studies have examined many factors that contribute towards the perception of second language (L2) sounds and, in some cases, the subsequent establishment of L2 phonetic categories. Phonetic, phonemic and acoustic factors among others have been determined to influence the learning of L2 sounds (Tees and Werker, 1984; Polka, 1991; Best, 1995; Flege, 1995). However, learning a third language, unlike learning a second language, may be influenced by additional factors attributed to the presence of two language systems in a bilingual instead of one language system in a monolingual (Cenoz & Valencia, 1994; Ardeo, 2000; Munoz, 2000; Bild & Swain, 1989). This hypothesis was based on the studies in lexical processing and word learning by bilingual and monolingual children showing that bilingual children are more adept at doing metalinguistic tasks and developing cognitive benefits (Peal and Lambert, 1962; Malakoff, 1992; Klein, 1995; Sanz, 2000 etc.). Numerous other studies looking at cognitive processing at a linguistic level have explored the concept of multilingual benefit in bilingual and multilingual children as well as adults which emphasized the fact that bilinguals with near native proficiency in both languages show evidence of multilingual benefit (Bialystok, 1992, 2001, 2004; see also review in Jessner, 2006). However, this concept of meta-linguistic awareness is not yet fully explored in the field of cross-language speech perception and the acquisition of new phonetic
categories by bilingual learners of an L3. The results of the present study extend the research of metalinguistic ability to the field of cross-language speech perception.

The study provides suggestive evidence that the bilinguals function at a level different from monolinguals in reorganizing the information gathered from the speech stream in order to deal with the high-demand task of acquiring new non-native speech contrasts. The results support the claims of previous studies on lexical processing with respect to the effects of multilingual benefit seen in acquiring a third language by bilinguals. For instance, the study by Klein (1995) showed that during lexicon acquisition of a language, multilinguals learned a higher number of lexical items than monolinguals. Klein based the explanation for this phenomenon on enhanced cognitive skills in multilinguals which helped them tease out the potential relevant data for resetting the particular parameter for the new language. Similarly in the present study, the results of the lateral non-native contrast in particular showed that bilinguals, with or without previous experience with the retroflex feature, displayed better perceptual performance than the monolingual group. The study reflects the effect of multilingual benefit which may be further explored as future research in speech perception.

The other factor that was explored in this study, feature generalization, did not show any effects. The exploration of this factor was based on the assumption that features, an important set of information contained in the established phonetic category, play a crucial role in developing new phonetic categories (Nosofsky, 1986, 1987; Kruschke, 1992; Jusczyk, 1989; Lively et al., 1993; Francis and Nusbaum, 2002). Few studies have looked at the feature generalization phenomenon in the field of cross-language speech perception. Polka (1992) and Harnsberger (1998) found results contrary to the expected results in this study that feature generalization will reveal better perceptual performance. Nevertheless, a study by McAllister, Flege & Piske (2002)
found in the data from a few subjects that experience with the duration feature did facilitate native American English speakers’ perception of Swedish vowels (phonemic lengthening), so that native AE speakers performed better than native Spanish speakers, who had no experience with duration feature. The difference between this study and others was that the subjects were ‘experienced’ learners of Swedish and used the target language often. Their study was suggestive of the hypothesis of feature generalization. Therefore, one factor that could explain the previous studies conflicting results is that these studies (Polka, 1991; Harnsberger, 1998) examined only naïve perceivers with speech perception experiment. The present study employed an extended approach which provided experience to the listeners through training, the next step from the previous studies where feature generalization was tested on only naïve listeners. The robustness of the high variability perceptual training program was assumed to be effective in providing sufficient linguistic exposure to the learners in order to bring forth any presence of the facilitating effects of these factors.

The present study was unable to provide the missing link between the studies where the feature generalization factor was explored but not found and the study which indirectly suggested the presence of feature generalization. Regardless of gaining from perceptual training with multiple talkers and high variability in stimuli tokens, Bengali-English speakers were not able to reach a native-like performance level. This language group’s perceptual performance in both testing areas of identification and discrimination was not indicative of any robust effects of feature generalization. Although, Bengali-English speakers performed significantly better than the monolingual group of American English speakers, they were unable to display an additive effect over the metalinguistic ability found in the Spanish-English speakers as well. Since both Bengali-English speakers and Spanish-English speakers performed near equally in most of the
testing measures, the performance of a bilingual group (SE) without any previous experience
with the retroflex feature negated any trace of the presence of feature generalization that may
have been possible. Thus, the present study confirmed the results of the earlier studies (Polka,
1992; Harnsberger, 1998; Tajima et al., 2008) that whether the listeners are naïve or received
limited implicit training, they are unable to generalize the phonetic features with which they have
had previous linguistic experience. Additionally, recent studies have shown that only phonemic
experience/ language experience over time facilitates the perception of non-native contrast and
that simultaneous bilinguals acquire native like phonetic categories which are developed fully by
adolescence (Silver, 2007; Sundara et al., 2006, 2007). This may explain the results of the current
study where no feature generalization was observed. The lack of extensive exposure of the
stimuli to the language groups is one of the explanations provided for results showing no support
for the feature generalization hypothesis. This possibility can be explored in future research work
by elongating the span of training and providing extensive exposure of the non-native contrast to
subjects. Extensive training may result in observance of effects of language experience in this
respect.

The third focus of the study concerned the role of the bilinguals’ particular perceptual
category inventory in the acquisition of non-native contrasts from a third language. Several
models have been developed to account for the influence of native perceptual categories on the
perception and acquisition of non-native speech sounds, including the Speech Learning Model
(SLM) and, of greater interest for this project, the Perceptual Assimilation Model (PAM). PAM
concerns the discriminability and learning of non-native contrasts based on the relationship
between the non-native and native category inventories. These relationships have been
encapsulated in several assimilation types which were adopted for this study. The results of the
present study assessed the extent to which initial patterns of assimilation change following limited training experience and measured the effects of the multilingual factor within the perceptual assimilation pattern shift. The results revealed a shift of perceptual assimilation types towards perceptual learning. On an average, the assimilation types SC and UU decreased and the assimilation types TC, CG and UC increased from the pre training phase to post training phase. However, the relationship of learning could not be deciphered as the PAM model does not provide a definitive ranking of UU, UC (categorical assimilations) and CG (within category assimilation) assimilation types with respect to learning. In addition, the multilingual effects were also observed within the results of perceptual assimilation tests. The results were suggestive of the positive perceptual learning seen among bilinguals which appeared to be higher than the meager amount of perceptual learning that took place among the monolingual group. This study provides a substantial empirical dataset which can be used to explore the issue of learnability ranking among the assimilation types as future research work.

Another area of future research that could be pursued would be to examine the rhotic segment (orthographic symbol: ﹪) of Malayalam which lacks definitive phonetic description. This sound, as discussed in Chapter 2, has been presented in previous literature with various articulatory descriptions (Kumari 1972, Sreedhar 1972, Asher & Kumari 1997, Kalackel 1985, Raja 1960, Bright 1998, Krishnamurti, 2003). A clear phonetic description of this particular segment of Malayalam consonant inventory is required. One method that can be used to describe the features of this rhotic segment is acoustic analysis. It is observed that most of the descriptions affirm the feature of retroflexion in this sound. Previous studies on Tamil liquids also report the occurrence of similar sound which includes features of retroflexion and central airflow (Narayanan et al 1996, 1999, McDonough & Johnson 1997). This type of research of the
Malayalam sound as well could provide a definitive phonetic description of the sound which could be compared to similar sound segment in other Dravidian languages and contribute to the knowledge of phonetic sounds of the Malayalam language.

**Conclusion**

This dissertation extends the research on cross-language speech perception and language acquisition to a new range of contrasts, different manners such as lateral, nasal, fricative and rhotic with alveolar – retroflex as the place distinction. The study is successful in determining whether current theories and established claims of multilingual benefit in the field of lexical processing were generalizable to the area of cross-language speech perception in regards to learning new non-native contrasts similar to that of leaning of the lexicon. The results suggested that the effects of multilingual benefit were prevalent in the learning of novel contrasts with limited training period. However, the claim on the effects of feature generalization received no support in the results, therefore, confirming the findings of the earlier studies.

In the area of cross-language perceptual classification, the study provides a range of empirical datasets for assessing the cross-language differences in the perceptual assimilation patterns observed before and after training. It points to open questions of ranking of assimilation types in terms of learnability that may strengthen the predictability of perceptual assimilation in determining the direction of learning. These concerns need to be addressed by the current models of perceptual assimilation. Beyond these general findings, the study also demonstrated the need for larger sampling size in order to get robust effects since the effects of multilingual benefit observed were merely suggestive in nature and seen along with lot of individual variations.

Overall, the results presented here lead to suggest a productive line of research for future work discussed in the above section. The caveats such as looking into the baseline talker intelligibility issue within training and testing; establishing learnability ranking of the assimilation patterns;
examining with exclusive focus on the effects of multilingual benefit were revealed during the progress of this study. These caveats will provide substantial insight and enhance the productivity of the future research.

In summary, the results supported only the hypotheses of multilingual benefit postulated in this study. The factor of multilingual benefit is brought forth and explored within the realm of cross-language speech perception field through the results of the study.
APPENDIX A
PRELIMINARY SCREENING

Subject Information form

Subject code (not to be filled by participant):

Name:

Age:

Email:

First language(s):

Language(s) you speak with your parents/grandparents/siblings:

Spoken proficiency in languages other than first language(s):

Language classes that you might have taken in Middle/high school:

Countries that you’ve visited (also give length of stay if more than 2 months):

History of any hearing impairment:

If participating in the study for credits specify name of course and Instructor:
Screening for Spanish Spoken Proficiency

This was a brief spoken test conducted at the time of screening the monolingual subjects as well as bilingual Spanish-English subjects. The subjects were asked to say out loud, at normal speaking rate, two phrases which are considered as tongue twisters in Spanish. Then they were asked whether they were aware of the meaning of the phrase or certain words in them. This brief proficiency test was conducted to assess the level of spoken proficiency as well as lexical knowledge of the Spanish language. In case of monolinguals, if the candidates were not able to maintain a normal speaking rate with less than 40% correct pronunciation (that is, they substituted English sounds like [t] instead of [r], [t] instead of [t], [g] instead of [γ] and dropped the [e] at the end of word “roque” etc.), they could safely be considered under the monolingual language group, provided they had exposure to no other language. In case of self-reported Bilingual candidates, high level of spoken proficiency was considered along with complete knowledge of lexical items and sentence structure.

Hay tres tristes tigres en un trigal.
[aɪ . ˈtres . tris . tįs . γre . se . nuŋ . tri . γal] IPA transcription

El perro de san roque no tiene rabo.
[el . pe . ro . ðe . sūn . ro . ke . no . tje . ne . ra . ßo] IPA transcription
First time interested participant email:

Thanks for the interest in the study!
You can see more details about the study at
http://web.csd.ufl.edu/langbrain/Other_Research.htm#_Perception_training_study_2

Which of the language groups would you put yourself into? Before I go ahead and recruit you, you'll have to provide me with some information on your language background.

At what age did you start learning your second language, if any? What other languages have you been exposed to and for how long? (Please specify middle/high school, college, trips to places etc.)
Also, did you work in any school or college projects involving a different language? Please let me know, if you’ve had any roommates or friends from whom you were exposed to a language other than English.
What is your major, minor and year at UF? Also, where did you learn about the study?
If you’ve any questions, please feel free to ask!

Thank you!

Divya
**APPENDIX B**  
**SUBJECT INFORMATION**

**Individual Subject Attributes**

<table>
<thead>
<tr>
<th>Name code</th>
<th>Age</th>
<th>Spanish Dialect</th>
<th>First Language(s)</th>
<th>Age of Exposure (L2)</th>
<th>Linguistic background</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE01</td>
<td>20</td>
<td>Venezuelan</td>
<td>English, Spanish</td>
<td>--</td>
<td>Portuguese - average proficiency, French - low proficiency</td>
</tr>
<tr>
<td>SE02</td>
<td>23</td>
<td>Panamanian</td>
<td>English, Spanish</td>
<td>--</td>
<td>exposure to Arabic</td>
</tr>
<tr>
<td>SE03</td>
<td>20</td>
<td>Salvadorian</td>
<td>Spanish</td>
<td>2yrs</td>
<td>English, 2 years of French</td>
</tr>
<tr>
<td>SE04</td>
<td>21</td>
<td>Cuban</td>
<td>Spanish, English</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE05</td>
<td>23</td>
<td>Argentinean</td>
<td>Spanish</td>
<td>4yrs</td>
<td>English, Portuguese- average proficiency</td>
</tr>
<tr>
<td>SE06</td>
<td>19</td>
<td>Colombian</td>
<td>Spanish</td>
<td>4yrs</td>
<td>English, a little French</td>
</tr>
<tr>
<td>SE07</td>
<td>20</td>
<td>Colombian</td>
<td>English</td>
<td>6yrs</td>
<td>Spanish</td>
</tr>
<tr>
<td>SE08</td>
<td>20</td>
<td>Mexican</td>
<td>Spanish, English</td>
<td>--</td>
<td>Introductory French and Portuguese</td>
</tr>
<tr>
<td>SE09</td>
<td>19</td>
<td>Cuban</td>
<td>Spanish</td>
<td>4yrs</td>
<td>English</td>
</tr>
<tr>
<td>SE10</td>
<td>20</td>
<td>Panamanian</td>
<td>English, Spanish</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE11</td>
<td>22</td>
<td>Cuban</td>
<td>English, Spanish</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE12</td>
<td>19</td>
<td>Salvadorian</td>
<td>Spanish, English</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE13</td>
<td>20</td>
<td>Nicaraguan</td>
<td>Spanish, English</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE14</td>
<td>19</td>
<td>Colombian</td>
<td>Spanish, English</td>
<td>--</td>
<td>2 years of French (Middle school)</td>
</tr>
<tr>
<td>SE15</td>
<td>19</td>
<td>Mexican</td>
<td>English, Spanish</td>
<td>--</td>
<td>a little French</td>
</tr>
<tr>
<td>SE16</td>
<td>20</td>
<td>Mexican</td>
<td>Spanish</td>
<td>4yrs</td>
<td>English</td>
</tr>
<tr>
<td>SE18</td>
<td>19</td>
<td>Cuban</td>
<td>Spanish</td>
<td>3 yrs</td>
<td>English</td>
</tr>
<tr>
<td>SE19</td>
<td>21</td>
<td>Mexican</td>
<td>Spanish, English</td>
<td>--</td>
<td>N.A</td>
</tr>
<tr>
<td>SE20</td>
<td>21</td>
<td>Colombian</td>
<td>English</td>
<td>2yrs</td>
<td>Spanish</td>
</tr>
<tr>
<td>SE21</td>
<td>20</td>
<td>Nicaraguan</td>
<td>Spanish, English</td>
<td>--</td>
<td>N.A</td>
</tr>
</tbody>
</table>
### Bengali-English Speakers

<table>
<thead>
<tr>
<th>Name Code</th>
<th>Age</th>
<th>City, State</th>
<th>First Language(s)</th>
<th>Age of exposure (L2)</th>
<th>Linguistic Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE01</td>
<td>28</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, Hindi</td>
</tr>
<tr>
<td>BE02</td>
<td>24</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>6 yrs</td>
<td>English, a little Hindi</td>
</tr>
<tr>
<td>BE03</td>
<td>27</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>3 yrs</td>
<td>English, formal education in Hindi</td>
</tr>
<tr>
<td>BE04</td>
<td>25</td>
<td>Kolkota, West Bengal</td>
<td>Bengali, English</td>
<td>--</td>
<td>English, Hindi, Sanskrit</td>
</tr>
<tr>
<td>BE05</td>
<td>30</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>3 yrs</td>
<td>English, formal education in Hindi</td>
</tr>
<tr>
<td>BE06</td>
<td>26</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>6 yrs</td>
<td>English, a little Hindi</td>
</tr>
<tr>
<td>BE07</td>
<td>32</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, Hindi</td>
</tr>
<tr>
<td>BE08</td>
<td>28</td>
<td>Kolkota, West Bengal</td>
<td>Bengali, English</td>
<td>--</td>
<td>English, Hindi, exposure to Assamese and Gujarati</td>
</tr>
<tr>
<td>BE09</td>
<td>32</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, a little Hindi</td>
</tr>
<tr>
<td>BE10</td>
<td>24</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>2 yrs</td>
<td>English, Hindi, exposure to Urdu and Kannada (3 months)</td>
</tr>
<tr>
<td>BE11</td>
<td>24</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, formal education in Hindi</td>
</tr>
<tr>
<td>BE12</td>
<td>24</td>
<td>Kolkota, West Bengal</td>
<td>Bengali, Hindi</td>
<td>--</td>
<td>English, formal education in Hindi, exposed to Telugu (6 months)</td>
</tr>
<tr>
<td>BE13</td>
<td>25</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>6 yrs</td>
<td>English, a little Hindi</td>
</tr>
<tr>
<td>BE14</td>
<td>27</td>
<td>Chandan Nagar, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, Hindi, exposure to Kannada (20 months)</td>
</tr>
<tr>
<td>BE15</td>
<td>27</td>
<td>Chakda, West Bengal</td>
<td>Bengali</td>
<td>3 yrs</td>
<td>English, a little Hindi and Nepali</td>
</tr>
<tr>
<td>BE16</td>
<td>29</td>
<td>Barddhaman, West Bengal</td>
<td>Bengali</td>
<td>5 yrs</td>
<td>English, Hindi, Sanskrit, German, a little Urdu</td>
</tr>
<tr>
<td>BE17</td>
<td>28</td>
<td>Dhaka, Bangladesh</td>
<td>Bengali</td>
<td>12 yrs</td>
<td>English, a little Hindi, Assamese and Oriya</td>
</tr>
<tr>
<td>BE18</td>
<td>30</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>3 yrs</td>
<td>English, a little Hindi and Assamese</td>
</tr>
<tr>
<td>BE19</td>
<td>28</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>6 yrs</td>
<td>English, Assamese, a little Hindi</td>
</tr>
<tr>
<td>BE20</td>
<td>32</td>
<td>Kolkota, West Bengal</td>
<td>Bengali</td>
<td>8 yrs</td>
<td>English, Hindi</td>
</tr>
</tbody>
</table>
### American English Speakers

<table>
<thead>
<tr>
<th>Name Code</th>
<th>Age</th>
<th>First Language(s)</th>
<th>Linguistic Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE01</td>
<td>23</td>
<td>English</td>
<td>Spanish - Middle/High school</td>
</tr>
<tr>
<td>AE02</td>
<td>20</td>
<td>English</td>
<td>2 years of Spanish</td>
</tr>
<tr>
<td>AE03</td>
<td>19</td>
<td>English</td>
<td>Spanish - Middle/High school</td>
</tr>
<tr>
<td>AE04</td>
<td>23</td>
<td>English</td>
<td>a little Spanish</td>
</tr>
<tr>
<td>AE05</td>
<td>19</td>
<td>English</td>
<td>Introductory French and Spanish</td>
</tr>
<tr>
<td>AE07</td>
<td>20</td>
<td>English</td>
<td>a little Spanish and French</td>
</tr>
<tr>
<td>AE08</td>
<td>19</td>
<td>English</td>
<td>a little Spanish</td>
</tr>
<tr>
<td>AE09</td>
<td>22</td>
<td>English</td>
<td>2 years of Latin, exposed to Spanish</td>
</tr>
<tr>
<td>AE10</td>
<td>19</td>
<td>English</td>
<td>1 year of Latin</td>
</tr>
<tr>
<td>AE11</td>
<td>19</td>
<td>English</td>
<td>2 years of Spanish(High school), 2years of Latin (Middle School)</td>
</tr>
<tr>
<td>AE12</td>
<td>20</td>
<td>English</td>
<td>1 year of Spanish (Middle School)</td>
</tr>
<tr>
<td>AE13</td>
<td>20</td>
<td>English</td>
<td>1 year of Spanish (Middle School)</td>
</tr>
<tr>
<td>AE14</td>
<td>19</td>
<td>English</td>
<td>3 years of Spanish (Middle/High school)</td>
</tr>
<tr>
<td>AE15</td>
<td>19</td>
<td>English</td>
<td>3 years of Spanish (High school)</td>
</tr>
<tr>
<td>AE16</td>
<td>22</td>
<td>English</td>
<td>a little Spanish and Italian</td>
</tr>
<tr>
<td>AE17</td>
<td>19</td>
<td>English</td>
<td>a little French, exposure to Spanish</td>
</tr>
<tr>
<td>AE18</td>
<td>20</td>
<td>English</td>
<td>2 years of Spanish (Middle School)</td>
</tr>
<tr>
<td>AE20</td>
<td>21</td>
<td>English</td>
<td>2 years of Spanish (High School), 1 year of Latin</td>
</tr>
<tr>
<td>AE21</td>
<td>22</td>
<td>English</td>
<td>2 years of Spanish (Middle School)</td>
</tr>
<tr>
<td>AE22</td>
<td>19</td>
<td>English</td>
<td>1 year of Spanish (Middle School)</td>
</tr>
</tbody>
</table>
APPENDIX C
EXPERIMENT INSTRUCTION SHEET

Instructions given to subjects on the first day of the experiment (Pretest Phase):

Session 1 - Instructions

Welcome to the Linguistics Laboratory. We appreciate your participation in this experiment and also hope you will find it interesting.

You will be participating in a study of the perception training of new sounds (consonants – for example, the first sound “l” in the word “lamp”) from another language. This training study will carry on for 8 sessions. Today is your first session. In this session, you will have to do four tasks.

Task 1  Familiarization:

In this task, you will listen to consonants from another language. Each consonant sound will be preceded and followed by a vowel sound (example : “ana” or “ulu”). Each consonant sound, regardless of the vowels, is associated with a symbol. You will hear a word like “ulu” and then see a symbol “ l ”. You do not have to click any response button for this task. All you’ve to do is learn to associate the consonant sound with the symbol. This may be difficult initially as the sounds and symbols are from another language. This will take no more than 6 minutes.

Task 2  Identification task:

In this test, you will listen to one word like “ana” and then you’ve to choose the best possible symbol that you identify the consonant sound with (example: after you hear “ana”, you click on “n”). There will be total of 8 consonant symbols (l, L, n, N, sh, S, r, R) to respond from. Again, use the knowledge gained from the familiarization task. Since some of the sounds will be new to you, they may be difficult for you to tell apart. However, even if you’re not entirely sure
which consonant symbol it is, just make the best choice possible and move on to the next sound. This will take about 15-20 minutes.

Task 3  Discrimination task:

In this test, you will hear two words in a row, separated by a short pause. These words are from another language, and each word consists of a consonant preceded and followed by a vowel sound. Your task is to determine whether the two consonants are the same consonant or different consonants. If the two words have the same consonant, click on the “same” button on the screen. If the two words have different consonants, click on the “different” button on the screen.

For example, you might hear something like the following:

ala  (pause)  ana

In this example, the consonant in the first word is “l” and the consonant in the second word is “n”. Since these are different consonants, you would choose the “different” button. This is a relatively easy example. However, in this test, you will be hearing words from another language. Some of these words may be difficult for you to tell apart. However, even if you are not entirely sure if the consonants are the same or different, just make the best choice possible and move on to the next set of words.

Task 4  Instructions for Perceptual Assimilation task:

In this task, you will listen to words from another language, similar to previous tasks. You must choose the best possible sound from your native language that closely corresponds to the sound given. Write the letter (or letters) best representing the speech sound on the sheet provided. Then compare the non-native sound with closely related sound from your own language. On the scale from 1-7, circle the number that you feel is appropriate in comparing the similarity of the non-native sound with that of your own language sound.
For Example, if you hear “ala” and your native language is American English, you may choose the closest sound from you language as being “l as in laugh”. If you choose this, write “l” on your answer sheet and then scale the similarity of the two sounds. How similar are they? If they sound exactly same to you, give the scale number “7- exactly same”. Similarly, you can use other options on the scale as well depending on your judgment.

1 2 3 4 5 6 7

Very different    exactly same

This is a paper-pencil task, so you’ll listen to the sound via headphones and record all your responses on the sheet provided.
APPENDIX D
PERCEPTUAL ASSIMILATION RESPONSE SHEET

Note: This is a sample page of the perceptual assimilation response sheet. In all, 80 responses were elicited in this task.

<table>
<thead>
<tr>
<th>Name:</th>
<th>Subject Code:</th>
<th>Task:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial No.</td>
<td>Letter symbol</td>
<td>Goodness Rating (1-7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Very different</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

136
APPENDIX E
RESULTS IN TABULAR FORMAT

Table E-1. Mean percent identification scores averaged over all four contrasts at the pretest and the posttest level.

<table>
<thead>
<tr>
<th>Language groups</th>
<th>Pretest scores</th>
<th>Posttest scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE</td>
<td>51</td>
<td>63</td>
</tr>
<tr>
<td>BE</td>
<td>61</td>
<td>80</td>
</tr>
<tr>
<td>SE</td>
<td>56</td>
<td>74</td>
</tr>
</tbody>
</table>

Table E-2. Mean d’ scores of AX discrimination test averaged over all four contrasts at the pretest and posttest level.

<table>
<thead>
<tr>
<th>Language groups</th>
<th>Pretest scores</th>
<th>Posttest scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE</td>
<td>0.85</td>
<td>1.3</td>
</tr>
<tr>
<td>BE</td>
<td>1.17</td>
<td>1.94</td>
</tr>
<tr>
<td>SE</td>
<td>0.95</td>
<td>1.64</td>
</tr>
</tbody>
</table>

Table E-3. Mean percent identification scores of contrasts spoken by trained-on talker (posttest) and new talker (generalization test).

<table>
<thead>
<tr>
<th>Language groups</th>
<th>Posttest scores</th>
<th>Generalization scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE</td>
<td>63</td>
<td>64</td>
</tr>
<tr>
<td>BE</td>
<td>80</td>
<td>82</td>
</tr>
<tr>
<td>SE</td>
<td>74</td>
<td>75</td>
</tr>
</tbody>
</table>

Table E-4. Mean percent scores of consonant identification test for all four contrasts by the bilingual groups (BE and SE) and the monolingual group (AE) at the pretest and the posttest levels.

<table>
<thead>
<tr>
<th>Language groups</th>
<th>Laterals</th>
<th>Nasals</th>
<th>Fricatives</th>
<th>Rhotics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>AE</td>
<td>47</td>
<td>55</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td>BE</td>
<td>65</td>
<td>90</td>
<td>59</td>
<td>80</td>
</tr>
<tr>
<td>SE</td>
<td>56</td>
<td>74</td>
<td>54</td>
<td>68</td>
</tr>
</tbody>
</table>

Table E-5. Mean d’ scores of AX discrimination test across all four contrasts, by the bilingual groups (BE and SE) and the monolingual group (AE) at the pretest and the posttest levels.

<table>
<thead>
<tr>
<th>Language groups</th>
<th>Laterals</th>
<th>Nasals</th>
<th>Fricatives</th>
<th>Rhotics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>AE</td>
<td>1.00</td>
<td>1.04</td>
<td>0.47</td>
<td>0.76</td>
</tr>
<tr>
<td>BE</td>
<td>1.58</td>
<td>2.77</td>
<td>0.90</td>
<td>1.38</td>
</tr>
<tr>
<td>SE</td>
<td>0.86</td>
<td>1.75</td>
<td>0.49</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Table E-6. Reanalysis results: Mean percent identification as well as AX discrimination scores averaged over all four contrasts at the pretest and the posttest level for the bilingual groups (BE and SE) combined versus the monolingual group (AE).

<table>
<thead>
<tr>
<th>Language factor</th>
<th>ID Test Score</th>
<th>AX Discrimination score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Monolingual</td>
<td>51</td>
<td>63</td>
</tr>
<tr>
<td>Bilingual</td>
<td>58</td>
<td>77</td>
</tr>
</tbody>
</table>

Table E-7. Mean percent correct response (averaged over contrasts) for identification of the three place-contrasts (laterals, nasals, fricatives) by bilingual (BE and SE) and monolingual (AE) groups at the pretest and posttest levels.

<table>
<thead>
<tr>
<th>Language groups</th>
<th>ID Test Score</th>
<th>AX Discrimination score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>AE</td>
<td>48</td>
<td>57</td>
</tr>
<tr>
<td>BE</td>
<td>59</td>
<td>77</td>
</tr>
<tr>
<td>SE</td>
<td>52</td>
<td>69</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Divya Verma Gogoi received her Ph.D. in linguistics from the University of Florida in 2010. During her time at UF, she worked as a teaching assistant for the Department of Linguistics and worked as a research assistant to Dr. Caroline Wiltshire. Her doctoral degree was supported by a grant from Language Learning: A Journal of Research in Language Studies. Before enrolling as a graduate student at UF, she received her B.A and then M.A in English from the Panjab University, Chandigarh, India in 1999 and 2001 respectively. Thereafter, she completed her M.Phil in Linguistics at the Central Institute of English and Foreign Languages, India (now The English and Foreign Languages University) in 2005 with specialization in Phonetics and Phonology. Her research interests are Phonetics and second/third language acquisition; more specifically, acoustic phonetics and cross-language speech perception and learning.