Crosslinguistic Influence and Crosslinguistic Interaction in Multilingual Language Learning

Edited by
Gessica De Angelis, Ulrike Jessner and Marijana Kresić

Bloomsbury Academic
An imprint of Bloomsbury Publishing Plc

B L O O M S B U R Y
LONDON • NEW DELHI • NEW YORK • SYDNEY

Bloomsbury Academic

An imprint of Bloomsbury Publishing Plc

50 Bedford Square
London
WC1B 3DP
UK

1385 Broadway
New York
NY 10018
USA

www.bloomsbury.com

BLOOMSBURY and the Diana logo are trademarks of Bloomsbury Publishing Plc

First published 2015

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN: HB: 978-1-4742-3585-3
ePDF: 978-1-4742-3587-7
ePub: 978-1-4742-3586-0

Library of Congress Cataloging-in-Publication Data
TO COME

Typeset by Deanta Global Publishing Services, Chennai, India
Printed and bound in Great Britain

6

Perceptual Training of Novel Speech Contrasts in L3 Acquisition: The Effect of Multilingual Benefit

Divya Verma Gogoi, James D. Harnsberger and Caroline Wiltshire[1]

1 Introduction

The globalization of the world has led to many scenarios whereby people with different native languages come face to face with each other and the need to communicate arises. Studies of second language acquisition (SLA) have looked into various factors that influence the process of acquiring a new language by adults who have only one language system (see Flege 1995; Kirsner et al. 1984; Grosjean 1992; Krashen 1981; Kroll and Tokowicz 2005; De Bot 1992). Until recently, research on SLA assumed that there was no distinction between the factors influencing the underlying learning process by an L2 learner and an L3 or an L5 learner. However, research in the field of third language acquisition or, in broader perspective, multilingualism, has made strides in the last few years (Cenoz, Hufeisen and Jessner 2001; Baker 2001; Bialystok 2001; Munoz 2000; Sanz 2000; Malakoff 1992; Cenoz 2001. Fewer studies have looked into third or additional language acquisition within specific areas of proficiency like phonological acquisition (Enomoto 1994; Werker 1986). Previous third or additional language acquisition research has suggested that being bilingual provides a positive influence for learners of third languages in attaining general proficiency (Cenoz and Valencia 1994; Gonzalez-Ardeo 2000; Munoz 2000; Bild and Swain 1989). In other words, studies examining language acquisition beyond L2 suggest that prior knowledge of non-native language(s) and previous learning experience significantly affect the language acquisition process.

In the present study, an experiment is set up to examine the effect of a multilingual benefit on learning a new language by adult bilinguals after providing perception training in a laboratory setting. The goal of the present study is to examine the effects of a multilingual benefit on the acquisition of a target language, Malayalam, by two bilingual groups, Bengali-English and Spanish-English speakers, and one monolingual group of American-English speakers. Within the bilingual groups, the

Bengali-English speakers were chosen since their L1, Bengali, is typologically closer (relative to Spanish) to Malayalam. The second bilingual group of Spanish-English speakers were chosen because their L1, Spanish, is typologically distant from the target language. By presenting the subjects with novel speech contrasts (retroflex 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 bi/multilingual are seen after the training period.

Before the introduction of the hypothesis, a brief discussion on bi/multilingual group classification is warranted. De Angelis (2007) has cited many studies where the term bilingual has been used to encompass multilingual individuals as well and vice versa (*The Concise Oxford Dictionary of Linguistics*, Matthews 1997; Myers-Scotton 2002; Grosjean 1992), including, for instance, definitions like 'bilingualism is the regular use of two (or more) languages, and bilinguals are those people who need and use two (or more) languages in their everyday lives.' (Grosjean 1992: 51). In addition to that, until recently researchers have used the term L2 to refer not only to the second language but also to the third language and beyond. Proponents of a distinction between the acquisition of a second and a third language have, in their studies, referred to third or additional language acquisition as L3, not considering the fact that it may be a third, fourth or fifth language. Most of these studies are based on a fractional view of bilingualism where two monolingual language systems were visualized within the bilingual mind as opposed to the holistic view of bilingualism (Grosjean 1992). The same holistic view can be extended to multilingualism where languages beyond L2 in the mind of a multilingual share an intact system as opposed to separate linguistic systems for each of the non-native languages (De Angelis 2007).

Looking at the 'bilingual' groups in the present study, the holistic view of bi/multilingualism was adapted. The subjects in this group were balanced bilinguals with regard to the L1/L2 (Bengali/English and Spanish/English) considered for this study. Some of the subjects were multilinguals and had other non-native linguistic information in their minds (all subjects self-reported low proficiency in their L3s) but these languages were typologically distant from the target language. For the purpose of this study, all bi/multilinguals were considered under one 'bilingual' group for two reasons: First, the hypothesis of the study concerned the general effect, when compared with monolinguals, of having two or more established languages on acquiring the speech contrasts of a third or additional language. Secondly, the other non-native languages that multilinguals in this group had were carefully considered not to have any phonemic features in their sound systems similar to those in the speech contrasts chosen for training, so as to avoid any crosslinguistic influence. The notion that there could be additional or different factors that may influence acquisition of languages beyond L3, while interesting, is beyond the scope of the present study.

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 unknown target language. Such results would argue for the positive effect of a multilingual benefit on third or additional language acquisition in the specific area of phonetics.

The study also aims to examine the changes in assimilation patterns of non-native speech contrasts by the subject groups at both the initial perceptual stage and post-training stage by adopting the assimilation types described in the perceptual assimilation model (PAM) proposed by Best and Tyler (2007), Best (1995).

The specific sound contrasts from Malayalam that are used as stimuli are lateral [l-ɭ], nasal [n̪-ɳ] and fricative [ʃ-ʂ]. The experiment set up for this study includes various tests of identification, AX discrimination and perceptual assimilation in order to observe the effects of a multilingual benefit on different levels of perception.

2 Previous research

Acquisition of more than two languages represents more of a global norm than an exception and, therefore, it is laudable that research in the field of third language acquisition and multilingualism has made progress in recent years (e.g. Bialystok 2001; Jessner 2006; De Angelis 2007). Third or additional language (L3) acquisition refers to the learning of a non-native language by an individual who has already acquired two (or more) languages either simultaneously (before puberty) or sequentially. The scenario of L3 acquisition, although quite similar to second language (L2) acquisition, still presents differences since, unlike L2 learners, L3 learners have more experience as language learners and have access to two linguistic systems when acquiring a third language (Hoffmann 2001; Cenoz 2003; Wrembel 2011). Most previous studies on L3 acquisition have looked into the question of whether bilingualism facilitates the acquisition of third language. The focus of these studies has been primarily to examine the level of general proficiency in the target language through various linguistic tasks like word awareness, grammar, speaking, reading and writing (studies typically conducted with children in classroom settings). Their results have supported the hypothesis that bilingualism facilitates the acquisition of the third language at those corresponding levels of linguistic structure (Bild and Swain 1989; Cenoz and Valencia 1994; Enomoto 1994; Klein 1995; Gonzalez-Ardeo 2000; Sanz 2000; Cenoz, Hufeisen and Jessner 2001; Bialystok 2001; Herdina and Jessner 2002; Cheung et al. 2010).

Of interest is whether this observed bilingual benefit also extends to the phonological/phonetic level of analysis, in which several other factors strongly influence the expressive and receptive abilities of L3 learners. Recent studies in L3 acquisition have also looked at the crosslinguistic influence of L1 and L2 languages on third language learning and found mixed results, showing L2 language influences, combined with an L1 transfer effect, in L3 phonology acquisition (Wrembel 2010, 2011; Llama, Cardoso and Collins 2010; Wunder 2011; Bono 2011). Other factors that may affect L3 acquisition are quite similar to the typical factors that affect acquisition of a second language, such as the age of acquisition and the frequency of L2 usage (Flege 1998; Flege, MacKay and Piske 2002); first language (L1) transfer of phonetic, phonemic or featural properties (Nosofsky 1986, 1987; Best, McRoberts and Sithole 1988; Pruitt, Jenkins and Strange 2006; Kruschke 1992; Lively, Logan and Pisoni 1993; Francis and Nusbaum 2002; Sundara, Polka and Genesee 2006;

Aliaga-Garcia and Mora 2008; So and Best 2010); the acoustic robustness of target non-native contrasts (Burnham 1986); socio-economic factors (Hammarberg 2001) and exposure to a greater variety of acoustic-phonetic features that compose perceptual categories (Werker et al. 1981; Tees and Werker 1984; Werker and Tees 1984; Strange and Dittmann 1984; Polka 1992). The existence of so many influences on L2 phonological/phonetic acquisition raises the question of exactly what in the experience of a bilingual may enhance the capacity to acquire a third language? A bilingual with a near-native proficiency in both languages may be aided by an expanded overall inventory of perceptual categories onto which L3 speech sounds may 'assimilate', resulting in more two-category or category goodness assimilations, to use the taxonomy of PAM.

The bilingual benefit may also be a function of an enhanced *metalinguistic awareness* of the systematic sound pattern differences between multiple languages. Metalinguistic awareness 'refers to the ability to focus attention on language as an object or to think abstractly about language and, consequently, to play with or manipulate language' (Jessner 2006: 42). According to Masny (1997), metalinguistic awareness is an indicator of what learners know about language through reflection on and manipulation of language (as stated in Jessner 2006: 43). Metalinguistic awareness here concerns the ability of the bilingual L3 learner to attend to the structural features of the language in order to abstract the knowledge of its distinguishing cues in a speech contrast, a task that demands certain cognitive and linguistic skills. 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 performing metalinguistic tasks and exhibit cognitive benefits from knowing two or more languages (Peal and Lambert 1962; Klein 1995; Munoz 2000; Sanz 2000; Cheung et al. 2010 etc.). Numerous studies looking at cognitive processing[2] at a linguistic level in bilingual and multilingual children and adults emphasize the fact that bilinguals with near-native proficiency in both L1 and L2 languages show evidence of metalinguistic awareness (Bialystok 1992, 2001, Bialystok et al. 2004; see also review in Jessner 2006).

Another focus of this study concerns the role of the bilinguals' particular perceptual category inventory in the acquisition of non-native contrasts from a third language. Previously, several 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) (Flege 1995) and, of greater interest for this project, PAM, since it predicts the assimilation patterns not only at the initial exposure stage but also beyond the first 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 that is consistently identified with a single native sound but differs in goodness of fit) and uncategorizable, either UC or UU (involving one/both non-native speech sounds that are not consistently identified with a particular native category). The present study uses these predictions in: (1) Determining change in assimilation types, if at all, following the training period and (2) measuring the extent of assimilation pattern change among subject groups for any effects of the multilingual benefit factor explored here.

3 Methodology

To document the existence of a phonological/phonetic bilingual benefit, two groups of bilingual learners were contrasted with a group of monolinguals in acquiring a set of third-language speech contrasts (not known to subjects) in a controlled setting. The use of two bilingual groups was necessitated by the particular need here for generalization: any observed differences between a bilingual group and a monolingual group could reasonably be the product of one or more of the many aspects of the perceptual assimilation process that results in cross-language differences. The use of two bilingual learner groups also permitted us to examine the relative benefits of different aspects of bilingual experience. Systematic differences were required in the phonological and phonetic structures of the two sets of bilingual learner L1/L2s relative to the L3 in question; the L3 contrasts also had to be perceptually challenging in order to avoid ceiling effects, while being inherently discriminable enough to actually observe learning following limited training time in the laboratory.

For these purposes, Malayalam was selected as the target L3, contributing three appropriate consonant contrasts for laboratory training: a lateral /l - ɭ/, nasal/n̪-ɳ/, and a fricative /ʃ-ʂ /contrast. Malayalam is a Dravidian language, spoken primarily in southern India. All three contrasts entail a common place of articulation distinction involving the retroflex feature. This feature was selected as a relevant systematic difference between the phonological/phonetic structure of the two bilingual learner groups: Bengali-English speakers and Spanish-English speakers. Bengali is an Indo-Aryan language spoken primarily in eastern India and Bangladesh. Bengali does not possess a retroflex lateral, nasal or fricative, although Bengali listeners have extensive experience with the retroflex place among their stop consonants, with three voicing types (voiceless, voiced, aspiration) and two phonation types (modal, murmured). Spanish, like Bengali, lacks close correspondents to the Malayalam retroflex consonants in this study; unlike Bengali, Spanish listeners have no equivalent experience with retroflex stops in multiple syllabic environments. American-English listeners served as the monolingual control group to these bilingual groups, selected because of the absence of both bilingual expertise and experience with retroflex consonants.

To compare monolingual and bilingual L3 learners, a training study was conducted involving identification, discrimination and perceptual assimilation measures at pre-training and post-training and in a generalization phase in order to fully explore any facilitating (or otherwise) effects of bilingualism. Identification and discrimination tasks are commonly employed in cross-language speech perception and learning studies. The perceptual assimilation metric allowed for a more direct examination of how the relationship between the L1/L2 and the L3 systems changes over training, adopting assimilation patterns provided by PAM and, secondarily, assessing the presence of a multilingual benefit, if any, shown through perceptual learning patterns.

3.1 Participants

Sixty adults were remunerated to participate in this study. All subjects self-reported normal hearing with no history of hearing or speech impairment. The participants were chosen based on their L1 and/or L2: Bengali-English (BE) bilinguals, Spanish-English (SE) bilinguals and American-English (AE) monolinguals. The participants in the BE bilingual group (n = 20) ranged in age from twenty to thirty-five years (M = 27.5 years) and those in the SE bilingual group (n = 20) ranged in age from eighteen to twenty-five years (M = 21 years). In the screening process, early bilinguals with an age of acquisition (AOA) for the L2 of lower than twelve years were preferred. If the AOA for L2 was higher than twelve years, then candidates only with a self-reported L2 proficiency level of *high* were enrolled. Keeping in mind that self-reported language proficiency may not be an accurate measure for screening potential subjects, certain precautionary measures were taken while recruiting the subjects for the two groups. The screening for bilingual BE participants involved assessing their level of spoken proficiency in English (here referring to Indian English, since most of the student population were graduate students staying in the United States for no more than five years) by screening for any noticeable signs of L1 transfer in consonantal segments (in which the influence of Bengali [L1] transfer would be more prominent in production) during conversation. More specifically, the Bengali-English participants were screened for any exposure to Dravidian languages, such as duration of residence in a south Indian region and, if such a factor was present, then to what extent (two subjects stayed for three months and six months in Kannada- and Telugu-dominant speaking regions, respectively). If potential candidates were found to have even minimal awareness of the retroflex laterals, nasals or fricatives from Dravidian languages, they were not recruited for this study. In the case of the SE bilingual group, all participants have lived in North America all or most of their lives, having moved to North America within the first six years of their lives. In both bilingual groups, the few multilingual subjects had self-reported low proficiency in their L3 language. On screening, these L3s were languages that were typologically (in terms of phonetic categories) distant from the target language. Moreover, the proficiency threshold hypothesis (Cummins 1979 as cited in De Angelis 2007) was followed where these L3s were not expected to influence the learning of target speech contrasts as the proficiency reported was low.

For the AE monolingual group, participants (n = 20) ranged in ages from eighteen to twenty-five years (M = 20.3 years). The participants of this control group self-reported exposure to languages other than English and even indicated conversational ability in another language or passive exposure to another language (for instance, a trip for a couple of months to another country). Self-rated proficiency levels in non-native language(s) represent a variable of which the control is limited. Previous studies (Werker 1986; Nayak et al. 1990, as discussed in De Angelis 2007) have equated low proficiency in a non-native language to no knowledge of a non-native language and have perhaps incorrectly classified subjects as monolinguals. De Angelis (2007) discussed this concern of ignoring self-reported prior non-native linguistic knowledge in several studies and concluded that since there is not much empirical evidence available on proficiency threshold levels in non-native language acquisition, only time and further research will provide an answer to suitable proficiency threshold levels. She also demonstrated by example how recruiting a pure monolingual would be difficult given current educational policies in most countries.

> If we come across a study with adult Italian L1 learners of German as an L2, for instance, we can safely assume that these subjects are third or additional language learners and not L2 learners for the simple reason that the study of foreign languages, usually French or English, has been compulsory in Italian schools for several decades. Italian L1 speakers could be true L2 learners of German only in the case in which they had failed to complete compulsory education in Italy, or if they were illiterate. Similarly, if a study examines English L1 learners of French as an L2 at a Californian University, it is reasonable to wonder whether these subjects are true L2 learners, as a large number of students in California, and the United States in general, study Spanish in high school. (De Angelis 2007: 7)

A scenario of this kind undoubtedly existed in this study as the subjects were students at the University of Florida. It was almost impossible to find pure monolingual subjects with no exposure to another language, especially in the age group of eighteen to thirty-five years, due to the curriculum in North American schools. To somewhat counter this factor, subjects, many of whom had to take Spanish as a required language class, were asked to read out two brief phrases in Spanish. This was done to determine if the subject had acquired any perceptual categories of Spanish. Only those participants were recruited as subjects whose speech did not provide any evidence of Spanish categories but in fact displayed complete phoneme transfer from American English. The subjects were asked to say out loud, at a normal speaking rate, two phrases considered to be 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 assessed the level of spoken proficiency as well as lexical knowledge of the Spanish language. If the candidates were not able to maintain a normal speaking rate with more than 40 per cent correct pronunciation (that is, they substituted English sounds like [ɹ] instead of [ɾ], [t] instead of [t̪], [g] instead of [ɣ] and dropped the [e] at the end of the word *roque*, etc.), they were considered under the monolingual language group, provided they had no exposure to any other language.

Hay tres tristes tigres en un trigal.

[ai̯ . t̪ɾes . tɾis . tes . t̪i . ɣɾe . se . nun̪ . tɾi . ɣal]

IPA transcription

El perro de san roque no tiene rabo.

[el . pe . ro . ð̪e . sãn . ɾo . ke . no . t̪ i̯e . ne . ra . βo]

IPA transcription

3.2 Elicitation material

The novel non-native speech contrasts for the study were taken from Malayalam. A member of the Dravidian language family, Malayalam was chosen as the target language for the experiment since the retroflex place of articulation is used widely for various manners of articulation (stops, fricatives, laterals, nasals), providing a flexible choice of speech contrasts: /l - ɭ/ (alveolar vs. retroflex lateral approximant), /n̪ - ɳ/ (dental/alveolar vs. retroflex nasal) and /ʃ-ʂ/ (palato-alveolar vs. retroflex fricative). These speech contrasts differ primarily in place of articulation[3]. One of the speech sounds in each speech contrast from Malayalam is marked with the retroflex feature /ɭ, ɳ, ʂ/, and those sounds are not found in Bengali, American English or Spanish.

The elicitation materials were recorded as non-words read in a carrier sentence by eight native speakers (five males and three females) of Malayalam, their ages ranging from twenty-eight to forty-five years. They were born in the state of Kerala, India, and have resided in India ever since. Other languages known to some of these speakers are Kannada, Tamil, Hindi and Indian English. For the stimulus materials, all speakers produced each of the six speech sounds from the three speech contrasts. 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 in every vowel context were recorded in order to achieve the best possible tokens to use for the stimuli.

The data were recorded and digitized at 44.1 kHz on a Marantz Digital recorder and transferred directly onto a personal computer. The non-words (in Malayalam script) were presented 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 familiarize themselves with the non-words. The data were subsequently segmented using PRAAT speech analysis software (version 5.0.46).

The intelligibility of the stimulus material was evaluated in the form of expert listener ratings, which were given by four native speakers of Malayalam through a consonant identification task in order to eliminate any erroneous or less-than-ideal tokens. Tokens that were consistently identified at 83 per cent to 100 per cent as the target sound were selected as stimuli. Three out of the thus selected tokens, which were closest in duration and mean fundamental frequency to each other, were chosen for each consonant in all three vowel contexts, resulting in a total of seventy-two stimuli (eight consonants × three vowel contexts × three tokens) per talker.

A preliminary experiment was conducted to examine the ease of discriminability of these contrasts by monolingual AE speakers, since the AE monolinguals acted as the control group in the main experiment. It was found in the AX discrimination task for this preliminary experiment that the /l-ɭ/ speech contrast was the least discriminable (d' = 1.13) followed by the nasal contrast /n̪-ɳ/ (d' = 1.6) and fricative contrast /ʃ-ʂ/ (d' = 2.02). Among the vowel contexts, it was most difficult to discriminate within the [i] context (d' = 1.5), closely followed by the [u] context (d' = 1.6), whereas the [a] context was the most easily discriminable (d' = 2.3). No ceiling or floor effects were found for any of the non-native speech contrasts.

Based on the results of the preliminary experiment as well as the native expert listener ratings, the stimulus materials for the perception training experiment were selected. For each talker, three physically different tokens per consonant and vowel context (eight consonants[4] × three vowel contexts × three tokens) were chosen. The production tokens of one talker (M08) were used solely for the familiarization task. Talker M04 was used for the pre-test and post-test as well as the discrimination and perceptual assimilation tests, since all of these tests have to be identical in stimuli and procedure to enable direct comparison. M04 was also included during training so that the testing (pre-test and post-test) would be on a trained talker. M07 was used for the generalization test only; thus M07 served as a novel talker, whose productions were not exposed to the subjects anytime during the pre-test or training phase. Data of six talkers (M01, M02, M04, M06, M09, M10) were used for the training sessions.

- Familiarization phase: one talker (M08) (8 consonants × 3 vowels × 2 tokens × 2 repetitions) 96 trials. ITI was 2 seconds.
- Pre-test phase
  - Pre-test: one talker (M04) (8 consonants × 3 vowels × 2 tokens × 6 repetitions) 288 trials.
  - Discrimination test: one talker (M04) (8 consonants × 3 vowels × 2 tokens × 4 orders) 192 trials. ISI 1.5 seconds.
  - Perceptual assimilation test: one talker (M04) (8 consonants × 1 vowel × 2 tokens × 5 repetitions) 80 trials.
- Training phase: two talkers per session (8 consonants × 3 vowels × 3 tokens × 2 repetitions) 288 trials per session. Over a total of six training sessions, productions from six talkers are used (M01, M02, M04, M06, M09, M10).
- Post-test phase
  - Post-test: one talker (M04) (8 consonants × 3 vowels × 2 tokens × 6 repetitions) 288 trials.
  - Discrimination test: one talker (M04) (8 consonants × 3 vowels × 2 tokens × 4 orders) 192 trials. ISI was 1.5 seconds.
  - Perceptual assimilation test: one talker (M04) (8 consonants × 1 vowel × 2 tokens × 5 repetitions) 80 trials.
- Generalization test: one talker (M07) (8 consonants × 3 vowels × 2 tokens × 6 repetitions) 288 trials.

Since all the subjects were not familiar with Malayalam orthography or vocabulary, it would be difficult to train them without some sound–symbol association. This kind of issue does not often arise in L2 training studies as the subjects are normally familiar with the vocabulary and orthography of the target language. In the current study, the question of orthography associated with the six 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, who have no appropriate symbol for representing the new sounds. A number of alternate methods of representing the sounds were considered before appropriate symbols were finalized. Using the original Malayalam script, IPA symbols, or arbitrary geometric shapes to represent the sounds would have taken the subjects much more time and effort to learn, resulting in decreased focus on the perception of sounds. In addition to that, the possibility of learners perceiving the sounds correctly but clicking on the wrong symbol could be detrimental to the validity of the results. A better alternative was taking letters from the English alphabet as representative symbols. This option was considered best as the subjects in all the language groups were familiar with the English alphabet and their corresponding English consonants. The lower and uppercase of the letter ‘n’ could be used for the /n̪/-/ɳ/ sound distinction so that whenever subjects hear a nasal, there would not be the basic confusion of finding the correct symbol out of the six symbols as would be the case with purely arbitrary representations. To help the subjects learn the symbols, the dental/alveolar sounds in the three sets of contrasts that were presumed to be assimilated to their L1/L2 category were assigned the lowercase letters, while the retroflex sounds in all three sets of contrasts were assigned the corresponding uppercase letters. 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 and *S* = retroflex fricative. During the 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 were to hear would not be English sounds but from a language not known to them.

3.3 Procedure

The experiment used high-variability perception training. A set-up involving a symbol familiarization task, pre-test, training, post-test and a generalization test was employed, adapting the procedure used by Lively, Logan and Pisoni (1993), Lively and Pisoni (1994). The effect of training was measured by comparing the performance in pre-test and post-test 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 a perceptual assimilation task were also employed during the pre-test and post-test phase of the experiment.

A minimum of a seven-hour gap was used between sessions to prevent any adverse effects of fatigue during consecutive training sessions. All training and testing took place in a sound-attenuated room equipped with individual computer stations (containing a keyboard, a CRT monitor and headphones) for subjects. Stimuli were presented to participants over a set of headphones. The software collected individual responses during all phases of the experiment. All subjects were tested and trained individually.

In the familiarization phase, participants were acquainted with the symbols associated with each consonant. A sound file was played while the corresponding symbol was displayed on the computer screen. The participants were asked to pay attention to both the speech stimulus and its corresponding symbol. They were informed that the words were from a language not known to them and that the arbitrary symbols represented different sounds. The non-words contrasting in consonantal place (e.g. [ili] vs [iLi]) were presented within the same vowel context in order to maximize the impact of the familiarization task by focusing it on the differences in the consonants.

During the pre-test phase, several tasks were administered. The first was a consonant identification task presenting the randomized stimuli over the headphones. The stimuli consisted of one talker (M04), whose productions were also used for training. The task was self-paced; that is, there was no ITI assigned and the subject moved on to the next trial when ready.

Secondly, a categorical AX discrimination task was used to measure listener sensitivity to non-native speech contrasts at the naïve exposure level. Categorical trials were employed in which the two *same* trials are repeated in order to maintain an equal number of *same* and *different* trials, which resulted in a set of sixteen trials per speech contrast.

The third task in the pre-test phase was the perceptual assimilation task, which examined the similarity of the speech sounds in stimuli to the listener's native speech sounds. In this task, subjects were asked to listen to each stimulus sound (provided in only one vowel context) and write down the perceptually closest speech sound in the native language(s) orthography. This assimilation task was an open-set test. Along with that, for each trial, subjects were asked to provide a category goodness rating of how similar the native speech sound was to the newly heard stimulus sound. The goodness rating scale consisted of 1 to 7 levels, where level 1 was to be chosen when the target sound was 'very different' from the sound closest to it in their respective native language(s). Level 7 was to be chosen when the target sound was 'exactly the same' as the sound closest to it. 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 both the pre-test and the post-test phases of the experiment in order to evaluate any changes in assimilation patterns as an indication of developing new phonetic categories.

The order of presentation for the various tasks during the pre-test phase was kept the same for all the subjects across language groups. Once the familiarization task was over, the pre-test phase began with the consonant identification pre-training test as the initial task. It was followed by the AX discrimination test. After a five-minute rest, 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 training phase consisted of six sessions of thirty-five minutes each. A consonant identification task with feedback on every trial was administered. If the response was correct, the message 'You are correct!' was displayed. If the response was incorrect, the message 'Incorrect!!! Please listen to it again' was displayed. After a 500 milliseconds interval, the sound file was replayed along with the correct symbol for reinforcement. The task was self-paced. Data gathered from six talkers were used as stimuli for the training part of the experiment. The participants heard only two talkers per training session. The sequence of presenting talkers remained the same throughout training. Therefore, subjects listened to each talker two times during the six-day training phase. In all, there were 288 trials (two talkers per session). The task was paused for sixty seconds after one block of trials (stimuli of one talker).

The post-test phase was identical to the pre-test phase. Finally, there was a generalization test, the design of which was identical to the consonant identification test except that the stimuli were taken from a talker who did not appear in training.

3.4 Data analysis

The results of the pre-test and post-test were analysed across speech contrasts and across language groups. The mean percentage scores for the consonant identification tasks (pre-test, post-test and generalization test) and d-prime scores for the discrimination tasks were obtained. Analysis of Variance (ANOVA)s were conducted separately for each task to test the hypothesis.

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'. 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 pre-test and the post-test discrimination test, difference scores (post-test value minus pre-test value) were obtained and 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: To arrive at the modal response for each sound, a matrix was developed that calculated the percentage of a particular response for that speech sound. Once the modal response for each of the sounds was determined, the general assimilation types were obtained. The resultant general assimilation types for each contrast (pre- as well post-test values) were either within category (WC) or between categories (BC). For the BC 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 uncategorizeable assimilation type (UU); otherwise it was termed as uncategorizeable-categorizeable assimilation type (UC). In the case of WC assimilation types, the ratings for the speech sounds in that contrast were submitted to T-tests 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), and otherwise single category (SC).

4 Overall results

Three different tests were employed to investigate the hypothesis mentioned in the introduction. Data from a total of sixty subjects, twenty from each language group, were analysed. The bilingual groups were analysed separately as Bengali-English and Spanish-English groups, and also merged to be analysed as a single bilingual group (n = 40) compared with the control monolingual group (n = 20). The perceptual performance of subjects in various tests like the identification tests (pre-test–post-test and generalization test), and the AX discrimination test was measured by comparing the performance at the pre-test level to that at the post-test level. The data were submitted to an independent samples T-test where the independent variable was multilingual benefit.

Table 6.1 represents the averaged accuracy scores for identifying all three non-native contrasts at the pre-test and post-test level by the bilingual groups (BE and SE) and the monolingual group (AE). The improved performance seen in subject groups after training was 18 per cent for the bilingual group and 9 per cent for the monolingual group across all three contrasts.

The results show a significant main effect for group (t (178) = −3.90, $p < 0.01$). The bilingual participants' perceptual performance in the identification test was significantly better than the monolingual participants' performance. In the case of the AX discrimination test results (Table 6.1), the dataset of the bilinguals' performance revealed that d' scores of the bilinguals were significantly different than those of the monolinguals (t (166) = −3.02, $p < 0.01$).

The results of the Identification task and AX discrimination task exhibit the effects of a multilingual benefit facilitating the acquisition of non-native contrasts by bilinguals.

In the case of the generalization test (identification task with a novel talker), these scores were compared to scores of the identification test (post-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 (Table 6.2). The generalization scores correspond to perceptual performance of subjects on contrasts

AQ:The terms "data set", "dataset", "data sets" and "datasets" are found in the book. Please note that "datasets" and "dataset" have been made consistent throughout the book as per predominant usage. Please confirm if this is okay.

Table 6.1 Mean percentage identification (ID) as well as AX discrimination scores averaged over contrasts at the pre-test and the post-test level for the bilingual groups (BE and SE) combined versus the monolingual group (AE)

	ID test score		**AX discrimination score**	
	Pre-test (%)	**Post-test (%)**	**Pre-test (d')**	**Post-test (d')**
Monolingual	48	57	0.70	0.97
Bilingual	55	73	0.91	1.52

Table 6.2 Mean percentage identification (ID) scores averaged over contrasts at the post-test and the generalization test level for the bilingual groups (BE and SE) combined versus the monolingual group (AE) in case of generalization test

	Generalization test score (ID task)	
	Post-test (%)	**Generalization (%)**
Monolingual	57	58
Bilingual	73	75

spoken by a novel talker (M07) and the post-test scores correspond to those of the trained-on talker (M04) with the same non-native contrasts. The difference between the post-test scores and generalization scores were submitted to the mixed-model ANOVA test. The bilinguals' performance showed no significant difference from that of monolinguals (t (58) = 0.54, n.s.).

4.1 Results in terms of individual language groups

In the case of the identification tests at the pre-test and post-test levels, the main effect for language groups was observed to be significant (F (2, 178) = 7.818, $p < 0.05$). Figure 6.1 shows the averaged accuracy scores of the identification test over the three contrasts by all language groups at the pre-test and post-test levels. All three listener groups improved from pre-test to post-test, with the most significant improvement in performance observed in the Bengali-English group, followed by the Spanish-English and American-English subjects. The percentage improvement of the bilingual groups was significantly higher: 19 per cent for the BE group and 17 per cent for the SE group. On the other hand, the percentage of improvement in perceptual performance seen in the AE group was 9 per cent.

The main effect of contrast showed no significant difference across all three groups, providing evidence for homogeneity within the three contrasts (F (2, 114) = 1.81, (p = n.s.) (p = 0.16). The interaction of language groups with the three contrasts was also observed to be significant (F (4, 114) = 5.81, $p < 0.05$). A post hoc analysis using Tukey's test was conducted. Both the BE and SE bilingual groups' performance was shown to be significantly different from that of the AE monolingual group (BE–AE [$p < 0.05$], SE–AE [$p < 0.05$]). However, no significant difference was seen in the performance scores between BE and SE, the two bilingual groups ($p > 0.05$,

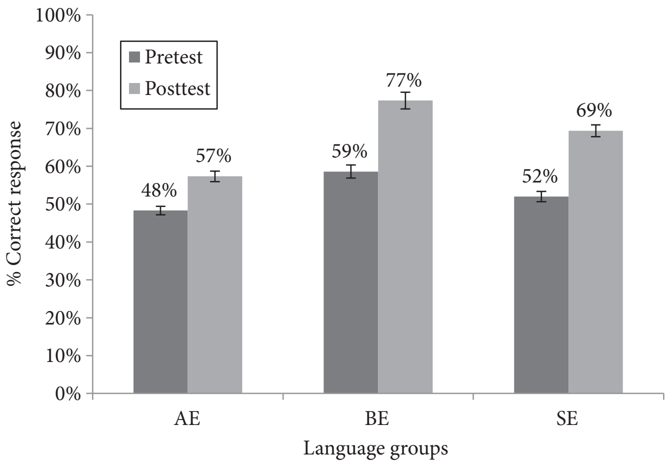

Figure 6.1 Mean percentage correct response for identification of the three place contrasts (laterals, nasals, fricatives) by bilingual (BE and SE) and monolingual (AE) groups at the pre-test and post-test levels.

p = 0.774). The post hoc results confirm the hypothesis formulated for the factor of multilingual benefit.

In the case of the AX discrimination test, the ANOVA results indicated marginal differences between language groups. Figure 6.2 represents the averaged d' discriminability scores over the three non-native contrasts at the pre-test and post-test levels. All three listener groups improved from pre-test to post-test, with the most significant increases in performance observed with the Bengali-English group, followed by the Spanish-English and American-English groups. The main effect for language groups was near significant (F [2, 57] = 2.81, $p = 0.06$); only when the bilingual groups were merged together as a single population and compared with the monolingual group was the comparison of mean discrimination scores significant as mentioned above. The interaction between the language groups and the three place contrasts was also observed to be significant ($F (4, 114) = 5.45, p < 0.05$).

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.

In the case of the generalization test (Figure 6.3), the results for the generalization test scores when compared with the post-test scores showed no significant difference ($F (2, 169) = 1.63$, n.s.). The main effect of the contrast factor was significant ($F (2, 112) = 8.88, p < 0.01$) but the interaction of language groups with the three contrasts did not reveal any significance (Language*Contrast: $F (4, 112) = 0.88$, n.s.). The area of testing for robust phonetic category development did not reveal any significant differences among the language groups in any of the analyses. At first, these results seem to reflect a surprising lack of generalization of perceptual learning through varied talkers during training to perceiving the same contrasts by a novel talker, contrary to findings of previous studies (Lively and Pisoni 1994; Clopper and Pisoni 2004). However, when

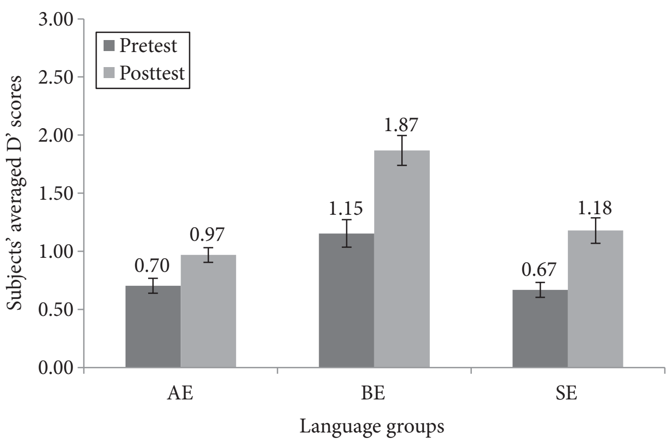

Figure 6.2 Mean d' scores from the AX discrimination tests for the three place contrasts by the language groups at the pre-test and post-test levels.

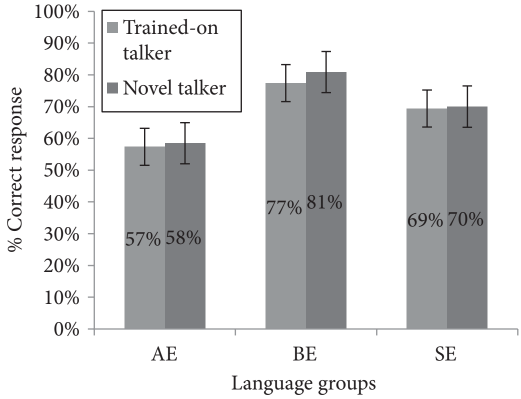

Figure 6.3 Mean percentage correct identification of contrasts spoken by trained-on talker (post-test) and novel talker (generalization test).

the perceptual performances of all groups in this task were analysed at the individual subject level, a large amount of within-group variability was found. Individual subjects' performances may have obscured any differences between language groups.

Another noteworthy point in this finding was the unexpected result that generalization scores were higher than post-test scores across all language groups. Previous studies have provided evidence of development of robust phonetic categories by L2 learners through generalization tests (Lively and Pisoni 1994; Clopper and Pisoni 2004). Based on their results, it was expected that subjects will perform either equally well or badly in identifying the stimuli from a novel voice and the trained-on voice, as was the case in the post-test ID test. 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 better performance being observed with the Bengali-English group, followed by the Spanish-English group and the American-English group.

4.2 Perceptual assimilation results

Examining the perceptual assimilation patterns elicited from the subjects provides an opportunity to examine the formation of new perceptual categories, if any, developed post-training. The perceptual assimilation task conducted at the pre-test and post-test levels first requires a more descriptive evaluation of its results. Based on PAM-L2 predictions (Best and Tyler 2007), we could expect to see a direction towards learning.

For this study, five major assimilation types – two-category (TC), single category (SC), category goodness (CG), uncategorizeable-categorizeable (UC) and both uncategorizeable (UU) – were used to evaluate the results. Learning was defined in terms of a shift from non-categorical assimilation types like SC or UU to assimilation types, such as CG or UC, that reflect a trend towards higher sensitivity of discrimination, even up to TC assimilation, which represents excellent discrimination of the non-native contrasts.

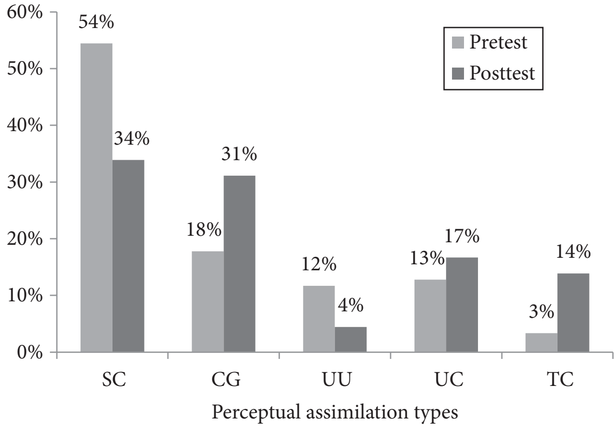

Figure 6.4 Mean percentage of assimilation elicited from all three language groups, pre-test versus post-test.

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.

Figure 6.4 shows the mean percentage of assimilation types drawn from subjects over all three language groups collectively. The results showed a decrease of 20 per cent in the SC assimilation type and 8 per cent in the UU assimilation type from pre-training to post-training levels. The TC and CG assimilation types displayed an increase of 11 per cent and 13 per cent, respectively. The UC assimilation type showed a modest increase of 4 per cent from pre-test to post-test. The overall change of assimilation types from showing less categorization to showing a higher sensitivity towards categorization provides evidence of learning taking place during the limited training period.

More importantly, the perceptual assimilation test was one of the areas where the effects of multilingual benefit on the acquisition of non-native contrasts could be examined. The overall perceptual assimilation results provide evidence of a definite shift towards learning these novel non-native contrasts in a very limited training period. Now the question arises as to whether any language-group differences were seen, revealing the effect of a multilingual benefit factor as explored in this study.

Figure 6.5 displays percentages of assimilations elicited from individual language groups at the pre-test and post-test levels. The bilingual groups, BE and SE, showed similar shifts in assimilation types from pre-test to post-test, unlike the monolingual group. The SC type of assimilation pattern was observed to have a lower percentage at the pre-test level for the BE (35 per cent) group, whereas the same assimilation type had a very high percentage of 53 per cent at the pre-test level for the SE group and 75 per cent for the monolingual AE group.

A sharp decline was seen in the SC type at the post-test level for all three language groups. An inverse trend was observed for the TC and in some cases for the CG assimilation type from pre-test to post-test. Both bilingual groups (BE 18 per cent,

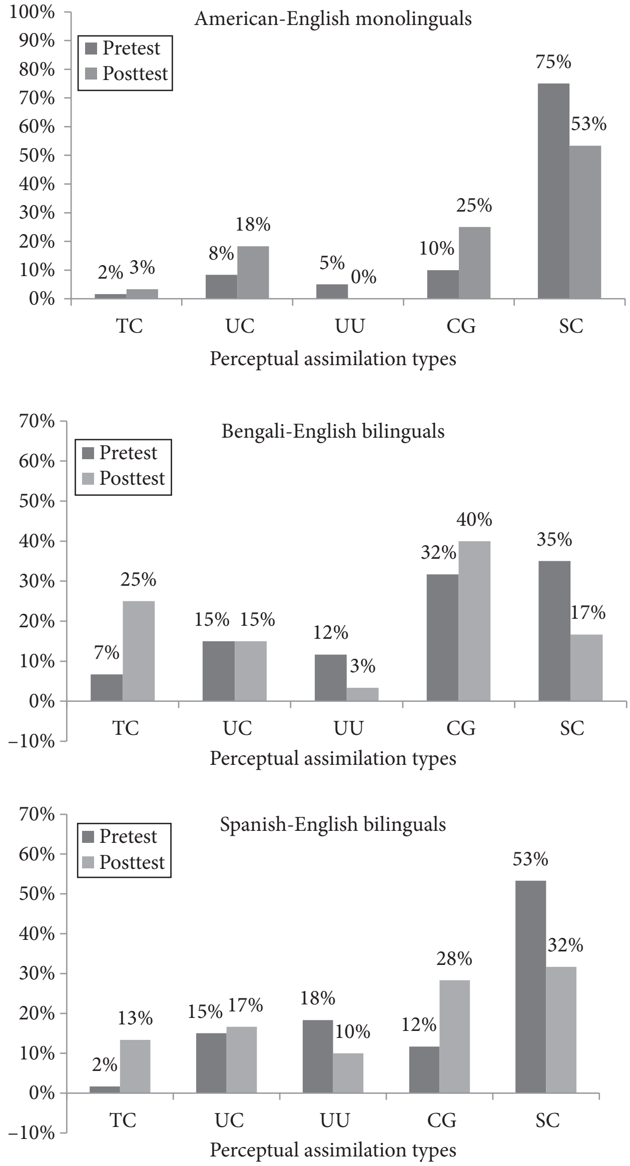

Figure 6.5 Percentages of assimilation types elicited from individual language groups at the pre-test and post-test levels are provided in three different charts.

SE 11 per cent) showed a high rise in the TC assimilation type at the post-test level, contrary to the monolingual group (AE 1 per cent). In the case of the UC assimilation type, both bilingual groups displayed a negligible increase (BE 0 per cent, SE 2 per cent) whereas the monolingual group (AE 10 per cent) showed a high increase in the UC type.

A statistical analysis (one-way ANOVA) on proportion change in assimilation types was also conducted in order to examine significance based on the individual language groups as well as the multilingual factor. The main effect for language did not show significance (F (2, 57) = 2.50, p = 0.09, n.s). However, pooling the dataset of BE and SE groups into one bilingual group and comparing it with the monolingual group showed a significant difference in the means (T-Test) of the two sets (t (58) = −2.22, $p < 0.05$). These results, again, are suggestive of the effect of a multilingual benefit on acquiring the non-native contrasts by the bilinguals as a group, regardless of their L1s.

5 Discussion

The experiment was designed to address a number of important questions concerning third or additional language acquisition and cross-language speech perception: Would a multilingual benefit facilitate the acquisition of novel non-native contrasts among bi/multilinguals? Can perceptual assimilation patterns be modified over a limited training period? Would these changes in assimilation patterns reflect differences between the bilingual and monolingual subjects in terms of acquisition of non-native contrasts?

The results of the current study are suggestive of the existence of a multilingual benefit. The results of the ID task (pre-test–post-test level) provided support for the hypothesis stated for the multilingual benefit factor. The results of the AX discrimination task only approach significance. This difference in results of the two dependent measures (ID and AX discrimination) could perhaps be because of the fact that the ID task was used during training, unlike the AX discrimination task.

Pooling the bilingual population and comparing it with the monolingual population revealed significant differences in performance in ID, AX discrimination and perceptual assimilation tasks. These results suggested that perhaps the effects of factors like a multilingual benefit require a larger sample size than average (n = 20) to emerge with robust results. These effects could be subtle and possibly obscured by individual subject variations. 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 a multilingual benefit. However, no study is without its caveats. There were some unexpected results. The interpretation of results and caveats are discussed herewith.

The hypothesis for a multilingual benefit was supported in identification task performance across all contrasts. Statistical significance was seen across language groups. In the ID task measure, the BE and SE groups individually identified the contrasts significantly better than the AE group. The evidence of a multilingual benefit influence emerged when the BE group showed no greater significant difference from the SE group, even though subjects from the BE group have an L1 typologically closer to Malayalam than Spanish. However, the comparison of mean d' difference scores of individual language groups was only marginally significant.

The study provides suggestive evidence that bilinguals function differently from monolinguals, facilitated by a multilingual benefit factor, in abstracting and reorganizing the information gathered from speech 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 a multilingual benefit (or loosely termed metalinguistic awareness) seen in acquiring a third or additional language by bilinguals (Bild and Swain 1989; Cenoz and Valencia 1994; Enomoto 1994; Klein 1995; Gonzalez-Ardeo 2000; Munoz 2000; Cheung et al. 2010). For instance, Klein (1995) showed that during lexicon acquisition of a language, multilinguals learnt a higher number of lexical items than monolinguals. Klein based the explanation for this phenomenon on enhanced cognitive skills in multilinguals that helped them tease out the potentially relevant data for resetting the particular parameters for the new language. Similarly, in the present study, the results in the ID task measure showed that bilinguals, with or without previous experience with the retroflex feature, displayed better perceptual performance than the monolingual group.

Some studies looking at metalinguistic awareness have not displayed clear and robust effects due to this factor but rather subtle effects limited to specific areas. It appears that the effects of metalinguistic awareness are revealed through linguistic phenomena that are established as very difficult to acquire. Brooks and Kempe (2013) observed the effects of metalinguistic awareness mediating the learning of Russian case-marking indirectly through non-verbal intelligence and auditory sequence learning predictors. The learning of Russian gender, however, was predicted by prior knowledge of languages with grammatical gender. In the study by Cheung et al. (2010) on Chinese-English bilingual children, crosslinguistic transfer occurs with speech perception and metalinguistic awareness in L2 reading measures but not as much in a vocabulary measure. In the current study as well, the results evidencing the effects of a multilingual benefit showed only trends towards its presence within each bilingual group, and only in certain contrasts (e.g. lateral non-native speech contrast in ID and AX discrimination task more than other contrasts). Perhaps this is due to the complex nature of the stimuli (three pairs of novel non-native contrasts with spectral properties to perceive and learn) as well as a very limited time period for implicit perceptual training (six sessions of thirty-five minutes each). 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 factor being explored would have had to be very strong. The fact that a large sample size is required for the weak effects of a multilingual benefit to emerge as robust needs to be emphasized here. The pooling of bilinguals into one group displayed a significant effect of the multilingual benefit across all tasks (ID, AX and perceptual assimilation). Additionally, a large degree of individual variation within the language groups pointed towards the need for a larger sample size. Individual variability could have resulted from uncontrolled factors such as the individual aptitude for learning. In order to see significant group differences with multilingual benefit as the determining factor, a larger sampling size may be required, especially with a limited laboratory training period, as in the current study.

5.1 Generalization test analysis

The results of the test of generalization during the post-test phase of the training were contrary to the prediction. It was expected that the post-test scores of identification would be higher than or equal to the generalization test scores, as seen in previous literature (Lively and Pisoni 1994; Clopper and Pisoni 2004). Moreover, if a multilingual benefit factor was 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 post-test 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 hypothesis. On the contrary, there were no group differences between bilingual and monolingual populations. Moreover, all three language groups displayed generalization scores (novel talker) that were higher than the post-test scores (trained-on talker). On analysing 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 single contrast was consistently identified accurately by all speakers. There was considerable variation in identification scores within groups and across contrasts. Therefore, one particular contrast could not be held responsible for these unexpected results.

An explanation for these surprising results could perhaps be the issue of *talker intelligibility*. The term talker intelligibility could be seen as a continuum where at one extreme, a talker produces '*clear*' hyper-articulated speech that enhances the intelligibility of the speech for the listener of the same language. The term 'clear' refers to speech that is distinct, where the acoustic/articulatory features are indicated clearly so as to be intelligible in adverse listening conditions, perhaps allowing the talker to enhance the distinctions between contrasts in a phonological space (Uchanski 2000; Smiljanic and Bradlow 2005, 2007). On the other end of the continuum, a talker produces 'plain' speech that may be less intelligible in the context of various kinds of noises. The plain speech refers to the normal rate of speech that a talker will produce in ideal listening conditions, that is, not hyper-articulating the sounds in order to make one's speech intelligible. This inherent intelligibility of an individual talker may be the factor affecting the results of the generalization test. Perhaps hyper-articulated sounds produced by the novel talker made it easier for all the subjects, across the language groups, to identify the stimuli more accurately than was the case with trained-on talker.

With the 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 indicative of a direction towards the development of robust phonetic categories corresponding to the non-native contrasts, since all groups fared better in generalization scores. The monolingual groups scored equally well as compared to their performance with the trained-on voice, with results similar to the bilingual groups. Contrary to the 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, Lively and Pisoni 1991; Lively, Logan and Pisoni 1993; Lively and Pisoni 1994; Bradlow et al. 1997). Many perceptual training studies where multiple talkers are introduced during training acknowledge the variations found in an individual talker's intelligibility as a confounding factor for the extreme variations in the results (Lively, Logan and Pisoni 1993; Lively and Pisoni 1994; Clopper and Pisoni 2004; Iverson, Hazan and Bannister 2005; Semiljanic and Bradlow 2007 etc.). However, this area of inherent talker intelligibility requires a deeper probe and should be looked at exclusively in future research work that can contribute substantially to enhancing the efficacy and accuracy of the high-variability perceptual training method.

5.2 Perceptual assimilation analysis

An overall shift in assimilation types from within-category non-discriminable types to more discriminable categorical assimilation types was observed across all language groups and all three contrasts. Noticeably, these results reflect the growth of learning for all three language groups. In addition to influencing the overall direction of learning, it was expected that the changes in assimilation types before and after training, if any, would reveal the effects of a multilingual benefit. The results of the study show support for this prediction. The learning was seen more in bilinguals than monolinguals. The results of the bilingual groups were comparable and showed a considerable decrease in SC (single category) and UU (uncategorizable) assimilation types and an increase towards more categorical/higher sensitivity for discrimination assimilation types such as TC (two-category) and CG (category goodness) assimilation types. The trend towards learning was also seen in the monolingual group but to a lesser extent. This indicates that learning took place even during the limited time period of training, even in the monolingual group.

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. According to PAM definitions of all five assimilation types, the following ranking can be assumed to indicate the direction of learning: SC < CG < UU < UC < TC. This placement of assimilation types implies that any shift in assimilation patterns towards TC during the course of training would evidence learning. Although the ranking of SC and TC assimilation patterns appears to be decisive on the continuum, this is not the case with CG, UU or UC. According to PAM, the CG assimilation type reflects the perception of a particular contrast as being within the same L1 phonological category, with one phone being considered as the good exemplar of the category and the other as the deviant exemplar of the same category. It restricts the placement of CG on the continuum closer to SC as related to TC. However, the assimilation type CG may be seen from 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 at the pre-training level, and over the course of training the learner 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, the learner still labels the modal response (from their L1/L2) for each sound of the contrast as the same; this is because in the current laboratory setting, the learners did not have new orthographic representations for the newly formed phonetic categories, which, if used as modal response, would reflect the TC assimilation type. The problem here lies with the orthography of the language. However, in real-life situations, 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. Therefore, in this light, CG-type assimilation during post-test does 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; with the UC type, the discrimination of the speech contrasts would be excellent. However, with PAM-specified definitions, the ranking of the UC type to the CG type cannot be ascertained since the CG type can allow for a wide range in the quality of the discrimination of the L2 contrast, from poor to good. Moreover, the UC assimilation type is not limited to just one scenario as explained in the PAM model. For example, the learner may assimilate one L3 phone consistently with L1 phonological category 'x' but assimilate the more deviant L3 phone to various L1 phonological categories: 'y', 'z' or even 'x'. In this scenario, the UC assimilation type cannot be interpreted to indicate successful learning since the L1 phonological category with which an L3 phone is fully assimilated overlaps with the other uncategorizable L3 phone. Like the UC type, the case of the UU assimilation type is also quite ambiguous when it comes to the placement on the learning continuum. 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 hypothesize a learning continuum that shows the assimilation types representing various stages of L2/L3 learning by adults. Figure 6.6 shows the hypothetical learning continuum with five different assimilation types representing different types of learning.

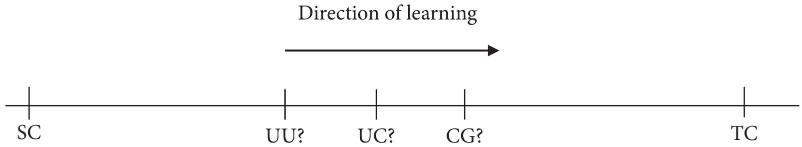

Figure 6.6 Hypothetical representation of the learning continuum displaying the ambiguity of ranking the UU, UC and CG assimilation types in order to assess the direction of learning.

The SC and TC assimilation types form the two extreme ends of the continuum. 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 the CG and UU types, even though CG assimilation is a within-category pattern, its goodness ratings of the deviant L3 phone exemplar reveal that the learner perceives the L3 phones as distinct. 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. The UU assimilation type can result from many scenarios that may not confirm evidence of learning. Therefore, CG can be ranked above the UU assimilation type.

Unlike the UU type, a shift towards the UC assimilation type at the post-training stage signifies that the learner has perhaps progressed from not distinguishing the L3 phones at all to consistently assimilating one L3 phone to one L1 phonetic category. This indicates progression towards learning and thus places the UC type above the UU type on the learning continuum. The overall perceptual assimilation results also validate this ranking as they show a decrease in the UU type and an increase in UC type at the post-training stage. Between the UC and CG assimilation types, no definitive ranking can be stated since both the types show an increase at the post-training stage.

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. It is necessary 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 confirming the learning continuum and extending the models of speech perception in the course of learning.

AQ: Please check the number usage for "Conclusion" and confirm if this is fine.

6 Conclusion

The study is successful in determining whether current theories and established claims of a multilingual benefit in the field of lexical processing were generalizable to the area of cross-language speech perception in regard to learning new non-native contrasts. The results suggested that the effects of a multilingual benefit were prevalent in the learning of relatively difficult novel contrasts within a limited training period. 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. The results revealed a shift of perceptual assimilation types towards perceptual learning. It points to open questions concerning the ranking of assimilation types that may predict the direction of learning. Also, the results were suggestive of the levels of positive perceptual learning seen among bilinguals, which appeared to be higher than the amount of perceptual learning that took place among the monolingual group. Beyond these general findings, the study also demonstrated the need for a larger sampling size in order to get robust effects, since the effects of the multilingual benefit observed in some of the measures were merely suggestive in nature and accompanied by a great deal of individual variation. Overall, the results presented here suggest a productive line of research for future work.

Notes

1 This chapter was partially supported by the Language Learning Dissertation grant awarded to Divya Verma Gogoi by the *Language Learning* journal.

2 For this study, the term *cognitive processing* is operationally defined as the act of abstraction of relevant information from a speech utterance, as a result of a higher level of a concept formation.

3 A fourth phonemic contrast /ɾ - ɻ/ (alveolar tap vs. retroflex approximant) from Malayalam was initially proposed for the study. The training data were collected using all four contrasts. However, in retrospect, it was considered best not to include this contrast in analysis for the following reasons: First, during the preliminary speech perception experiment, the rhotic speech contrast, although not showing a ceiling effect, was found to be the most discriminable of the four contrasts (63 per cent) as opposed to the lateral contrast, which stood at approximately chance level (50 per cent). With this much ease of discriminability, the probability of observing the subtle effects of a multilingual benefit would have been minimal. This possibility was supported by the pre-test ID test scores, which were well above chance level (AE 59 per cent, BE 67 per cent, SE 67 per cent). Secondly, close examination of the phonetic features of this contrast reveal that it lacks the overall homogeneity with the stimuli of this study. The other three speech contrasts differed only in their place of articulation, whereas the rhotic speech contrast differed in manner as well as place. Perhaps this provided the contrast with inherent extra salient acoustic features, which may have facilitated the ease of discriminability.

4 The number of consonants in this equation (and further equations for the training) is indicated as eight since the fourth speech contrast, rhotic, was also included for the training task. However, the gathered data on the additional contrast were not included for this study and will not be further reported on. For details about the exclusion, see endnote 3.

References

Aliaga-García, C. and J. C. Mora (2008), 'Assessing the effects of phonetic training on L2 sound perception and production', in A. S. Rauber, M. A. Watkins and B. O. Baptista (eds), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, 10–27, Florianópolis: Federal University of Santa Catarina.

Baker, C. (2001), *Foundations of Bilingualism and Bilingual Education*, 3rd edn, Clevedon: Multilingual Matters.

Best, C. T. (1995), 'A direct realist perspective on cross-language speech perception', in W. Strange (ed), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech*, 167–200, Baltimore, MD: York Press.

Best, C. T. and M. D. Tyler (2007), 'Nonnative and second language speech perception: Commonalities and complementarities', in O.-S. Bohn and M. J. Munro (eds), *Language Experience in Second Language Speech Learning*, 13–45, Amsterdam: John Benjamins.

Best, C. T., G. W. McRoberts and N. M. Sithole (1988), 'Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants', *Journal of Experimental psychology: Human perception and Performance*, 14 (3): 345–60.

Bialystok, E. (1992), 'Selective attention in bilingual processing', in R. Harris (ed.), *Cognitive Processing in Bilinguals*, 501–514, Amsterdam: North Holland.

Bialystok, E. (2001), 'Metalinguistic aspects of bilingual processing', *Annual review of Applied Linguistics*, 21: 169–81.

Bialystok, E., F. Craik, R. Klein and M. Viswanathan (2004), 'Bilingualism, aging and cognitive control: evidence from the Simon task', *Psychology and Aging*, 19 (2): 290–303.

Bild, E. R. and M. Swain (1989), 'Minority language students in a French immersion programme: Their French proficiency', *Journal of Multilingual and Multicultural Development*, 10: 255–74.

Bono, M. (2011), 'Crosslinguistic interaction and metalinguistic awareness in third language acquisition', in G. DeAngelis and J.-M. Dewaele (eds), *New Trends in Crosslinguistic Influence and Multilingualism Research*, 25–52, Exeter: Short Run Press.

Bradlow, A. R., D. B. Pisoni, R. Aakhane-Yamada and Y. Tohkura (1997), 'Training Japanese listeners to identify English/r/ and /l/: IV some effects of perceptual learning on speech production', *Journal of the Acoustical Society of America*, 101 (4): 2299–310.

Brooks, P. J. and V. Kempe (2013), 'Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness', *Memory and Cognition*, 41: 281–96.

Burnham, D. K. (1986), 'Developmental loss of speech perception: Exposure to and experience with a first language', *Applied Psycholinguistics*, 7: 207–40.

Cenoz, J. (2001), 'The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition', in J. Cenoz, B. Hufeison and U. Jessner (eds), *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, 8–24, Clevedon: Multilingual Matters.

Cenoz, J. (2003), 'The additive effect of bilingualism on third language acquisition: A review', *The International Journal of Bilingualism*, 7 (1): 71–87.

Cenoz, J. and J. F. Valencia (1994), 'Additive trilingualism: evidence from the Basque country', *Applied Psycholinguistics*, 15: 195–207.

Cenoz, J., B. Hufeison and U. Jessner, eds (2001), *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistc Perspectives*, Clevedon: Multilingual Matters.

Cheung, H., K. Chung, S. W. L. Wong, C. McBride–Chang, T. Penney and C. S. Ho (2010), 'Speech perception, metalinguistic awareness, reading, and vocabulary in Chinese-English bilingual children', *Journal of Educational Psychology*, 102 (2): 367–80.

Clopper, C. G. and D. B. Pisoni (2004), 'Effects of talker variability on perceptual learning of dialects', *Lang Speech*, 47 (3): 207–39.

Cummins, J. (1979), 'Linguistic interdependence and the educational development of bilingual children', *Review of Educational Research*, 49 (2): 222–51.

De Angelis, G. (2007), *Third or Additional Language Acquisition*, Clevedon: Multilingual Matters.

De Bot, K. (1992), 'A bilingual production model: Levelt's "speaking" model adapted', *Applied Linguistics*, 13 (1): 1–24.

Enomoto, K. (1994), 'L2 perceptual acquisition: The effect of multilingual linguistic experience on the perception of a "less novel" contrast', *Edinburgh Working Papers in Applied Linguistics*, 5: 15–29.

Flege, J. E. (1995), 'Second language speech learning: Theory, findings and problems', in W. Strange (ed), *Speech Perception and Linguistic Experience: Issues in Cross–Language research*, 233–77, Baltimore: York Press.

Flege, J. (1998), 'The phonetic study of bilingualism', *Ilha do Desterro*, 35: 19–28.

Flege, J., I. MacKay and T. Piske (2002), 'Assessing bilingual dominance', *Applied Psycholinguistics*, 23: 567–98.

Francis, A. L. and H. C. Nusbaum (2002), 'Selective attention and the acquisition of new phonetic categories', *Journal of Experimental Psychology: Human Perception and Performance*, 28 (2): 349–66.

Gonzalez–Ardeo, J. (2000), 'Engineering students and ESP in the Basque Country: SLA versus TLA', in J. Cenoz, B. Hufeisen and U. Jessner (eds), *Looking Beyond Second Language Acquisition: Studies in tri- and Multilingualism*, 75–95, Tübingen: Stauffenburg.

Grosjean, F. (1992), 'Another view of bilingualism', in R. J. Harris (ed.), *Cognitive Processing in Bilinguals*, 51–62, Amsterdam: North Holland.

Hammarberg, B. (2001), 'Roles of L1 and L2 in L3 Production and Acquisition', in J. Cenoz, B. Hufeison and U. Jessner (eds), *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistc Perspectives*, 21–41, Clevedon: Multilingual Matters.

Herdina, P. and U. Jessner (2002), *A Dynamic Model of Multilingualism*, Clevedon: Multilingual Matters.

Hoffmann, C. (2001), 'Towards a description of trilingual competence', *International Journal of Bilingualism*, 5 (1): 1–17.

Iverson, P., V. Hazan and K. Bannister (2005), 'Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/–/l/ to Japanese adults', *Journal of the Acoustical Society of America*, 118 (5): 3267–78.

Jessner, U. (2006), *Linguistic Awareness in Multilinguals: English as a Third Language*, Edinburgh: Edinburgh University Press.

Klein, E. C. (1995), 'Second versus third language acquisition: is there a difference?' *Language Learning*, 45 (3): 419–68.

Kirsner K., M. C. Smith, R. S. Lockhart, M. L. King and M. Jain (1984), 'The bilingual lexicon: Language–specific units in an integrated network', *Journal of Verbal Learning and Verbal Behavior*, 23: 519–39.

Krashen, S. (1981), *Second Language Acquisition and Second Language Learning*, Oxford: Pergamon Press.

Kroll, J. F. and N. Tokowicz (2005), 'Models of bilingual representation and processing: Looking back and to the future', in J. F. Kroll and A. M. B. Degroot (eds), *Handbook of Bilingualism: Psycholinguistic Approaches*, 531–53, New York: Oxford University Press.

Kruschke, J. (1992), 'ALCOVE: An exemplar-based connectionist model of category learning', *Psychology Review*, 90: 22–44.

Lively, S. E. and D. B. Pisoni (1994), 'Training Japanese listeners to identify English /r/ and /l/. III: Long–term retention of new phonetic categories', *Journal of the Acoustical Society of America*, 96 (4): 2076–87.

Lively, S. E., J. S. Logan and D. B. Pisoni (1993), 'Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories', *Journal of the Acoustical Society of America*, 94 (3): 1242–55.

Llama, R., W. Cardoso and L. Collins (2010), 'The influence of language distance and language status on the acquisition of L3 phonology', *International Journal of Multilingualism*, 7 (1): 39–57.

Logan, J. S., S. E. Lively and D. B. Pisoni (1991), 'Training Japanese listeners to identify English /r/ and /l/: A first report', *Journal of the Acoustical Society of America*, 89 (2): 874–886.

Malakoff, M. (1992), 'Translation ability: A natural bilingual and metalinguistic skill', in R. Harris (ed.), *Cognitive Processing in Bilinguals*, 515–30, Amsterdam: North Holland.

Masny, D. (1997), 'Linguistic awareness and writing: Exploring the relationship with language awareness', *Language Awareness*, 6 (2–3): 105–18.

Matthews, P. H. (1997), *The Concise Dictionary of Linguistics*, Oxford: Oxford University Press.

Munoz, C. (2000), 'Bilingualism and trilingualism in school students in Catalonia', in J. Cenoz and U. Jessner (eds), *English in Europe: The Acquisition of a Third Language*, 157–78, Clevedon: Multilingual Matters.

Myers–Scotton, C. (2002), *Contact Linguistics*, Oxford: Oxford University Press.

Nayak, N., N. Hansen, N. Krueger and B. McLaughlin (1990), 'Language–learning strategies in monolingual and multilingual adults', *Language Learning*, 40 (2): 221–44.

Nosofsky, R. M. (1986), 'Attention, similarity, and the identification categorization relationship', *Journal of Experimental Psychology: General*, 115: 39–57.

Nosofsky, R. M. (1987), 'Attention and learning processes in the identification and categorization of integral stimuli', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15: 700–8.

Peal, E. and W. Lambert (1962), 'The relation of bilingualism to intelligence', *Psychological Monographs*, 76: 1–23.

Polka, L. (1992), 'Characterizing the influence of native language experience on adult speech perception', *Perception and Psychophysics*, 52 (1): 37–52.

Pruitt, J. S., J. J. Jenkins and W. Strange (2006), 'Training the perception of Hindi dental and retroflex stops by native speakers of English and Japanese', *Journal of the Acoustical Society of America*, 119 (3): 1684–96.

Sanz, C. (2000), 'Bilingual education enhances third language acquisition: evidence from Catalonia', *Applied Psycholinguistics*, 21: 23–44.

Smiljanic, R. and A. Bradlow (2007), *Clear Speech Intelligibility: Listener and Talker Effects*, 661–4, *ICPhS XVI*, Saarbrucken 6–10 August 2007.

Smiljanic, R. and A. Bradlow (2005), 'Production and perception of clear speech in Croatian and English', *Journal of the Acoustical Society of America*, 118 (3): 1677–88.

So, C. K. and C. T. Best (2010), 'Cross–linguistic perception of non–native tonal contrasts: Effects of native phonological and phonetic influences', *Lang Speech*, 53 (2): 273–93.

Strange, W. and S. Dittmann (1984), 'Effects of discrimination training on the perception of /r–l/ by Japanese adults learning English', *Percept Psychophys*, 36 (2): 131–45.

Sundara, M., L. Polka and F. Genesee (2006), 'Language–experience facilitate discrimination of /d–ð/ in monolingual and bilingual acquisition of English', *Cognition*, 100: 369–88.

Tees, R. C. and J. F. Werker (1984), 'Perceptual flexibility: maintenance or recovery of the ability to discriminate non-native speech sounds', *Canadian Journal of Psychology*, 38: 579–90.

Uchanski, R. M. (2000), 'Clear speech', in D. B. Pisoni and R. E. Remez (eds), *The Handbook of Speech Perception*, 207–35, Malden, MA: Blackwell Publishing.

Werker, J. F. (1986), 'The effect of multilingualism on phonetic perceptual flexibility', *Applied Psycholinguistics*, 7: 141–56.

Werker, J. F. and R. C. Tees (1984), 'Phonemic and phonetic factors in adult cross-language speech perception', *Journal of the Acoustical Society of America*, 75 (1): 1866–78.

Werker, J. F., J. H. V. Gilbert, K. Humphrey and R. C. Tees (1981), 'Developmental aspects of cross-language speech perception', *Child Development*, 52: 349–53.

Wrembel, M. (2010), 'L2-accented speech in L3 production', *International Journal of Multilingualism*, 7 (1): 75–90.

Wrembel, M. (2011), 'Cross-linguistic influence in third language acquisition of VOT', in W.-S. Lee and E. Zee (eds), *Proceedings of the 17th International Congress of Phonetic Sciences. 17-21 August 2011, Hong Kong*, 2157-60, Hong Kong: City University of Hong Kong.

Wunder, E. M. (2011), 'Crosslinguistic influence in multilingual language acquisition: Phonology in third or additional language acquisition', in G. De Angelis and J.-M. Dewaele (eds), *New Trends in Cross-linguistic Influence and Multilingualism Research*, 105–128, Exeter: Short Run Press.