Spring 1-1-2013

Listener Adaptation to Non-Native Speech and the Limits of its Generalization

Stephanie Anne Landblom
University of Colorado at Boulder, stephanie.landblom@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/ling_gradetds

Part of the Linguistics Commons

Recommended Citation
Landblom, Stephanie Anne, "Listener Adaptation to Non-Native Speech and the Limits of its Generalization" (2013). Linguistics Graduate Theses & Dissertations. 21.
https://scholar.colorado.edu/ling_gradetds/21

This Thesis is brought to you for free and open access by Linguistics at CU Scholar. It has been accepted for inclusion in Linguistics Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
LISTENER ADAPTATION TO NON-NATIVE SPEECH
AND THE LIMITS OF ITS GENERALIZATION

By

STEPHANIE ANNE LANDBLOM

B.A., University of Colorado, 2008

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
Of the requirement for the degree of
Master of the Arts

Department of Linguistics

2013
This thesis entitled:
Listener Adaptation to Non-Native Speech and the Limits of its Generalization
Written by Stephanie Anne Landblom
Has been approved for the Department of Linguistics

____________________________
Dr. Rebecca Scarborough

____________________________
Dr. David Rood

Date________

The final copy of this thesis has been examined by the signatories, and we Find that both the content and the form meet acceptable presentation standards Of scholarly work in the above mentioned discipline.

IRB protocol #: 12-0705
Previous research has shown that listeners undergo a process of adaptation to individual speakers by learning indexical information phonetically encoded in the speech signal and using that information to help in subsequent speech processing (Nygaard & Pisoni, 1998). Such adaptation processes can generalize across speakers and can apply to non-native speech as well (e.g. Weil, 2001). Listeners with enough exposure to several speakers from the same language background are able to process speech of a new speaker (of the same language background) more accurately (Bradlow & Bent, 2008).

This study looks to expand this research by testing whether generalization can extend to speakers of different language backgrounds who share similarities with the speaker a listener adapted to. To test this, listeners were exposed to English sentences and narratives recorded by several Gujarati speakers. They were then post-tested on a series of English sentences from either a Kannada or a Russian speaker. Though different, both Kannada- and Gujarati-accented English fall into the category of Indian English for the average American listener. This was designed to test how categorical dialect representation could help with the adaptation/generalization process. The Russian speaker shared a salient feature with the Gujarati speakers, namely the substitution of /v/ for /w/. This tested whether listeners could adapt to salient features across unrelated accents.

Both reaction time and accuracy scores were measured in the post-test and compared between the Gujarati listeners and a control group who heard only native-accented American English speakers. There were no significant differences in response times, but the Gujarati-trained listeners were significantly more accurate in both the Russian and the Kannada post-test, with the Gujarati trained listeners being more accurate in the Kannada post-test than in the Russian post-test. The findings in this
study suggest that listeners start adapting to feature alternations and can generalize this to other speakers of similar language backgrounds or to speakers that have certain features in common in their L2.
ACKNOWLEDGEMENTS

I would like to thank everyone who helped me out during this process. Without each of them, their help and support, this would have hardly been possible. Firstly, I would like to thank Dr. Rebecca Scarborough for being chair of my committee. Her suggestions and guidance as I figured out my exact questions and her expertise in experimental design were all extremely helpful in the process of figuring out how to design this study.

I would also like to thank Dr. David Rood and Dr. Maria Thomas-Ruzic for sitting on my committee and giving me helpful advice on how to structure my paper and background ideas in which I could ground my research.

Thanks go to Will Styler who generously donated his time and technological savvy to helping me finalize the experiment so I could finally start testing subjects.

I would also very much like to thank all the people who helped me find both speakers and participants. Without their help, it would have been impossible to even begin. I would also like to thank my speakers for taking time out of their schedule to come into the lab for their recordings. I also owe gratitude to any of my friends who were willing to give me insights about any particular varieties or features that I wanted to look at in my study.

I would like to thank all my participants who were willing to give me a half hour of their time to complete my experiment. Last but not least, I would like to very much thank my family who was there to give me emotional support when it was needed.

This research was funded by the Research Fund of the Linguistics Department at the University of Colorado.
CONTENTS

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW.........................................................1
  1.1 Introduction.........................................................................................................1
  1.2 Literature Review...............................................................................................3
  1.3 Goals of the Present Study..................................................................................9
  1.4 The Background Languages.............................................................................11
  1.5 Predictions and Hypotheses.............................................................................14
  1.6 Overview of the Study.......................................................................................15

2 EXPERIMENTATION..................................................................................................16
  2.1 Methods................................................................................................................16
    2.1.1 Listeners.......................................................................................................16
    2.1.2 Talkers.........................................................................................................16
    2.1.3 Materials and Set Up...................................................................................17
    2.1.4 Procedure.....................................................................................................20
  2.2 Results..................................................................................................................21
  2.3 Discussion..............................................................................................................28
  2.4 Limitation of This Study.....................................................................................34
  2.5 Implications..........................................................................................................35

REFERENCES..............................................................................................................38

APPENDIX

  I Sentence Verification Stimuli...............................................................................40
  II Transcription Sentences.....................................................................................42
  III Narratives............................................................................................................43
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gujarati and Kannada Consonants</td>
</tr>
<tr>
<td>2</td>
<td>Vowels of Gujarati and Kannada</td>
</tr>
<tr>
<td>3</td>
<td>Russian Consonants</td>
</tr>
<tr>
<td>4</td>
<td>Russian Vowels</td>
</tr>
<tr>
<td>5</td>
<td>Combinations of Training and Post-test Groups</td>
</tr>
<tr>
<td>6</td>
<td>Individual Sentences with Faster Response Times in the Kannada Post-test</td>
</tr>
<tr>
<td>7</td>
<td>Individual Sentences with Faster Response Times in the Russian Post-test</td>
</tr>
</tbody>
</table>
FIGURES

Figure

1. Post-test Accuracy Scores for Gujarati and AmE Listeners........................22
2. Post-test Reaction Times for Kannada and Russian Post-test..................23
3. Post-test Mean Response Times for Gujarati and AmE Listeners..............24
4. Accuracy Scores in the Training Phase for Each Training Group.............28
5. Mean RTs in the Training Phase for Each Training Group......................28
CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

Throughout the course of a conversational exchange a listener is confronted with an almost unending amount of variation in the speech signal. This variation is caused by many different factors, among those being differences in the physiological make up of individual speakers’ vocal tracts which leads to differences in pitch, in the quality of voice, and in the realization of individual sounds. The speaker’s current emotional and physical states are additional factors which can affect other aspects of speech such as the rate of speaking or the amplitude. In addition, speakers often exhibit their own idiosyncrasies as well and incorporate other deviations in speech from what the listener may expect or produce himself. These deviations can be caused by differences in regional dialect, sociolect, etc. In addition to all the variation produced by the speaker, there are almost always different background noises and varying acoustic environments of the space that can also affect the listener’s perception. Taken into account, all these factors make the number of possible variations in speech seem potentially infinite. Despite all this variation in the realization of spoken speech, individuals most often seem to have little difficulty in overcoming this variability and recovering the underlying message. This may happen in what could be a period of adaptation to individual speakers and tends to happen very quickly. An exception to this, however, can occur when a speaker encounters non-native accented speech. Non-native speech has the potential to deviate more drastically from what a listener is used to finding in most native speech variations. These variations are generally said to arise from the effects of a speaker’s first language (L1) on their second language (L2) and these variations have the potential to make the adaptation process more noticeable since it takes more effort on the listener’s part to process the speaker’s intended message.
This study aims to expand on previous research that has looked at listener adaptation, especially to non-native speech. Specifically, is based on studies that have demonstrated that processing times are slower when confronted with non-native accented speech, even if the intelligibility level is high (Munro & Derwing, 1993; Schmid & Yeni-Komshian, 1999). Other studies have also shown that, even though it may take some extra time and effort, a listener can adapt to the speech of a non-native speaker, which means the processing times will decrease and intelligibility level will increase, even if the accent stays the same (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Weil, 2001). This process can happen relatively quickly and processing times have been shown to decrease in as little as under one minute (Clarke & Garrett, 2004).

Two types of adaptation processes have been investigated; namely speaker specific adaptation and multiple speaker adaptation. It has been shown that when exposed to a single speaker of a specific linguistic background, a listener will adapt to that individual speaker (Bradlow & Bent, 2008; Clopper & Pisoni, 2004; Nygaard & Pisoni, 1998). Adaptation to a single speaker does not help in intelligibility or dialect categorization for other speakers from that same speech community (Bradlow & Bent, 2008; Clopper & Pisoni, 2004) (although it has been suggested that gaining enough exemplars of different phonetic contexts from a single speaker could potentially actually have benefits (Weil, 2001)). It has been shown that when speakers are exposed to multiple speakers (generally from three to five) of a single language background then the knowledge they gain from those speakers can generalize to a new speaker (Bradlow & Bent, 2008; Clopper & Pisoni, 2004). That is, if listeners are confronted with a new speaker of the same language background then their intelligibility scores will be higher than a listener who has only been exposed to one speaker of that same language background. Finally, it has also been demonstrated that this generalization does not extend to speakers of other language backgrounds, but is restricted to speakers of a specific L1 (Bradlow & Bent, 2008; Weil, 2001). This study takes these
findings into consideration is designed to expand on them to investigate how flexible the listener’s generalization process may be.

1.2 LITERATURE REVIEW

Previous studies have been able to show that upon hearing non-native speech processing times can increase. In their study, Munro and Derwing (1993) investigated the ‘costs’ of having a foreign accent. Specifically, they were measuring whether a foreign-accented utterance was intelligible, how much difficulty listeners had understanding it, and how long it took them to process. To do this, they presented listeners with forty sentences in a sentence verification task. One group of listeners heard sentences spoken in English by native Mandarin speakers, and another group heard the same sentences spoken by native speakers of American English. They asked the listeners to verify each sentence and then to transcribe it by hand. After hearing the set of sentences once, the listeners heard them two more times and were asked to give comprehensibility ratings (that is, whether the listener perceived the utterance to be difficult to understand) as well as to degree of accentedness ratings. This allowed them to look at the perceptual dimensions of intelligibility, comprehensibility and degree of accent in their study. What they found, which is relevant to this study, was that response times were significantly slower by about 50 milliseconds on average for the listeners who verified the sentences spoken by the non-native English speakers. This was one of the first studies to confirm that non-native speech led to delays in processing times.

Schmid and Yeni-Komshian (1999) also investigated the effects of non-native speech on processing times. In their experiment, listeners were asked to judge whether sentences contained a mispronunciation. Listeners heard a series of English sentences spoken by a native Spanish, Tamil and American English speaker. They were instructed to hit the spacebar on a keyboard as soon as they identified a mispronunciation. Their results showed that listeners were quicker and more
accurate in identifying mispronunciations by the native speaker than by the non-native speakers.

Listeners were also more accurate at identifying mispronunciations in mildly to moderately accented non-native speakers than in speakers with a heavier accent.

Despite the challenges and processing delays caused by this type of variation, listeners have been shown to be able to latch on to indexical properties of the speaker. Abercrombie (1967) describes indexical properties of a language as those features which index certain identifying characteristics of a speaker. These characteristics fall into three categories. The first category indicates membership in a group, which includes traits such as geographical origin, gender and social groups. The second category encompasses the idiosyncratic properties of individual speakers. The third category includes the short term, fluctuating states of a speaker, encoding information such as the present emotional or psychological states. While these features are said to be distinct from linguistic features, they are all conveyed simultaneously through the speech signal. Listeners can use the indexical information they learn to help in the perception and processing of subsequent speech produced by that speaker.

Nygaard and Pisoni (1998) devoted their work to demonstrating this process. They performed three different experiments in which they trained listeners to learn the voices of ten different speakers. They found that listeners trained on individual spoken words performed better on post-test word recognition tasks. Listeners trained on sentence length utterances performed better at recognizing words in full utterances. Listeners trained on sentence length utterances, however, did not do better at recognizing individual words in the post-test. From their findings they argue that “the perceptual learning of voices facilitates the analysis of the linguistic content of the signal. Listeners who learned to attend to talker-specific attributes of the speech signal were able to use that information to aid in the recovery of the linguistic content in the acoustic speech signal” (Nygaard and Pisoni, 1998, p. 372). This was one of the first studies to suggest that both indexical and linguistic properties of speech could be useful in speech recognition.
Clarke and Garrett (2004) devoted their study to investigating how quickly the delay in processing times could be overcome in the case of non-native speech. They carried out three separate experiments in which participants were asked to determine whether the visual probe they were presented with matched the last spoken word in a sentence. Listeners were exposed to sixteen English sentences and one group heard an American English speaker, while the other group heard a native Spanish speaker, except in the third experiment where the listeners heard a native Mandarin speaker. Their results showed that listeners could adapt to the non-native speech of a single individual in under a minute. Processing times were initially longer for the group of listeners exposed to non-native speech, but this delay in processing decreased significantly after brief exposure to the point of being almost identical to the response times of the control group.

In their study on native speech adaptation, Kraljic and Samuel (2006) showed that perceptual learning on one speaker, on one set of contrasting phonemes, could generalize both to a new speaker and to another phoneme pair and that this could be done in a short time with a relatively small set of stimuli. They designed a lexical decision task in which some participants were exposed to stimuli that contained phonetic tokens that were ambiguous between /t/ and /d/. Some listeners heard this ambiguous phoneme in words which contained /d/ and others heard it in words which contained /t/. In a post-test they were asked to judge VCV sequences in which they could judge whether the consonant was a /t/ or a /d/. Listeners who had heard the ambiguous phoneme in /d/ words were more likely to judge this to be a /d/ and listeners who heard it in /t/ words were more prone to judge it to be a /t/. They also found that listeners generalized that to a different set of phonemes, that being /p/ versus /b/. If listeners were used to interpreting an ambiguous sound as a /d/ then they were also more likely to judge an ambiguous bilabial stop as a /b/. They judged it to be a /p/ if they were trained to hear a /t/. Not only did this categorization map onto a new set of phonemes, but it also generalized to a new speaker that was also featured in the post-test. This
gives evidence that generalization can occur in a relatively short time, as the listeners were trained on sets of words. It also shows that phonological features can be generalized. In this case the feature of voicing was being generalized in one set of phonemes and was then mapping onto another set.

Weil (2001) was one of the first to explore how adaptation to one non-native speaker could generalize to another speaker of the same linguistic background of the first speaker. He builds his hypothesis on previous research done on adaptation to individual speakers. He cites Goldinger’s (1996) findings which show that listeners perform better when they hear a familiar voice, but also perform better on a novel voice when that voice has some similarities to the trained talker than when it has fewer similarities. From this, he proposes that “generalization effects might be more pronounced with FAS (foreign accented speech) because there would be phonetic similarities may be more salient to the listener. If accented talkers share many indexical qualities (i.e., awkward prosody, ambiguous consonant production, etc.) encoding these characteristics for one talker should generalize to others” (Weil, 2001, p. 7). Since speakers of the same language background should share many of the same characteristics in a second language, learning these properties should allow a speaker to generalize from one speaker to the next. He predicts that listeners trained on one Marathi speaker will show effects of generalization when they hear a different Marathi speaker. He also predicts that training with a Marathi speaker will lead to no training effects (outside of any practice effects listeners gain by participating in the experimental tasks) when the listeners encounter a Russian speaker, as the generalization would be accent specific.

Weil (2001) relates adaptation of non-native accented speech to exemplar theory. He draws on Goldinger (1998) to state that listeners store individual exemplars of each speech token in their memory. When a speech token is first heard, it is compared to all the tokens stored in the listener’s
mind to match it to the closest token. By gaining more tokens, or exemplars, a listener can more easily identify a new incoming token when it is similar to one (or many) that are already stored. By learning the similar speech patterns of a specific group of non-native speakers, a listener can build an inventory of tokens and then more easily match a new token to the existing repository. After gaining a sufficient number of tokens, a listener will process the speech more easily, accurately, and with faster response times than a listener who has had less experience with the particular speech group.

To test his predictions, participants were brought in on five different days, the first and last of these days consisting of a pre- and post-test respectively. The three days in between were spent completing various tasks which gave the listeners more exposure to the Marathi speaker in the training phase. His findings showed that training with a Marathi speaker did lead to better perception of that same Marathi speaker. It also actually led to some small improvements in perception of the Russian speaker, which he suggests are largely practice effects. There were also improvements when listeners heard a second Marathi speaker in the post-test but it was dependent on the task in that it was stronger for tasks that tested perception at the sentence level as opposed to at the level of an individual word.

Bradlow and Bent (2008) further investigated this talker-independent adaptation to non-native speech, specifically adaptation to non-native speakers with high intelligibility levels. Participants were divided into four different training conditions. The group in the first condition heard five different Chinese accented speakers of English in the training tasks, and a novel Chinese or Slovakian speaker in the post-test. In the second condition, the listeners were trained on a single talker, and were later tested on that same talker. Condition three was also a single-talker condition, but the listeners were sub-grouped to hear different talkers with different levels of baseline intelligibility. There was also a control group who heard American English speakers and a fifth
condition that received no training at all. Training consisted of the listeners transcribing sets of sentences over the span of two days. The researchers made four findings in their study. First they determined that there were practice effects. Participants improved to an extent just by gaining familiarity with the task. They also found that listeners adapted to a single speaker, but when exposed to a novel speaker in the post-test there was no benefit gained outside of the practice effects. Finally, they discovered that when the group trained in the multiple talker condition heard a novel speaker in the post-test, they performed identically to the group that heard the same single speaker in both the training and the post-test tasks. This means that the listeners were able to adapt, independently of any single speaker, to the Chinese accent. Gaining enough exposure to multiple speakers of this particular accent allowed listeners to better comprehend a novel speaker of that same accent.

Clopper and Pisoni (2004) showed that perceptual learning of indexical information (specifically dialects) can generalize. In their study they trained American English listeners to recognize six different U.S. dialects. Listeners completed three training sessions with either one speaker from each dialect region or with three speakers. In each training session they received feedback as to whether their categorization was correct. This was followed by a testing phase, in which listeners received no feedback. This, in turn, was followed by a generalization phase with new speakers. Listeners who had been trained on a single speaker for each dialect were able to more accurately categorize the dialect of each individual speaker in the testing phase. Listeners trained with three speakers were able to more accurately categorize the novel speakers in the generalization phase than the listeners from the single speaker condition. This shows that there can be generalization effects with three speakers of differing dialectical degree after having been trained on limited stimuli.
1.3 GOALS OF THE PRESENT STUDY

The goal of this thesis is to expand on this pre-existing research by testing how flexible generalization may be or what its limitations are. It looks at whether salient segmental features could play a role in the adaptation and generalization process. In order to test this, listeners were trained on a group of speakers that shared Gujarati as a first language. Emphasis was put on the labial velar approximant /w/ in English, which was often substituted by speakers and replaced by the bilabial approximant /ʋ/ which can be perceived to have an English /v/-like quality. Many of the stimuli that listeners were exposed to contained this /w/-/ʋ/ alternation. This was followed by a Russian speaker in the post-test who shared this feature with the Gujarati speakers. The purpose behind this was to test whether listeners were adapting to segmental features, and whether this could extend and generalize from one non-native accent to another unrelated accent that shares a similar feature.

Another group of listeners were exposed to a Kannada speaker in the post-test. This specific Kannada speaker did not exhibit the /w/-/ʋ/ alternation in the sentences he produced. He did, however, share other features similar to those of the Gujarati speakers, such as the use of retroflex stops and relatively short voice onset times (VOT) for word-initial voiceless stops. This helped examine the categorical level at which listeners were adapting and avoided generalization based on that feature if that proved to be significant. It was designed to answer the question as to whether listeners were updating their category of ‘Indian English’ or whether they were sensitive enough to the differences to build a new ‘Gujarati-accented English’ category.

I chose specifically to look at two varieties of Indian English since it would allow me to look at whether there could actually be generalization to a speaker of a different L2 if they spoke two varieties that tend to be categorized by the average American listeners as the same variety. Most American speakers recognize one kind of uniform English spoken by Indian speakers. This tends to
be refuted by many Indian speakers, who place more emphasis on the existence of multiple dialects which are based on regional origin. Some of these varieties have strong enough idiosyncrasies that they can be characterized and stereotyped based on specific, salient features. Both of these accounts seem to have some truth to them. Since India is such a multilingual society, there are many languages which have transfer effects on a speaker’s English. However, in his studies, Masica (1976) notes that some Indian speakers still had an accent that could be classified as ‘Indian’ but that did not reveal their state or general region of origin. This non-regional variety was not (as could be expected) based on a variety of British English (or any other native variety of English). Instead, there were specific traits that marked this variety of being distinctly Indian. This has since been promoted as General(ized) Indian English (GIE) and is often considered the target variety of English that is taught in schools. It is a variety that does not contain too many regionalisms, and therefore is one that should be widely understood throughout the country.

Although the presence of features will vary drastically from speaker to speaker, some that are often mentioned as being unique to Indian English are the use of retroflex as opposed to alveolar stops, allophonic variation between [w] and [u], and lack of voiceless aspirated stops, which tend to be realized without aspiration. Vowel qualities and quantities can differ as well as include segments such as a long, monophthong /o:/ and /e:/.

Although there have been few comparative studies on the different varieties of Indian English, there has been some work done on individual varieties. Most researchers claim that many of the differences lie in different prosodic patterns, a feature which is carried over from the speaker’s L1. However, other differences have been found as well, such as differences in vowel quality and in the realization of other consonantal segments. Wiltshire and Harnsberger (2006) begin to look at some of the L1 transfer effects in Indian English to try to determine which characteristics of different
Indian English varieties are based on GIE and which ones are transferred from a speaker’s L1. In their study, they compare Tamil English with Gujarati English. They found evidence of both a presence of an over-arching Indian English, as well as transfer effects specific to each language background. The transfer effects they found included special intonation patterns and back vowels specific to Gujarati speakers. The retroflex approximant was found to have been adopted into English from the Tamil phonological inventory. The features of Generalized Indian English that they found in both groups of speakers included the presence of retroflex stops, free variation between [w] and [u], as well as a distinction between voiced and voiceless stops with voiced stops being voiced throughout. This last point is significant since Tamil does not make this distinction.

1.4. THE BACKGROUND LANGUAGES

This study makes the underlying assumption that the phonology of a speaker’s L2 is systematically influenced by the phonology of his L1. This means that speakers with the same linguistic background will have very similar properties in their L2. There are three different background languages being used in this particular study, those being Gujarati, Kannada and Russian. The phonologies of each language will be discussed in this section to determine what the differences are between the three languages and where some cause of variation in the L2 (English) could lie.

As discussed above, Gujarati- and Kannada-accented English should have some features in common, both being varieties of Indian English. However, the fact that they are two different languages should lead to some variation as well. Gujarati and Kannada are both languages spoken in India and are the state languages of Gujarat and Karnataka respectively. They belong to separate language families with Gujarati belonging to the Indo-Aryan family and Kannada being a Dravidian language. While their roots are quite separate from each other, the fact that the Indo-Aryan languages have been in contact for so long with the Dravidian languages means that the languages
have adopted some similarities. These similarities, as well as the differences are shown in the

Tables 1 and 2 below:

**Table 1: Gujarati and Kannada Consonants**

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Retroflex/Postalveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiceless</td>
<td>p</td>
<td>t</td>
<td>t̂</td>
<td>c</td>
<td>k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirated</td>
<td>pʰ</td>
<td>tʰ</td>
<td>tʰ</td>
<td>cʰ</td>
<td>kʰ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced</td>
<td>b</td>
<td>d</td>
<td>ď</td>
<td>ǰ</td>
<td>ǧ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breathy</td>
<td>bʰ</td>
<td>dʰ</td>
<td>dʰ</td>
<td>jʰ</td>
<td>gʰ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiceless</td>
<td>f</td>
<td>s</td>
<td>ʃ/ʃ̊</td>
<td></td>
<td>h̊</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced</td>
<td>z/z̊</td>
<td>ɦ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td>η</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>ľ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>u</td>
<td></td>
<td></td>
<td>y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap</td>
<td>ɾ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


**Table 2: Vowels of Gujarati and Kannada**

<table>
<thead>
<tr>
<th>Gujarati</th>
<th>Kannada</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>æ</td>
<td>ø</td>
</tr>
<tr>
<td>ɑ</td>
<td>ɔ</td>
</tr>
</tbody>
</table>


As can be seen in the preceding tables, Kannada and Gujarati share many segmental features. Both languages have a large inventory of plosives, which contrast in five different places as well as in voicing and aspiration. Kannada, with a set of six different fricatives, has a larger inventory than Gujarati, which has a maximum of three depending on the dialect. The languages do share six similar vowel qualities, with Gujarati having a total of eight. Kannada speakers make a length distinction for these vowels. Gujarati speakers make a distinction between nasal, breathy and oral vowels. The phonotactic structures of each language vary as well, with Kannada being more restrictive in syllable

---

1 Normal font = Found in both Gujarati and Kannada; *Italicized Bold* = Found in Gujarati; *Single Underline* = Found in Kannada
structure than Gujarati. Kannada does not permit consonant clusters at the beginning of syllables (except now in a few loan words which often receive an epenthetic vowel), and does not permit codas. Gujarati, on the other hand, allows codas and allows up to two consonants to occur in a cluster in both the onset and the coda.

Russian differs from the two Indian languages in many ways. First of all, there are no retroflex sounds in Russian. There are only contrasts involving three different places of articulation in plosives and two different voicing contrasts (voiced and voiceless). However, for each plosive, there is a phonemically palatalized contrast as well. This gives Russian a total inventory of twelve stops. Most other consonants also make a palatalized versus non-palatalized contrast. These include nasals, fricatives, the trill and lateral. Russian has an inventory of only six different vowels with no contrasting length or nasalization features. It also has much more complex syllable structure, allowing up to five consonants in a cluster in either the onset or the coda. The Russian phonemic inventory is reflected in the tables below:

**Table 3: Russian Consonants**

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Dental/Alveolar</th>
<th>Post-Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td>p b</td>
<td>t d</td>
<td></td>
<td>k g</td>
<td></td>
</tr>
<tr>
<td>Palatalized</td>
<td>pʲ bʲ</td>
<td>tʲ dʲ</td>
<td></td>
<td>kʲ gʲ</td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td>f v</td>
<td>s z</td>
<td>f j z j</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Palatalized</td>
<td>fʲ vʲ</td>
<td>sʲ zʲ</td>
<td></td>
<td>c z</td>
<td>x</td>
</tr>
<tr>
<td>Affricates</td>
<td>ts</td>
<td>tf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palatalized</td>
<td>rʲ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palatalized</td>
<td>lʲ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From http://web.phonetik.uni-frankfurt.de/L/L2008.html based on Haile (1959) and Jones and Ward (1969)
Table 4: Russian Vowels

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>ɛ</td>
<td>ɔ</td>
</tr>
<tr>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

From http://web.phonetik.uni-frankfurt.de/L/L2008.html based on Halle (1959) and Jones and Ward (1969)

1.5. PREDICTIONS AND HYPOTHESES

I predict that the listeners who are exposed to the Gujarati speakers in the training phase will perform more quickly and more accurately than those who hear the American English speakers in both the Russian and the Kannada post-tests. I would also predict that there would be a difference in the Kannada post-test and the Russian post-test between the Gujarati-trained listeners. Specifically, I predict that the Gujarati-trained listeners will respond faster and more accurately in the Kannada post-test than the American English-trained listeners. Since Indian English constitutes one category for many listeners, hearing and adapting to the Gujarati speakers will help update that category. The fact that both Gujarati- and Kannada-accented English will have some features in common will allow these listeners to process speech from another language background more easily, even if these similarities are not quite as salient, and even though there will be differences between the speakers. This prediction is largely based on research that has shown that learning indexical information can help a speaker with subsequent speech processing (Nygaard & Pisoni, 1998). Because of this, adaptation to one accent should be able to generalize to another separate accent that belongs in the same representational dialectical category.

I also predict that the Gujarati-trained listeners will perform more quickly and accurately in the Russian post-test than the control group. The fact that both the Russian and the Gujarati speakers share a salient feature, the /w/~/u/ alternation will help the Gujarati-trained listeners in this regard. This prediction is based on the fact that it has been shown that listeners can adapt to
speakers and then generalize specific features of that person’s speech to a new speaker, and in a relatively short amount of time. However, since the two non-native accents will only be sharing this feature, and because the speakers will represent two distinct categories of accents for the listener, the gains made in this post-test will be smaller than in the Kannada post-test.

1.6. OVERVIEW OF THE STUDY

In order to test these predictions, a study was designed that would allow listeners to be exposed to Gujarati-accented English and then to test accuracy rates and response times in a post-test with either a Russian speaker or a Kannada speaker. This study consists of a training phase and a post-test phase. In both phases, listeners participated in a variety of tasks including, sentence verification, narratives and transcription tasks. There were two different training groups. One trained on hearing Gujarati speakers while a control condition group heard American English speakers. Half of the participants in each condition heard the Russian speaker in the post-test and the other half heard the Kannada speaker. This made for a total of four listening groups.
CHAPTER 2
EXPERIMENTATION

2.1 METHODS

2.1.1 Listeners
Twenty four native speakers of American English participated in this experiment for a total of six participants in each group. About a third of the group of listeners consisted of undergraduate students enrolled in an introductory linguistics course. The group of participants was also composed of other members of the community who varied in age range and profession. All reported normal hearing at the time of testing. Only one reported having a close contact from India. The rest had no close contacts in the U.S., had not had an extensive stay in South Asia, and were not familiar with any of the languages spoken in that region. The undergraduate participants were offered extra credit in compensation for their time, and the other participants were offered six dollars for the half hour they spent completing the five tasks. Each speaker was placed into one of the four testing groups upon arrival to the Phonetics Lab.

2.1.2 Talkers
Three native speakers of Gujarati participated in recording the training phase of the experiment. Two of the speakers were male and one was female. Each speaker had a unique speaking background. The female speaker, who was twenty nine years old at the time, had been in the United States for three years and had been learning English since she was three years old. English was reported to be a language spoken at home for her, along with both Hindi and Gujarati. Impressionistically her English seemed to have the lowest degree of accentedness. The first male speaker was her husband. He was thirty three years old, had been in the United States for the past six years, and had begun to learn English at the age of seven. Gujarati was the only language spoken at home, but he had also learned Hindi in his childhood. The third speaker seemed to have the strongest accent of the three. He also had the most diverse language background. He was born in Tamil Nadu and from childhood began learning
Gujarati and Saurashtra, a language closely related to Gujarati that had been brought down South from Gujarat. He lived in Gujarat for part of his childhood and also spent time in Mumbai. During his time in these areas, he spoke not only Gujarati, but Hindi and Marathi as well. He had been in the United States for one and half years and had been learning English since he was sixteen, which made for a total of about fourteen years as he was thirty years old at the time of recording.

To create a control condition, a group of American English speakers was chosen to mirror the speakers in the Gujarati training condition. There were a total of three speakers, one female and two males. Each had been living in Colorado for at least six months, and each had grown up in the United States with only American English as their first language.

The post-test speakers consisted of two native speakers of Kannada, and one native speaker of Russian. The native speakers of Kannada were both male, were both from Bangalore and had both started learning English as the age of five and had been in the United States for twenty months. They were both twenty five years of age at the time of recording. Neither had had a previous stay abroad in another English speaking country. Both had also learned Hindi, but neither spoke it as a first language. After recording the post-test sentences, the speaker who produced the fewest /w/-/o/ alternations was chosen to be the post-test speaker. The Russian speaker was a twenty four year old female who had been in the United States five years and had begun to learn English at the age of sixteen.

All speakers came to the Phonetics Lab for the recordings and all were offered a pro-rated rate of twelve dollars an hour compensation for their time.

2.1.3 Materials and Set Up

For the production phase of the experiment, recordings took place in the sound attenuated booth in the Phonetics Lab, located in the Linguistics Department at the University of Colorado. All speakers used a Shure Sm10A head-worn microphone. Speakers in both the control and the Gujarati
training groups were asked to record two different sets of sentences and five narratives along with their accompanying True/False type questions about each narrative. The first set of sentences was for two different sentence verification tasks (for sentences used, see appendix I). These were recorded in two sittings. Some of the sentences were taken from Brown (1999), which was based on Munro and Derwing (1995). Other sentences were newly invented and some of the borrowed sentences were edited to contain more tokens of the /w/ phoneme. There were a total of forty sentences, twenty for each task. Out of these forty sentences, half contained at least one occurrence of /w/. Twenty-one of these sentences had the value of True, while nineteen had the value of False. The false sentences contained examples of semantically anomalous sentences as well as those that could be judged to be false based on cultural and world knowledge. Some examples can be seen below:

*A calendar informs you when it is noon.*
*Airplanes fly in the sky.*
*All women have long hair.*
*Hot is the opposite of cold.*

The second set of sentences was for the transcription task. These were both statements and questions that the participant did not have to answer. These were largely taken from Bansal (1976) but were edited to be longer and to contain more /w/ tokens. Some of the sentences were invented to contain more /w/ tokens in environments where the speaker was most likely to produce the /u/ substitution. There were a total of twenty three sentences each participant had to transcribe, and out of these twenty three, eighteen contained at least one occurrence of /w/. The five narratives were taken from various grade school readers and featured texts about nature, such as about animals, Mount Everest and weather. Each narrative was one to two paragraphs long. These texts were edited in areas to contain more /w/ tokens and one to three questions were invented for each text for the participant to respond to after listening to each narrative. In all of these tasks, the phoneme alternation occurred in syllable-initial position, in both word-initial and word-medial positions such as in winter and always, with
most of the tokens occurring in word-initial positions. There were also a few words in which it occurred in a cluster, e.g. words such as swim.

Each sentence or narrative was put into its own file. The sound files for the sentence verification tasks were cut to the very start of the sound and then measured for duration to the millisecond in order to be able to accurately measure response times from the end of the stimulus. Some files were manipulated to contain more /u/ substitution tokens when necessary.

Recordings for the post-test also took place in the sound booth. Speakers were asked to record one set of sentences to be used in the post-test. There were a total of twenty three sentences, all of which contained at least one token of the /w/ phoneme. These were also put into individual sound files, cut to the very beginning, and measured for duration.

These recordings formed the experiment for the perception phase. The experiment was designed using Psychopy software. There were no visual stimuli; everything the participants were exposed to was solely auditory. A participant was placed into one of four groups upon arrival in the lab. These were based on the training condition as well as the post-test to which the listener was exposed. The listener heard either the American English speakers or the Gujarati speakers in the training phase. Then they were exposed to either the Kannada or the Russian speaker in the post-test phase.

Table 5: Combinations of Training and Post-test Groups

<table>
<thead>
<tr>
<th>Group nº</th>
<th>Training Language</th>
<th>Post-test Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gujarati</td>
<td>Kannada</td>
</tr>
<tr>
<td>2</td>
<td>Gujarati</td>
<td>Russian</td>
</tr>
<tr>
<td>3</td>
<td>American English</td>
<td>Kannada</td>
</tr>
<tr>
<td>4</td>
<td>American English</td>
<td>Russian</td>
</tr>
</tbody>
</table>

The participants listened to the speakers in the sound booth through a pair of headphones and used either a button box or a keyboard to log their responses.
2.1.4 Procedure

Each participant came to the Phonetics lab located in the Linguistics Department at the University of Colorado. Everyone was asked to read and sign the consent form and then to fill out a short survey regarding their linguistic background, whether they had any hearing problems, and whether they had any contacts with any speakers from India or whether they had spent some time there. These questions were designed to determine whether any of these factors could have played a role in affecting the outcome and designed to confirm that there would be no factors that could affect the results. After the participant had filled out these forms, they went into the sound booth and were given a short overview of what tasks we were going to do. I told them they would hear several speakers, but no emphasis was put on this point, and that they would be completing a series of five different tasks. I then gave instructions to the first task which was the first sentence verification task. This was followed by the narrative task, which was followed by the transcription exercise. The last training sentence verification task was after the transcription exercise, which was followed by the post-test sentence verification task. Between each task, there was a slight pause where I switched the participant from the button box onto the keyboard (or vice versa) and gave them a short set of instructions for each task. The only time the participant did not have a pause in between tasks was between the fourth task of the training phase and the post-test. Both of these tasks were the same (sentence verification) so before the fourth task began, and after explaining that the next two tasks were sentence verification tasks, I informed each participant that these two tasks would be run consecutively.

In each sentence verification task, the listeners heard twenty statements and were asked to judge if they were true or false by pushing the appropriate button on the button box. They were told that the statements were not meant to trick them and that they should answer each as quickly and as accurately as possible. The narrative task was completed after the first sentence verification task. The
listeners were instructed on how to use the keyboard. They then heard five narratives. Each narrative was followed by a short pause which was, in turn, followed by one to three True/False statements. The listener’s answer was input on the keyboard. This was followed by the transcription task. The listener was given a pen and paper and instructed to write each sentence down. They were informed that each sentence would be played only once and asked to write down everything they heard as best as they could. They then had to push any key on the keyboard to move on to the next sentence. This task was followed by the final sentence verification task in the training phase, which was immediately followed by the sentence verification task for the post-test.

2.2 RESULTS

Overall, accuracy rates were quite high for both the post-test groups. There were some slight differences in accuracy between the Gujarati-trained listener group and the American English (AmE) trained listener control group. A paired by-items t-test was performed on the number of correct responses. This indicated that the Gujarati listeners performed more accurately in the post-test than the AmE listeners. The Gujarati listeners had entered a correct response for 267 of the 276 questions for a correct response rate of 96.74%. Those in the AmE control condition answered 256 of the questions correctly for a correct response rate of 92.75%. This does show to be a significant difference with $t(45)=2.12, p<0.05$.

More paired by-item t-test analyses were carried out to investigate the effects of training group on the individual post-tests. These showed that the Gujarati training group was significantly more accurate than the AmE group in the Kannada post-test, with only one incorrect answer. The Gujarati listeners answered 137 questions correctly out of a total of 138 with a correct response rate of 99.28%. The AmE group answered 130 correctly with a correct response rate of 94.2%. This also shows to be significant with $t(22)=2.61, p < 0.05$. 
The Gujarati listeners were also slightly more accurate in the Russian post-test, but not to as such a high degree as in the Kannada post-test. The Gujarati listeners logged a total of 130 correct responses out of 138 for a correct response rate of 94.2%. The control group answered a total of 126 out of the 138 correctly for a correct response rate of 91.3%. This difference, however, is not significant with \( t(22) = .89, p = .38 \). Accuracy results shown by training group in the two post-tests are shown in Fig. 1:

![Fig. 1: Post-test Accuracy Scores for Gujarati and AmE Listeners in both post-tests](image.png)

There was also found to be a main effect for post-test group. Overall, participants in the Kannada post-test were more accurate with a total accuracy score of 267 out of 276 for a correct response rate of 96.74%. The total accuracy score for the Russian post-test was 235, which is a correct response rate at 85.14%. This was significant at \( t(22) = 4.36, p < .05 \).

Overall, post-test reaction time results showed that the Gujarati-trained listener group was on average slightly faster than the American English trained listener control group. Reaction times were first culled by removing outliers that were one standard deviation above the total mean reaction time. Then a paired by-items t-test analysis was performed on the post-test reaction time scores. The Gujarati listeners logged in their answers at mean reaction time rate of 544ms with a standard deviation of
328.36 whereas the AmE group answered in an average of 550ms with a standard deviation of 311.03. Given the high variability, however, there was no significant effect found in the paired t-test with $t(45)=-0.65$, $p=.519$, meaning that the Gujarati listeners were not significantly faster than the AmE listeners.

There was a main effect for the post-test group. After the data was culled, it was found that participants in the Russian post-test were faster with an average RT of 474ms with a standard deviation of 323.23, and those in the Kannada post-test had a mean RT of 611ms with a standard deviation of 319.8. This was found to be significant at $t(22)=3.34$, $p<.05$. This effect can be seen in Fig. 2:

![Overall Mean RTs in Post-Tests](image)

More importantly, however, the two post-test groups needed to be examined separately from each other. Both a paired by-items t-test and an unpaired by-subject t-test were performed on these results. For the Kannada post-test the Gujarati listener group was slightly slower with a reaction time mean of 613ms and a standard deviation of 309.24. The AmE group scored an average reaction time of 609ms with a standard deviation of 298.18. The paired t-test does not reach significance with $t(22)=.242$, $p=.811$. The unpaired t-test also fails to reach significance with $t(274)=1.36$, $p=.18$. 
The Gujarati listener group was slightly faster than the AmE listener group in the Russian post-test with a mean response time of 467ms and a standard deviation of 333.43. The AmE listeners, on the other hand had a mean response time of 482ms with a standard deviation of 313.14. This was also not found to be of statistical significance with $t(22)=-.62, p=.54$ for the paired t-test and with $t(274)=-.21, p=.84$ for the unpaired test. Post-test RTs by post-test and training group are shown in Fig. 3:

**Fig. 3: Post-test Mean Response Times for Gujarati and AmE Listeners in both post-tests**

There was some variability in performance across the different individual sentences of the post-test. For both post-tests the AmE listeners had faster RTs for eleven of the twenty three sentences and the Gujarati listeners had faster RTs for twelve. Many of the mean RTs of each sentence for each group were fairly close in value which led to there being no significant differences between the two training groups in terms of reaction times. There were certain sentences, however, that did differ more drastically. In the Kannada post-test the Gujarati listeners had a total of four sentences in which they had a mean RT of 200ms or more faster than the AmE group. The AmE group had three sentences with mean RT times faster (also by 200ms or more) than the Gujarati listeners. In the Russian post-test, the Gujarati listeners had faster mean RTs for five of the sentences, and the AmE listeners had faster mean...
RTs for three of the sentences. These differences are reflected in the tables below. The two training groups had no large differences between each other in the other sentences.

**Table 6: Individual sentences with faster (over 200ms) response times in the Kannada post-test**

<table>
<thead>
<tr>
<th>Kannada Post-test</th>
<th>Gujrati Listeners</th>
<th>AmE Listeners</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence:</strong></td>
<td><strong>Order:</strong></td>
<td><strong>RT in ms (Acc):</strong></td>
</tr>
<tr>
<td>Some people will win the lottery</td>
<td>1</td>
<td>861 (6)</td>
</tr>
<tr>
<td>A triangle has twelve sides</td>
<td>4</td>
<td>474 (6)</td>
</tr>
<tr>
<td>The weatherman gives you the forecast</td>
<td>10</td>
<td>326 (6)</td>
</tr>
<tr>
<td>Winnie the Pooh is a bear.</td>
<td>23</td>
<td>706 (6)</td>
</tr>
<tr>
<td>Some young women are fathers</td>
<td>3</td>
<td>788 (6)</td>
</tr>
<tr>
<td>Whales are large land animals.</td>
<td>8</td>
<td>574 (5)</td>
</tr>
<tr>
<td>Norway is a country in Africa.</td>
<td>19</td>
<td>926 (6)</td>
</tr>
</tbody>
</table>

**Table 7: Individual sentences with faster (over 200ms) response times in the Russian post-test**

<table>
<thead>
<tr>
<th>Russian Post-test</th>
<th>Gujrati Listeners</th>
<th>AmE Listeners</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence:</strong></td>
<td><strong>Order:</strong></td>
<td><strong>RT in ms (Acc):</strong></td>
</tr>
<tr>
<td>Some people will win the lottery</td>
<td>1</td>
<td>540 (6)</td>
</tr>
<tr>
<td>Fish swim in ice cold Coco-Cola</td>
<td>2</td>
<td>301 (6)</td>
</tr>
<tr>
<td>Whales are large land animals.</td>
<td>8</td>
<td>280 (6)</td>
</tr>
<tr>
<td>Taiwan is a city in France.</td>
<td>11</td>
<td>696 (5)</td>
</tr>
<tr>
<td>The US is larger than Switzerland</td>
<td>15</td>
<td>310 (4)</td>
</tr>
<tr>
<td>Some young women are fathers</td>
<td>3</td>
<td>639 (6)</td>
</tr>
<tr>
<td>Frogs write wonderful newspaper articles.</td>
<td>17</td>
<td>655.45 (6)</td>
</tr>
</tbody>
</table>
As can be seen from the tables above, for both post-tests the very first sentence the listeners were exposed to varied the most in mean response time between the AmE and the Gujarati listeners. The sentence that had the largest difference between the two training groups in both post-test conditions was the very first sentence. Both groups of listeners were slower than the total average in both post-tests, with the AmE group having a much slower RT than the Gujarati listeners. In the Kannada post-test there was a difference of 366ms between the two training groups. The RT of the AmE listeners was 1226, which was a total of 605ms over their average for the post-test. The Gujarati listeners’ mean RT for this sentence was 861, which was 250ms over their total mean RT. This makes for a difference of 366ms. The difference in response time was accompanied by a difference in accuracy. Four of the AmE listeners answered this correctly, while all six of the Gujarati listeners answered this question correctly.

In the Russian post-test the AmE listeners’ mean RT for the first sentence was 1112ms which deviates from their total average by 675ms. The mean RT for the Gujarati listeners was 540ms, which only higher than their average by 134ms. The two groups differ in RT by 572ms. Again, for this sentence four out of six of the AmE listeners answered this question correctly while all six of the Gujarati listeners answered it correctly.

As can be seen from the preceding two tables, while there is a fairly drastic difference in the first sentence there are not many other sentences that have large response time differences between the two training groups. In some cases, one of the training groups answered well below their average RT, while the other group answered only slightly below or slightly above (e.g. sentence 10 in the Kannada post-test). In other cases one group logged a high response time above their mean, while the other group had an answer that did not deviate too drastically from their mean (e.g. sentence 17 in the Russian post-test). It is also apparent from these tables that both groups had particular answers for
which they performed more quickly than the other group. The fact that there is so much variability makes it difficult to see any patterns that could indicate that one training group is performing consistently more quickly than the other group, making it difficult to see any response time trends that could be indicative of adaptation.

Because there were no significant differences in response times between the two training groups in the post test, it is worth trying to determine if there was any adaptation in the training phase for the Gujarati speakers. This can help determine whether listeners were able to adapt in the training phase or whether there was something preventing adaptation, which was therefore not able to carry over into the post-test. In order to do this, both the first and second training sentence verification tasks can be examined to see if there are any differences in either response times, accuracy rates or both from the first task compared to the second. Examining RTs and accuracy rates in these two tasks indicated that accuracy improved for the Gujarati listeners, while RTs increased in the second task. Accuracy stayed the same and RTs increased for the AmE listeners between the two tasks. In the first sentence verification task (which, as stated above, was the very first task in the training phase) the Gujarati-trained listeners had a mean RT of 611ms with a standard deviation of 374.43 and an accuracy score of 205 out of 240 (at 85.42%). They had a mean RT of 626ms with a standard deviation of 340.47 and an accuracy of 222 out of 240 (91.67%) in the second sentence verification task (which was the fourth and last training task).

The AmE listeners started the first sentence verification task with a mean RT of 408ms and a standard deviation of 300.14 and an accuracy score of 216 of 240 for an accuracy rate of 90%. They scored a mean RT of 491ms with a standard deviation of 286.81 in the second verification task and had an overall accuracy score of 213/240 which put their accuracy rate in this task at 88.75%. These differences in accuracy scores and response times are reflected in the following Figs. 4 and 5:
2.3 DISCUSSION

Overall, reaction times in this study fail to be a good indicator that perceptual adaptation and generalization have occurred. In many of the post-test sentences reaction times between the two
training groups are not too different from each other. There are a few sentences that differ drastically, but it is not predictable as to which group will be faster in a certain sentence. The only case where this is the exception is in the first sentence where the Gujarati-trained listeners are consistently faster than the AmE listeners. This difference, however, is quickly overcome. Another indicator that reaction times in this study are not the best measure of adaptation is the fact that both the Gujarati listener group and the AmE listeners are slower in the second sentence verification task than in the first task. The AmE listeners do not have to adapt to any speech above and beyond the individual talkers in the training phase. By the second task they should have gained enough practice and familiarity with the speakers that would cause RTs to fall in this last task. This shows, rather, that there are other factors affecting reaction times to which listeners are sensitive. These could include semantic or syntactic complexities of certain sentences that increase the average response time. Variability in response times between listeners seems to be quite high.

Measures of accuracy, on the other hand, seem to be a much better indicator of adaptation. The first sign that this may be true is the fact that the Gujarati-trained listeners, at an accuracy rate of 85.42%, perform much more poorly in the first sentence verification task of the training phase than the listeners in the AmE control condition who perform at a rate of 90%. This increases drastically for the Gujarati listeners in the second sentence verification task who reach 92.5%, whereas the rate for the AmE listeners falls very slightly to 88.75%. The Gujarati trained listeners were more accurate in both post-tests than were the AmE listeners. This difference does show to be significant which could indicate that the Gujarati listeners were able to adapt even if their response times did not decrease.

Not only does the accuracy rate of the Gujarati listeners improve from the first sentence verification task to the second in the training phase, but it also continues to improve in the post-test, where they score an accuracy rate that is higher than either of the two in the training phase. Accuracy
scores of the Gujarati-trained listeners are very high in the Kannada post-test, with an almost perfect score. The fact that there is a continuing trend of improvement combined with such a high accuracy score in this post-test lends support to my hypothesis that the Gujarati-trained listener group will perform more accurately than the AmE listener control group upon hearing a Kannada-accented speaker. This finding could suggest that listeners are using representations previously indexed as Indian English, including Gujarati-accented data they were exposed to in the training phase, and are applying it when listening to a Kannada speaker. The application of the knowledge listeners gained from the Gujarati speakers then can help them process the speech of the Kannada speaker, which improves their accuracy scores quite drastically.

In the Russian post-test there is again a difference in accuracy. The Gujarati trained listeners performed more accurately in the Russian post-test than did the AmE trained listeners. They also had slightly faster response times. If, as stated above, accuracy is a better measure of adaptation than response times, then this also provides evidence for my next hypothesis that exposure to salient features in one non-native accent can generalize to another, unrelated accent since there is a significant difference in levels of accuracy between the AmE listener control group and the Gujarati-trained listener group.

For the Gujarati trained listener group, there was a difference between accuracy results in the Russian post-test and the Kannada post-test. The Gujarati listeners had an almost perfect accuracy score in the Kannada post-test, which was not the case in the Russian post-test, in which they answered eight questions incorrectly. There was also a larger degree of accuracy between the Gujarati listeners and the AmE listeners in the Kannada post-test as opposed to the Russian post-test. In the Kannada post-test there was an accuracy score difference of seven points, with a score difference of only four in the Russian post-test. This would indicate that the Gujarati listeners adapted more successfully to the
Kannada speaker than to the Russian speaker, which supports my prediction that the adaptation to Russian will be of a lesser degree since speakers will have been able to generalize one feature and will not be able to use the representation of indexed information they built up from hearing the Gujarati speakers. Because Russian accented English constitutes a distinct category from the Indian speakers, listeners do therefore not index anything helpful for the Russian speaker.

One aspect that needs to be taken into consideration for the Russian post-test is the study of reaction times by Clarke and Garrett (2004) who showed that initial increases in processing times can be quickly overcome. It’s possible that the AmE listeners overcame their initial delay in response time within a matter of the first few sentences. This could lend more support to the earlier suggestion that measuring and averaging response times is not the most accurate indication of whether a listener has adapted to a specific accent, or to a specific feature of that accent. The fact that the Gujarati-trained listener group performed better in accuracy could be a more reliable indication that their training helped their later perception of the Russian speaker.

There are a couple of potential reasons as to why the Gujarati trained listeners performed more accurately in both of the post-tests than the control group. Firstly there could have been a general attention effect. Listeners could have become used to paying more attention and listening more closely, which would both be reasonable requirements in hearing non-native accented speech in general. The amount of attention they paid to the task could be the reason behind the higher accuracy scores than for the control group, who would have not have needed to pay as close attention. As had been mentioned in the Methods section, a portion of the listeners did come from an introductory Linguistics course, which means these listeners could have been predisposed to pay more attention to small phonetic details more so than the average population.
However, the higher accuracy scores could have also been caused by an adaptation process in which listeners are beginning to adjust their phonemic categories upon hearing in the non-native speech patterns. It could be that, by hearing and learning the systematic idiosyncracies of non-native speech, a listener’s phonemic categories begin to shift or become more flexible in order to be able to accommodate for the variation. However, as a first step in adapting to a non-native accent, a listener has to first be able to comprehend the underlying linguistic message and can then start to explicitly learn the substitutions and alternations a particular speaker or speech group may produce. For example, a listener hears the sentence *Friday is a day of the week* with *week* pronounced as [uik]. First the listener is able to equate [uik] with [wik], and after enough examples with a similar substitution in other words the listener could be able to analytically learn this particular substitution, so every time the speaker produced a [u] the listener will know to interpret it as a /w/. Because this is happening analytically and fairly consciously, response times are not going to start decreasing immediately. However, since the listener is beginning to adapt to the variations in order to better recover the underlying linguistic message, his accuracy scores can increase. Not all features will be quite as salient as this /w~/u/ alternation, and linguistically naïve listeners will most likely not be able to remark on or become consciously aware of features such as differences in prosody, VOT times and shifts in vowel qualities. Despite the fact that listeners may not notice their adjustment to these features as clearly, the mechanisms behind these shifts could function similarly.

Response times, as opposed to accuracy scores, were not found to have been affected by training with the Gujarati speakers. A decrease in response times could be dependent on having developed representations with categorical indexical markings. Strong dialectical indexation would allow listeners to quickly be able to access a lexical category to help process similar incoming speech which would lead to a decrease in response times. In order to form a dialectical category that is easier for the listener to access and identify, enough of the patterns have to be stored and classified. This
could easily require more time and exposure to a set of speakers than this study provided which means that response times would not have been affected.

Another issue that needs to be addressed is the fact that the Russian post-test is significantly faster than the Kannada post-test. There are a couple of potential reasons as to why this could be the case. It is possible that the listeners in general have already developed a better a priori dialectical category for Russian speakers, having had more previous relevant experience with Russian accented speakers than they have had for Kannada accented speakers. Having already developed this category could have led to more ease of understanding accompanied by a decrease in response times. However, it could have also been caused by the fact that the Russian speaker was less accented than the Kannada speaker. This would have caused there to be fewer differences between her speech and American English than there was between the Kannada-accented English and American English. I would predict that this could have been the driving factor in allowing for faster response times in the Russian post-test.

Finally, this leaves the question as to why accuracy improves in the training group while response times do not. As predicted above, it is possible that there are two different mechanisms in the adaptation process. The first of these involves a mechanism in which the listener learns the specific substitutions used in the speech of a particular speech community. This could be analytical at first, and the listener is probably explicitly aware of some of the deviations from native-accented speech. Once a listener begins to adapt and shift phonemic categories in response to one speaker, he can generalize that shift onto another speaker, in the same way that was shown in previous studies (Kraljic & Samuel, 2006). This would allow for an increase in accuracy since listeners could then better interpret the underlying linguistic message. This does not, however, necessarily have to lead to faster response times since the analysis of speech is still slow and analytical. There may also be other non-native features that a listener still has to adapt to, which would keep response times relatively high. Also, it was previously
proposed that response times would only truly decrease after strong dialectical category indexation was formed, which would occur over a longer time period than the phonemic category shift (which could happen relatively quickly). It could be that category indexation occurs only after listeners learn the phonetic patterns of a specific group (i.e. after categories have already shifted). It would be interesting for future studies to compare the response times of this study to a similar adaptation study that takes place over a longer period of time to see if response times do start to decrease over a longer time period and a greater amount of exposure.

2.4 LIMITATIONS OF THIS STUDY

There are several aspects of this study that could have prevented any findings or trends from being as strong and clear as they could have been. One area that could have been problematic is the amount of exposure. Many of the studies that have investigated adaptation to non-native speech had required the participants to come in repeatedly over a period of days before they completed the post-test. This experiment took half an hour for the participant to complete, so the participants received much less exposure. These previous experiments often looked at intelligibility through tasks such as transcription exercises, whereas mine was designed to look at the adaptation and generalization at a finer level by measuring reaction times, so it was assumed that differences in reaction times could be measured over a shorter duration of time. This does not seem to have been the case however. Also, there have been no previous studies measuring response times in the process of generalization. It could be that intelligibility is a better indicator of generalization than speed.

This was also a fairly small study in terms of sample size. This could cause any trends to not be apparent with only six participants in each group.

Another aspect that could have been problematic was the third Gujarati speaker. Most adaptation studies have used speakers of one particular language background. This speaker, however, was from a
different area than the other two and was bilingual in both Saurashtra and Gujarati. Because this study did make assumptions on how a speaker’s L1 can affect the realization of his L2, this could have played a role, and his speech could have had different properties than the other speakers. This could have introduced too much variability for the listeners to be able to form a good Gujarati category or to be able to latch on to any indexical properties of the speakers. Another potential issue with this speaker is that he had a high degree of accent and he had the lowest degree of accuracy of all three speakers. This could be indicative that his overall intelligibility is lower. Bradlow and Bent (2008) have shown that listeners can adapt more quickly to speakers with higher levels of intelligibility. Because the exposure to non-native speech was so short, it could have been more effective to use a speaker with a higher intelligibility level.

2.5 IMPLICATIONS

One aspect this study looks at is the level of dialectical categorical representation. It is still unknown whether listeners are building new, distinct categories in the generalization process or whether they are updating categories that already exist for the listener. The findings that the Gujarati listeners were able to significantly improve through the Kannada post-test could be a step in answering this question as it suggests that listeners are updating an existing category.

This study poses some interesting questions for exemplar theory. As stated above, Weil (2001) couches his research in exemplar theory, but does not look at reaction times. Citing Goldinger (1996, 1998), he claims that reaction times should decrease with more exposure to a specific non-native speech variety. This decrease in reaction times is caused by more exemplars of non-native speech tokens. The more there are the easier it is to access to compare the input to an already existing token. It is interesting to note that there is a significant increase in accuracy rates in my findings, and that this is accompanied by only a little amount of decrease in reaction times. This could potentially present a
problem for exemplar theory as a decrease in reaction times is supposed to be part of the adaptation process, or it could be indicative of how much exposure a listener needs to be able to quickly access a token for comparison. It may be the case that a decrease in reaction times requires more stimulus exposure, which, as discussed above, may correspond with forming indexical categories based on dialect.

In line with the above, these findings also have the potential to help us understand the generalization process by giving an indication as to how flexible it can be. In this way, it takes a step towards answering questions about how many speakers a listener needs to be exposed to before generalization can occur and whether exposure to one group of speakers can help in the perception and processing of the speech of another group of features.

Finally, the study of adaptation to non-native speech could have implications for improving communication across different varieties of English. This is especially relevant today, where according to Crystal (2003) English has been able to reach global status in that it has been officially recognized by many nations by either having an official status in a country or being taught as a mandatory subject in schools. Today there are more L2 and foreign language speakers of English than there are native speakers, and the number of L2 speakers is growing at a faster rate than the number of native speakers. India in particular has a relatively large English speaking population with about one third of the country’s population being able to hold a conversation in English with about twenty percent using the language on a regular basis. Despite the fact that English is spoken in both the United States and India, these two varieties differ from each other on different linguistic levels, which can impede communications. Anything done to understand where these miscommunications could be coming from and how they could be minimalized could be of benefit for international communication. The fact that accuracy scores were higher for the Gujarati-trained listeners could be particularly useful here since it suggests exposure
to any variety of Indian English can help overcome these communication barriers caused by phonological differences of the two varieties. Such studies can have practical implications in different environments, such as in university classrooms where a number of teaching assistants come from diverse language backgrounds and interact extensively with native speakers of American English. The area of international business is another setting in which this could be relevant since people of many different backgrounds have to work together to achieve a common understanding. Research, such as that carried out in this study, can start to indicate how flexible a listener can be in understanding non-native speech. Indications that a listener can actually be quite flexible and come to adapt to different non-native speech patterns can lead to training programs that help listeners better be able to understand non-native accented speech. If it is found that listeners cannot adapt easily in being able to perceive non-native speech, training programs might rather want to focus on the production aspects of speaking a second language. These findings suggest that listeners can be quite flexible and quickly improve accuracy in understanding a non-native speaker, which suggests that perceptual learning training programs could be beneficial in these exchanges.
REFERENCES


Appendix I: Sentence Verification Stimuli

Training Task 1:

A brick weighs more than a feather.
A calendar informs you when it is noon.
Airplanes fly in the sky.
Bears buy dinner at restaurants.
Bears wear socks and shoes to school.
California is west of New York.
Children learn to talk before they learn to drive.
Doctors and nurses sometimes work in hospitals.
Horses are bigger than foxes.
In the winter, gloves keep your hands warm.
Many children like to play outside.
On rainy days, there are clouds in the sky.
People smile when they're feeling sad.
People usually walk on their hands.
Saturday is part of the weekend.
Some people like to eat pizza.
Women give birth to babies.
You can buy fruit at the grocery store.
You can watch news on television.
Weasels lay big purple eggs.

Training Task 2:

A raw egg will break when you drop it.
All women have long hair.
Carrots are one type of purple fruit.
Carrots grow on apple trees.
Cars need orange juice to run.
Dictionaries contain word definitions.
Friday is a day of the week.
Hot is the opposite of cold.
May is a month in the winter.
Most boys wear blue wigs.
Most people work in a pharmacy.
People wear shoes on their feet.
People wear warm fur coats when it is hot.
Some people believe in God.
The moon is the brightest light in the daytime sky.
To get your roof repaired, you should call an author.
Vegetarians like to eat hamburgers.
Wild lions make good, friendly pets.
You can always see the sun at night.
Your father is younger than you.
Post Test:

Some people will win the lottery.
Fish swim in ice cold Coca-Cola.
Some young women are fathers.
A triangle has twelve sides.
There are twenty-four hours in a day.
Orange juice tastes like watermelon.
Swiss cheese has holes.
Whales are big land animals.
The woman a man marries becomes his wife.
The weatherman gives you the forecast.
Taiwan is a city in France.
Ice is made of frozen water.
Wolves usually are covered in fur.
Polar bears have white fur.
The United States is larger than Switzerland.
Wine is made from celery.
Frogs write wonderful newspaper articles.
Norway is a country in Africa.
Dogs wag their tail when they’re happy.
Some people eat sandwiches for lunch.
Russia is very warm in the winter.
London is located in the western part of China.
Winnie the Pooh is a bear.
Appendix II: Transcription Sentences

Has the boss found his way to work yet?
That’s a fun dress that you’re wearing.
I always feel so tired when I’m trying to work.
Forget about him, we’ll just go to the farm by ourselves.
I would like it if you were to wear the blue suit.
We dined at the hotel before turning in for the night.
We export a lot of different teas.
The woman was dressed in a wool sweater.
Which is the man we were with?
William was living out west for awhile.
I’ll try not to worry, but when will you be back?
We went out to eat on this past weekend.
This upcoming week will pass pretty quickly.
I was once told you shouldn’t wear white after Labor Day.
He believes his remarks to be quite witty.
Do you plan on attending the wedding on Wednesday?
It sure is windy out today.
Do you know the way to campus?
We will try to win some money.
He would always wander away from us.
I’m living with my brother and we fight a lot.
I often buy all my fruit at that store.
Let’s have some peas with our dinner.
Do you have a pen so I can fill out this form?
She’s a professor of mine who failed her nephew in that class.
Appendix III: Narratives

Narrative 1:

What’s the difference between African wild dogs and the dogs we know as pets? For one thing, African wild dogs, which live in Africa, south of the Sahara desert, only have four toes, while domestic dogs and wolves have five. But you wouldn’t want to count for yourself, because these are truly ferocious animals.

Wild dogs are not somebody's domestic dogs that ran away and didn't find their way back, although some people often used to think that," explain researchers, who study these animals in the African country of Botswana. "They are actually Africa's wolf, and just like wolves, they do not make good pets. They need to be out in the wild doing what they are supposed to be doing – freely ranging many miles every day and hunting to find the food they need to survive and feed pups."

Questions:
African wild dogs make good pets.
These dogs hunt to find food.
African wild dogs live in Africa, south of the Sahara.

Narrative 2:

At an elevation of more than 29,000 feet above sea level, Mount Everest is the world’s tallest mountain. Mount Everest is relatively young, having been formed only 60 million years ago, and it is still growing a few millimeters every year! Mount Everest is in the mountain range called the Himalayas. Different countries and languages have different names for this mountain, each coming from a different source. In English, Mount Everest is named after Sir George Everest, the first person to record the height and location of the mountain.

Questions:
Mount Everest is the world’s fifth tallest mountain
Mount Everest is located in the Himalayas.

Narrative 3:

When we think of winter, we often imagine sledding, icicles and snowball fights. But think about this: winter is when earth’s north pole is tilted farther away from the sun than at any other time of the year. This is called the winter solstice, and the days just before Christmas are the shortest of all. During this time, the sun shines more directly on the lower half of the earth, or Southern Hemisphere. While we in the United States have winter, Australia, South Africa and the southern parts of South America will have summer. Because the weather continues to cool for about another month after the shortest days, the coldest weather actually arrives after the winter solstice. For most of the Northern Hemisphere, January and February are typically the coldest months.
Questions:
It’s winter in the US while it’s summer in Australia.
The coldest weather in the Northern Hemisphere occurs before the winter solstice.

Narrative 4:
The weather forecast tells us what the temperature and air or wind conditions are likely to be outside in the near future. There is a wide variety of weather, from sunny to stormy and warm to cool. It can be rainy or cloudy or windy. Listening to or watching the weather forecast can help us be prepared so we know what to wear. If weather conditions will be severe, like snow, a storm or a hurricane, getting information ahead of time can help us prepare so we can stay safe.

Questions:
Weather forecasts are always accurate.
Listening to weather forecasts can be helpful.

Narrative 5:
Wind brings us weather because it blows clouds from one place to another. Therefore, it is helpful to know from which direction the wind is blowing. A wind vane can provide this information so meteorologists know what is coming. Weather forecasts are not always right, but they get more accurate all the time.

Questions:
Wind can bring weather.
Wind vanes are used to measure rain.