UC Berkeley

Title

Phonetic Accommodation to Non-Native English Speech

Permalink

https://escholarship.org/uc/item/7gq1s5n4

Author

Liu, Qiu Ting

Phonetic Accommodation to Non-Native English Speech

Qiu Ting Liu

Advisor: Professor Keith Johnson

University of California, Berkeley

Spring 2017

Table of Contents

Phonetic Accommodation to Non-Native English Speech

Abstract

Phonetic accommodation is the process in which a speaker becomes more phonetically similar to his or her interlocutor over the course of a conversation. This experiment investigates phonetic accommodation in the English speech of Mandarin speakers after exposure to a model speaker who shares their language background. The results show that when including phones, tasks, and conditions as dependent variables, there are statistically significant differences across tasks and phones. Phonetic accommodation is observed in all shadowing tasks and the effect remains in post-shadow task in some dependent variables. The social manipulation of this study is only statistically significant in formant durations and word-final consonant clusters durations and the pattern suggests that subjects who were in the condition that encouraged speakers to achieve a closer social distance with the interlocutor accommodated more than subjects in the condition that were encouraged to be native-English like. This experiment thus contributes to the understanding of phonetic accommodation in second language English speakers under different conditions.

1 Introduction and Background

1.1 The issues

Speech is the vocalized form of communication, which people use to interact and share thoughts with others on a daily basis. In the process, speakers are often exposed to phonetic variants that are different from their own, and they may pick up phonetic features of their interlocutors, resulting in changes in their pronunciations temporarily over the course of a conversation. There are multiple reasons why a person would consciously or subconsciously change his or her accent. For example, speakers, when moving to a new environment and immersing in a different accent for an extensive period of time, may modify their speech and acquire features that are distinct in the dialect subconsciously. Alternately, speakers may understand that their accents give away their social status or language background and may modify their speech consciously to their advantage under certain situations.

Phonetic accommodation, also known as phonetic adaptation, is the process by which a talker acquires phonetic features from an interlocutor's speech (Babel 2012). Linguists have sought explanations for this phenomenon utilizing different mechanisms, including psychological and sociopsychological perspectives, to account for this phenomenon. This study investigates phonetic accommodation observed in second language speakers of English under different social situations in order to understand whether this phenomenon is observed in L2 English speakers.

1.2 Models for Phonetic Accommodation

The automatic interactive alignment account proposed by Pickering and Garrod (2004) attempts to account for all types of linguistic alignment, including phonetic accommodation. The account is developed in part from the researchers' beliefs that dialogue is the basic form of language use, and that every level of linguistic representation – semantics, syntax, and phonology – is tightly connected within an individual and between interlocutors. In daily communication, interlocutors constantly interact with each other and exchange information at every level of linguistic representation. In order to facilitate conversation, speakers make their productions reach an implicit common ground, at all levels of the linguistics representations. While speakers prime themselves with their interlocutors, speakers also self-monitor and employ a repair mechanism, which ensures that they maintain the common ground. The self-monitoring mechanism checks if one can interpret the interlocutor's expression and if the speaker fails to do so, it reformulates the utterance so that it will lead to the established implicit common ground. The process that speakers acquire their interlocutor's phonetic features during the conversation is what observed as phonetic accommodation. In short, the automatic interactive alignment model assumes that the convergence at all levels of linguistic representations of speakers are automatic and at a subconscious level.

Goldinger (1998) proposes an exemplar-based model suggesting that phonetic accommodation is due to an automatic cognitive reflex of the system. The model posits that detailed traces are stored in the lexicon and can be retrieved during later perception. Goldinger's (1996) research assessed both explicit and implicit memory for spoken words and demonstrated that voice attributes of spoken words were retained in memory. In a follow up study, Goldinger (1998) conducted a lexical shadowing task showing that low frequency words and unique voices activate fewer traces in the lexicon because speakers had encountered fewer cases of these. As a result, speakers are able to retrieve the information more easily and would be more likely to converge.

Although these two psycholinguistic models differ with respect to whether phonetic accommodation is the result of between-speaker phonetic negotiation or the activation of episodic memory, both the interactive alignment account and the exemplar-based account argue that phonetic accommodation is inevitable and automatic. In contrast, the Communication Accommodation Theory (CAT) is a sociopsychological model that explores how social context facilitates phonetic accommodation (Shepard et al. 2001). The theory attempts to explain speech modifications by focusing on the language behaviors of interpersonal and intergroup interactions. The main argument is that language is a tool used by individuals to achieve a desired social distance between the speaker and his or her interlocutor, which speakers can manipulate. As a result, speakers manipulate their speech consciously to reach the desired social distance. There are four strategies individuals could employ, and which one a speaker chooses reflects their attitude toward their interlocutors. Convergence refers to the idea that when speakers want to have a closer social distance with their interlocutors, they would adapt, whether linguistically, paralinguistically, or non-verbally, to behave more similar to the interlocutor. If, on the other hand, individuals want to accentuate their differences from their interlocutor, they may make their speech sound dissimilar to that of interlocutors, which is a process known as divergence. Maintenance refers to the phenomenon when individuals decide to continue their own speech

without converging or diverging their speech despite interacting with an interlocutor. The last strategy is speech complementarity, which is a speech modification that accentuates valued sociolinguistic differences. For example, men would adjust their tones to make them sound more masculine when conversing with women and women take on more feminine sounding forms in mixed-sex conversations than in same-sex dyads (Shepard et al. 2001, p.35). This model therefore differentiates itself from the psycholinguistic models by arguing that speakers have a choice in how to modify features in their speech in order to manipulate their social distance with their interlocutors.

Studies of phonetic accommodation typically study two distinct varieties of a single language (Smiljanić & Bradlow 2011). Smiljanić and Bradlow (2011) extend the study of phonetic accommodation to second language (L2) accent and intelligibility. Their model addresses the interaction between intelligibility and accentedness and the role of phonetic accommodation in the interaction. The model assumes that the primary purpose of speech communicative, and therefore phonetic accommodation will be observed if speakers realize that doing so could enhance comprehension. The hypothesis associated with this theory is that because L2 learners are relatively inexperienced in noting the salient features in their L2, they should not gain as much as an advantage in improving intelligibility of different speaking style of speeches comparing to native talkers. The data the researchers obtained suggested that accentedness had no significant impact on intelligibility by native and non-native English speakers. Rather, it was the type of speech that mattered more for intelligibility. The intelligibility of non-native talkers' speech was enhanced when they heard clear speech; whether speakers had accents did not matter as much.

1.3 Previous Studies on Phonetic Accommodation

The automatic interactive alignment account and the Communicative Accommodation Theory have been corroborated by various studies using different methods and measurements. This section outlines studies that explore and further extend the models mentioned above.

A variety of studies further explored the automatic theories by performing non-interactive production studies (Nielsen 2008; Babel 2009; Babel & Bulatov 2012). Nielsen (2008) investigated how allophonic variations were imitated by manipulating VOT and found evidence that both prove and disprove automaticity of accommodation. In her study, she lengthened VOT in one task and reduced VOT in another task to understand whether phonetic accommodation would be observed the same way. The lengthened VOT task suggested that there was accommodation and generalizability at an abstract feature level, but the shortened VOT task showed that imitation was not automatic because convergence was not observed. The lack of accommodation could be due to the fact that shortening the VOT of the voiceless consonant /p/ could endanger the contrast between voiced and voiceless stop consonants. Thus, Nielsen's experiment demonstrated that accommodation was observed in phonemes and could be determined by multiple factors, such as contrast in phonemes. Babel (2009) focused on vowel imitation, selecting words with the five vowels /i/ /æ/ /ɑ/ /o/ /u/ as stimuli for participants to repeat in a shadowing task. Although convergence in formant values was found in all words, the

two low vowels /ɑ/ and /æ/ demonstrated stronger imitation effects than non-low vowels in the F1 dimension, which she speculated could be due jaw movements as a result of the height differences. In another study, Babel and Bulatov (2012) examined whether fundamental frequency (f0) was a feature that would be imitated under non-interactive production studies. The researchers set up a single-word shadowing task, in which one group of subjects were presented with unfiltered speech while another group of subjects were exposed to the same speech that had been filtered below 300Hz so f0 was eliminated. An acoustic analysis of subjects' f0 revealed that subjects exposed to the unfiltered speech accommodated their f0 while the subjects in the other group did not. This study thus indicates that f0 is a feature that could be accommodated.

The social theory of phonetic accommodation is also well-studied and well-supported. Production experiments have found that speakers' implicit social attitude can determine the extent of convergence. Natale (1975) measured the mean vocal intensity of speakers that participated in interview tasks and found that speakers in general changed correspondingly to match the interviewer's vocal level and that the degree of convergence was positively associated with the social desirability of the speaker. Similarly, Babel (2012) included visual stimuli in the experiment and obtained results that were in favor of the CAT. In a shadowing task, Babel presented either a Caucasian or an African American's face on the screen while exposing subjects to audio stimuli. By analyzing the data together with the scores from the IAT (implicit association task) completed by participants prior to the experiment, Babel (2012) concluded that female participants converged more to the model talker's vowels when they rated the talker as more attractive whereas male participants' vowels diverged from those of more attractive talkers.

Some other studies inspected the possibility that speech imitation could be not only influenced by speakers' attitudes toward interlocutors but also due to other social factors such as gender and the role a speaker plays in a conversation (Pardo et al. 2012; Pardo et al. 2010; Namy et al. 2002; Bilous and Krauss 1988). Pardo et al. (2012) conducted a phonetic accommodation experiment among college roommates where pairs of male roommates were recorded four times over the course of a year and found that phonetic convergence was related to the strength of the interlocutors' relationship. The stronger the self-reported relationship was, the more accommodation was observed. In another experiment, Pardo et al. (2010) recruited twelve pairs of unacquainted talkers to complete a conversational task and instructed receivers in the pairs to imitate the talkers' speech. By analyzing perceptual measures of phonetic convergence from the AXB Perceptual tests as well as acoustic measurements, including speech rates and vowel formants, the researchers concluded that the sex of the pair of talkers and well as the talker's role influenced the degree of convergence. In particular, male converged more than female talkers did, and talkers in the conversation maintained a faster speech rate than receivers (Pardo et al. 2010). However, Namy et al. (2002) examined phonetic accommodation in laboratory settings by speakers of different gender and obtained data that suggested otherwise. Their data revealed that during shadowing tasks, female speakers accommodated more than males and the degree of convergence was stronger when the model speaker was male. Bilous and Krauss(1988) investigated subjects' behaviors such as speech productivity and laughter from conversations in held among same-gender and mixed-gender dyads, and showed that males and females demonstrated different patterns. Specifically, females diverged on frequency of back-channels and laughter while males converged on the two variables. The researchers proposed that social

power could play a role in determining the outcome of convergence. Subjects that were of the less powerful status converged more toward their superior gender.

As this section shows, ample studies demonstrate that phonetic accommodation is a phenomenon that can be found at a segmental and phonetic feature level. Also, there is not one feature that one can use to be able to predict the outcome of phonetic accommodation. Factors such as conditions, gender, and social attitudes can contribute to the degree of phonetic accommodation. The next section will focus on the studies done to explore the relationship between second language learning and phonetic accommodation.

1.4 L2 Learning and Phonetic Accommodation

Second language acquisition, or L2 learning, refers to people learn a new language after a native language has been established. In the process, learners may need to acquire aspects in syntax and phonology that are different from their native language. When learners need to acquire new sounds that do not exist in their native phonetic inventory, they confront problems with producing new phones or contrasting two phonemes due to the lack of an analogous phonemic distinction in their native language. The difference in phonologies can result in problems with intelligibility, as distinctions that should be made, either in vowels or consonants, are not present. Understanding how L2 speakers accommodate to a model speaker of the same or different language background thus can shed light on to the motivations or causes for phonetic accommodation. Researchers have started to not only study how native speakers of English modify their speech during conversations but also investigate how L2 learners cope with the exposure of an interlocutor's speech.

Some researchers have investigated the accommodation observed among speakers of different language background. Kim (2012) tested the impact linguistic distance between interlocutors has on the extent of phonetic accommodation on monosyllabic words, disyllabic words, and sentences. Monolingual native American-English talkers were recruited as subjects and were assigned to one of the three conditions: Northern dialect of American English, non-Northern dialect, and speech produced by female fluent Korean speakers of English. The results suggested that dialect mismatch and L1 mismatch between participants and the model talker in the passive auditory exposure did not inhibit phonetic convergence; participants converged toward the model speakers regardless of their linguistic distance and negative attitudes toward foreigners. Also, there was generalizability in subjects' speech, especially in the monosyllabic and disyllabic word in participants' VOT.

However, other studies show that convergence by native and nonnative English speakers to nonnative speech of English is not automatic. Kim et al. (2011) performed a XAB perception test in which listeners judged whether a token "X" was more similar to one of the tokens (A or B) produced by speakers that were put were in different-L1 or different-dialect pairs. The results showed greater convergence for same-dialect pairs, which suggested that there is a relationship between language distance and phonetic accommodation. The researchers posited that the desire to achieve a higher level of intelligibility may encourage convergence while attention demand and processing load involved in nonnative speech production may inhibit alignment in speech.

To summarize, this study showed that linguistic distance plays a role in phonetic accommodation, contrary to Kim (2012).

In accordance with Kim et al. (2011)'s study, Hwang et al. (2015) also supported the non-automatic theory of phonetic accommodation, suggesting that phonetic accommodation is not obligatory but is motivated and shaped by pragmatic needs. The researchers recruited Korean speakers of English who conversed with a monolingual American English talker and performed a spontaneous imitation task. They investigated the vowel /æ/ and coda consonant /b/, as these segments do not exist in Korean, and Korean speakers tend to produce them as [e] and [p], respectively. During the spontaneous production task, Korean speakers only produced more English-like phonetic segments after they had just heard the segment produced by the English speaking American interlocutor but the segments were not produced the same in the post-experiment task. However, when participants were required to produce words that were minimal pairs and only contrasted in the target segments, their production became more English-like. Thus, the results obtained from these two experiments point to the fact that non-native speakers of English can accurately produce English phonemes when it enhances intelligibility.

Similarly, Zajac (2013) also observed phonetic accommodation in L2 speech, but she proposed that the degree of convergence might be determined by speakers' attitude toward accent L2 speech. Zajac recruited Polish learners of English to perform an auditory naming task, in which they heard either the speech of a native English speaker or a native Polish speaker fluent in English. The results showed that participants imitated their vowel durations in both conditions. However, in the non-native speaker condition, subjects did not converge to the model vowel duration. Zajac proposed that this was because participants in the non-native condition realized that the speech was from a non-native speaker of English so they did not converge their speech so as to distance themselves from foreign-accented English. In other words, the researcher suggests that the phenomenon of converging and diverging in her data could be explained by the CAT.

There have also been studies that investigate phonetic accommodation in target languages that are not English. Smith (2007) analyzed the speech of ten native speakers of French in spontaneous conversations when conversing either with a native or a non-native speaker of French. Subjects in the non-native condition adjusted their f0 to a bigger range and segmental modification to use a more emphatic style but did not adjust their speech rate or duration. This study indicates that native speakers selectively implement accommodation to non-native speakers, which may not be the most beneficial to non-native speakers.

1.5 The Present Study: English and Mandarin

The present study chooses Mandarin speakers of English as participants because there are different vowels and consonants in Mandarin and English. Because some segments in English do not exist in Mandarin speakers' native phonemic category, Mandarin speakers may fail to produce segments in English with enough contrast. As a result, their speech may sound hard to understand, or be labeled as having an accent.

The features that differ between English and Mandarin mainly occur in tenseness of vowels and voicing in consonants as well as the presence or absence of consonant clusters. In terms of vowels, there is no /ɪ/ and /ʊ/ in Mandarin. As a result, speakers do not contrast between /i/ and /ɪ/ and /u/ and /ʊ/ and often confuse with "eat" and "it" or "fool" and "full" (Chang 1987, p.311). /æ/ is another vowel that does not exist in Mandarin and therefore learners may confuse this vowel with other vowels such as /e/, /ɑ/, or /ʌ/.

For final consonants, Mandarin has nasal consonants only. Since this is the case, Mandarin speakers may feel unnatural to produce stop consonants /b/, /d/, and /g/ and would either omit the consonant or insert a vowel after the final stop consonant to make an extra syllable. Word-initial and word-final consonant clusters can also be a problem for Mandarin learners of English since Mandarin does not have this feature. The adjustment made by learners is to insert a vowel in between the consonants to make an additional syllable.

The fact that English and Mandarin differ in several segmental features enables us to investigate the change in features in the English speech of Mandarin speakers. This study will thus investigate how vowels and consonants change in the English spoken by Mandarin L1 speakers after hearing the English of a model talker with their same language background in an attempt to understand whether accommodation could be observed in L2 speech and the role model speaker's speech plays in determining the outcome of subjects' productions.

2 The Experiment

2.1 Questions and Hypothesis

Although a copious amount of studies have addressed the issue of phonetic accommodation under different conditions and speakers from various language backgrounds, there have only been a few studies that explore the accommodation observed in speakers whose first language is not English after presenting them with English speech of the speaker's same background. Based on the CAT and previous literature on phonetic accommodation to nonnative speech, the current experiment asks whether second language learners of English would converge their English phonetically to interlocutors who share mutual language background under different social conditions. By manipulating social conditions, the experiment tests if social attitudes could determine the likeliness and extent of phonetic accommodation. To test this question, the study sets up a production experiment with two conditions, a condition under which participants should be prone to accommodate to the interlocutor and another condition in which participants should be less likely to converge. The prediction of the automatic alignment model is that all subjects converge to the interlocutor's speech but the extent of phonetic accommodation will depend on the conditions they are in, as predicted by the CAT. Namely, speakers will converge more if they feel psychologically close with the interlocutor.

2.2 Experimental Design

2.2.1 Experimental Procedure

This experiment followed the procedure of previous phonetic accommodation experiment, which was consisted of three tasks: a pre-test, a shadowing task, and a post-shadowing task (Nielsen 2008; Babel 2009; Babel & Bulatov 2012). A baseline task was used to test the baseline of participants' speech. The shadowing task, during which participants were asked to repeat words after hearing productions by the model speaker, exposes subjects to the condition. The post-shadowing task was used to gather information about participants' speech and observed any difference in their speech from their production from the baseline. The three tasks added up to about thirty minutes. The experiment was built in Opensesame and was run on the Berkeley Phonetics Machine.

To simulate subjects' attitudes under different situations, participants were assigned to one of two conditions in the shadowing task. In the "professor" condition, designed to encourage speakers to accommodate, subjects were introduced to the model speaker as a professor from Asia, with whom the subject would want to familiarize herself. The setting was described as an office hour meeting, where the professor was pronouncing words that subjects were asked to repeat. In the "employer" condition, subjects were introduced to the same model speaker as an employer from a company the student was applying to work for. The qualifications the employer was looking for included having resided in the U.S. for an extensive period of time and understanding the American culture well. During the shadowing task as well as the post-shadow task, a photo of a woman dressed according to the condition was presented on the screen. A woman wearing a robe and writing on a chalkboard was presented for the "professor" condition and a woman in business attire was used in the "employer condition". The post-shadow task was introduced as a presentation task, in which subjects were asked to present the words to their interlocutor. During the post-shadow task, the picture of the woman remained on the screen while the words were shown.

Participants were invited to Dwinelle 53, in the UC Berkeley PhonLab, where they sat in a sound-attenuated booth and read instructions for the task on a computer screen. In order to better motivate subjects to imagine the setting, a questionnaire asking about physical characteristics of the lady shown on the screen was presented, and we told them that they would be required to answer the questions at the end of the experiment.

Participants were asked to complete a language background survey prior to the experiment and a post-experiment survey after the experiment had been done. The language background survey made sure that the participants fulfilled the experiment requirements and the post-experiment survey consisted of questions concerning participants' opinions on the speaker's accentedness, their familiarity with the words used in the experiment, and the questions that were shown to the participants before the shadowing task.

2.2.2 Participant Population

The participant population is L1-Mandarin L2-English speakers. To eliminate gender as a confounding factor, this experiment only examined the speech of female participants. Under these criteria, female Mandarin speaking international students studying at UC Berkeley were recruited. Also, for proficiency uniformity, participants were screened to make sure that they had been residing in the United States for at least a year.

Twenty subjects in total were recruited and were randomly assigned to one of the two conditions. Subjects had an average age of 21.4 years and on average had been living in the United States for 3.7 years. All participants were recruited through invitations and received $10 as compensation for their time.

2.2.3 Stimuli
2.2.3.1 Model Speaker

The model speaker is staff at UC Berkeley and had resided in China for 23 years before moving to the United States 30 years ago. She speaks Mandarin as her first language and started learning English at the age of 20. The productions were recorded with an AKG C3000 microphone at a sampling rate of 22050Hz. The recording session took place in the UC Berkeley PhonLab, in Dwinelle 53. The model speaker read two word lists, which had a total of 180 words, twice with each word presented individually on the screen. The recording of one of the word list that sounded clearer was selected as the stimuli to be played in the experiment.

2.2.3.2 Word Lists

Two word lists, each of which consisted of 90 words, were used in the three tasks that made up of the experiment. One word list consisted of 47 words and 43 filler words while the other word list had 46 target words and 44 filler words. All words were either monosyllabic or disyllabic. The first word list was used in the baseline task and the second list was used in the shadowing task. The word list for the third task was composed of all the target words from the two word lists. This way, we can not only investigate whether the productions were different in the shadowing task and in the post-shadow tasks but also whether the accommodation effect was generalized to words that subjects had not heard the model speaker produce. In other words, the word list helps to see whether there was generalizability of the features that speakers were exposed to. The tasks and word list correspondence is shown below (Table 1).

Task	Task 1 (Baseline)	Task 2 (Shadowing)	Task 3 (Post-shadow)
Wordlist	1 (47 target words, 43 fillers)	2 (45 target words, 45 fillers)	1 (47 target words) & 2 (45 target words)

Table 1: Word Lists and Tasks Correspondence

The same words were presented in the two conditions and all word lists were repeated twice for each task The words were generated using a Perl interface, which allowed the researcher to input segments and gave all the words that contained the segments in the SUBTLEX lexicon. The researcher made the final decisions of which words to be included.

2.2.4 Data Measurement

This experiment targeted the differences between English and Mandarin. For vowels, the experiment investigated the English monophthongs [i] [ɪ] [ɛ] [æ] [ɑ] [ʌ] [ʊ] [u] and diphthongs [ei] [ai]. These vowels include different height and frontness, which can shed light on to the different extent of accommodation determined by phones. As Babel's (2009) data shows, low vowels are more likely to accommodate than non-low vowels. Another reason for including both tense and lax vowels is to investigate whether different patterns can be observed from vowels contrast in tenseness, as Mandarin does not have lax vowels in its inventory. Formants (F1 and F2) of these vowels were measured to see if speakers modified their productions of these vowels upon hearing the model speech and whether, as previous research suggested, some vowels were accommodated more than others. The duration of vowels, from the beginning of formants to where F2 ends, was also measured.

Durations of consonants in consonant clusters and word-final stop consonants were also measured, since Mandarin lacks these features. Word-initial consonant clusters [sp] [sk] [st] and word-final consonant clusters [st] [pt] [kt] were included. Durations, from the beginning of the closure to the end of release, of word-final stop consonants (i.e [b] [p] [d] [t] [g] [k]) were also included.

2.2.5 Data Analysis

Measurements were taken from all the target words. Before extracting quantitative measures, the recordings were forced aligned in the Berkeley Phonetics Machine using the aligner developed by the University of Pennsylvania Linguistics Lab (M.L Yuan 2008). The forced-aligned textgrids were hand-corrected so that vowels were marked with the beginning of formants as onset and where F2 ends as offset. The boundaries of consonants were marked at the beginning of closure as onset and the end of release as offset. Two Praat scripts took measurements from the acoustics and the aligned textgrid files. One script measured formant values of the vowels at vowels' midpoints, while the other script obtained the duration of the segments.

The post-shadow task was divided into two tasks and investigated separately. This was due to the word list design, where half of the wordlist was drawn from the wordlist used in the shadowing task and the other half was from the baseline task. Measuring words that subjects were directly exposed to and not exposed to should reflect whether accommodation was generalized in the post-shadow task. The rest of the paper will refer to words that appear in post-shadow task and were also presented in the shadow task as post-shadow(a) task and words that subjects were not exposed to the audio stimuli and were identical to the baseline task as post-shadow(b) task.

One of the objectives of the experiment is to evaluate how a participant's productions change after being exposed to the speech of the model talker. Because there are two values corresponding to each vowel (F1 and F2), Euclidean distance (Equation 1) was used to compare each participant production to the model talker. The equation for finding the Euclidean distance is shown below.

$$\sqrt{(subject\ F1 - model\ speaker\ F1)^2 + (subject\ F2 - model\ speaker\ F2)^2}$$

Equation 1: the Euclidean distance model to calculate distance between F1 and F2 values of subjects and model speaker

The data for durations was normalized prior to any investigation. All the duration values were put onto a log scale and values over two standard deviations were eliminated, to make the data more normal. The log-duration then was used in data analysis.

3. Results

Since the study asks whether second language learners of English would converge their English phonetically to interlocutors who share mutual language background under different social conditions, the expected result is that there is phonetic accommodation in shadowing and post-shadow task and the amount of accommodation is determined by the condition, in which subjects in the "professor" condition would show more accommodation than if they were in the "employer" condition.

Subjects were asked to complete a survey once they had finished the experiment. The survey showed that on a scale of 1 to 5, where 1 was strongly disagree and 5 was strongly agree, subjects in general were able to understand the productions of the model speaker (average 3.8/5) and did not find the model speech as heavily accented (3.1/5). They also rated the words in the experiment as words that they would use often (4.0/5).

3.1 Vowels

3.1.1 Formant
Monophthongs [i], [ɪ], [ɛ], [æ], [ɑ], [ʌ], [ʊ], [u] and diphthongs [ai] and [ei] were analyzed. F1 and F2 were extracted and transformed before distance was calculated before comparing them with the model speakers' formant values by using a Euclidean distance model. An Anova test of a mixed effect model, which controls subjects as random factors (Table 2), suggests that the difference across phones (p-value < 0.001) and tasks (p-value < 0.001) is statistically significant. The interaction between phone and task is also significant (p-value = 0.009). Condition and the interaction between condition and task are not statistically significant.

	Chisq	Df	Pr(>Chisq)
Phone	96.446	9	< 2.2e-16 ***
Condition	1.932	1	0.166
Task	28.472	3	2.892e-06 ***
Phone:Condition	9.373	9	0.404
Phone:Task	47.089	27	0.009 **
Condition:Task	1.614	3	0.656

Phone:Condition:Task   17.109   27   0.928

Table 2: Anova result of mixed effects for predicting formant difference by including phone, condition, task as fixed effects and controlling subjects as a random effect

To understand the interaction between phone and task, Figure 1 is plotted with difference between subjects' productions and model value on the y-axis. A F1F2Distance of 0 means that the subjects' productions are identical to the model value.
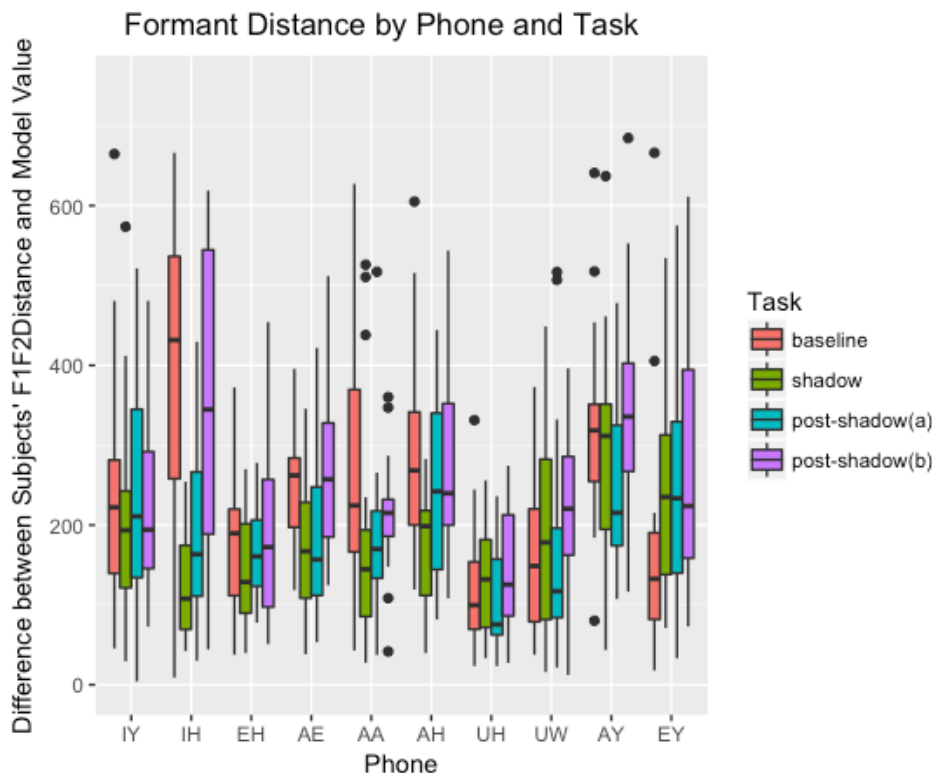

Figure 1: Formant difference between subjects and model speaker categorized by phone and task

The plot shows that the difference between subject's productions and model productions across phones varies greatly. Despite this variation, most phones show the same trend, in which the distance between participants and the model talker decreases greatly in the shadowing task. The pattern is strongest in low vowels [æ] (AE) and [ɑ] (AA), as well as the high vowel [ɪ] (IH). Vowels [ʊ] (UH), [u] (UW), and [ei] (EY) demonstrate the opposite pattern, where the distance increases in the shadowing task.

Since [ɪ] demonstrates an unusually obvious accommodation, a comparison with its tenseness counterpart, the high front vowel [i], is plotted (**Figure 2**) in order to understand the reason behind the significant accommodation in the shadowing task. The boxplot shows the averages of productions of [ɪ] and [i], with the error bar showing the standard error from the averages.
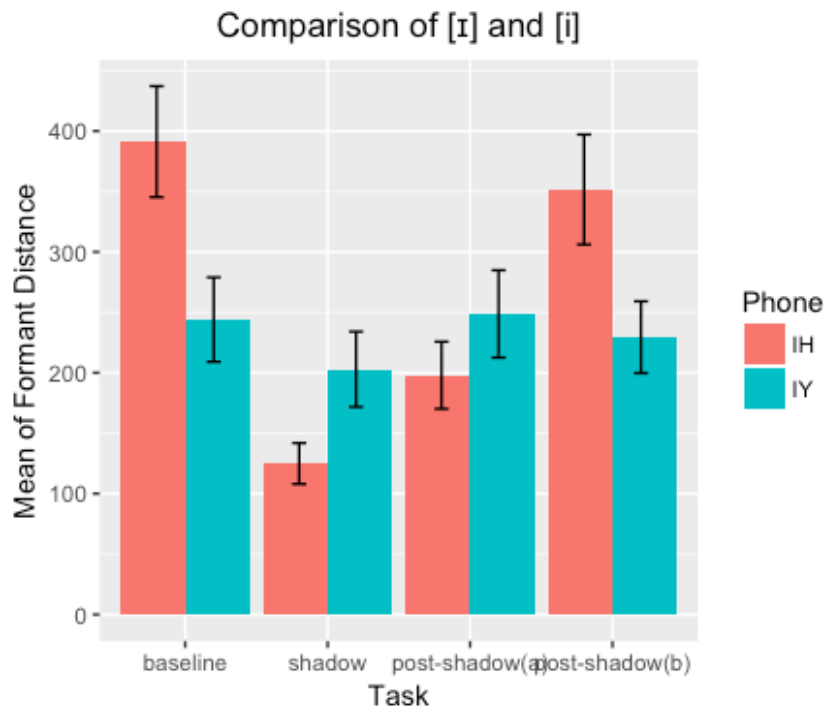
Figure 2: Comparison of [i] and [ɪ] by plotting mean of formant difference on the y-axis and standard error as error bars

The figure reveals that the two vowels demonstrate different trends and [ɪ] has more variability. The formant distance between subjects' productions and model speaker remains relatively the same throughout the tasks for [i], which is in accordance with Babel's result that non-low vowels accommodate less. The trend for [ɪ] is interesting in that there is much variability, shown by the error bar, in productions in the baseline and post-shadow(b) task and the distance from model speaker is farther away. One possible explanation of the variability is that since [ɪ] is not a phone that exists in Mandarin, subjects produce the vowel differently: some produce it the same as [i] while others notice the difference and produce more [ɪ] like. The distance and variability decreases significantly in the shadowing task, suggesting that subjects are willing to adopt a new way of pronouncing the vowel since they do not have a set pronunciation. Thus, the accommodation seen in [ɪ] could be due to a second language acquisition issue, in which speakers are not sure about how to produce a phone that does not exist in their native phonetic inventory.

Followed by the analysis of formant distance categorized by phone and task is an investigation of conditions. Wilcoxen tests were conducted in order to determine whether the two conditions demonstrate different trends in their population mean ranks across tasks.
These tests reveal that in the "professor" condition, there is significant difference between baseline productions and post-shadow(b) productions (p-value = 0.046) as well as between post-shadow(a) and post-shadow(b) task (p-value = 0.03). The productions of words in shadow task are not significantly different from those in the post-shadow(a) task. This result suggests that subjects modified their speech to match the model speech and that the effect remained in the later task.

The tests for productions in the "employer" condition reveal similar results. The difference between baseline and shadow task (p-value = 0.001) and between the two word lists in the post-shadow task (p-value = 0.002) is significant. There was no significant difference between productions in post-shadow(a) and post-shadow(b). The difference between baseline and shadowing tasks suggests that subjects accommodated their formant values to that of the model speaker, while the non-significant difference between the shadowing and post-shadow(a) tasks suggest that they were able to maintain the speech even when they were not immediately repeating after the speech. However, the difference between the two post-shadow tasks reveal that subjects do not generalize the changes they make in the vowels in the words that they hear to the words that they have not heard. The test results are summarized in Table 3 below.

	"professor" condition	"employer" condition
Baseline vs. Shadow	0.046**	0.001***
Baseline vs. Post-shadow(b)	0.016**	0.954
Shadow vs. Post-shadow(a)	0.199	0.459
Post-shadow(a) vs. Post-shadow(b)	0.03 **	0.002***

Table 3: Summary of Wilcoxen test results by condition and task

From the test results, it can be seen that the only difference between the two conditions is at the comparison between baseline and post-shadow(b), where productions in the "professor" condition is statistically significant different while those in the "employer" condition are not. In order to understand how different productions in post-shadow(b) task are, the averages of all the productions categorized by tasks and conditions are plotted (Figure 3). The graph can be interpreted the same way as the previous one, where the distance between subjects' productions and model value is shown and a zero on the y-axis indicates no difference between the productions.
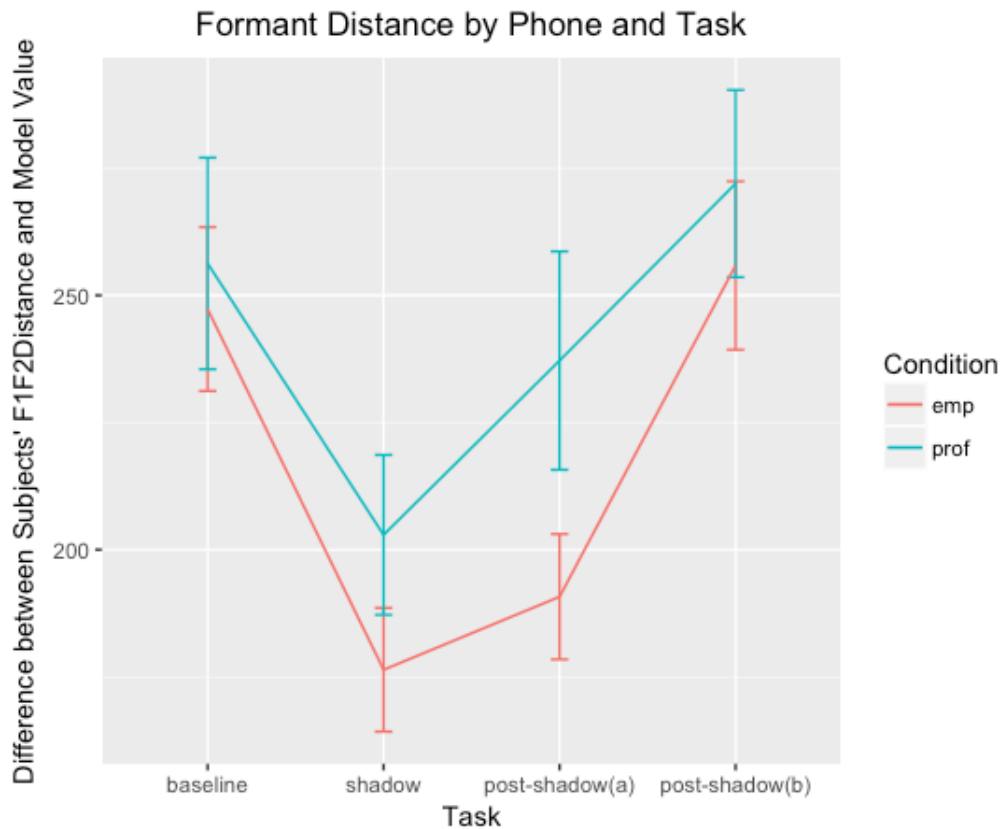
Figure 3: Formant difference between productions of subjects and model speaker categorized by condition and phone with error bars showing standard error

The visualization shows that the two conditions have the same trend. Both conditions have a significant decrease in distance in shadow task and increasing distance in the later post-shadow tasks. From looking at this plot, the difference between the baseline and post-shadow(b) tasks in two conditions can be explained. In both conditions, the means of productions diverge from the model value. The significant difference found between productions in baseline and post-shadow(b) task in the "professor" condition is due to divergence, where the values in post-shadow(b) are further away than those in the baseline task. Thus, the difference observed in post-shadow(b) task in the "professor" condition is in fact due to the divergence from the model value, but not generalizability from post-shadow(a).

3.1.2 Duration

The duration of vowels were also analyzed. An Anova test of the mixed effect model (Table 4) gives statistical significance to the visualized trends. The result shows that there are significant effects of task (p-value < 0.001) and phone (p-value < 0.001). Difference across conditions is not significant but the interaction between condition and task (p-value = 0.009) is. The interaction between task and phone (p-value < 0.001) is also significant.

	Chisq	Df	Pr(>Chisq)
Condition	1.630	1	0.202
Task	256.116	3	< 2.2e-16 ***
Phone	574.744	9	< 2.2e-16 ***
Condition:Task	11.443	3	0.010 **
Condition:Phone	6.837	9	0.654
Task:Phone	244.136	27	< 2.2e-16 ***
Condition:Task:Phone	11.675	27	0.995

Table 4: Mixed effects for predicting difference between vowel duration of subjects and model speaker

Figure 4 shows the differences between subject and model log duration categorized by phone and task. A zero on the y-axis indicates that there is no difference between the values of subject and model speaker. A negative value means that subject values are shorter than the model value while a positive value means subject values are longer than the model value.
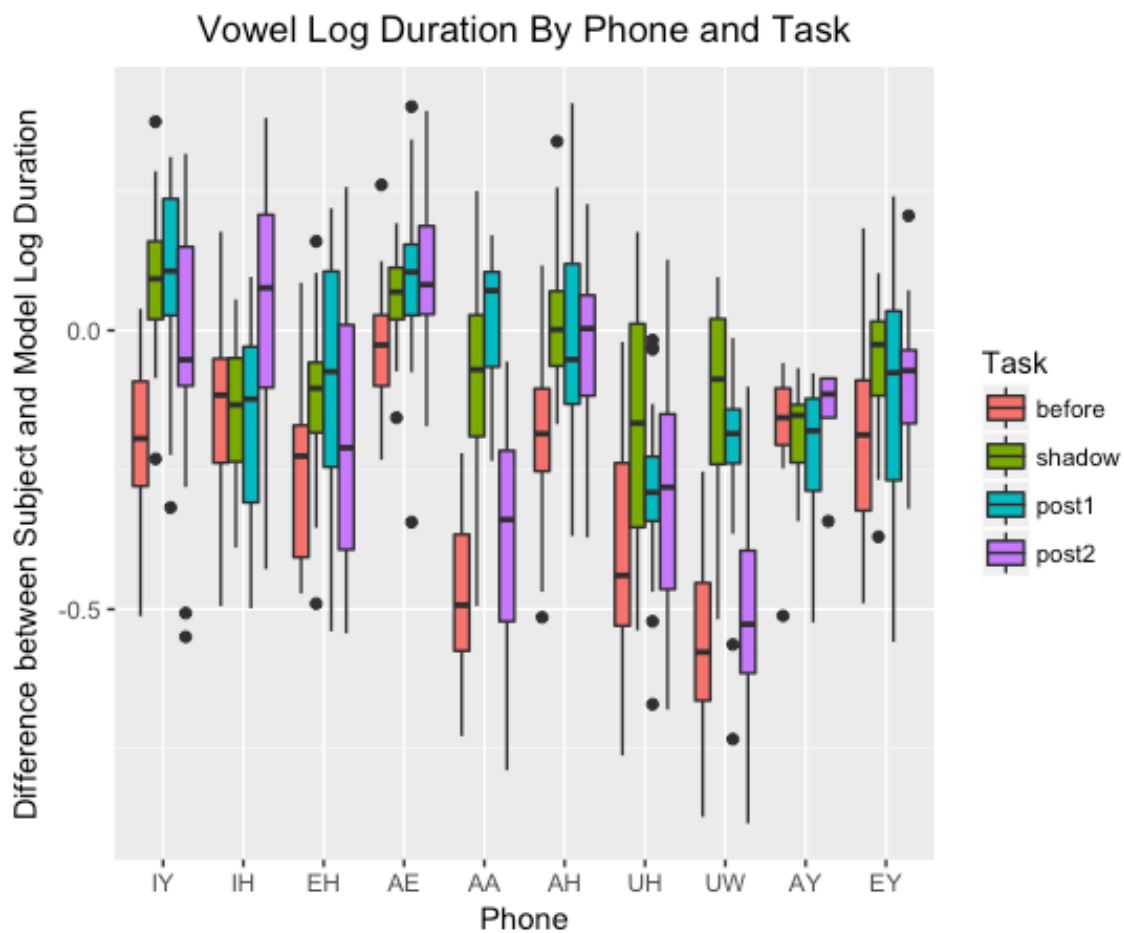

Figure 4: Difference between vowel log duration of subjects and model speaker categorized by phone and task

There is in general a big difference between productions in the baseline task and shadowing task and the plot confirms the medians are different across phones, and the amount of convergence differs by vowel. It seems that there is more convergence in the shadowing task if the baseline productions are farther away from the model value, as exemplified by [ɑ], [u], and [ʊ]. Vowels that have values already close to the model production, such as [æ] and [ai] have less convergence. The productions in post-shadow(a) and post-shadow(b) vary greatly but overall productions in post-shadow(b) seem to return to the baseline productions.

Difference across conditions and tasks are compared in order to understand whether condition makes a difference in productions. Figure 5 shows that there are differences between the baseline task, post-shadow(a) task, and post-shadow(b) task. The medians of productions in shadowing tasks in both conditions are similar. In the "employer" condition, accommodation is observed in the shadowing task and the difference between subject and model log duration increases in post-shadow(a) and falls back more in post-shadow(b) task. In the "professor" condition, the accommodation is observed in the shadow task and remains at about the same level in both post-shadow tasks. The different trends are in line with the hypothesis that subjects in the "employer" condition would want to maintain their speech while subjects in the "professor" condition are prone to accommodating to their interlocutor in order to facilitate social distance.
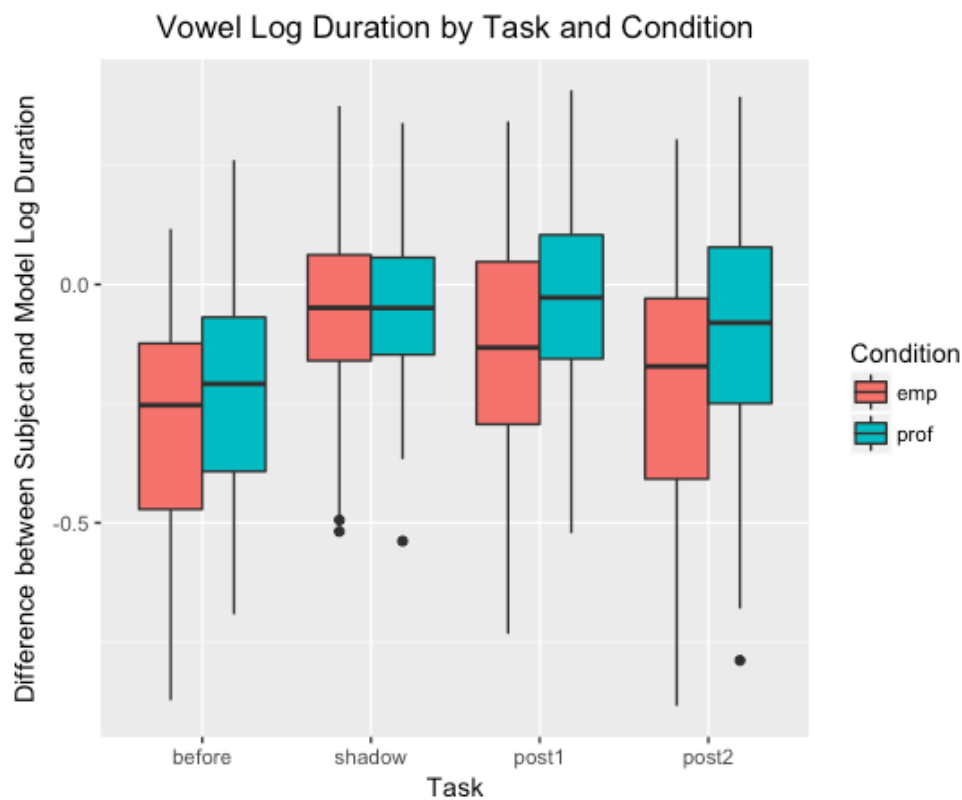

Figure 5: Difference between subject and model log duration categorized by condition and task

3.2 Consonants

3.2.1 Word-final Stops Duration

Durations of six word-final stop consonants ([b], [p], [d], [t], [g], [k]) were analyzed. An Anova test of a mixed effect model (Table 5) demonstrates there are significant differences across tasks (p-value < 0.001) and phones (p-value < 0.001). Condition has no significant effect whether alone or in interaction with other dependent variables.

	Chisq	Df	Pr(>Chisq)
Condition	3.343	1	0.068
Task	48.083	3	2.044e-10 ***
Phone	141.652	5	< 2.2e-16 ***
Condition:Task	3.448	3	0,328
Condition:Phone	2.820	5	0.728
Task:Phone	15.236	15	0.435
Condition:Task:Phone	8.189	15	0.916

Table 5: Mixed effects for predicting difference between word-final stop duration of subjects and model speaker

Figure 6 below shows the difference between normalized durations of subjects' consonant durations and the model speaker's, by phone and tasks. Similar to vowel duration, there is an obvious difference between productions in the baseline and shadowing task. The difference in the post-shadow tasks in general increases and the values move toward baseline productions.
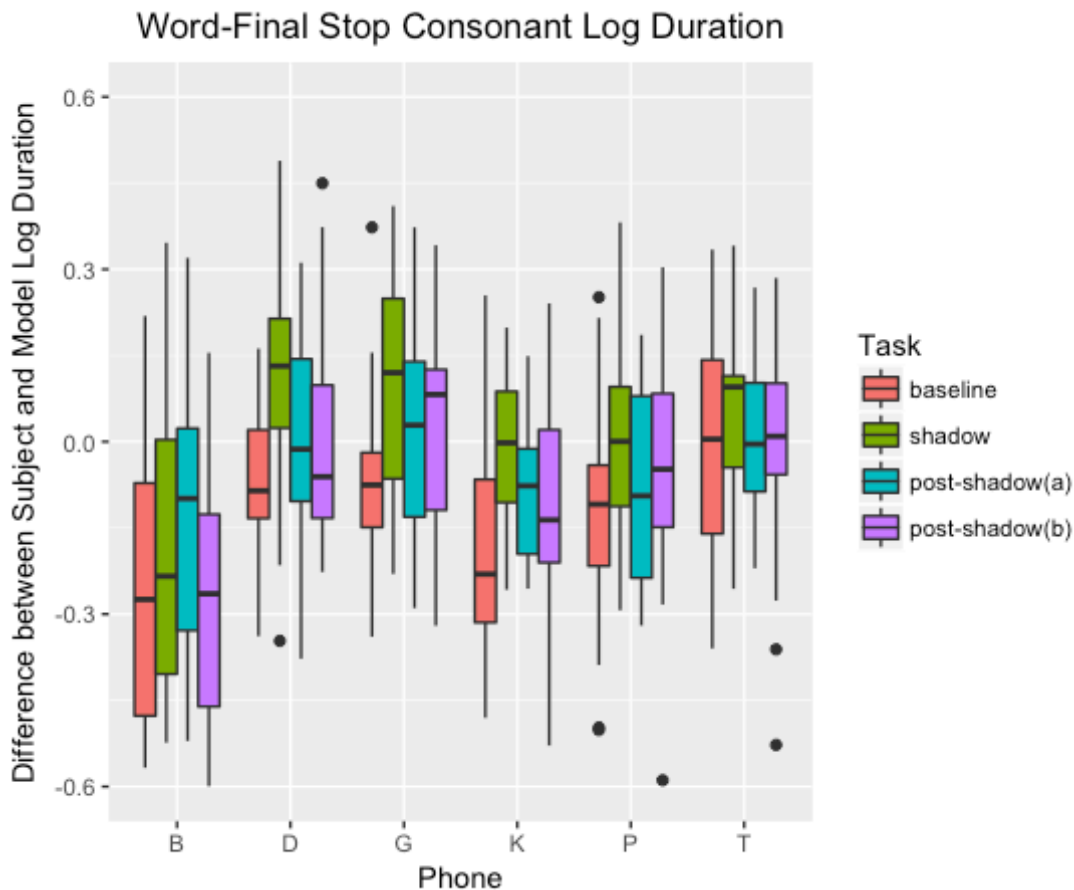
Figure 6: Difference between word-final stop log duration of subjects and model speaker categorized by phone and task

As one of the hypothesis concerns the role condition plays in the outcome of predicting values across tasks, Figure 7 compares medians categorized by conditions over tasks. The plot confirms the result from the Anova test that there are no significant differences across conditions. There is a consistent difference between medians in two conditions across tasks and both conditions have the same trend.
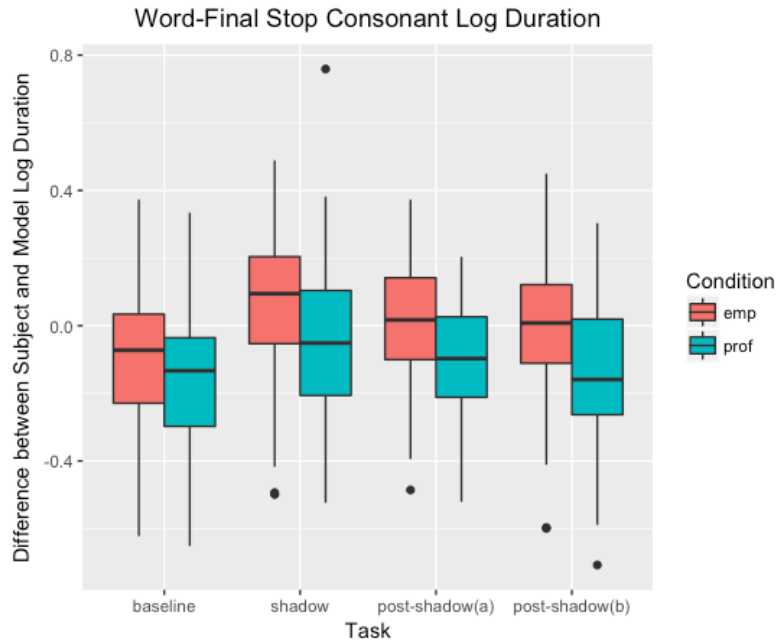
Figure 7: Difference between word-final stop log duration of subjects and model speaker categorized by condition and task

3.2.2 Word-Initial Consonant Clusters Duration

Each segment in the word-initial consonant clusters ([sp], [st], [sk]) was analyzed. An Anova test of a mixed effect model (Table 6) suggests that differences across tasks (p-value < 0.001) and phones (p-value < 0.001) are significant. The difference across conditions is not significant but the interaction between condition and phone is (p-value < 0.001).

	Chisq	Df	Pr(>Chisq)
Condition	0.311	1	0.578
Task	130.081	3	< 2.2e-16 ***
Phone	1149.927	5	< 2.2e-16 ***
Condition:Task	5.213	3	0.157
Condition:Phone	24.678	5	0.000***
Task:Phone	5.488	15	0.987
Condition:Task:Phone	11.829	15	0.692

Table 6: Mixed effects for difference between word-initial consonant cluster log duration of subjects and model speaker

Figure 8 shows the difference between the log duration of consonant clusters spoken by subjects against those of the model speaker, separated by phone and task. The fricatives are grouped to the left and the stops are shown in the right of the plot.
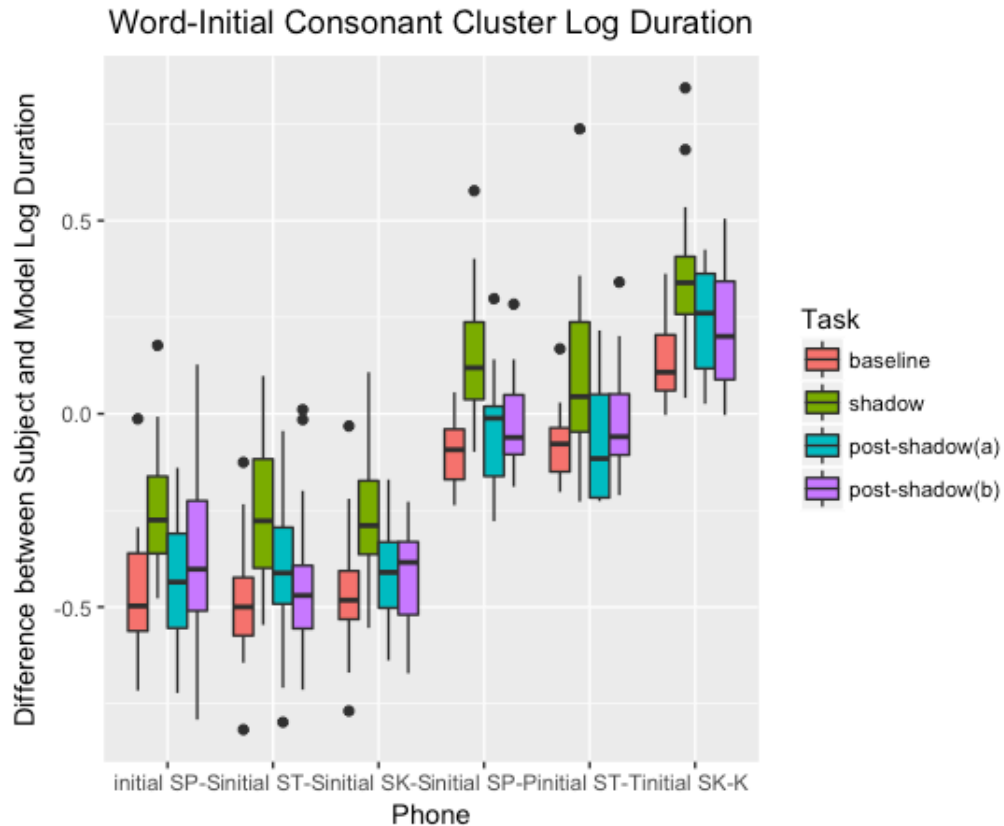

Figure 8: Difference between word-initial consonant cluster log duration of subjects and model speaker categorized by task and phone

The differences across phones can be seen in Figure 8, which shows that phones with different manner of articulation have different medians. Namely, fricatives in the subjects' speech were a lot shorter than the model's, while subjects' stop productions were close to the model value. Despite the difference, the trends across tasks are in line with the previous results. The productions in the shadowing task are closer to zero than the baseline productions, demonstrating that the subjects modified their speech after hearing the model speech. It is possible that subjects might have treated the consonant clusters as a single unit instead of two separate segments. If this were the case, the consonant cluster would demonstrate patterns that are similar to what is observed from word-final stop consonants and the accommodation pattern would be observed more clearly.

Figure 9 shows differences between cluster durations across tasks and conditions. On top of observing the same result as the previous graph, where accommodation is observed in the shadowing task and productions in post-shadow tasks become more different from the model

value, the graph also helps to clarify the patterns across conditions. The productions in the two conditions have the same trend and demonstrate no significant difference in medians.
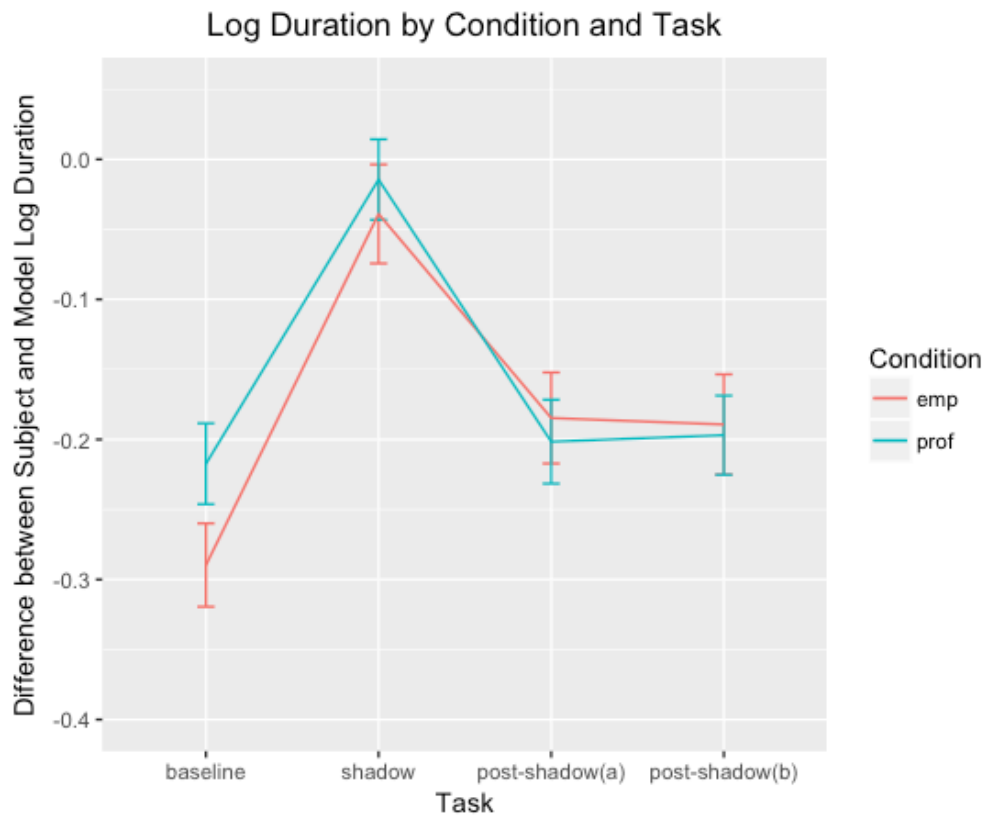

Figure 9: Averages of difference between the word-initial consonant cluster duration of subjects and model speaker categorized by condition and task with error bars showing the standard error

3.2.3 Word-Final Consonant Clusters Duration

The duration of each segment in the word-final consonant clusters ([pt], [st], [kt]) was measured. An Anova test of a mixed effect model (Table 7) suggests that distances across tasks and phones are significant (p-value < 0.001). Condition is not significant, but the interaction between condition and task (p-value = 0.004) and condition and phone (p-value = 0.001) is.

	Chisq	Df	Pr(>Chisq)
Condition	1.398	1	0.237
Task	54.738	3	7.808e-12 ***
Phone	288.792	5	< 2.2e-16 ***
Condition:Task	13.459	3	0.004 **
Condition:Phone	19.595	5	0.001 **
Task:Phone	20.467	15	0.155

Condition:Task:Phone   13.956   15   0.529

Table 7: Mixed effects for predicting difference between word-final consonant cluster log duration of subjects and model speaker
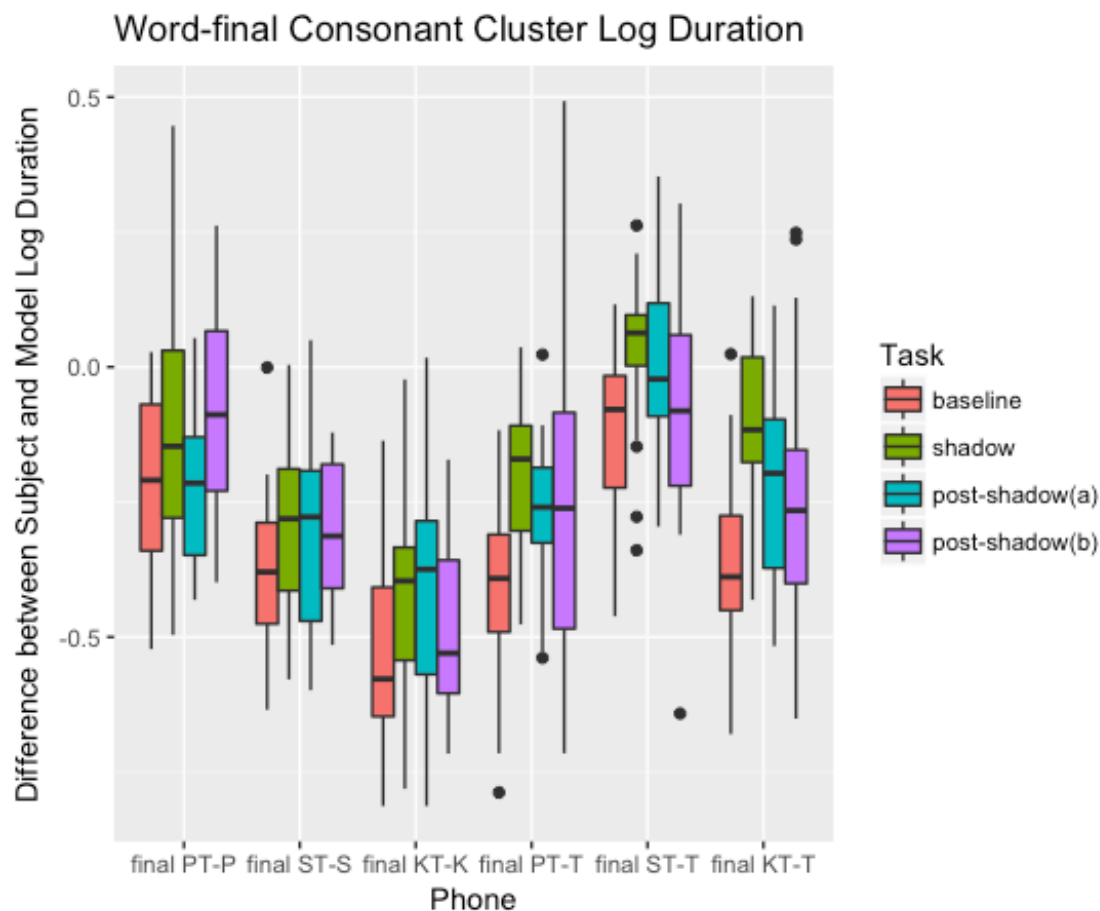

Figure 10: Difference between word-final consonant cluster log duration of subjects and model duration categorized and phone and task

Figure 10 confirms the difference in distance across phones and tasks. The baseline productions are all negative values, meaning that the subject durations are shorter than the model duration. The medians and ranges for each phone are different from one another. Productions in the shadowing task all become closer to zero, indicating that the difference between the subject values and model value decreases. The productions in the post-shadow task show different patterns of converging or diverging depending on phones, but the values in general return to the baseline state.

Since the interaction between condition and task is significant, Figure 11 shows the difference between the subjects' productions and model values on the y-axis categorized by condition and task is used to understand the differences among groups. The plot reveals that there is a consistent difference across means in conditions in the baseline, shadow, and post-shadow(a) task, but a different pattern is seen in post-shadow(b) task, where the median from the "employer" condition moves away from zero and becomes more similar to the baseline value while the median from "professor" condition remains at about the same level as that in post-shadow(a) task. In fact, the difference between subject and model values is the smallest in post-shadow(b) task. This observation supports the hypothesis that subjects in the "professor" condition will be prone to accommodate to the model speaker and the subjects in the "employer" condition will try to maintain their speech.
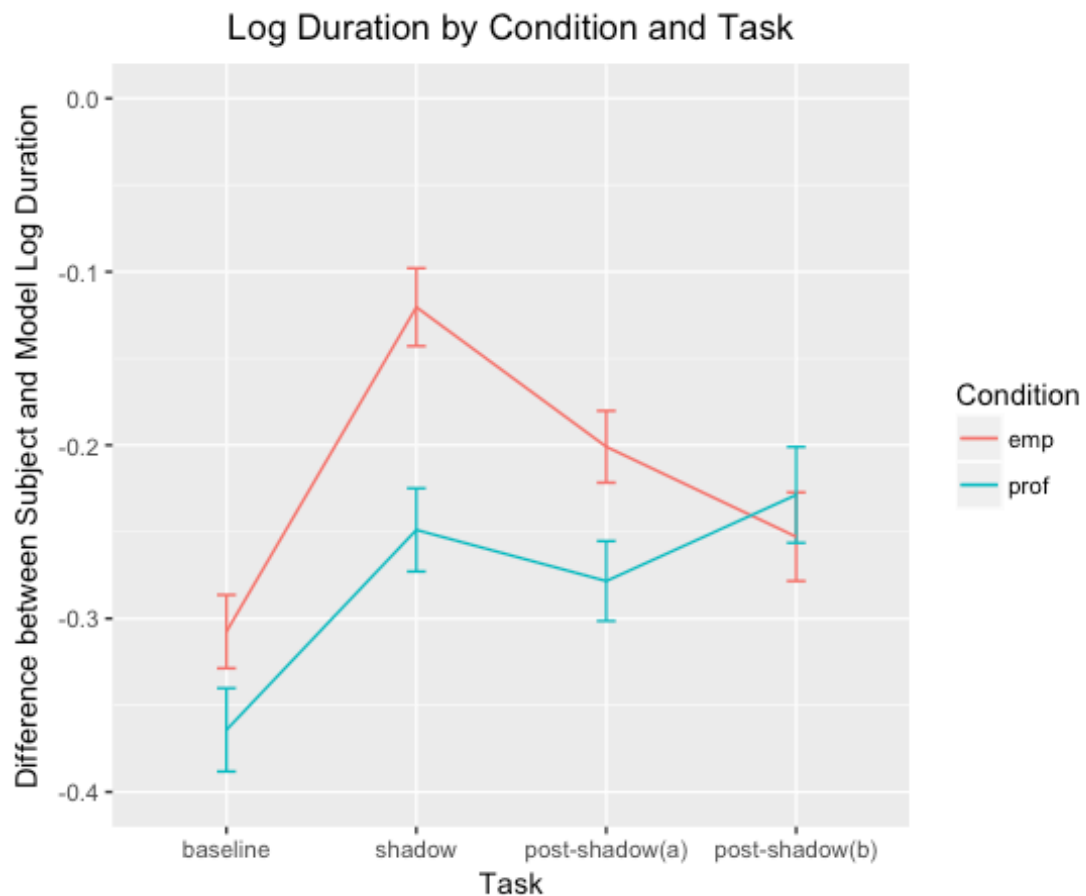

Figure 11: Averages of difference between word-final consonant cluster log duration of subject and model speaker categorized by condition and task with error bars showing the standard error

4. Discussion

The results from the data demonstrate interesting patterns and further explore the models by looking at the difference in effect that phones, tasks, and conditions have in determining the outcome of phonetic accommodation. The results corresponding to the hypotheses are summarized in the table 8 below.

Measurement	Results	Hypothesis
Vowel Formants	Task 2.892e-06 *** Condition:Task 0.656	1. Accommodation in shadowing and post-shadow tasks: Supported 2. Amount of Accommodation Depends on Condition: No significant result
Vowel Durations	Task < 2.2e-16 *** Condition:Task 0.010 **	1. Accommodation in shadowing and post-shadow tasks: Supported 2. Amount of Accommodation Depends on Condition: Supported
Word-final Stops Durations	Task 2.044e-10 *** Condition:Task 0.328	1. Accommodation in shadowing and post-shadow tasks: Supported 2. Amount of Accommodation Depends on Condition: No significant result
Word-Initial Consonant Clusters Duration	Task < 2.2e-16 *** Condition:Task 0.157	1. Accommodation in shadowing and post-shadow tasks: Supported 2. Amount of Accommodation Depends on Condition: No significant result
Word-Final Consonant Clusters Duration	Task 7.808e-12 *** Condition:Task 0.004 **	1. Accommodation in shadowing and post-shadow tasks: Supported 2. Amount of Accommodation Depends on Condition: Supported

Table 8: Results summary

The significance of difference across task in all the dependent variables measured in this study support the automatic alignment theory, which proposes that phonetic accommodation is an automatic, subconscious level (Pickering & Garrod 2004). The difference between the productions of subjects and model values significantly reduced in the shadowing task. The formant values, which show no significant difference between productions in the shadowing task and post-shadow(a) tasks suggest that the accommodation effect is maintained for at least a brief period.

Also, there are robust differences across phones in all the dependent variables measured. In accordance with Babel (2009)'s results, significant differences across phones are observed in formant values. In particular, low vowels [æ] and [ɑ] are accommodated more than non-low vowels such as [ʊ], [u], and [ɛ]. One exception, in which huge change in subjects' productions are observed, is [ɪ] in the dataset. The wide range in subjects' baseline productions could explain the seemingly huge accommodation. Because [ɪ] does not exist in the Mandarin inventory, subjects have different ways of producing the vowel; some subjects produce more [ɪ] sound

while others' productions are more similar to [i]. Thus, the speculation is that as subjects do not have a standardized [ɪ] sound, they are more willing to accommodate to how the model speaker produces the sound. In the case of vowel durations, the same observation does not seem to apply. [æ] and [ai] are among vowels that are least accommodated comparing to other vowels and vowels [ɑ] and [u] are accommodated. A possible explanation is that when the subjects' productions are farther away from the model value, they are more likely to be aware of the difference and therefore make changes accordingly. When the difference is too small, subjects do not notice the difference and as a result less accommodation is observed.

On the other hand, the prediction that degree of convergence produced by subjects should depend on the condition they are put in does not have strong support. The interaction between condition and task is only significant in formant duration and word-final consonant clusters duration. For these two dependent variables, post-shadow(b) seems to be the task that is making the difference, where productions in the "professor" condition remains a closer distance with the model value while productions in the "employer" condition tend to return to their baseline values. This observation supports the Communication Accommodation Theory, which states that speakers will be more likely to accommodate when they want to achieve a closer social distance with the interlocutor (Shepard et al. 2001). In this case, because the subjects in the "professor" condition are primed with the idea that they want their interlocutor to like them, they continue to accommodate to the model value in the post-shadow task.

The patterns in the data also shed light on phonetic accommodation in L2 English speakers. The accommodation found in the shadowing and post-shadowing tasks suggests that speakers accommodate in their L2 language. Also, the language background of the model speaker plays a role in determining the extent of subjects' accommodation. The fact that subjects did not accommodate in the post-shadowing tasks in the condition that encourages subjects to speak in proficient American English (the "employer" condition) suggests that they are aware of the accent of the model speaker and modify accordingly. Thus, this study supports that the theories for phonetic accommodation can also account for productions of L2 speakers.

5. Conclusions

By analyzing vowel formants, vowel durations, durations for word-final stops and consonant clusters, this experiment sheds light on some of the factors relating to phonetic accommodation. First, direct exposure contributes significantly to the amount of accommodation. All phonetic variables measured from the shadowing task are closer to the model speaker, when compared to the baseline and post-shadow tasks. Also, the effect of phone identity is significant. The height of vowels seems to play a role in determining the extent of accommodation in formant values while the difference between speakers' baseline productions and the model's seems to predict the amount of accommodation to vowel duration. Last but not least, condition in this experiment does not have a robust effect in post-shadow tasks; only half of the dependent variables show difference in conditions. However, for the variables that do show a significant difference across the interaction between condition and task, the trend supports the hypothesis that predicts that subjects in the "professor" condition accommodated more than subjects in the "employer" condition.

Despite these findings, there are a few limitations in this experiment. One of the shortcomings of this experiment is the extent of social manipulation in this study. Due to the laboratory setting and the tasks that involve subjects repeating the model speaker individual words for a hundred and eighty words, the condition introduction and design place subjects in an unnatural interaction. This might inhibit subjects' motivations and decisions for accommodation. Another factor that could have been more carefully considered is the choice of model speaker. Although the model speaker recruited for this experiment is a native Mandarin speaker and started learning English in her twenties, she is fluent in English. As a result, the response from subjects about the accentedness and intelligibility of the model speech varied greatly between subjects. Some subjects had difficulty understanding some of the words they heard while others felt the model's speech was native-like.

There are a few extensions that could be made with the limitations in order to better understand phonetic accommodation in second language speakers under social settings. For the condition designs, a normal interaction with a model interlocutor or another form of shadowing such as having listeners listen to narratives instead of repeating word by word after the model speaker could be used in order to achieve a better stimulation in subjects' willingness to accommodate. Future studies could also investigate the role level of accentedness of the model speaker plays, by including a heavy accented speaker and a proficient second language speaker of English, in phonetic accommodation.

References

Babel, M. (2009). Phonetic and social selectivity in phonetic accommodation. *Doctoral dissertation, University of California at Berkeley, Berkeley, CA*.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177-189.

Babel, M., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech*, *55*(2), 231-248.

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same-and mixed-gender dyads. *Language & Communication*, *8*(3-4), 183-194.

Chang, J. (1987). Chinese speakers. *Learner English, 310-315*.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of experimental psychology: Learning, memory, and cognition*, *22*(5), 1166.

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language*, *81*, 72-90.

Kim, M. (2012). *Phonetic accommodation after auditory exposure to native and nonnative speech* (Doctoral dissertation, NORTHWESTERN UNIVERSITY).

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory phonology*, *2*(1), 125-156.

M. L. Jiahong Yuan, "Speaker identification on the SCOTUS corpus," in Proceedings of Acoustics, 2008, pp. 5687–5690.

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*, *21*(4), 422-432.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, *32*(5), 790.

Nielsen, K. (2008). Word-level and Feature-level Effects in Phonetic Imitation, University of California, Los Angeles.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, *40*(1), 190-197.

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics*, *72*(8), 2254-2264.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(02), 169-190.

Shepard C.A., Giles H., and Le Poire B. A. (2001). Communication Accommodation Theory. In *The New Handbook of language and Social Psychology*, ed. Robinson W. P, and Giles H., 33-56. John Wiley & Sons Ltd.

Smiljanić, R., & Bradlow, A. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness. *The Journal of the Acoustical Society of America*, *130*(6), 4020.

Smith, C. L. (2007). Prosodic accommodation by French speakers to a non-native interlocutor. In *Proceedings of the XVIth international congress of phonetic sciences* (pp. 313-348).

Zając, M. (2013). Phonetic imitation of vowel duration in L2 speech. *Research in Language*, *11*(1), 19-29.

Appendix

Word List 1

Target words

abrupt	act	addict	adopt
ape	attempt	bed	bet
bid	bit	breast	cooked
fake	feast	food	foot
hub	hut	lab	lap
looked	map	pick	pig
rib	rip	root	rude
scar	school	script	shook
skate	ski	snug	spa
speak	spend	spy	stay
steep	step	style	tab
taste	tug	waste	

Fillers

allow	account	flower	passport
abuse	home	preview	protest
airline	firework	juice	legend
Asia	drink	beyond	copper
change	smooth	weather	something
chase	cheese	yoga	dashboard
pursue	watch	thirty	tooth
reform	beach	appoint	lawyer
short	sharp	treasure	garage
waffle	year	yummy	yourself
waiter	layout	update	

Word List 2

Target Words

absorbed	aced	bankrupt	booked
boot	cracked	dab	dropped
hate	jade	least	leg
loop	neat	overcooked	peck
peg	post	rug	scan
scooter	shoot	sipped	skin
sky	soaked	soup	span
special	speed	spend	spoke
steal	stick	stiff	study
swept	tag	tap	tape
toast	took	tub	tube
wood			

Fillers

across	advice	airflow	assign
background	between	boring	bother
chain	chestnut	choice	clause
counter	dolphin	ensure	feather
flashcard	flour	flush	garden
gesture	judge	louder	major
margin	massage	nation	noise
oyster	pathway	queen	quick
shadow	smooth	strange	thirty
triangle	version	vision	waffle
weekly			