Predicting Discrimination Accuracy by Assimilation Pattern: How Do Mandarin Speakers Discriminate English Vowels?

Yuan Chai

University of Colorado at Boulder, yuan.chai@colorado.edu

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

Part of the Phonetics and Phonology Commons

Recommended Citation
https://scholar.colorado.edu/ling_gradetds/66

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.
Predicting Discrimination Accuracy by Assimilation Pattern: How do Mandarin Speakers Discriminate English Vowels?

by

Yuan Chai

B.A., Beijing Normal University, Beijing, 2015

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 Arts
Department of Linguistics
2017
This thesis entitled:
Predicting Discrimination Accuracy by Assimilation Pattern: How do Mandarin Speakers Discriminate English Vowels?
written by Yuan Chai
has been approved for the Department of Linguistics

Prof. Rebecca Scarborough

Prof. Mans Hulden

Prof. Barbara Fox

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 # 16-0217
Chai, Yuan (M.A., Linguistics)

Predicting Discrimination Accuracy by Assimilation Pattern: How do Mandarin Speakers Discriminate English Vowels?

Thesis directed by Prof. Rebecca Scarborough

The Perceptual Assimilation Model (PAM) claims that two types of non-native contrasts are easiest to discriminate: when two contrastive sounds are assimilated into two different native categories; when one sound is categorized into a native category while the other is not (Best, 1991). However, some studies have found the latter one is sometimes worse discriminated than the former (Guion et al., 2000). The current study asks why such disagreement emerges and postulates that the problem lies in PAM’s definition of categorization. PAM does not allow sounds to be assimilated into more than one category (multiple-category assimilation), regarding them as “uncategorized” sounds. In addition, existing studies testing PAM mostly adopted arbitrary categorization criteria, making the results ungeneralizable. This study proposes a new non-arbitrary criterion, allowing multiple-category assimilation: An L2 sound is categorized if it is labeled as the same L1 category significantly above chance.

The subjects of the study were 30 first-year middle school students from China and 30 English native speakers. The Mandarin subjects participated in three experiments: discriminating English contrasts /i/-/ɪ/, /eɪ/-/æ/; assimilating those English vowels into Mandarin categories; and rating their goodness in the categories. The English subjects participated in the discrimination test only. The purpose is to test whether the English group discriminates English contrasts better than the Mandarin group.

The discrimination accuracies of the four contrasts are in a hierarchy of /i/-/ɪ/ = /ʊ/-/u/ > /eɪ/-/æ/ = /ɛ/-/æ/). Comparisons with the assimilation test results indicate
that contrasts with no overlap in the native assimilation categories (/ɪ/-/ɪ/, /ʊ/-/ʊ/) are more discriminable than those with overlap in native assimilation categories (/ɛɪ/-/ʌ/, /ɛ/-/æ/). The new categorization criterion converts PAM’s Uncategorized-Categorized contrasts into overlapped or non-overlapped contrasts, explaining the disagreements on the discrimination of “uncategorized” sounds. This study also contributes to L2 phonetic acquisition. L2 educators can predict what L2 contrasts are challenging by how learners assimilate L2 into L1, then design discrimination exercises targeting at those difficult contrasts specifically.
Dedication

This paper is dedicated to my parents, Shufan Chai and Hong Jiang, who have always been supporting me to pursue my dreams.
Acknowledgements

I want to attribute my deepest thanks to my advisor, Prof. Rebecca Scarborough. She has given me tremendous help through the whole experiment design, data interpretation, and paper-writing process. Her advice improved the logic, organization, and language of this paper, and guided me to dig out a bigger picture from the results. What I have learnt from her will be the guidance of my future study. Her support motivates me to proceed towards my goal of being a phonetician.

I would like to thank my committee members, Prof. Mans Hulden and Prof. Barbara Fox. Prof. Hulden was my instructor of phonetics and phonology courses. He introduced me into the field of phonetics. This thesis was inspired from the final project of that course. Prof. Fox was my instructor of semantics and pragmatics course. Her encouragement motivated me to take the initiative of writing a thesis. This thesis would have been impossible without their instructions and supports.

Many thanks to Story Kiser, who produced the English stimuli for the experiments; Dongmei Ma and Hong Jiang, who helped recruit and run subjects in China. I would also like to attribute my thanks to my dearest friends, Yiding Zhao, who gave me unconditional support throughout my entire thesis-writing process and my life so far; and Shihong Ye, who has been my research buddy since college and helped me with my thesis defense.
Contents

Part I Introduction ........................................................................................................................................... 1

Part II Theoretical Background ......................................................................................................................... 3

2.1 Why Assimilation May Predict Discrimination ......................................................................................... 3

2.2 How Assimilation May Predict Prediction ................................................................................................. 4

2.3 The Significance of Using Assimilation to Predict Discrimination ......................................................... 6

2.4 Research Gap ............................................................................................................................................... 7

2.4.1 Disagreements and Drawbacks in Existed Studies Examining PAM .................................................... 7

2.4.2 Drawbacks of PAM ............................................................................................................................... 9

2.4.3 Language-Wise Motivation .................................................................................................................. 10

Part III Current Study .................................................................................................................................... 13

3.1 Research Questions .................................................................................................................................... 13

3.2 Hypotheses .................................................................................................................................................. 13

Part IV Experiments ....................................................................................................................................... 15

4.1 Participants .................................................................................................................................................. 15

4.2 Experiment 1: English Vowel Discrimination .......................................................................................... 16

4.2.1 Stimuli .................................................................................................................................................... 16

4.2.2 Procedure ............................................................................................................................................. 17

4.2.3 Evaluation ........................................................................................................................................... 18

4.3 Experiment Two: Assimilation identification .......................................................................................... 18

4.3.1 Stimuli .................................................................................................................................................... 18
Tables

Table 1. The Assimilation options for Each Target English Vowel ......................... 19
Table 2. English Non-Words and Possible Mandarin Options ................................. 20
Table 3. Mean F1, F2, and Duration of Vowel ......................................................... 22
Table 4. Descriptive Data of Euclidean Distance Between Sounds Within Contrasts .......................... 24
Table 5. Discrimination Accuracy and Reaction Time by English and Mandarin Groups .................................................................................................................. 27
Table 6. F Statistics of Discrimination Score and Reaction Time ANOVA .............. 30
Table 7. Assimilation Proportion and Goodness of English into Mandarin .......... 31
Figures

Figure 1. Perceptual Assimilation Model (PAM) ................................................................. 5
Figure 2. Vowel Chart of English and Chinese Vowels ...................................................... 23
Figure 3. Comparison of Duration Difference of Different Contrasts .......................... 25
Figure 4. Comparison of Discrimination Accuracy Between Groups ...................... 28
Figure 5. Comparison of Reaction Time Between Groups ........................................... 30
Figure 6. Assimilation Type of Each Contrast ................................................................. 33
Figure 7. Assimilation Categories of /ɹ/-/w/ and /s/-/θ/ .................................................. 36
Part I Introduction

When listening to a non-native language, people sometimes have the experience that two sounds that are different to native speakers sound the same to them. Such a phenomenon indicates that people’s perception of non-native speech sounds is different from their perception of native sounds. Some non-native sounds are discriminable while others are not. However, among those non-discriminable contrasts, their degrees of discrimination difficulty are different from each other. Researchers have been striving to find ways to predict the degree of intelligibility of non-native sounds. The Perceptual Assimilation Model (PAM) (Best, 1991) proposed that the discrimination of a contrast can be predicted by whether and how the sounds in the contrast were assimilated into native phonetic categories. This study aims to test PAM under the circumstance of Mandarin native speakers discriminating English vowels. The study will first illustrate how Mandarin speakers discriminate English vowel contrasts and assimilate those English vowels into Mandarin vowel categories. Secondly the study will examine whether the discrimination pattern is predicted by the assimilation pattern as PAM predicted.

Another motivation of this study is that existing studies examining PAM have disagreements on how well people can discriminate non-native contrasts involving sounds that are not categorized into any native phonetic categories. Different studies yielded different conclusions. Some studies found that such contrasts were perfectly discriminated (Tyler et al., 2014) while some found they were unintelligible (Guion and Flege, 2000). Bearing such disagreements in mind, I examined such uncategorized sounds that emerged in this study, and tested how well contrasts involving uncategorized sounds were discriminated. The results provide some insights for the debate over the predicted discrimination of uncategorized contrasts, and suggest some modifications to the model.
The results of this study will facilitate L2 phonetic acquisition. Predicting discrimination accuracy is of interest because it suggests the causes of discrimination accuracy and indicates a way to improve discrimination accuracy. Once it is proved that the discrimination accuracy of L2 contrasts can be predicted by how L2 sounds are assimilated into L1 categories, and L2 sounds assimilated into the same L1 categories are less discriminable, explicit training can be designed to distinguish such L2 sounds from their nearest L1 sounds. Such training may help L2 learners establish separate phonetic categories for the less discriminable L2 sounds. Thus, the ultimate goal of this study is to use the results to help speakers “dissimilate” L2 sounds from L1 sounds and improve their L2 perception ability.
Part II Theoretical Background

2.1 Why Assimilation May Predict Discrimination

This study proposes that discrimination accuracy can be predicted by how L2 sounds are assimilated into L1 categories. A theory called Native Language Magnet (NLM) (Kuhl, 1992) proposes that people’s perception of speech sounds is influenced by their native phonetic categories.

NLM theory (Kuhl, 1992) explains how people become committed to native language phonetic categories while ignoring non-native phonetic differences. The underlying mechanism is the perceptual magnet, referring to the idea that the prototype of a phonetic category functions as a magnet pulling the variants near the prototypical exemplar together. As a result, the perceptual distance between tokens surrounding the prototype decreases while distance between tokens at the boundary of two categories increases. Such distorted perception helps listeners to distinguish phonetic differences between categories while ignoring differences within single categories.

Kuhl et al (2008) expanded NLM to suggest that people go through four phases during the development of the perceptual magnet. Newborn babies are at Phase 1 when they can distinguish the phonetic categories of any language. Infants move into Phase 2 at 6-10-month-old when a “prototype” is developed for each phonetic category according to the distributional pattern of the language input. The prototype functions as a “magnet” and distorts the perceptual distance between phonetic units as described earlier. At Phase 3, the improved native language perceptual ability allows infants to detect word boundaries and make connections between meaning and sound. At Phase 4, the neural structure becomes highly committed to the native language phonetic system, which explains why it is difficult for adults to learn a new language.
Thus, for a highly committed monolingual speaker, when he is exposed to a foreign language, it is very possible that he will compare the foreign sound with existing native speech sounds. Consequently, when he is trying to discriminate a foreign contrast, how the components of the foreign contrasts are assimilated into native phonetic categories will influence how well the speaker discriminates the contrast.

2.2 How Assimilation May Predict Prediction

The Perceptual Assimilation Model (PAM) specifically attempts to predict the perception accuracy of non-native contrasts (Best, 1991; Best & Tyler, 2007). According to PAM, the accuracy of discrimination depends on how people assimilate non-native speech sounds into their native phonetic categories. If the contrasting sounds are assimilated into two different categories, the pattern is classified as “Two Categories (TC)”. If the contrasting sounds are assimilated into the same category and have equal distance to the prototype, the pattern is “Single Category (SC)”, while if the contrasts are assimilated into the same category but differ in their distance to the prototype of the category, the pattern is called “Category Goodness (CG)” since the two sounds differ in their category goodness. There are also foreign phones that cannot be assimilated. If one member of a contrast can be categorized while the other is not, the pattern is “Uncategorized-Categorized (UC)” ; if both phones are not categorized, the pattern is called “Uncategorized-Uncategorized (UU)”. Any other phones are perceived as pure noises, and are labeled as “Non-Assimilable (NA)”. Those assimilation types are illustrated in Figure 1.
Among those six types of assimilation, TC contrasts should be discriminated with near-native accuracy because listeners could distinguish them as two different native phonemes. UC contrasts should also be well discriminated since one of the sounds falls into a native category while the uncategorized one falls on the edge of the native phonological space. CG contrasts will be better discriminated than SC contrasts because the difference in goodness gives the listener a clue to distinguish them. The discrimination accuracy of UU contrasts depends on the phonetic difference between two sounds and the difference in their nearby native categories (Best & McRoberts & Goodell, 2001, p. 777; Tyler et al., 2014, p. 6). The performance of NA discrimination depends on their articulatory differences (Best, 1995). Thus, the theory predicts a hierarchy of those types of contrast: TC = UC > CG > SC, with UU and NA lying somewhere in between.

Faris et al. (2016) postulated that the discrimination of UU contrasts depended on what types of uncategorized sounds they were and further divided uncategorized sounds into three types: focalized, clustered, and dispersed.
uncategorized sounds. For a given L2 sound, if there were several L1 sounds as its assimilation candidates, and none of those candidates were selected greater than the threshold proportion of categorization, but one of the candidates was selected significantly greater than the chance level, that L2 sound is named “focalized uncategorized sound”. Under the same situation, except that multiple candidates were selected significantly greater than the chance level, that L2 sound is named “clustered uncategorized sound”. If none of the candidates were selected significantly greater than the chance level, that L2 sound is named “dispersed uncategorized sound”. Faris et al. (2016) further postulated that how overlapped the L1 labels are affected UU discrimination. Empirical studies are required to test their postulation.

2.3 The Significance of Using Assimilation to Predict Discrimination

If assimilation is proved to be able to predict discrimination accuracy, then it is possible to use the assimilation result to improve discrimination accuracy. Flege (1995) proposed a Speech Learning Model (SLM). One of SLM’s hypotheses is that it is possible for L2 speakers to create a separate category for a given L2 sound. L2 sounds do not have to be assimilated into L1 categories. The establishment of a phonetic category is mainly affected by three factors: 1) the perceptual similarity between the L2 sound and the nearest L1 sound; 2) the age of starting to learn L2 with a native exposure; 3) the length of learning and use of L2 within a native environment. Flege and colleagues show that the more perceptually different two sounds are, the earlier the learning is (Flege & Munro & MacKay, 1995; Flege & MacKay & Meador, 1999; Flege et al., 2006), and the longer one learns L2 (Cebrian, 2005; Flege & Bohn & Jang, 1997; Flege et al., 2006), the more likely it is that a separate phonetic category would be established for that L2 sound.
As PAM suggested, two sounds being assimilated into the same L1 category creates discrimination difficulty. Thus, for such a contrast, knowing what L1 category those two given L2 sounds are assimilated into, it is possible to find another way to help establish a separate category: explicit training to discriminate that L2 sound from the nearest L1 sounds. Consequently, speakers will be able to discriminate that L2 sound from L1 sounds and other L2 sounds better, and produce them in a more native-like fashion. The goal of using assimilation to predict discrimination is to free L2 learners from L1 assimilation and establish a new category for that L2 category.

In summary, the three aforementioned speech perceptual models are related to each other and address speech perception from different levels. NLM explains PAM's hypothesis that non-native contrasts assimilated into the same category are less discriminable than contrasts assimilated into different categories. SLM shows the next stage of PAM: how speakers can “dissimilate” L2 sounds from L1 sounds and establish a separate phonetic category for each L2 sound. Thus, using assimilation to predict discrimination is reasonable based on the Native Language Magnet theory, and is useful for building new categories for non-discriminable L2 sounds to facilitate speech learning.

2.4 Research Gap

2.4.1 Disagreements and Drawbacks in Existing Studies Examining PAM.
Several empirical studies have been conducted to examine the relationship between discrimination and assimilation. Most of them confirm the discrimination hierarchy of TC (Two-Category) > CG (Category Goodness) > SG (Single-Category) (Best & McRoberts & Goodell, 2001; Tyler et al, 2014; Ghaffarvand Mokari & Werner, 2017; Guion et al., 2000).
However, disagreements arise when contrasts involving “uncategorized” sounds emerged. Some studies found contrasts involving both “uncategorized” and “categorized” sounds were as well-discriminated as TC contrasts while some found such contrasts were as non-discriminable as CG or SG contrasts. Tyler et al. (2014) conflated TC and UC type contrasts as “cross-boundary” contrasts and tested the discrimination accuracy hierarchy of contrasts of different assimilation types. Tyler et al. found that there was great variation among individuals in the categorization. Thus, instead of designating an assimilation type to each contrast, they found the assimilation type of each contrast by individual, and compared the discrimination score among different assimilation types. The result suggested that such “cross boundary” contrasts are better discriminated than CG and SG.

Ghaffarvand Mokari and Werner (2017) tested how Azerbaijani speakers assimilated Standard Southern British English (SSBE) vowels into Azerbaijani, and how the assimilation type was related to discrimination of SSBE vowel contrasts. They tested 11 SSBE contrasts, among which emerged TC, UC, SG, and UU type assimilations. The result suggested that in terms of discrimination, TC = UC > SG = UU, which was in line with PAM’s predictions on TC, UC, and SG discrimination hierarchy. However, the deficiency of Mokari and Werner’s study was that the assimilation and goodness rating results were drawn from a previous study using a different group of subjects (Ghaffarvand Mokari & Werner, 2015). Since discrimination and assimilation patterns may vary with different populations, even if they share the same L1, using the assimilation results of one group to predict the discrimination of another group is not appropriate. Consequently, the support that Mokari and Werner (2017) provide to PAM is not fully solid.

Guion et al. (2000), on the other hand, contradict PAM’s predictions. They studied how Japanese speakers assimilated English consonants /b s t v w ɻ θ/ into Japanese and whether the assimilation pattern predicted the discrimination
accuracy of English consonants as PAM suggested. Among the contrasts they tested, /ɻ/-/w/ and /s/-/θ/ were both UC type contrast. However, /ɻ/-/w/ was better discriminated than /s/-/θ/, contradicting PAM’s prediction that any UC contrast will be as well discriminated as TC (two category) contrasts. The explanation Guion et al. (2000) gave was that for the /s/-/θ/ contrast, /θ/ was most frequently labeled as Japanese /s/ or /ɸ/ while English /s/ was categorized as Japanese /s/. In other words, the possible Japanese categories that English /θ/ can be assimilated into overlapped with the consistent Japanese category that English /s/ was assimilated into. Thus, although /θ/ is uncategorized while /s/ is categorized, they overlapped in the subjects’ L1 phonetic categories, making them indistinguishable.

2.4.2 Drawbacks of PAM. The disagreements in how well UC and UU type contrasts were discriminated could result from the deficiencies in the Perceptual Assimilation Model itself. Although PAM provides a way to predict the discrimination of non-native contrasts, it is not a perfect model. First, it does not specify how to classify whether a sound is “categorized”. Some studies (Harnsberger, 2001; Ghaffarvand Mokari & Werner, 2017) used the criterion that if a given sound was designated with the same label over 90% of time, that sound was categorized. Other studies used 75% (Guion et al., 2000), 70% (Best & McRoberts & Goodell, 2001; Guion et al., 2000; Strange et al., 2001; Bundgaard-Nielsen et al., 2011; Tyler et al, 2014), and 50% (Bundgaard-Nielsen et al., 2011) as thresholds. However, all those criteria were mostly arbitrary. There is no gold standard for which criteria should be used under which circumstances and why.

The other drawback of PAM is that it only considers single-category assimilation. If a given sound is not consistently assimilated into a single category, it is not categorized at all. However, as Escudero and Boersma (2002) pointed out, multiple-category assimilation did occur when speakers’ L1 had a larger vowel/consonant inventory than L2. They found that Dutch speakers assimilated
Spanish front vowels, which have a two-way contrast (/i-e/), into any of the three Dutch front vowels (/i-ɪɛ/). The current study considers whether such multiple-category assimilation will also occur when L1 has a smaller vowel inventory than L2. The author hypothesizes that speakers may assimilate the same L2 sounds into different L1 categories which are phonologically close to each other.

Regarding to those two drawbacks, the current study proposes a new way of deciding whether a given sound is categorized and what category it is categorized into. Instead of using an arbitrary threshold of proportion, this study looks for whether a given L2 sound is assimilated into any L1 categories more often than chance and if so, how many such L1 categories there are. If a given L2 sound is assimilated into only one L1 category by this criterion, there is a single-category assimilation between the L1 and L2 sound. If more than one L1 category is more consistent than chance in the assimilation test, that L2 sound is assimilated into multiple L1 categories. If none of the L1 categories is selected above chance, that L2 sound is regarded as uncategorized. This study borrows the term “focalized” from Faris et al. (2016): L2 sounds that are consistently assimilated into a single L1 category are called “focalized” assimilated sounds. The advantage of these criteria is that they will be language-general and experiment-general. The advantage of introducing the notion of multiple-category assimilation into PAM is that the assimilation pattern of PAM “uncategorized” sounds will no longer be ignored, which is helpful for predicting the discrimination of contrasts involving such “uncategorized” sounds.

2.4.3 Language-Wise Motivation. This study uses Mandarin (L1) – English (L2) as target languages because studies have shown that Mandarin speakers do have difficulties discriminating English vowels that do not exist in the Mandarin vowel inventory (Wang, 1997, 2002, 2006). The relationship between L1-L2
assimilation and L2 discrimination can be tested as a way to explain the discrimination difficulties in L2 perception.

The second reason for looking at Mandarin and English is that previous studies testing how Mandarin speakers discriminate and assimilate English contrasts are either not thorough enough or deficient in methodology. Sun and Heuven (2007) tested the assimilation pattern of English vowels into Mandarin by Mandarin speakers. However, there were deficiencies in the data they collected. Sun and Heuven (2007) covered a wide range of English vowels: /i: i eɪ æ u: u ʊ ɔː aʊ ə ɛ aɪ ʊə ɔɪ/. They used discrimination scores obtained from another study (Jia et al, 2006) to try to predict their assimilation patterns. The result of that study was not exactly as PAM predicted. The /ɛ/-/æ/ contrast, which was classified as a UC contrast and expected to be well discriminated, received a poor discrimination score. The /æ/-/ɑː/ contrast, which was classified as SC contrast and expected to be poorly discriminated, received a good score. The explanation the authors provided for the UC contrast with the poor score was that /ɛ/ and /æ/ are phonetically similar. As in the /s/-/θ/ contrast in Guion et al.’s test (2000), the assimilation categories of English /ɛ/ and /æ/ into Chinese vowel categories overlapped. However, the study failed to provide an explanation for the mismatch between SC contrasts and high discrimination accuracy. This may be because Sun and Heuven (2007) were using different subjects as the source of the assimilation patterns and the discrimination performance. Although they were all Mandarin speakers, there may still be dialect variation, and the educational and intellectual backgrounds of the two groups of subjects may also vary. Thus, it is not surprising to see that the performance of Sun and Heuven’s subjects cannot predict the performance of Jia et al’s subjects. In the present study, the discrimination, identification, and rating scores will all be drawn from the same subject group.
Lai (2010) did a similar study and similarly maintained that the results did not fully meet the prediction of PAM. The evidence was that the /æ/-/ɛ/ contrast was a UU contrast, where both vowels were most frequently labeled as /e/ in Mandarin, but this contrast received a higher discrimination score than other UC contrasts. However, the evidence Lai (2010) provided is not sufficient to undermine PAM for methodological reasons. First, the criterion for determining whether a vowel is categorized or not was very strict: only if a vowel was labeled by the same category over 90% of the time was it called “categorized”. However, the assimilation task required transcribing the English vowels as Mandarin vowels, which yielded a larger variety in the assimilation options compared to multiple choice, so 90% is too stringent for such a task. If Lai adopted a 70% criterion, for example, /æ/ and /ɛ/ would both be categorized as Mandarin /e/. Furthermore, Lai did not measure the category goodness of the English vowels in the Mandarin categories. It is possible that although /æ/ and /ɛ/ were categorized into the same category, the category goodness was different. If so, the higher discrimination score than for UC and SC contrasts would be justifiable according to PAM.
Part III Current Study

3.1 Research Questions

The disagreement among existing studies testing PAM mainly focuses on the place of UC and UU type contrasts on the discrimination hierarchy. Because PAM does not consider the possibility of multiple-category assimilation, it is treating some multiple-category assimilated sounds as uncategorized sounds, ignoring those “uncategorized” sounds’ position in the L1 phonological system. That is probably why PAM's prediction on UC type contrasts failed in some of the previous studies. Thus, the current study tests whether the multiple assimilation concept will solve the problems caused by the “uncategorized” sounds in PAM. In addition, existing studies using Mandarin and English as target languages were not thorough enough in terms of experiment design and data analysis. Thus, it is worthwhile to redesign and redo the experiments in order to find out more accurately how Mandarin speakers perceive English vowels. With the above two motivations, this study proposes three questions: 1) Do Mandarin speakers discriminate English vowel contrasts /i/-/ɪ/, /eɪ/-/ɪ/, /ɛ/-/æ/, and /ʊ/-/u/ with different accuracies? 2) If so, can the differences be predicted by the assimilation of those English vowels into Mandarin categories? 3) What is the position of contrasts involving multiple-category assimilated sounds and uncategorized sounds on the discrimination hierarchy?

3.2 Hypotheses

Based on the results of previous studies, my hypotheses to those research questions are as follows: 1) Mandarin speakers will discriminate those four English contrasts differently; 2) The discrimination of TC, CG, and SG type contrasts will
follow the prediction of PAM, ranking in TC > CG > SG; 3) The difficulty of contrasts involving multiple-category assimilated sounds will be affected by the overlap among the L1 assimilation categories of the contrastive L2 sounds. If the L1 categories that two sounds are assimilated into are the same or overlap, the discrimination will be poor. If the L1 categories they assimilated into are non-overlapping, the discrimination will be good.
Part IV Experiments

In order to test the relationship between discrimination accuracy and assimilation pattern, three experiments were conducted: 1) an English vowel discrimination test to obtain the accuracy hierarchy between contrasts; 2) an identification test to decide how English vowels are assimilated into Mandarin vowels; 3) a goodness rating test to figure out the “fitness” of the English vowels in the Mandarin vowel category that they are assimilated into.

4.1 Participants

30 native Mandarin speakers were recruited among first year middle school students in Qiqihaer, Heilongjiang Province, China. Those Mandarin speakers’ exposure to English was mainly from English class in school, which mainly focused on English reading and writing, and was taught by non-native English speakers. Thus, the subjects had minimal exposure to native spoken English and could be considered as near-monolinguals of Mandarin. In addition, all the subjects were from the same dialect region of China.

This study also recruited 30 native speakers of American English from introductory linguistics classes at the University of Colorado, Boulder. The purpose of involving English speakers in this study is to compare their scores on the English vowel discrimination test to the scores of the Mandarin speakers. For a given vowel contrast, if Mandarin speakers’ discrimination scores are significantly lower than the English speakers’, Mandarin speakers can be regarded as having difficulties discriminating that vowel contrast.
4.2 Experiment 1: English Vowel Discrimination

This experiment aimed to test how well Mandarin speakers discriminated English vowel contrasts /i/-/ɪ/; /eɪ/-/ɪ/; /ɛ/-/æ/; and /ʊ/-/u/ using AXB method.

4.2.1 Stimuli. The target vowel contrasts for the test were /i/-/ɪ/; /eɪ/-/ɪ/; /ɛ/-/æ/; and /ʊ/-/u/. The reasons for selecting those four contrasts were 1) one member of each pair (/ɪ, æ, u/) was not included in the Mandarin vowel inventory; 2) the researcher conducted a pilot study prior to this research, testing the discrimination accuracy of the above four vowel contrasts. The result of the pilot study shows that there is a significance difference between them, namely /i/-/ɪ/ > /ɛ/-/æ/ > /ɪ/-/eɪ/ > /ʊ/-/u/. It is reasonable to assume that those four contrasts would continue to elicit a discrimination hierarchy in this study.

The vowels were elicited in non-word minimal pairs. The reason for not using meaningful words is that meaningful words differ in word frequency, and word frequency has been shown to influence discrimination: a high frequency word is usually more intelligible than a lower frequency one (Howes, 1957). This study aimed to test the effect of assimilation on discrimination and eliminate the influence of word frequencies. Non-words, which do not have a frequency difference, serve this purpose. /i/, /eɪ/, /ɛ/, /æ/, /ʊ/, and /u/ were elicited in a t_d environment. The environment was the same for all the vowels because previous studies suggested that consonant context would affect the discrimination accuracy of foreign sounds (Strange et al., 2001).

The discrimination test was presented in AXB mode, following the procedure of the perception test by Hojen and Flege (2006). The pitches of A, X, and B were different so that the subjects had to make decisions by the acoustic quality properties of the sounds despite the interference of pitch difference. The F0 of the target words were either falling or rising. In order to manipulate the pitch of the
target words, they were elicited in carrier sentences “I say ____.” and “Can I say ____?” The first carrier sentence yielded a target word in falling F0 while the second rising F0. The F0 of the AXB triple was either “neutral” or “incongruent”. “Neutral” meant F0 did not indicate whether the adjacent pair was the same word or not. For example, A (rising F0) X=A (falling F0) B (rising F0). “Incongruent” meant F0 was the same between different words. For example, A (rising F0), X=A (falling F0), B (falling F0). The stimuli were produced by a female from Colorado who was a native speaker of American English. Each target word was repeated 10 times with falling F0 and 10 times with rising F0. In order to ensure that speakers were able to tolerate the variations in the realizations of the same word, AA/BB were all different tokens of A/B.

The experiment consisted of four blocks. Each block contained 16 different trials for each contrast. Within each block, each type of AXB combination occurred 4 times (AAB, ABB, BBA, BAA). The chance of each token occurring at each position was equal. Half of the trials had neutral F0 and the other half had incongruent F0. There were 256 (4 contrasts * 4 combinations * 4 repetitions * 4 blocks) trials in this experiment in total.

4.2.2 Procedure. Both Mandarin and English speakers participated in this experiment. In this AXB discrimination test, subjects listened to three words in a sequence, and were then asked to decide whether the first two or the last two words were the same. They were asked to press “F” on the keyboard if they considered the first two words the same; “J” if the last two were the same. Mandarin speakers participated in this experiment in a quiet conference room of a local middle school. Due to time and venue limit, subjects participated in groups of ten, working simultaneously on separate laptops. English speakers conducted this experiment in the sound booth in the phonetics lab of University of Colorado, Boulder individually.
For both groups, sounds were played through headphones and the instructions were presented on a computer screen using the Psychopy program (Peirce, 2007).

4.2.3 Evaluation. The keyboard responses of participants were recorded. The discrimination accuracy was calculated as the number of times that the participants picked out the same tokens correctly divided by the total number of trials. In other words, the discrimination accuracy was represented by the proportion of correct answers in the overall test. The reaction time of each response was measured from the end of the third token in the AXB triple to the moment when a participant pressed a key.

4.3 Experiment Two: Assimilation identification

This experiment tested how Mandarin speakers categorized English vowels into Mandarin vowel categories.

4.3.1 Stimuli. The English stimuli were the same set used in Experiment 1. Because Mandarin does not have obstruents in codas, tVd English non-words were clipped into tV. At the same time, only target English words with a low F0 were included in this test in order to match tone 4 (falling tone) in Mandarin.

There are disagreements on the vowel inventory of Mandarin Chinese. This study adopted the inventory listed in the book Modern Mandarin (Huang & Liao, 1997) which claims there are 7 monophthongal vowels in Mandarin /a, o, x, ε, i, u, y/, 9 diphthongs /ai, ei, au, ou, ia, iε, ua, uo/, and 4 triphthongs /iau, iou, uai, uei/. I assumed that subjects would very unlikely to assimilate a front vowel into a back vowel category, a high vowel into low vowel category, or an unrounded vowel into a rounded vowel category, or vice versa. Based on that assumption, I narrowed the possible assimilation options for English /i, ɪ, eɪ, æ/ into four and the options for
English /u, o/ into two. The correspondences between English vowels and their possible counterparts in Mandarin were listed in Table 1.

<table>
<thead>
<tr>
<th>English vowel</th>
<th>Mandarin vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>/i/, /ie/, /ei/, /ai/</td>
</tr>
<tr>
<td>/u/</td>
<td>/i/, /ie/, /ei/, /ai/</td>
</tr>
<tr>
<td>/ei/</td>
<td>/i/, /ie/, /ei/, /ai/</td>
</tr>
<tr>
<td>/e/</td>
<td>/i/, /ie/, /ei/, /ai/</td>
</tr>
<tr>
<td>/æ/</td>
<td>/i/, /ie/, /ei/, /ai/</td>
</tr>
<tr>
<td>/u/</td>
<td>/u/, /ou/</td>
</tr>
<tr>
<td>/o/</td>
<td>/u/, /ou/</td>
</tr>
</tbody>
</table>

Table 1. The Assimilation options for Each Target English Vowel

These correspondences explain why t__d was selected as the particular environment in Experiment One. First, all tV combinations in tone four except /tei/ were actual Mandarin words. And this was the syllable environment that could elicit the maximum number of real Mandarin words. Using real Mandarin words might better invoke subjects’ memory of the corresponding Mandarin pronunciation.

4.3.2 Procedure. Only Mandarin speakers participated in this experiment. During the experiment, participants listened to the English non-words. Then they were shown several Mandarin words that they could probably categorize the English words as, and asked to choose which word they thought they just heard by pressing the number of that option on the keyboard. The options were shown in Table 2. During the experiment, the Mandarin candidates were presented in both orthography and pinyin (the transcription system of Mandarin). The experiment consisted of eight different tokens for each English vowel. There were 56 trials (7 vowels * 8 tokens) in total. The instructions and trials was presented on a computer screen. The stimuli were played through headphones. The subjects made responses using a keyboard.
Table 2. English Non-Words and Possible Mandarin Options

<table>
<thead>
<tr>
<th>English Non-Word</th>
<th>Mandarin Words (Presented in order of Mandarin character – pinyin transcription – IPA transcription)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ti/</td>
<td>1. 替 (tì) /tì/  2. 帖 (tiè) /tie/  3. tèi /tei/  4. 太 (tài) /tai/</td>
</tr>
<tr>
<td>/ti/</td>
<td>1. 替 (tì) /tì/  2. 帖 (tiè) /tie/  3. tèi /tei/  4. 太 (tài) /tai/</td>
</tr>
<tr>
<td>/tei/</td>
<td>1. 替 (tì) /tì/  2. 帖 (tiè) /tie/  3. tèi /tei/  4. 太 (tài) /tai/</td>
</tr>
<tr>
<td>/tei/</td>
<td>1. 替 (tì) /tì/  2. 帖 (tiè) /tie/  3. tèi /tei/  4. 太 (tài) /tai/</td>
</tr>
<tr>
<td>/tæ/</td>
<td>1. 替 (tì) /tì/  2. 帖 (tiè) /tie/  3. tèi /tei/  4. 太 (tài) /tai/</td>
</tr>
<tr>
<td>/tu/</td>
<td>5. 兔 (tù) /tu/  6. 透 (tòu) /tou/</td>
</tr>
<tr>
<td>/tʊ/</td>
<td>5. 兔 (tù) /tu/  6. 透 (tòu) /tou/</td>
</tr>
</tbody>
</table>

4.4 Experiment Three: Exemplar Goodness Rating

Like Experiment Two, this experiment only involved the Mandarin participant group. The subjects were asked to rate the similarity between a given English word and several Mandarin words that they could be categorized as. The purpose of this experiment was to reveal the goodness of a given English vowel in a given Mandarin category through rating scores.

4.4.1 Stimuli. The English stimuli in Experiment Three were the same set used in Experiment Two. The Mandarin stimuli were the Mandarin words used in Experiment Two as possible assimilation options for the English words. Mandarin stimuli were produced by the author, a native Mandarin speaker coming from the same dialect region as the participants. Each Mandarin word was repeated 10 times.

4.4.2 Procedure. Subjects listened to one English non-word and one Mandarin word in each trail. Then they were asked to rate how similar those two words were using a scale from 1 (very different) to 9 (very similar). The order of English and Mandarin stimuli were balanced. Each English non-word and Mandarin word was tested by eight different tokens. There were 192 trials (5 English vowels * 4 Mandarin options * 8 tokens + 2 English vowels * 2 Mandarin options * 8 tokens) in
this experiment in total. As in the previous two experiments, the instructions and questions was presented on a computer screen using the Psychopy program. Subjects listened to the stimuli through headphones and responded by keyboard.
Part V Description of Stimuli

5.1 Formant Frequency and Duration of Stimuli

The F1 and F2 of the English vowels in Experiment One stimuli, and the F1, F2, and duration of the Mandarin vowels used in Experiment Three were measured. The F1 and F2 of monophthongs were measured at vowel midpoint. In diphthongs, the nucleus and off-glide were measured separately. The results are presented in Table 3 and plotted in Figure 2.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>F1 (Hz)</th>
<th>F2 (Hz)</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>English</td>
<td>/i/</td>
<td>8</td>
<td>402.625</td>
<td>31.268</td>
</tr>
<tr>
<td></td>
<td>/l/</td>
<td>8</td>
<td>542.75</td>
<td>26.174</td>
</tr>
<tr>
<td></td>
<td>/e/ in /æ/</td>
<td>8</td>
<td>527.75</td>
<td>19.002</td>
</tr>
<tr>
<td></td>
<td>/u/ in /æ/</td>
<td>8</td>
<td>455</td>
<td>36.249</td>
</tr>
<tr>
<td></td>
<td>/e/</td>
<td>8</td>
<td>738.625</td>
<td>20.085</td>
</tr>
<tr>
<td></td>
<td>/æ/</td>
<td>8</td>
<td>863.75</td>
<td>31.089</td>
</tr>
<tr>
<td></td>
<td>/u/</td>
<td>8</td>
<td>574.875</td>
<td>30.866</td>
</tr>
<tr>
<td></td>
<td>/a/</td>
<td>8</td>
<td>407.75</td>
<td>29.060</td>
</tr>
<tr>
<td>Mandarin</td>
<td>/i/</td>
<td>8</td>
<td>356.25</td>
<td>21.907</td>
</tr>
<tr>
<td></td>
<td>/i/ in /ie/</td>
<td>8</td>
<td>463.75</td>
<td>13.446</td>
</tr>
<tr>
<td></td>
<td>/i/ in /ie/</td>
<td>8</td>
<td>626</td>
<td>30.393</td>
</tr>
<tr>
<td></td>
<td>/e/ in /ei/</td>
<td>8</td>
<td>698.875</td>
<td>36.541</td>
</tr>
<tr>
<td></td>
<td>/i/ in /ei/</td>
<td>8</td>
<td>491.875</td>
<td>37.479</td>
</tr>
<tr>
<td></td>
<td>/a/ in /ai/</td>
<td>8</td>
<td>1024.5</td>
<td>60.321</td>
</tr>
<tr>
<td></td>
<td>/i/ in /ai/</td>
<td>8</td>
<td>829.5</td>
<td>78.569</td>
</tr>
<tr>
<td></td>
<td>/u/</td>
<td>7</td>
<td>494.857</td>
<td>16.077</td>
</tr>
<tr>
<td></td>
<td>/a/ in /ou/</td>
<td>8</td>
<td>678.5</td>
<td>28.835</td>
</tr>
<tr>
<td></td>
<td>/u/ in /ou/</td>
<td>8</td>
<td>570.125</td>
<td>27.885</td>
</tr>
</tbody>
</table>

Table 3. Mean F1, F2, and Duration of Vowels
The English speaker who produced the stimuli had very front [u]s and [ʊ]s.

Figure 2. Vowel Chart of English and Chinese Vowels

Before processing the discrimination and assimilation results, I calculated the Euclidean spectral distance and vowel duration difference between every pair of contrastive sounds and compared those two factors across the four contrasts. The reason for comparing the Euclidean spectral distance between contrastive sounds was to examine the hypothesis that the larger the acoustic distance between two sounds, the better the contrast will be discriminated. According to PAM, perceptual similarity is different from spectral similarity, and perceptual similarity is a better predictor than spectral similarity. This study tests the relationship between spectral similarity and discrimination results, then compares this relationship with the relationship between perceptual similarity and discrimination.
The reason for comparing the vowel duration difference is that Flege & Bohn & Jang (1997) suggested that subjects distinguished tense-lax contrasts by vowel length rather than spectral difference. The hypothesis was that the larger the vowel duration difference between two sounds, the better those two sounds would be discriminated. According to PAM, vowel duration is not as good a predictor as perceptual similarity either. The reason is that the similarity between a native and non-native sounds is decided by multi-dimensional factors rather than the vowel length itself. This study is going to test the relationship between vowel duration and discrimination accuracy in order to decide whether vowel duration or assimilation pattern is a better predictor of discrimination accuracy.

5.2 Comparison in Vowel Distance

The Euclidean distances between the sounds within contrasts were computed by the following formula. Suppose there is a token of /i/ and a token of /ɪ/. The equation of computing the spectral distance between /i/ and /ɪ/ is:

\[ \text{Euclidean Distance (i-ɪ)} = \sqrt{(F1(i) - F1(ɪ))^2 + (F2(i) - F2(ɪ))^2} \]

There were 8 tokens for each vowel in the discrimination test. Thus each contrast had 64 samples of Euclidean distance. The mean and standard deviation of the spectral distance within each contrast are presented in Table 4.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/-ɪ/</td>
<td>64</td>
<td>518.032</td>
<td>54.242</td>
</tr>
<tr>
<td>/u/-eɪ/</td>
<td>64</td>
<td>290.408</td>
<td>69.976</td>
</tr>
<tr>
<td>/e/-æ/</td>
<td>64</td>
<td>310.847</td>
<td>45.822</td>
</tr>
<tr>
<td>/u/-ʊ/</td>
<td>64</td>
<td>198.057</td>
<td>54.922</td>
</tr>
</tbody>
</table>

\(^1\) The F1 and F2 of /eɪ/ is the average of the nucleus /e/ and the offglide /ɪ/

Table 4. Descriptive Data of Euclidean Distance Between Sounds Within Contrasts
The Euclidean distances within contrasts were submitted to a one-way ANOVA. The Levene’s F test reveals that the homogeneity of variance assumption is not met (p = 0.001). The Welch’s F test is used: F(3, 138.749) = 386.187, p < 0.001, indicating that at least two contrasts had different duration differences. The post-hoc tests using Games-Howell tests indicate that the Euclidean distances of contrasts are in an order of /ɪ/-/ʊ > /ɪ/-/æ = /ɛ/-/æ > /u/-/ʊ.

5.3 Comparison in Vowel length

The differences in vowel length between the two sounds in each contrast were computed by the following equation using /ɪ/-/ʊ contrast as an example:

\[ \text{Length Diff (i~ʊ)} = \text{Length(ɪ)} - \text{Length(ʊ)} \]

As mentioned previously, each vowel was repeated 8 times, creating 64 samples of vowel length difference. The mean and standard deviation of the vowel length of each sounds in each contrast are illustrated in Figure 3.

![Vowel Length Comparison](image)

**Figure 3. Comparison of Duration Difference of Different Contrasts**
The vowel length differences between sounds within contrasts were submitted to a one-way ANOVA. The Levene’s F test reveals that the homogeneity of variance assumption is not met (p < 0.001). The Welch’s F test is used: F(3, 135.688) = 61.799, p < 0.001, indicating that at least two contrasts had different duration differences. The post-hoc test using Games-Howell test shows that there is no significant difference between /ɪ/-/eɪ/ and /ɛ/-/æ/. However, both of them have a larger duration difference than /i/-/ɪ/ while /i/-/ɪ/ has a significantly larger duration difference than /u/-/ʊ/. The duration differences within contrasts are in an order of /i/-/eɪ/ = /ɛ/-/æ/ > /i/-/ɪ/ > /u/-/ʊ/.
Part VI: Experiment Results

6.1 Experiment One: Discrimination Test.

The discrimination test examined how English and Mandarin speakers discriminate /i/-/ɪ/, /ɪ/-/eɪ/, /ɛ/-/æ/, and /u/-/ʊ/. Table 5 presents the accuracy and reaction time for the discrimination of each English vowel contrast by English and Mandarin speakers respectively. The discrimination accuracy is the percentage of trials where a subject made correct choice out of all trials. The reaction time is calculated from the end of stimuli to the moment when subjects made a response.

<table>
<thead>
<tr>
<th>L1</th>
<th>Contrast</th>
<th>N</th>
<th>Discrimination</th>
<th>Reaction time</th>
<th>Reaction time logged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean (%)</td>
<td>Mean (s)</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td>English</td>
<td>/i/-/ɪ/</td>
<td>30</td>
<td>97.81</td>
<td>0.583</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>/ɪ/-/eɪ/</td>
<td>30</td>
<td>97.97</td>
<td>0.550</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>/ɛ/-/æ/</td>
<td>30</td>
<td>97.86</td>
<td>0.611</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td>/u/-/ʊ/</td>
<td>30</td>
<td>98.33</td>
<td>0.570</td>
<td>0.168</td>
</tr>
<tr>
<td>Mandarin</td>
<td>/i/-/ɪ/</td>
<td>30</td>
<td>93.75</td>
<td>1.671</td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>/ɪ/-/eɪ/</td>
<td>30</td>
<td>71.41</td>
<td>2.483</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>/ɛ/-/æ/</td>
<td>30</td>
<td>75.26</td>
<td>2.525</td>
<td>1.097</td>
</tr>
<tr>
<td></td>
<td>/u/-/ʊ/</td>
<td>30</td>
<td>92.86</td>
<td>1.566</td>
<td>0.496</td>
</tr>
</tbody>
</table>

Table 5. Discrimination Accuracy and Reaction Time by English and Mandarin Groups

The results are divided by language and contrast, creating eight groups in total (4 contrasts * 2 languages). The discrimination accuracies and reaction times of those eight groups were submitted to a two-way ANOVA test, examining whether contrast and native language affect discrimination results.

6.1.1 Discrimination Results.

6.1.1.1 Effect of Language. A two-way between (language: Mandarin vs. English) by within (4 contrasts) ANOVA revealed that on average, English speakers
discriminate English vowels more accurately than Mandarin speakers \((F(1, 58) = 71.827, \text{PRE} = 0.553, p < 0.0001)\).

6.1.1.2 Effect of Contrasts. 12 planned comparisons were conducted to compare the effect of contrast within each language group. Using Bonferroni adjustment, the alpha level was adjusted to 0.004 \((0.05/12)\) in comparisons. The comparison statistics are shown in Table 6. As for the discrimination score, within English group, there is no significant difference between any two contrasts. Within Mandarin group, /i/-/ɪ/ and /u/-/ʊ/ contrasts receive significantly higher discrimination scores than /ɛ/-/æ/. There is no significant difference between /i/-/ɪ/ vs. /u/-/ʊ/ or /ɛ/-/æ/ vs. /ɛ/-/æ/. The discrimination accuracies for each contrast by each language group are illustrated in Figure 4.

![Figure 4. Comparison of Discrimination Accuracy Between Groups](image)

6.1.1.3 Effect of Euclidean distance and Duration. In order to examine whether the variations in the Mandarin speakers’ discrimination scores among contrasts were affected by the Euclidean distance and duration difference, the Mandarin speakers’ discrimination scores were regressed on the Euclidean distances, duration differences, and their interactions in a linear mixed model. The result shows that a larger duration difference predicts a lower discrimination score,
given $F(1, 26) = 4.84$, $p < 0.05$. Euclidean distance does not have a significant influence on the discrimination score, given $F(1, 26) = 0.116$, $p > 0.05$. And the effect of duration is not significantly influenced by the Euclidean distance of that contrast, given $F(1, 26) = 0.476$, $p > 0.05$.

6.1.2 Reaction Time.

6.1.2.1 Effect of Native Language. A two-way between (language: Mandarin vs. English) by within (4 contrasts) ANOVA revealed that on average, the reaction time for English speakers in the discrimination test is significantly shorter than for Mandarin speakers ($F(1, 58) = 70.074$, PRE = 0.547, $p < 0.0001$).

6.1.2.2 Effect of Contrast. 12 planned comparisons were conducted to compare the effect of contrast within each language group. Using Bonferroni adjustment, the alpha level was adjusted to 0.004 (0.05/12) in comparisons. Within the English group, /ɪ/-/eɪ/ had a significantly shorter reaction time than /ɛ/-/æ/ ($F(1, 58) = 9.973$, $p = 0.003$). There is no evidence of a significant difference between either of other two groups. Within the Mandarin group, /i/-/ɪ/ and /u/-/ʊ/ contrasts had significantly shorter reaction times than /ɛ/-/æ/ and /æ/-/æ/. There was no significant difference between /ɛ/-/eɪ/ vs. /ε/-/æ/ or /ɛ/-/æ/ vs. /ɛ/-/æ/. The mean reaction times are graphed in Figure 5. The F and p statistics of each planned discrimination accuracy score and reaction time comparison are presented in Table 6.

![Reaction Time Graph](image-url)
Table 6. F Statistics of Discrimination Score and Reaction Time ANOVA

6.1.2.3 Effect of Euclidean distance and Duration. In order to examine whether the variations in the Mandarin speakers’ reaction times among contrasts were affected by the Euclidean distance and duration difference, the Mandarin speakers’ reaction times were regressed on the Euclidean distances, duration differences, and their interactions in a linear mixed model. The result shows that a larger duration difference predicts a longer reaction time, given $F(1, 26) = 7.053$, $p < 0.05$. Euclidean distance does not have a significant effect on the reaction, given $F(1, 26) = 1.206$, $p > 0.05$. And the effect of duration difference on the reaction time
is not significantly influenced by the Euclidean distance of the contrast, given F(1, 26) = 1.605, p > 0.05

6.2 Experiment Two and Three: Identification and Rating Tests.

In the identification test, if a subject selected, for example, the Mandarin word [ti] when hearing English [ti], he/she was assimilating English /i/ (E/i/) into Mandarin /i/ (M/i/). For every target English vowel, the proportion that it was assimilated into each Mandarin vowel category was calculated. For example, 90.8% for the E/i/-M/i/ pair means that out of all the identifications that participants made for E/i/, 90.8% of time they selected M/i/ as the most similar Mandarin vowel to E/i/.

The mean goodness of each English vowel in each Mandarin vowel category was also calculated. For example, 7.504 for E/i/-M/i/ pairs means that on a scale of 1-9, where 1 means very different and 9 means very similar, E/i/ received a 7.504 score regarding to its similarity to M/i/. Both proportion scores (in percentage) and goodness scores (in italics) are presented in Table 7.

<table>
<thead>
<tr>
<th>Mandarin(M)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>English(E)</td>
<td>M/i/</td>
<td>M/ic/</td>
<td>M/ei/</td>
<td>M/ai/</td>
<td>M/u/</td>
<td>M/ou/</td>
<td></td>
</tr>
<tr>
<td>E/i/</td>
<td>90.83</td>
<td>8.75</td>
<td>0.42</td>
<td>0.00</td>
<td>20.00</td>
<td>80.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.504</td>
<td>4.775</td>
<td>3.617</td>
<td>3.208</td>
<td>4.942</td>
<td>5.779</td>
<td></td>
</tr>
<tr>
<td>E/i/</td>
<td>10.00</td>
<td>35.83</td>
<td>51.67</td>
<td>2.50</td>
<td>98.33</td>
<td>1.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.371</td>
<td>4.738</td>
<td>5.942</td>
<td>4.358</td>
<td>7.083</td>
<td>4.667</td>
<td></td>
</tr>
<tr>
<td>E/ei/</td>
<td>15.83</td>
<td>22.08</td>
<td>60.83</td>
<td>1.25</td>
<td>3.600</td>
<td>4.17</td>
<td></td>
</tr>
<tr>
<td>E/e/</td>
<td>0.83</td>
<td>4.17</td>
<td>17.50</td>
<td>77.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.600</td>
<td>4.133</td>
<td>5.250</td>
<td>6.983</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E/a/</td>
<td>0.00</td>
<td>4.17</td>
<td>17.50</td>
<td>77.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.354</td>
<td>3.396</td>
<td>4.358</td>
<td>7.338</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Assimilation Proportion and Goodness of English into Mandarin
6.2.1 Criterion of Categorization. Instead of using an arbitrary threshold when deciding whether one sound is categorized or not, this study adopted a new method, which has been mentioned in the literature review. For a target sound, if there is only one label received above chance proportion, that sound is focalized-categorized. If there are more than one label received above chance proportion, that sound is multiple-uncategorized. If there is no label having a significant above chance proportion, that sound is uncategorized. For English /i eɪ æ/, the chance level is 25% because each of them all have four options to assimilate into. For English /u ʊ/, the chance level is 50% because each of them have two options to assimilate into.

The advantage of this criterion is that the threshold of categorization is no longer arbitrary and multiple-category assimilation becomes possible. Using this criterion, in this study, English /i/ is categorized into Mandarin /i/, $F(1, 29) = 1011.113, p < 0.0001$. English /eɪ/ is categorized into Mandarin /ei/, $F(1, 29) = 83.652, p < 0.0001$. English /i/ is multiple-uncategorized into /iɛ/ and /ei/ significantly above chance (English /i/ to Mandarin /iɛ/: $F(1, 29) = 8.659, p = 0.006$; English /i/ to Mandarin /ei/: $F(1, 29) = 36.171, p < 0.0001$). English /ɛ/ and /æ/ are both categorized into Mandarin /ai/, $F(1, 29) = 180.983, p < 0.0001$. English /u/ is categorized into Mandarin /ou/, $F(1, 29) = 3752.175, p < 0.0001$.

6.2.2 Assimilation Patterns. The assimilation patterns of the attested English contrasts are as follows: /u/-/ʊ/ is a Two-Category (TC) contrast. /ɪ/-/i/ is a Multiple-Focalized contrast (MF) with /i/ being categorized into Mandarin /i/, and /ɪ/ being assimilated into both Mandarin /iɛ/ and /ei/. /ɪ/-/eɪ/ is also a Multiple-Focalized contrast (MF) with /i/ being categorized into two Mandarin categories, and /eɪ/ being consistently categorized into Mandarin /ei/. /ɛ/ and /æ/ are assimilated into the same
category /ai/. The average goodness rating scores of /ɛ/ and /æ/ by each person were submitted to a paired t-test. The result shows that there is no significant difference in goodness score between /ɛ/ and /æ/. F(1, 29) = 3.179, p = 0.085, suggesting that /ɛ/-/æ/ contrast makes a Single Category (SC) assimilation. The assimilation types of those contrasts are illustrated in Figure 6.

1Solid line indicates that the sound is categorized into a single category.
2Dashed line indicates that the sound is assimilated into multiple categories.

Figure 6. Assimilation Type of Each Contrast
Part VII General Discussion

Firstly, English speakers discriminate all the four target English vowel contrasts with generally equal accuracy and equal speed since as adults, speakers are supposed to be fully sensitive to their native phonetic categories. The evidence is that 1) English speakers obtain higher discrimination accuracy than Mandarin speakers in the discrimination accuracy of /ɪ/-/eɪ/, /ɛ/-/æ/, and /u/-/ʊ/, and 2) English speakers have shorter reaction times than Mandarin speakers in the discrimination of all four contrasts.

Secondly, the result of the discrimination test shows that Mandarin speakers do discriminate the four English contrasts with different accuracies. The discrimination accuracies are in a rank of /ɪ/-/eɪ/ = /u/-/ʊ/ > /ɪ/-/eɪ/ = /ɛ/-/æ/. The fact that the reaction times by Mandarin speakers are in the order of /ɪ/-/eɪ/ = /u/-/ʊ/ < /ɪ/-/eɪ/ = /ɛ/-/æ/ provides additional support for the discrimination rank. They together, suggesting that Mandarin speakers are able to distinguish /ɪ/-/eɪ/ and /u/-/ʊ/ better than /ɪ/-/eɪ/ and /ɛ/-/æ/. Thus, the first purpose of this study, which was to describe how well Mandarin speakers discriminate those four target vowel contrasts, is met.

Thirdly, the results indicate that Euclidean spectral distance or vowel duration difference are not accurate predictors of the discrimination accuracy. Firstly, neither does the Euclidean distance hierarchy nor does the duration difference hierarchy align with the discrimination accuracy hierarchy among contrasts. The Euclidean spectral distance rank is /ɪ/-/eɪ/ > /ɛ/-/æ/ > /u/-/ʊ/. Because the discrimination score of /u/-/ʊ/ is equal to /ɪ/-/eɪ/ and higher than /ɛ/-/æ/, it is not the case that a larger spectral distance guarantees a better discrimination. The duration difference order is /ɪ/-/eɪ/ = /ɛ/-/æ/ > /ɪ/-/eɪ/ > /u/-/ʊ/. However, the discrimination test yielded exactly an opposite rank, with the contrasts that had the least duration difference receiving the highest discrimination scores. Secondly, the regression of discrimination score and reaction time on duration difference shows
that a larger duration surprisingly hurts the discrimination accuracy. Meanwhile, the Euclidean distance does not have a significant influence on discrimination accuracy. Thus, there is not sufficient evidence to support the hypothesis that a larger durational difference between sounds in a contrast accurately predicts better discrimination.

Fourthly, the results of the assimilation test address the second research question: how do Mandarin speakers assimilate the seven target English vowels into Mandarin categories? Figure 5 illustrates how each attested English vowel is assimilated into Mandarin categories and what assimilation type those contrasts represent.

Lastly, with the rank of assimilation type being created, whether the results of this study match the predictions of PAM, in other words, whether assimilation is a good predictor of discrimination, could be tested. The answer is yes but with modifications of PAM. Using assimilation types for each contrast, the discrimination rank of the attested contrasts is TC (Two Category) = MF (Multiple-Focalized) (/i/-/ɪ/) > MF (/ɛ/-/ei/) = SC (Single Category). PAM’s prediction that if two sounds cross a native phonetic category boundary, then the contrast between those two sounds will be well discriminated is true given the fact that TC is more accurately discriminated than SC.

The multiple-assimilation contrasts, which are newly introduced into PAM by this study, do show up in the experiment. English /ɪ/ is multiple-uncategorized into Mandarin /ie/ and /ei/. The two contrasts involving English /ɪ/ are both Multiple-Focalized (MF) assimilation type. Their discrimination scores are significantly different from each other. The difference between those two contrasts is whether there is an overlap between the assimilation categories of the component sounds. The /i/-/ɪ/ contrast is better discriminated than /ɛ/-/ei/. For the /i/-/ɪ/ contrast, there is no overlap between the Mandarin categories that /i/ and /ɪ/ are assimilated into. On
the contrary, the assimilation categories of /ɪ/ and /eɪ/ overlap on Mandarin /ei/. A similar situation has been found in the aforementioned Guion et al. (2000)’s study: /ɹ/ in /ɹ-/w/ and /θ/ in /s-/θ/ were both “uncategorized” using a 75% categorization threshold. However, using the criterion of this study, those two “uncategorized” sounds were actually multiple-assimilated into Japanese categories. The other two “categorized” sounds were actually focalized-assimilated into a single Japanese category. The assimilation categories of those four sounds are illustrated in Figure 7, which shows that there was no overlap in the assimilation categories of /ɹ/ and /w/ while there was an overlap in the assimilation categories of /s/ and /θ/.

As expected, Guion et al. suggested that /ɹ-/w/ was significantly better discriminated than /s-/θ/. The results of the current study and Guion et al.’s study both suggest that for contrasts involving multiple-category assimilation, the overlap in the L1 assimilation categories predicts the discrimination accuracy of the L2 contrasts. MFs (Multiple-Focalized) with non-overlapped L1 assimilation will be well-discriminated like Two Category contrasts. MFs with overlapping L1 assimilation will be poorly-discriminated like Single Category contrasts.
Part VIII Conclusions

This study proposes that assimilation is a useful tool in predicting discrimination. For two given L2 contrastive sounds, if there is an overlap in the L1 categories that they are assimilated into, they will very likely to be poorly discriminated; if there are no such overlaps, they will very likely to be significantly better discriminated. This finding suggests that the native language magnet effect also influences L2 speech perception. Naïve L2 learners perceive L2 speech sounds by locating L2 sounds in their L1 phonological system.

More specifically, this study suggests some modification to the PAM methodology. Firstly, this study proposes a new criterion for deciding whether a given sound is categorized or not. Instead of using an arbitrary threshold of the categorization proportion, this study suggests counting the number of categories that are assimilated into above chance. A target sound is focalized-assimilated if it is assimilated into only one category above chance; multiple-assimilated if more than two categories are selected above chance; uncategorized if none of the categories are selected above chance. One of the advantages of this criterion is that it is not experiment-specific but can apply to any study testing PAM.

The other contribution of this study is that it incorporates multiple-category assimilation into PAM. The study results suggest that for contrasts involving one single-category assimilation sound and one multiple-category assimilation sound, if there is no overlap between the assimilation categories of two sounds, the discrimination of that contrast will be as good as a Two Category contrast. If there is an overlap between the assimilation categories, the discrimination is expected to cause more difficulty.

Finally, this study also has pedagogical significance. Knowing the degree of overlap in L1 assimilation categories predicts discrimination difficulty, L2 educators can easily assess what L2 contrasts will be challenging for discrimination
by investigating how L2 learners map L2 sounds into their L1. It will probably be helpful to design exercises differentiating L2 sounds from their L1 assimilation categories in order to create new phonetic categories for such L2 sounds, turning Single Category or overlapped Focalized-Multiple contrasts into TC contrast. L2 phonetic learning will consequently become more effective.
Bibliography


Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America, 139*(1), EL1-EL5. [https://doi.org/10.1121/1.4939608](https://doi.org/10.1121/1.4939608)


