Perceptual and Quantitative Analysis of Mandarin Loanword Phonology: Focusing on English Coda Nasal /m/

by

Xinyi Zhang

Yun J. Kim
Advisor

Linguistics

Yun J. Kim
Advisor

Weihua An
Committee Member

Hiram Maxim
Committee Member

2024

Perceptual and Quantitative Analysis of Mandarin Loanword Phonology: Focusing on English Coda Nasal /m/

by

Xinyi Zhang

Yun J. Kim
Advisor

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Linguistics

2024

Abstract

Perceptual and Quantitative Analysis of Mandarin Loanword Phonology: Focusing on English Coda Nasal /m/
By Xinyi Zhang

Crosslinguistically, loanword adaptation processes tend to align with phonotactic constraints, yet mismatches are also observed. At English coda, there are /m/, /n/, and /ŋ/, while only /n/ and /ŋ/ are allowed at Mandarin coda. The discussion on the adaptation of English coda nasal /m/ into Mandarin has primarily focused on the prenasal vowel quality, but the role of postnasal consonant features is less explored (Heish et al., 2008, Huang & Lin, 2016). This study examines whether postnasal consonant features complement prenasal vowels in explaining the adaptation process, while also investigating the potential interplay of other phonological, prosodic, and morphological factors in this process. Using a forced-choice paradigm, the study evaluates 105 English words with /Vm(C)/ sequences to determine the relative likelihood of two adaptation strategies—Vowel Epenthesis (VE) and Nasal Switch (NS)—with 10 bilingual Mandarin-English speakers. In this context, Vowel Epenthesis (VE) refers to listeners adapting /m/ by inserting an epenthesized vowel after /m/, whereas Nasal Switch (NS) involves replacing /m/ with /n/ or /ŋ/, which are phonemes available in Mandarin. The findings indicate that both prenasal vowels and postnasal consonants are critical in shaping phonological cues. Longer prenasal vowels and tense vowels are associated with increased Vowel Epenthesis, aligning with Mandarin's prosodic requirements for syllable timing. Conversely, bilabial postnasal consonants show a stronger tendency for Nasal Switch. Morphological boundaries also play a role, with increased favor of VE often used to take the morphological and mental division of a word into phonology. To analyze these effects, a Logistic Regression model was employed, incorporating six identified variables and their interactions. The results show that local phonological cues, such as the bilabialness and voicing of the immediate postnasal consonant, have stronger effects than prosodic or other non-phonological cues at word-level. Importantly, the study refines prior hypotheses that Vowel Epenthesis may only happen for /m/+[-labial] to /m/+[-bilabial]. These findings demonstrate that listeners weigh both local and word-level cues, balancing immediate cues with overarching linguistic structures. In this process, speakers rely on multiple cues rather than a single factor when adapting foreign phonology.

Perceptual and Quantitative Analysis of Mandarin Loanword Phonology: Focusing on English Coda Nasal /m/

By

Xinyi Zhang

Yun J. Kim
Advisor

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Linguistics

2024

Acknowledgements

First and foremost, I would like to express my deepest gratitude to all the professors on my committee for their invaluable guidance throughout the process of writing this Honors Thesis. I am especially grateful to Prof. Kim, my committee chair, for her mentorship and support in experimental phonology. She not only introduced me to the field of phonology but also opened doors to research that have shaped my academic path. I will always remember the numerous meetings and insightful advice she provided, not just for this thesis, but for all of my linguistic projects at Emory. Her time, dedication, and the way she helped me navigate both scholarly and personal challenges have been crucial in guiding my aspirations. I would also like to thank Prof. Maxim, whose ability to identify areas where I relied too much on intuition has been immensely helpful. He has been one of the longest-standing influences in my linguistic journey. I especially appreciate his guidance, both academic and personal, and feel very fortunate for the mentorship he provided ever since my first semester during the SURE program, which helped redirect me towards a path in scholarly research. Lastly, I am grateful to Prof. An for pushing me to think more statistically. Despite my initial fear that I might not do well, his guidance ultimately led us to a much stronger statistical explanation for this study.

Secondly, I want to thank all of my participants and speakers who recorded my stimuli for making this thesis possible. Especially Shirley and Kiki, my best friends, who were always by my side during my studies, brainstorming ideas with me and offering unwavering support.

Thirdly, I want to thank my parents for making my journey at Emory possible and giving me the opportunity to explore the world.

For someone who means so much to me, thank you for reassuring me in moments of doubt, celebrating every victory with me, and walking beside me through every twist and turn.

Lastly, I would like to extend my heartfelt gratitude to Emory Linguistics for the invaluable advice and opportunities I have received. The incredible support and close connections with the professors here have been truly inspiring. They gave me the confidence to take the most significant step in my life toward linguistics, reassuring me that I am not only capable of succeeding but also of embracing and enjoying the wonderfully “nerdy” spirit of the field.

To me from four years ago, I bet you never imagined standing here today. Thank you for holding on, for choosing what you love even when it meant facing struggles you should not have had to endure. I am so proud of you for achieving what once seemed impossible, for proving to yourself that you are capable of so much more than you ever dared to dream. Hopefully, you are now ready to not be who you should be, but who you can become. So, are you ready for the exciting journey ahead?

Table of Contents

1 Introduction **1**

2 Determinants of Loanword Adaptation **3**
2.1 Phonemic Quality . . . 3
2.2 Syllable structure . . . 3
2.3 Prosodic Feature . . . 4
2.4 Orthography . . . 4

3 Mandarin **5**

4 Methodology **7**
4.1 The Participants . . . 7
4.2 The Stimuli . . . 7
4.3 The Experiment . . . 9
4.4 The Questionnaire . . . 9

5 Results **10**
5.1 Postnasal consonant quality . . . 10
5.1.1 Postnasal consonant being [+labial] . . . 10
5.1.2 Postnasal Consonant being [+voice] . . . 13
5.2 Prenasal Vowel Quality . . . 14
5.2.1 General Prenasal Vowel Duration . . . 15
5.2.2 Tenseness and Laxness of the Prenasal Vowels . . . 16
5.2.3 Frontness of the vowels . . . 17
5.3 Syllabic Structure . . . 18
5.3.1 Syllable Counts . . . 18
5.3.2 Balancing Syllabic Duration . . . 19
5.4 Linguistic Boundary . . . 20
5.4.1 Morphological Boundary . . . 20
5.4.2 Syllabic Boundary (Postnasal consonant in the onset of the following syllable) . . . 21
5.5 Orthography . . . 22
5.6 Summary . . . 23
5.7 Modelling . . . 24

6 Conclusion and Discussion **28**
6.1 Hypothesized Outcomes and Conclusion . . . 28
6.2 Future Direction . . . 28

7 Conclusion **30**

Perceptual and Quantitative Analysis of Mandarin Loanword Phonology: Focusing on English Coda Nasal /m/

Xinyi Zhang

1 Introduction

In the development of human culture, contacts between different languages are inevitable. Loanwords help to borrow concepts from one language to another, allowing at least partial assimilation from the source language. Loanwords refer to words that are borrowed from one language and incorporated into another. When some sounds in the source language are missing from the target language, speakers may use what they have in the target phonological inventory to fill in the blanks of those that are missing. There are multiple strategies for adapting loanwords. One approach is insertion, where sounds are added to fit the phonotactic constraints of the borrowing language. For example, when hearing [ebzo], the Japanese would adapt it to [ebuzo] (Dupoux et al., 1999). Another strategy is deletion, where sounds that do not fit the target language, or sounds that are less salient are removed, as seen in the adaptation of English "all right" [ɔl ɾaɪt] to Japanese [oːɾai] (Shoji & Shoji, 2013). Additionally, loanwords can be adapted by maximizing perceptual similarity, where foreign sounds are replaced with phonetically similar native sounds. For example, "rug [ɹʌg]" can potentially be adapted to Korean in two ways: /ɹʌgɨ/ or /ɹʌk/. The fact that /ɹʌgɨ/ is the preferred adaptation demonstrates the tendency to make the English output and Korean input more similar (Kang, 2003).

Given the variety of strategies for adapting loanwords, one might wonder why it is not simply possible to match each sound in the source language with its exact counterpart in the target language. However, this direct mapping is often not feasible due to the differences in sound inventories between languages. For example, English has one voiceless alveolar fricative /s/, but Korean has two: lenis /s/ and fortis /s*/. When this happens, speakers typically map the foreign sound to the most perceptually similar sound in the target language. There are two main theories on how to evaluate this similarity. The phonological stance suggests that what is important is the sound's phonological categorization, rather than its phonetic details. Therefore, one-to-one mapping is not guaranteed (Hyman, 1970; Danesi, 1985; Paradis & LaCharité, 1997). For example, the English /p/ is adapted to the aspirated /p^h/ in Mandarin regardless of its aspiration in English. In contrast, the perceptual stance argues that phonetic details do matter (Silverman, 1992; Peperkamp & Dupoux, 2002). With this theory, aspirated /p^h/ and unaspirated /p/ in English, though both voiceless, would be adapted differently into Mandarin, as aspiration is a significant perceptual cue that differentiates the two sounds.

Building on the idea outlined in this chapter, it becomes clear that the adaptation of loanwords involves much more than simply filling in phonetic gaps with the closest available sounds in the target language. While the earlier discussion addressed the general mechanisms and monitoring systems involved in adapting loanwords, the discussion will now turn to a more detailed analysis of the various cues that impact this adaptation.

2 Determinants of Loanword Adaptation

2.1 Phonemic Quality

When discussing consonant quality, post-nasal devoicing (PND) is frequently mentioned in the literature (Solé et al., 2010; Beguš, 2019; Stanton, 2016). Researchers suggest that PND either enhances perceptual differences (Stanton, 2016) or arises as a side effect of other sound changes (Beguš, 2015). Additionally, anti-formants of labial sounds are typically lower than those of corresponding non-labial consonants, which exhibit higher anti-formants, as evidenced in Lithuanian (Jaroslavienė, 2019). Anti-formants are frequency bands in the speech signal where the energy is reduced or absent, typically associated with resonant structures like the nasal cavity. Higher anti-formants suggest higher energy and thus larger intensity. This could create an intensity gap with the surrounding sounds, which may be more or less perceptible depending on the context, with some gaps potentially being easier to recognize than others. Vowel quality, including features such as tenseness, height, backness, and roundedness, plays a crucial role in phonological processes, shaping both epenthesis patterns and phonation characteristics. Vowel epenthesis is a phonological process in which a vowel is inserted between consonants, typically to facilitate easier pronunciation or to conform to phonotactic rules of a language. Kang (2003) hypothesizes three environments where vowel epenthesis can occur, one of which involves the tenseness of the pre-final vowel epenthesis is more likely if the pre-final vowel is tense.

2.2 Syllable structure

The syllabic features of the target language can influence how loanwords can be borrowed. For example, Japanese does not allow consonant clusters except for palatalization at the onset, and no consonant clusters are allowed at the coda position. Consequently, vowels are often epenthesized to break down consonant clusters at the coda position. For example, English word "text" is adapted as /tekisuɯto/, where the /kst/ sequence in the coda of the

English word is resolved by epenthesizing /ɯ/ and /o/ (Kubozono, 1989).

2.3 Prosodic Feature

Prosodic features can provide another layer of cues above perceptions. For example, Mandarin exhibits syllable timing where maintaining equal syllable duration is crucial (Mok, 2009). Mandarin coda nasals vary in duration, leading to differences in vowel selection and syllabic structure depending on the type of nasal. Specifically, Mandarin /ŋ/ is about twice as long as /n/, and /ŋ/ in a Vŋ context is four times longer than V (Chen, 1972, 1975) . This is evident by the fact that the Mandarin transliteration of the English surname "King" as jīn.ēn (/tɕin.ən/), where the velar /ŋ/ is split into two syllables because of its overly long duration (Li, 2008). This strategy reflects Mandarin's prosodic need to maintain rhythmic regularity, where nasal durations are actively adjusted by vowel selection and syllabic structure to ensure consistent syllable timing. Duration itself also serves as a cue for adaptation. For instance, in Korean, the English /s/ with a longer duration is adapted as the fortis Korean /s*/, while shorter durations correspond to the lenis Korean /s/ (Kim & Curtis, 2000). The singleton /s/ in *song*, which is longer, is adapted to the tense /s*/ as in [s*oŋ]. /s/ in *star*, which is in a consonant cluster and thus shorter, is adapted as a lenis /s/ ([sɨtʰa]).

2.4 Orthography

Another factor that has an effect on loanword adaptation is orthography. Vendelin and Peperkamp (2005) demonstrate that online adaptation is significantly influenced by the presence or absence of written representation. Their study highlights the impact of orthography on the perception of English vowels /i/, /u/, /ɛ/, /æ/, and /ɪ/ in an experiment for French speakers to perceive English. Moreover, English orthography predicts the adaptation of Korean vowels, especially the Korean vowel /ɛ/ (mapped to graphemes 에 and 애), which initially had distinct phonetic representations /ɛ/ and /æ/ before merging into /ɛ/. The participants only behave differently for adapting English /ɛ/ and /æ/ when they see the

English orthography offered (Daland et al., 2015).

Transitioning from a general discussion of loanword adaptation strategies, the focus now shifts to Mandarin, where the discussion would showcase how these mechanisms play out in the context of Mandarin's phonological system.

3 Mandarin

Mandarin Chinese features five vowel phonemes: /i/, /y/, /u/, /ə/, and /a/. High vowels in Mandarin contrast in terms of [back] and [round], whereas the mid and low vowels do not (Heish et al., 2008). The language's consonant inventory includes 22 sounds, among which three are nasals: /m/, /n/, and /ŋ/. Notably, only /n/ and /ŋ/ can appear in the coda position, while /m/ is restricted to the onset. In contrast, English allows all three nasals, /m/, /n/, and /ŋ/, to appear in the coda position. There are no consonant clusters allowed in Mandarin.

Mandarin syllable structure can be described using the CGVX template, where C represents an initial consonant, G a glide, V a vowel, and X an optional final consonant or the second part of a long vowel or diphthong (Duanmu). In Standard Mandarin, if X is a consonant, it must be a nasal, specifically /n/ or /ŋ/. Focusing on how English nasals are brought into Mandarin, much research is condensed on the relationship between the prenasal vowel to the nasal and the word as a whole. Though speakers have faithfulness towards mapping nasals to their exact counterparts in Mandarin, not always, the adapted nasal and the preceding vowel can form an acceptable pair in Mandarin. For example, in the word *Monte Carlo*, even if the postvocalic nasal in *Monte* in an /n/, there is no such pair as /ɒn/, as a non-high vowel is not paired with a back nasal in Mandarin, and thus adapters may opt out for /ɒŋ/. The salient features of the vowel are thus the determiner of the phonotactically conflicting situations (Heish et al., 2008).

As for /m/, since it cannot appear at coda at all, when adapting English words with a coda /m/ to Mandarin, two main strategies emerge to accommodate the phonotactic constraints of Mandarin: (1) replacing /m/ with /n/ or /ŋ/, leveraging the shared nasal features of these phonemes in both languages, or (2) inserting a vowel after /m/, thereby moving it from the coda to the onset position, which aligns with Mandarin's phonological rules.

The consistency of the frontness/backness to the coronalness/dorsalness of the nasal is established to determine which nasal to use if the speaker chooses to replace /m/ with /n/ or /ŋ/ (Heish et al., 2008). For example, in *compost* /ɒm/, the prenasal vowel is a back vowel, thus the /m/ is adapted with a dorsal nasal /ŋ/, as in kang.po.si.te /ɒŋ/. As for *jam* [æm], since the prenasal vowel is a front vowel, the adapted form is with a coronal nasal, as in zhan [æn].

A significant determinant of the method of adaptation lies in the prosodic features of English and Mandarin. Mandarin is more sensitive to syllabic and moraic duration than English (Huang & Lin, 2019). Adapting /m/ to /n/ or /ŋ/, or inserting a vowel after /m/, affects the duration of syllables and moras in distinct ways. Mandarin speakers are likely to notice these changes. Focusing on m+nonlabial consonant clusters, as that is the prominent group where both nasal switch and vowel epenthesis happen, Huang and Lin (2016) found that monolingual speakers of Chinese would prefer switching /m/ to /n/ or /ŋ/, as avoiding vowel epenthesis to preserve the number of syllables they hear from the English stimuli (Huang & Lin, 2016), while bilingual speakers would prefer vowel epenthesis after /m/ as the labialness and nasalness of /m/ are preserved.

Though previous research reports that non-labial consonants are the only cases in which vowel epenthesis would happen, the specific reasons are not given. Although evidence from English is limited, post-nasal devoicing and lower formants of non-labial consonants, as mentioned before, suggest that postnasal non-labial consonants may possess features that significantly alter listener perception and undergo changes due to phonotactic constraints. The perceptual differences between labial and non-labial consonants could play a crucial role in their unique adaptability in loanword phonology.

Acknowledging these gaps, this study aims to test if postnasal consonant quality also influences the adaptation of English coda nasal /m/ to Mandarin, adding to the previously discussed prenasal vowel quality. Additionally, it seeks to reveal the factors involved in adapting coda nasal /m/s from English to Mandarin and if they interplay to influence this process.

4 Methodology

4.1 The Participants

10 participants who are bilingual speakers (L1=Mandarin, L2=English) who are studying in America as international students were recruited. They are all ethnically Chinese, with an average age of 21 years old.

4.2 The Stimuli

105 English words (96 stimuli, 9 filler words) were selected as the stimuli of this study. All English words contained a coda /m/ (e.g. lambda, bomb, impossible). The nine filler words are listed to filter out unqualified participants for this study. They are loanwords that are widely spoken and written in Chinese and serve as a threshold of fluency in Mandarin (e.g., California, pie, romantic). The other 96 stimuli words were carefully chosen based on the the features that are suggested to be influential in the past research.

Table 1: Summary of Potential Variables

Category	Details
Syllable Counts	One-syllable: 12; Two-syllable: 63; Three-syllable: 16; Four-syllable: 6
Mono-morphological	Mono-morphological: 16; Non-mono-morphological: 81
Resyllabification	Yes: 12; No: 85
Postnasal Consonant	
Labial-ness	[+labial]: 31; [-labial]: 60; No postnasal consonant: 6
Voicedness	[+voice]: 54; [-voice]: 36; No postnasal consonant: 7
Prenasal Vowel	
Height	High: 30; Mid: 40; Low: 27
Frontness	Front: 51; Mid: 30; Back: 16
Tenseness	Tense: 20; Lax: 77
Familiarity	Frequency counted by Baidu; Officially-documented translation: Yes/No

Two female native speakers of American English were recruited to record all 105 stimuli three times each. The audio was cleaned up using Praat and the clearest audio was selected for each stimulus. A pilot study was conducted using these stimuli to check the recordings' quality. The pilot study reported problems in the recordings, including the influence of the recorder's other native language, speed of utterance, clarity, etc.

The official stimuli, in order to avoid the problems detected from the pilot study, were recorded by a 22-year-old female native speaker of American English from the San Francisco Bay Area. She was instructed to read the list of stimuli at a clear, natural pace, with each word repeated twice. The recordings were made in a soundproof room, and all stimuli were standardized to 70 dB.

The average duration of the stimuli is 0.8s (sd=0.201), and they range from 0.2s to 1.8s. The average pitch of the stimuli is 202.2hz (sd=21.7), ranging from 96.6hz to 279.0 hz. The average intensity of the stimuli is 78.2 dB (sd=1.4), ranging from 74.8dB to 80.7dB.

Table 2: Summary of Acoustic Measures

Measures	Average	SD
Average Durations (s)	0.8	0.2
Duration Range (Min: Max) (s)	0.2: 1.8	
Average Pitch (Hz)	202.2	21.7
Pitch Range (Min: Max) (Hz)	96.6: 279.0	
Intensity (dB)	78.2	1.4
Intensity Range (Min: Max) (dB)	74.8: 80.7	

4.3 The Experiment

The experiment was generated using Gorilla (Pavliscak, 2021). It contains two major tasks: Typing Tasks and Multiple Choice. In the Typing Task, participants listened to the audios of the stimuli and then typed what they heard in Pinyin without labelling the tone. In the Multiple Choice Task, participants listened to a recording and chose which of the two possible options (one using Vowel Epenthesis and the other using Nasal Switch) was more similar to what they heard. The options were offered in a random sequence, and participants clicked 1 or 0 on the keyboard to make their choice. Before each section, an example page was provided to demonstrate how it works.

In the pilot study, participants found the Typing Task challenging and time-consuming, which led to confusion. Including half of the stimuli as Typing Task or Multiple Choice Task questions made the Typing Task excessively lengthy (the pilot study contained 10 typing tasks, whereas the actual word list would comprise 52). Consequently, a subset of 20% of the word list was chosen for the Typing Task. To ensure an unbiased selection, this subset needed to reflect the distribution of each variable in the full dataset, which includes 12 variables. Using Principal Component Analysis, an optimized subset of 20 words was selected.

4.4 The Questionnaire

The questionnaire consists of five sections: Ethnographic Questions, Introduction, Sample Questions, Typing Tasks, and Multiple Choices. The Ethnographic Questions gather self-reported years of English education and official English test scores to differentiate between deficient and fluent bilingual speakers. The Introduction provides a brief overview of the study's purpose without explicitly mentioning the targeted sound or what participants can expect during the study.

5 Results

The results will be presented through a discussion of the categories including postnasal consonant quality, prenasal vowel quality, syllabic structure, linguistic boundaries, and orthographic factors. It will address the significance of the adaptation choices, the consequential influence of each variable, and the justification for it. For simplicity, the choice of Vowel Epenthesis is labelled (VE), while the choice of Nasal Switch is labelled (NS).

5.1 Postnasal consonant quality

5.1.1 Postnasal consonant being [+labial]

Regarding the features of postnasal consonants, the primary question to address is whether the true determinant of Vowel Epenthesis (VE) lies in non-labialness, as previous studies have only examined the /b/ and /d/ contrast. To investigate this, a test was conducted to determine whether labials and non-labials indeed influence adaptation choices differently. An analysis of the stimuli based on whether the postnasal consonant is labial reveals that among 63 instances with a [-labial] postnasal consonant, 49 exhibit a statistically significant preference for VE (p-value = 0.0000001), which initially seems to support the claims of earlier studies. However, it is important to note that the argument asserting that VE occurs only when the postnasal consonant is non-labial also implies that, if Nasal Switch and Vowel Epenthesis are the only adaptation options, a labial postnasal consonant should exhibit a significant preference for NS.

Table 3: Summary of Nasal Duration and Choice Outcomes for Labial and Non-Labial Postnasal Consonants

Feature	**Choice**	**P-Value**	**Nasal Duration (s)**	**P-Value of Diff.**
[+labial]	NS	0.39	0.120	0.28
[-labial]	VE	0.0000*	0.129	

In practice, among 23 cases where the postnasal consonant is [+labial], 12 show a preference for NS, but this result is not statistically significant (p-value = 0.4). Furthermore, when examining the duration of /m/ in these two environments, the duration of /m/ before labials is 0.120s, which is not significantly different from its duration before non-labials (0.129s, p-value = 0.3). These findings suggest that labialness alone may not be sufficient to explain the observed differences in adaptation strategies.

Upon examining the [+labial] data, it was observed that variations primarily occur in cases where the postnasal consonant is specifically labiodental, such as /f/ or /v/. These account for 6 out of the 10 insignificant [+labial] cases. Furthermore, among the 9 cases where the postnasal consonant is labiodental, 7 exhibit a significant preference for VE (p-value = 0.008). This prompted an analysis to determine whether labiodentals should be excluded from the broader category of labials, leaving bilabials as the true determinant underlying the observed differences in adaptation choices.

To address this, a two-part question was posed: (1) Do labiodentals exhibit similar behavior to non-labials, given the significant VE outcomes observed in both groups? (2) Is there a significant in-group difference between labiodentals and bilabials?

Table 4: Summary of Nasal Duration and Choice Outcomes for Labiodental and Non-Labial Postnasal Consonants

Feature	**Choice**	**P-Value**	**Nasal Duration (s)**	**P-Value of Diff.**
[+labiodental]	VE	0.01*	0.128	0.83
[-labial]	VE	0.0000*	0.129	

First off, the average duration of /m/ in front of labiodentals is 0.128s, which is not significantly different from that of /m/ in front of non-labials, which is 0.134s (p-value = 0.8373). Add on the fact that the outcome of these two groups are also not significantly different, it is safe to conclude that labiodentals and non-labials have similar qualities on this issue.

Table 5: Summary of Nasal Duration and Choice Outcomes for Labiodental and Bilabial Postnasal Consonants

Feature	**Choice**	**P-Value**	**Nasal Duration (s)**	**P-Value of Diff.**
[+labiodental]	VE	0.01*	0.128	0.14
[+bilabial]	NS	0.0000*	0.112	

Second, the postnasal consonant being labiodental consistently yields significant VE outcomes, while bilabials result in significant NS outcomes. The choice outcomes between these two groups are therefore significantly different. Analysis of nasal duration does not indicate a significant difference, with the duration of /m/ before labiodentals being 0.128 seconds and before bilabials being 0.112 seconds (p-value = 0.14). Nevertheless, as the focus of this study is on the choice of adaptation method, the lack of significant durational differences does not preclude the possibility that labiodentals and bilabials perform differently in this context. Promising evidence supports the claim that labiodentals and bilabials may influence adaptation strategies differently.

Table 6: Summary of Nasal Duration and Choice Outcomes for Bilabial and Non-Bilabial Postnasal Consonants

Feature	**Choice**	**P-Value**	**Nasal Duration (s)**	**P-Value of Diff.**
[+bilabial]	NS	0.0000*	0.112	0.01*
[-bilabial]	VE	0.0000*	0.134	

To further verify whether being bilabial significantly influences the adaptation manner, the data reveal that out of 20 cases where the postnasal consonant is bilabial, only 2 yield significant VE outcomes, which is significantly lower than the 62 out of 76 cases observed when the postnasal consonant is non-bilabial (p-value = 0.000001). Additionally, the average duration of /m/ before bilabials is 0.112 seconds, which is significantly shorter than the 0.134 seconds observed before non-bilabials (p-value = 0.0106). These findings indicate that the postnasal consonant being bilabial or not exerts a significant influence, with Vowel Epenthesis occurring significantly less frequently before bilabials. Therefore, it can be concluded that the determinant of whether VE occurs lies in the distinction between bilabials and non-bilabials, rather than the broader labial versus non-labial contrast suggested in previous studies. This refines the distinction to a more specific one.

5.1.2 Postnasal Consonant being [+voice]

Table 7: Summary of Nasal Duration and Choice Outcomes for Voiced and Voiceless Postnasal Consonants

Feature	**Choice**	**P-Value**	**Nasal Duration (s)**	**P-Value of Diff.**
[+voice]	VE	0.09	0.140	0.0000001*
[-voice]	NS	0.0000*	0.106	

An analysis of whether the voicing of the postnasal consonant influences the results reveals significant findings. The average duration of /m/ when the postnasal consonant is [+voice] is 0.140s, significantly higher than the duration of /m/ when the postnasal consonant is [-voice], which measures 0.106s (p-value = 0.0000001). This indicates that the voicing of the postnasal consonant plays a crucial role in the articulation of /m/, suggesting that voiced postnasal consonants contribute to longer durations. However, postnasal consonant being [+voice] does not play a significant role in shaping the choice. Out of 76 cases where the postnasal consonant is voiced, 53 yields significant VE outcome, which is not significant to the proportion when postnasal consonant is voiceless (21 out of 36 cases, p-value=0.09). When the postnasal consonant is voiced, the duration of the nasal is significantly lengthened, although the direct influence of voicedness on the adaptation choice is minor.

5.2 Prenasal Vowel Quality

Prior research indicates that vowel position plays a crucial role in determining whether the nasal consonants /n/ or /ŋ/ are substituted for /m/ (Hsieh & Kenstowicz, 2008). Vowel epenthesis is particularly relevant in the context of long vowels or diphthongs in English (Huang & Lin, 2016). Furthermore, findings suggest that the contrast between tenseness and laxness in English prenasal vowels, along with nasalization, influences the selection of the geminate variant in syllable morphology (Huang & Lin, 2019). In this section, three analysis were conducted focusing on: 1) the general duration of prenasal vowels in relation to selection; 2) the distinction between tense and lax vowels; and 3) frontness of the vowels.

5.2.1 General Prenasal Vowel Duration

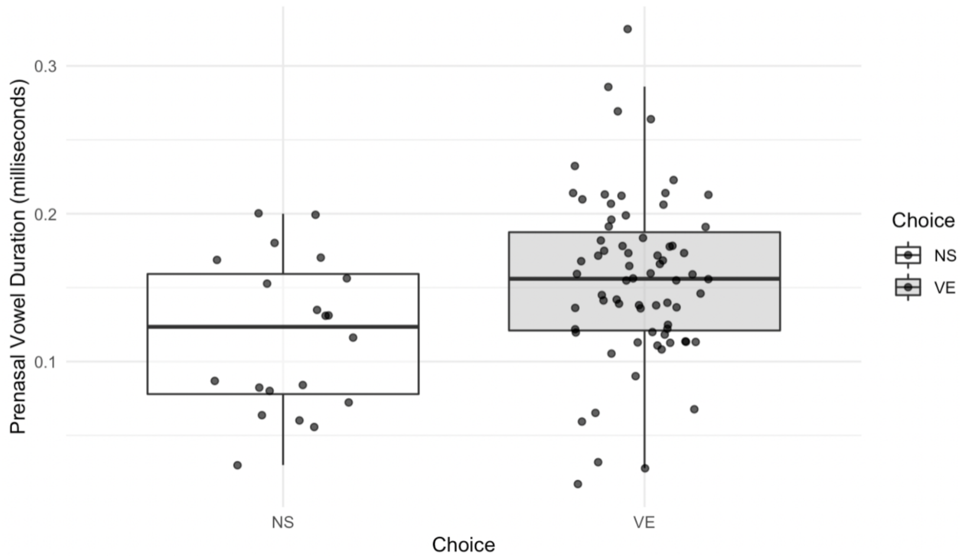

Figure 1: Distributions of Prenasal Vowel Duration for VE and NS

A preliminary analysis was performed to investigate the correlation between the duration of prenasal vowels and the duration of /m/. Figure 1 demonstrates that prenasal vowels leading to VE tend to have longer durations than those resulting in NS. Specifically, the average duration of prenasal vowels that generate VE outcomes is 0.156s, significantly exceeding the average duration of 0.118s for those generating NS outcomes (p-value = 0.002). For example, the prenasal vowel in *volume* /vɒlum/ has a duration of 0.3s and has significant VE outcome. While the prenasal vowel in *ambassador* /æmbæsədə/ has a much shorter duration of 0.03s, and it has a significant NS outcome. A specific discussion of the effect of prenasal vowel duration will be offered with the effect of syllabic duration.

5.2.2 Tenseness and Laxness of the Prenasal Vowels

Table 8: Summary of Nasal Duration and Choice Outcomes for Tense and Lax Prenasal Vowels

Feature	**Choice**	**P-Value of Diff.**	**Nasal Duration (s)**	**P-Value of Diff.**
[+tense]	VE	0.004*	0.129	0.4
[-tense]	VE		0.130	

In the analysis of 18 cases with tense prenasal vowels, all cases resulted in a significant VE outcome. In contrast, among 76 cases with lax prenasal vowels, 45 resulted in a significant VE outcome. A two-proportion z-test reveals that these two proportions are significantly different, indicating that tense prenasal vowels lead to a notably higher incidence of VE outcomes compared to lax prenasal vowels (p-value = 0.004). A comparison of the durations of prenasal vowels based on their tenseness and laxness was conducted. The average duration of tense prenasal vowels is 0.193s, significantly exceeding the average duration of lax prenasal vowels, which is 0.140 seconds (p-value = 0.000013). However, the average duration of /m/ after a tense vowel is 0.129s, which is not significantly shorter than that of /m/ after a lax vowel, which is 0.130s (p-value = 0.4).

The prenasal vowel in *ass<u>u</u>me* /əs<u>u</u>m/ is a tense one, thus it has significant VE outcome. In contrast, the prenasal vowel in *c<u>a</u>mp* /k^h<u>æ</u>mp/ is lax, and it has significant NS outcome.

Huang and Lin (2019) suggest that because English tense and lax vowels are mapped to different categories in Mandarin, they may influence the perception of nasals surrounding them in English. With this in mind, it is observed that tense vowels are significantly associated with the choice of VE, although their influence on the duration of /m/ is not significant. This is still consistent with Kang (2003) which suggests that vowel epenthesis is more likely to happen if the prior vowel is tense. While the duration of tense vowels is significantly longer than that of lax vowels, the difference may not proceed onto the duration of /m/.

5.2.3 Frontness of the vowels

The vowels in the dataset are either front (50), mid (30), or back vowels (16). Each group has 34, 16, 14 cases of significant VE, respectively. A Chi-Square test reveals that there is no significant difference among these three proportions ($\chi^2 = 5.556$, $df = 2$, p-value = 0.062). A one-way ANOVA is used to see if the average duration of /m/ is different after each of these three groups. Respectively, they each have an average nasal duration of 0.127s, 0.138s, and 0.120s. The ANOVA test reveals that the differences between group means are not statistically significant (p-value = 0.37). Even when central vowels are categorized as back vowels, the difference between proportions of significant VE outcomes between the two groups remain statistically insignificant (p-value = 0.37), as does their effect on nasal duration (p-value = 0.29). These findings collectively suggest that the frontness of prenasal vowels is unlikely to be a determining factor in the specific process examined in this study.

It might be initially surprising that the frontness of the prenasal vowels, a variable discussed at length in the previous paper, did not create a significant effect on either the choice or the duration of the nasal. This may be explained by the fact that the scope of this study does not involve how Nasal Switch picks alveolar or velar nasals, which, according to past research, is directly related to the frontness of prenasal vowels. Instead, it focuses on whether Nasal Switch would happen at all under specific situations. The effect of prenasal vowel frontness likely stays at the level of articulatory location, but not the level of prosody.

5.3 Syllabic Structure

5.3.1 Syllable Counts

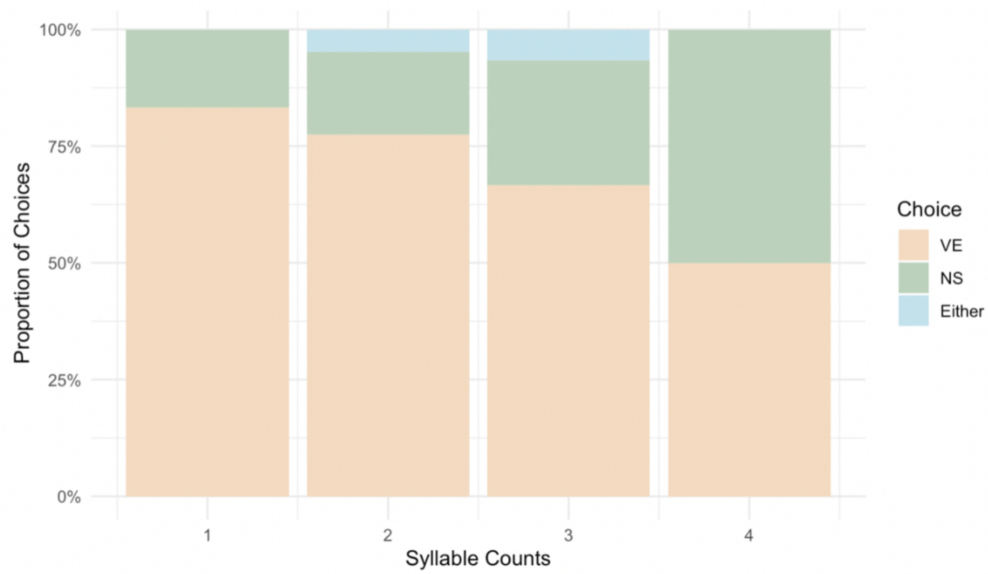

Figure 2: Proportions of Different Choices for Different Syllable Counts

The word list has words with one to four syllables, and, as shown in Figure 2, as the syllable counts increase, the proportion of VE decreases. This suggests that words with fewer syllables may be more likely to exhibit vowel epenthesis in their Mandarin-adapted form.

The trend of words with more syllables are more likely to choose VE over NS when adapting /m/ may be explained by the relationship between word length and the complexity of the adapted form. An increased syllable count implies a higher probability of consonant clusters occurring within individual syllables, compared to single-syllable words. In such cases, more additional vowel epenthesis may be required to maintain phonotactic well-formedness. If one were to satisfy all positions where VE may happen for a long word, that would inadvertently disrupt the syllabic resemblance between the adapted form and the original word compared to shorter words with fewer places to be possibly fixed. This suggests that bilinguals, when faced with longer source words, may also be more inclined to preserve the structure of the original word and minimize further epenthesis to maintain phonological similarity, even if this means deviating slightly from a perfect syllable match. These findings resonate with the broader understanding of adaptation strategies and syllable structure preferences in bilingual phonology.

5.3.2 Balancing Syllabic Duration

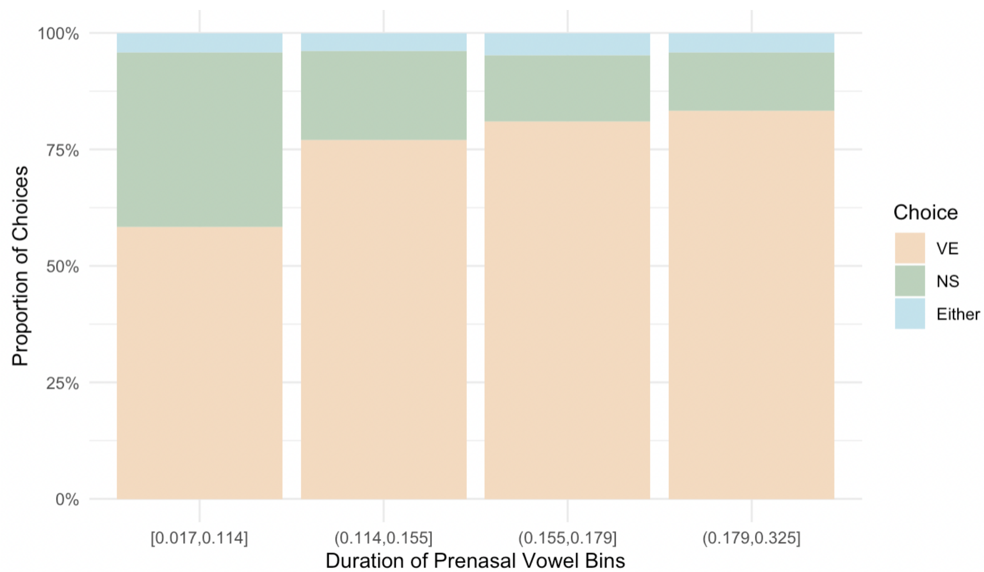

Figure 3: Proportions of Different Choices for Different Syllable Duration

The graph shows that as the duration of the prenasal vowel increases, the proportion of VE tends to increase while the proportion of NS decreases.

Another point discussed in the previous papers is the role of prenasal vowels and the selected nasal in regulating syllabic duration to achieve consistency across a word (Mok, 2009; Li, 2008). It can be posited that a longer prenasal vowel would be paired with a shorter nasal (/n/ or /ŋ/) in cases of NS, and vice versa. This raises the question: If a vowel is excessively long, could the addition of a nasal (/n/ or /ŋ/) further extend the syllable's duration, leading speakers to prefer Vowel Epenthesis instead to break the syllable into two and mitigate the overall length?

As suggested, Mandarin is a syllable-timed language (Mok, 2009). The duration of the prenasal vowel is a significant factor influencing both the duration of the nasal and the adaptation choice. Specifically, as the duration of the prenasal vowel increases, there is a notable shift towards more VE than NS. This effect may be closely related to the regulation of syllabic duration across a word (Huang & Lin, 2015; Li, 2006). In other words, when nasal switching occurs, a longer prenasal vowel is paired with a shorter nasal (likely /n/), and a shorter prenasal vowel is paired with a longer nasal (likely /ŋ/). As a result, the duration of the adapted /VN/ sequence remains balanced, neither overly long nor short. However, if the prenasal vowel is too long, adding an additional nasal to it may only be excessive. VE may break the original /VN/ into two syllables, thus the vowel and the nasal can appear in different syllables and not add to each other's duration, resulting in a structure like V+mV (or VN+mV, if assimilation occurs). This suggests the presence of a secondary process that regulates duration across syllables. The Typing Task data indicate that participants selected different vowels for epenthesis, even for the same word (e.g., Pinyin *me* or *mu*). This could suggest that participants are actively seeking different vowels, and thus of different durations to mitigate the duration of the /mV/ syllable - evidence that they are balancing the similar duration of /VN/ and /mV/ and thus indeed put the syllable timing in mind.

5.4 Linguistic Boundary

The analysis here aims to explore if morphological and syllabic separation would influence the participants' parsing of the relationship of /m/ to its surroundings.

5.4.1 Morphological Boundary

In a sample of 76 instances where the postnasal consonant occurs in a different morpheme, 47 cases yield a significant VE outcome. Conversely, among the 26 instances in which the postnasal consonant is not located in a different morpheme, 12 cases also exhibit a significant VE outcome. A two-proportion z-test indicates that the proportion of significant VE outcomes associated with postnasal consonants in a different morpheme is significantly higher than that observed when such consonants are in the same morpheme with /m/ (p-value = 0.003). For example, *em-barrass* /ɪmbæɹəs/, where /m/ and /b/ are in different morphemes, a significant proportion of VE outcome is yieled. As for *rampage* /ɹæmpʰeɪdʒ/, /m/ and /pʰ/ are within the same morpheme, and a significant proportion of NS is generated.

When the postnasal consonant is not in the same morpheme as the coda nasal /m/, there is significantly more VE than NS. The fact that participants demonstrated a tendency to use more VE in such cases suggests that people may mentally separate words based on their knowledge of English morphology and take that to their phonological perception. In other words, when /m/ appears at the end of the first morpheme and the postnasal consonant at the start of the next, participants may recognize this boundary, thereby emphasizing the presence and features of /m/ as distinctive to the postnasal consonant, and use the epenthesized vowel as a borderline between the /m/ and the postnasal consonant. According to Huang and Lin (2015), the best way to preserve the labial and nasal qualities of /m/ is through vowel epenthesis. Changing /m/ to a different nasal could result in a loss of alignment between the original word's mental representation and its adapted form. This may explain the observed preference for VE in this scenario.

5.4.2 Syllabic Boundary (Postnasal consonant in the onset of the following syllable)

Out of 13 cases where the postnasal consonant shares the same coda with /m/, 7 cases yield a significant VE outcome. In contrast, out of 76 cases where the postnasal consonant is in the onset of the following syllable, 52 yield a significant VE outcome. A two-proportion z-test shows that the difference between these proportions is not statistically significant (p-value = 0.16). Therefore, the position of the postnasal consonant at the onset of the next syllable does not significantly influence the choice. For example, for *prime.ly*, where /m/ and /l/ are not in the same coda, and *dreams*, where /m/ and /z/ are in the same coda, Vowel

Epenthesis happens for both cases.

Interestingly, syllabic boundaries did not produce the same effect as the morphological boundary. In fact, vowel epenthesis essentially disrupts the syllabic position that /m/ occupies. If a syllabic boundary exists between /m/ and the postnasal consonant, /m/ would be in the coda, which is the premise of this study. Adding a vowel after /m/ would shift it to the onset, which does not meaningfully distinguish the syllabic positions of /m/ and the postnasal consonant. Thus, while morphological boundaries prompt vowel epenthesis to emphasize certain features of /m/, syllabic boundaries do not have the same impact on adaptation choices in this context.

5.5 Orthography

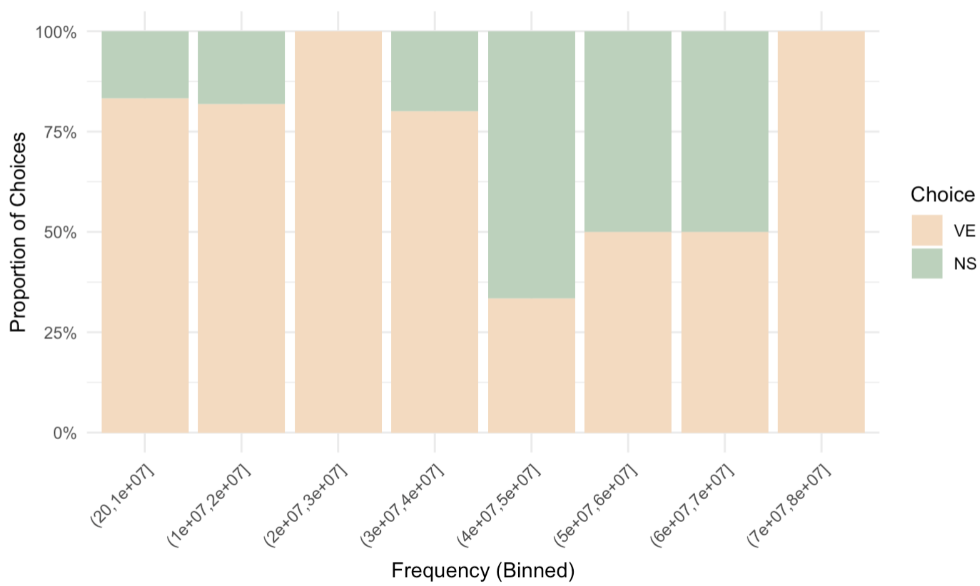

Figure 4: Proportions of Different Choices for Different Levels of Word Frequency

A bar graph illustrating the relationship between the frequency of a word on Chinese searching engines and the proportion of each adaptation strategy does not reveal a clear trend toward an increasing preference for either the VE or NS outcome. This suggests that familiarity with a word, that is, recalling of its orthography, does not necessarily influence the choice of adaptation strategy. While awareness of the presence of /m/ may lead individuals to prefer preserving its features through vowel epenthesis rather than replacing it, this does not translate into a consistent preference for VE over NS.

The influence of orthography is evident among certain individuals who consistently prefer VE over NS, particularly with more frequently encountered words. However, orthography does not have a significant influence on the adaptation choice. This phenomenon can be seen in instances where individuals struggle to accurately apply standard Pinyin rules, such as writing *am ba sa der* for "ambassador," where "am" is not permissible in Pinyin. They may access orthography to fail to pair the spelling of the words to a perfect sound in Mandarin. Additionally, some individuals may apply vowel epenthesis indiscriminately to all consonant clusters in the source English words, as seen in the representation *a er mu fu o* for "armful", showing that they recognized all consonant clusters in the words and focused solely on the existence of sounds but not the prominency of them. Furthermore, there are cases where individuals encounter consonants in spelling that either do not exist in pronunciation Or if the spelling is a consonant while the corresponding pronunciation is a vowel. For example, "gymnastics" is phonetically represented as *zhi yi mu na si di ke si* in Pinyin, where the original sound of "y" corresponds to the phoneme /ɪ/, yet it is reflected as "y" in the Pinyin spelling.

5.6 Summary

A summary of the previous findings is presented here. Table 9 includes all variables that either influence the choices, the duration of /m/, or both significantly.

Table 9: Summary of Influential Variables

Variable	**Choice**	**P-Value**	**Inf. Duration**	**P-Value**
[+*bilabial*]	NS	0.0000001^{***}	Negative	0.01^{*}
[+*voice*]	VE	0.09	Positive	0.0000001^{***}
Prenasal Vowel Duration (longer)	VE	0.002^{**}	NA	NA
[+*tense*]	VE	0.00000001^{***}	Negative	0.4
Syllable Count (more)	NS	NA	NA	NA
Separate Morpheme	VE	0.03^{**}	NA	NA

5.7 Modelling

In summary, there are six variables in total, with some influencing the duration of /m/ and, in turn, affecting the choice. The graph below shows that as the duration of /m/ increases, the likelihood of selecting the VE choice also increases, and vice versa.

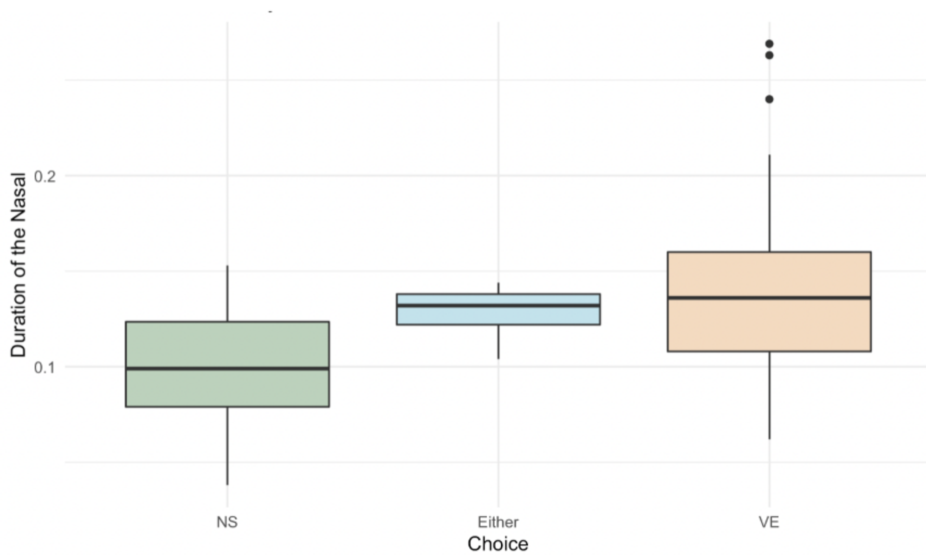

Figure 5: Distribution of the Duration of the Nasal across Different Types of Choices

Based on these interactions, I proposed three different models using Logistic Regression and validated the most optimal one. The first model only captures the individual effects of the six variables. Tense vowels tend to be longer than lax ones. However, with that durational difference in the prenasal vowel, the significant influence on the duration of the nasal is not obvious. Therefore, though prenasal vowel duration is dependent with [+tense], it is with doubt whether the interaction effect between prenasal vowel duration and tenseness is meaningful. The second model is thus proposed adding that interaction effect to the six individual variables. Additionally, the voicing of the postnasal consonant and its bilabial articulation can be considered as having an interactive effect, since these features combine to define key characteristics of the consonant, though they are not directly dependent. The last model will thus be testing if this interaction effect should be captured as well, adding to the interaction effect between prenasal vowel duration and tenseness, and the individual effect of the six variables.

The outcome of the models is a binary categorical variable called "Choice", where one level is "NS" and one level is "VE".

Model 1: No interaction effect

Variable	**coef**	**std err**	**z**	$P > \|z\|$	**[0.025**	**0.975]**
Intercept	-1.7576	2.712	-0.648	0.517	-7.072	3.557
Syllable_counts	0.1738	0.858	0.203	0.840	-1.508	1.856
C_is_bilabial	-2.1598	0.615	-3.514	0.000***	-3.364	-0.955
C_is_voiced	1.2210	0.569	2.145	0.032*	0.105	2.337
C_diff_morph	0.3237	0.501	0.646	0.518	-0.658	1.305
Duration_prenasal_vowel	13.5864	9.175	1.481	0.139	-4.396	31.569
Prenasal_Tenseness	-0.1405	1.295	-0.108	0.914	-2.680	2.399

Table 10: Regression output of Model 1 (* for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$)

Model 2: With interaction effect between vowel tenseness and vowel duration

Variable	**coef**	**std err**	**z**	$P > \|z\|$	**[0.025**	**0.975]**
Intercept	-2.1466	2.803	-0.766	0.444	-7.640	3.347
Syllable_counts	0.2337	0.861	0.271	0.786	-1.455	1.922
C_is_bilabial	-2.1402	0.617	-3.472	0.001***	-3.349	-0.932
C_is_voiced	1.2408	0.574	2.162	0.031*	0.116	2.366
C_diff_morph	0.3244	0.505	0.643	0.520	-0.665	1.314
Duration_prenasal_vowel	15.6641	10.015	1.564	0.118	-3.964	35.292
Prenasal_Tenseness	1.9141	3.935	0.486	0.627	-5.799	9.627
Duration_prenasal_vowel: Prenasal_Tenseness	-11.4030	19.724	-0.578	0.563	-50.062	27.256

Table 11: Regression output of Model 2 (* for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$)

Model 3: With interaction effect between both vowel tenseness and vowel duration, as well as between bilabialness and the voicing of the postnasal consonant

Variable	**coef**	**std err**	**z**	$P > \|z\|$	**[0.025**	**0.975]**
Intercept	-2.6471	2.965	-0.893	0.372	-8.458	3.164
Syllable_counts	0.2848	0.873	0.326	0.744	-1.426	1.996
C_is_bilabial	-2.2922	0.768	-2.983	0.003**	-3.798	-0.786
C_is_voiced	1.4960	0.808	1.852	0.064	-0.087	3.079
C_diff_morph	0.2572	0.528	0.487	0.626	-0.777	1.292
Duration_prenasal_vowel	17.0749	10.338	1.652	0.099	-3.188	37.338
Prenasal_Tenseness	1.8945	3.892	0.487	0.626	-5.734	9.523
Duration_prenasal_vowel: Prenasal_Tenseness	-11.5708	19.557	-0.592	0.554	-49.903	26.761
C_is_bilabial:C_is_voiced	0.4336	0.775	0.560	0.576	-1.085	1.952

Table 12: Regression output of Model 3 (* for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$)

A report of the MSE and log-likelihood of the three models is provided below:

Table 13: Model Performance Metrics

Model	**Log-Likelihood**	**MSE**
Model 1	-21.384	0.07208
Model 2	-21.233	0.07155
Model 3	-21.048	0.07211

Among the three models, Model 1 has the lowest Log-Likelihood and the second lowest MSE. Model 3 has the highest Log-Likelihood but the highest MSE. Meanwhile, Model 2 has the lowest MSE and the second highest Log-Likelihood. In summary, Model 2, the one that covers all individual effects and the interaction effect between prenasal vowel duration and tenseness, balances a good fit (second-best Log-Likelihood) with the best predictive accuracy (lowest MSE), making it the best overall model.

The outcome “VE” is coded as “1”, and the outcome of “NS” is coded as “0”. Among the predictors, the postnatal consonant being bilabial significantly decreases the likelihood of Vowel Epenthesis (p-value = 0.001), with an odds ratio of 0.12, indicating a strong negative effect. Conversely, a voiced consonant significantly increases the likelihood of Vowel Epenthesis (p-value = 0.031), with an odds ratio of 3.46, suggesting a substantial positive effect. This is consistent with the results discussed above.

Among the insignificant coefficients, there is a clear distinction in the size of the coefficients. Predictors with large coefficients, such as the duration of the prenasal vowel (15.66) and its interaction with prenasal tenseness (-11.40), suggest potentially meaningful relationships with the outcome (VE vs. NS) despite their insignificance. These findings may indicate underlying patterns obscured by variability or limitations in sample size, requiring further investigation in future research. Conversely, predictors with small coefficients, like syllable counts (0.23) and morphological boundaries (0.32), indicate weak associations and are likely to have minimal direct impact, as suggested by their high p-values. The small coefficients suggest that individuals may rely less on these variables when making their final decision, especially in the presence of other influencing factors. This trend implies a stronger focus on local phonological cues, with less emphasis on non-phonological or macro-level word features in determining outcomes.

In the chosen model, the interaction effect between the bilabialness and the voicedness of the postnatal consonant is not captured, indicating that these two variables operate individually to create a better influence on the choice. In other words, whether a consonant is bilabial or voiced does not appear to influence each other’s effects on the most optimal decision-making process. On the other hand, the chosen model captures the interaction effect between prenasal vowel duration and its tenseness, suggesting that people are likely to notice that tenseness influences the duration of prenasal vowels. However, the insignificant p-value for the interaction effect suggests that participants may prioritize the individual effects of prenasal vowel duration and tenseness over their interaction. This finding indicates that the influence of tenseness on prenasal vowel duration might serve as a secondary reference in this process. Once participants perceive the individual effects of prenasal vowel duration or tenseness, the interaction effect may function as a means of contextualizing the source of the duration differences, with tenseness contributing as part of the explanation.

6 Conclusion and Discussion

6.1 Hypothesized Outcomes and Conclusion

This study is based on the general hypothesis that features of the postnasal consonant, beyond just the quality of the prenasal vowel, also play a significant role in the adaptation of the English coda nasal /m/ to Mandarin. The results of the study report that bilabialness and voiceness of the postnasal consonants do influence the decision-making process. Consequently, it is hypothesized likely that multiple factors interact to influence the feature of /m/ and the adaptation choices. Results show that factors such as syllabic duration, syllable counts, and morphological boundaries are also at play. As tested, this study revealed that the significant feature of /m/ that reflects the factors in the surrounding environment is the duration of the coda nasal /m/, which also comes into place with other non-phonological factors.

6.2 Future Direction

One limitation of this study is that, due to the initial inclusion of a large number of variables, it was challenging to ensure that the distribution of variables—both across different categories and within the same category—was balanced. Upon examining the outcomes of various variables and their combinations, it became apparent that sometimes the sample sizes for each group varied significantly (e.g., 13 vs. 76). This disparity in sample sizes could raise concerns about the validity of significance tests. Additionally, since the study seeks to address findings from a previous paper suggesting that NS is generally more prevalent, we intentionally included more cases where VE was expected than NS in order to assess their significance. This decision may introduce a bias in the results, as the overrepresentation of VE cases could affect the overall significance of the findings, potentially skewing the comparison between NS and VE.

Further control of specific variables is necessary to draw more definitive conclusions. For instance, while there appears to be a trend suggesting that syllabic count may influence the adapted choice, it is not yet possible to assert this relationship with certainty. To gain a clearer understanding, it would be useful to control for the number of consonant clusters in a word, which shows all possible positions that permit vowel epenthesis in a word. By doing so, it would be easier to see if in cases where there are more syllables and possible consonant clusters allowing for vowel epenthesis, VE will be less preferred.

Another key limitation of this study is the lack of reverse verification, which could be addressed in future research. Specifically, the study does not have access to data that would allow for testing whether the variables' effect would be reflected in the adapted forms. For example, while it is hypothesized that a longer prenasal vowel would lead to more vowel epenthesis as a way of balancing syllabic duration across a word, future studies could collect voice data of the adapted forms so that the duration of each syllable can actually be measured. This would require recording speakers who are native in both English and Mandarin, allowing for reverse verification of the hypothesis and a more robust analysis of the findings.

Previous study mentioned an interesting contrast in the adaptation patterns between monolingual and bilingual Chinese. Monolingual speakers displayed greater sensitivity to syllable count equivalence between English words and their adapted forms, often favoring nasal substitution to avoid additional syllables. In contrast, bilingual speakers showed heightened sensitivity to the phonetic features of /m/, opting for vowel epenthesis to preserve its nasal and bilabial characteristics. However, due to limitations in the current study's setup, this comparison could not be captured. Future research aims to address this gap by incorporating monolingual Mandarin speakers residing in China to investigate their adaptation patterns more thoroughly.

7 Conclusion

In conclusion, the adaptation of English coda /m/ to Mandarin is influenced by a range of phonological, prosodic, and morphological factors. Both prenasal vowels and postnasal consonants play crucial roles in shaping key phonological cues. Additionally, listeners weigh both local and word-level cues when making adaptation decisions, balancing immediate phonological features with broader linguistic structures. The findings also suggest that speakers rely on multiple cues—rather than any single cue—in adjusting to foreign phonology, highlighting the complexity of phonological adaptation processes.

References

[1] Danesi, M. (1985). *Loanwords and phonological methodology.* Canada: Didier.

[2] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568–1578. https://doi.org/10.1037//0096-1523.25.6.1568.

[3] Beguš, G. (2019). Post-nasal devoicing and the Blurring Process. *Journal of Linguistics*, 55, 689–753.

[4] Hyman, L. (1970). The role of borrowings in the justification of phonological grammars. *Studies in African Linguistics*, 1(1), 1-48.

[5] Jaroslavienė, J. (2019). A brief analysis of the Lithuanian and Latvian nasal sonorants and their velar allophones. In ***Збрник матералв** III **круглого столу Сучасн тенденц фонетичних дослджень*** (p. 33). Institute of the Lithuanian Language.

[6] Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, 20(2), 219–273. https://doi.org/10.1017/S0952675703004524.

[7] Kandagor, M., & Sawe, S. K. (2014). Resyllabification of loan words in Kalenjin phonology. Retrieved from http://ir-library.kabianga.ac.ke/handle/123456789/574.

[8] Kim, T.-E. (2014). Preservation and deletion in Mandarin loanword adaptation. *International Journal of Chinese Linguistics*, 1(2), 214-243. https://doi.org/10.1075/ijchl.1.2.02kim.

[9] Kubozono, H. (1989). Japanese Tone Structure. *Language and Speech*, 32(4), 373-380. https://doi.org/10.1177/002383098903200405.

[10] Lehiste, I., & Fox, R. A. (1992). Perception of Prominence by Estonian and English Listeners. *Language and Speech*, 35(4), 419-434. https://doi.org/10.1177/002383099203500403.

[11] Peperkamp, S., & Dupoux, E. (2002). A typological study of stress 'deafness'. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 203–240). Berlin; New York: De Gruyter Mouton. https://doi.org/10.1515/9783110197105.1.203.

[12] Paradis, C., & Tremblay, A. (2009). Nondistinctive features in loanword adaptation. In *Loan Phonology* (pp. 211–224). https://doi.org/10.1075/cilt.307.09par.

[13] Shoji, S., & Shoji, K. (n.d.). Vowel epenthesis and consonant deletion in Japanese loanwords from English. *Proceedings of the Annual Meetings on Phonology*. University of South Carolina and Clemson University. https://doi.org/10.3765/amp.v1i1.16.

[14] Silverman, D. (1992). Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology*, 9(2), 289-328.

[15] Solé, M.-J., Hyman, L. M., & Monaka, K. C. (2009). More on post-nasal devoicing: The case of Shekgalagari. *UC Berkeley Phonology Lab Annual Reports*, 5. https://doi.org/10.5070/p70gk0g18d.

[16] Stanton, J. (2016). Predicting distributional restrictions on prenasalized stops. *Natural Language and Linguistic Theory*, 34, 1089–1133.

Appendix A: Stimuli Used in the Study

This appendix lists the stimuli and filler words used in this study, formatted in five columns for clarity.

Stimuli

1. volume	2. bomb	3. dreamed	4. dreams	5. assume
6. clamshell	7. scramble	8. primely	9. thumb	10. creams
11. warmth	12. condemn	13. armguards	14. amgarn	15. calmly
16. dumbs	17. camp	18. farmhouse	19. triumph	20. lambda
21. hamstring	22. armrack	23. tramcar	24. dreamt	25. omelet
26. ramshackle	27. ramjet	28. gamble	29. stems	30. timely
31. camrose	32. summed	33. uniformness	34. rampage	35. ample
36. hamlet	37. broomtail	38. firmware	39. hamster	40. armful
41. omnivores	42. camgirl	43. succumb	44. slimness	45. combo
46. dreamless	47. omniarch	48. chamber	49. amtrac	50. swimsuit
51. calmness	52. somnolent	53. clemson	54. dreamful	55. blameless
56. dumbly	57. harmful	58. solemn	59. dumbfound	60. stem-root
61. crumble	62. campfire	63. grumpy	64. crimson	65. grimgridder
66. teamwork	67. grimful	68. clumsy	69. columned	70. assumption
71. hymnbook	72. gumdrop	73. something	74. omnipotent	75. alumni
76. chumship	77. whimsical	78. remnant	79. chemtrail	80. emperor
81. bombard	82. accompany	83. comfort	84. locumship	85. circumradius
86. randomly	87. symphony	88. embrace	89. gymnastics	90. impossible
91. umbrella	92. circumduct	93. ambassador	94. embarrassed	95. circumvent
96. lambda				

Filler Words

1. California	2. pie	3. sandwich	4. email	5. aspirin
6. laser	7. London	8. romantic	9. mosaic	

Appendix B: Sample Questions from the Experiment

These texts are shown in Mandarin Chinese in the real experiment:

Typing Task

You will hear a word in English. Please use Pinyin to write down the sound that can adapt what you hear to Mandarin Chinese (no need to label the tones)

Please note, some sounds you hear may not be perfectly expressed by Pinyin, write down the closet Pinyin spelling if that happens.

Please type down your answer in the blank below and submit your answer to go to the next question. Please believe in your intuition. There is no need to spend too much time on each question. We value your first impression.

*Sample question:

Audio: dreamt

Sample answer: juan te

Multiple Choice Task

In this section, you will still listen to a word in English. You will choose a more similar form to the audio from the two options offered on the screen.

Press "1" on the keyboard if you think the option on the left side of the screen is more optimal, and press "0" on the keyboard if you think the option on the right is more optimal. You will automatically go to the next question as you click on the keyboard. Please believe in your intuition. There is no need to spend too much time on each question. We value your first impression.

* Sample options:

Audio: lambda

Options: 1. lan mu da 0. lan da

Appendix C: Background of the participants

Participant	Native Language	Age	Hometown	Current Location	English Proficiency*
A	Chinese	21	Wuxi, China	Santa Barbara, CA	8/10
B	Chinese	21	Jiaxing, China	Atlanta, GA	8/10
C	Chinese	21	Hangzhou, China	Atlanta, GA	9/10
D	Chinese	22	Xiamen, China	Atlanta, GA	8/10
E	Chinese	21	Zhangjiagang, China	Santa Barbara, CA	8/10
F	Chinese	21	Wuxi, China	Baltimore, MD	5/10
G	Chinese	22	Shizuishan, China	Lafayette, IN	6/10
H	Chinese	22	Hohhot, China	Washington, DC	7/10
I	Chinese	20	Tianjin, China	Santa Barbara, CA	8/10
J	Chinese	22	Beijing, China	Atlanta, GA	8/10

Table 14: Participant Information (*self-reported English proficiency)