Cue Integration in Mandarin Adaptation of English Coda /m/: Probabilistic and Perceptual Patterns

Xinyi Zhang
Cornell University

1 Introduction

Loanword adaptation provides insights into how listeners make decisions when presented with multiple phonetic and structural cues while mapping non-native segments to native categories. This setting naturally raises a question about how listeners use and weigh these cues: When multiple cues are present in loanword perception, do listeners treat them as simultaneously available inputs and depend on the most prominent one, or do they integrate them hierarchically within a probabilistic system? This paper addresses these issues by looking at how English coda nasal /m/s are adapted by Mandarin listeners. Mandarin permits /m/ at the onset but prohibits it at the coda, whereas English allows /m/ in both onsets and codas. This forces repair when English words with coda /m/s are adapted into Mandarin. As a result, English words with coda /m/ cannot be mapped faithfully and must be repaired. Two common repair options are (i) vowel epenthesis (VE), in which a vowel is inserted so that /m/ can be resyllabified as an onset (e.g., lambda→ [lan.mu.ta]; rum → [lɑŋ.mu])), and (ii) nasal switch (NS), in which /m/ is replaced with a licit Mandarin nasal coda, yielding an output with coda /n/ or /ŋ/(e.g., lambda → [lan.ta]; Angora → [an.kə.la]).

Prior works have mostly focused on how preceding vowel quality affects the adaptation choice. Hsieh et. al. (2009) argue that frontness/backness of the prenasal vowel systematically aligns with the place of the chosen nasal when the Nasal Switch happens, such that front vowels favor a coronal /n/ while back vowels favor a dorsal /ŋ/. For example, in compost /ɑm/, the prenasal vowel is a back vowel, thus the /m/ is adapted with a dorsal nasal /ŋ/, as in kang.po.si.te /ɑŋ/. As for jam [æm], since the prenasal vowel is a front vowel, the adapted form is with a coronal nasal, as in zhan [æn] (Hsieh et al., 2009). Huang and Lin (2016) argue that the adaptation strategy is sensitive to whether the coda /m/ is followed by a non-homorganic consonant, with VE associated with a non-labial postnasal consonant. In their experiment, the post-/m/ consonant environment is tested primarily through a /b/ versus /d/ contrast (e.g., Columbia → [kɤː.lun.pi.ja] vs. Walmsley → [wei.muː.sɤː.liː]).

What is still missing is an explicit account of how listeners combine prenasal vowel and postnasal consonant cues in the adaptation of coda /m/. Since prior work largely relied on a /b/-/d/ contrast, it remains relatively unclear whether the effect of place on adaptation strategies generalizes to other consonants. This study therefore looks at a wider range of post-/m/ consonants and evaluates how postnasal consonant properties interact with prenasal vowel quality in predicting the choice between VE and NS. Second, VE and NS are ultimately repairs forced by Mandarin syllable structure and phonotactics. /m/ is restricted to the onset position, and Mandarin does not permit consonant clusters under its (C)(G)V(X) syllable structure (Duanmu, 2000). These suggest that, apart from local cues, non-local, language-wide well-formedness pressures may also shape listeners' adaptation strategy. A perceptual stance in loanword phonology (Silverman, 1992; Peperkamp & Dupoux, 2002), which suggests that borrowers are sensitive to subphonemic acoustic cues in the source signal, also motivates tracking phonetic properties of the target segment itself. The duration of the target /m/ was included as a phonetic factor, following previous work arguing listeners' sole reliance on target duration in making adaptation choices (e.g., Kim & Curtis, 2000; Shirai, 1999). This predictor allows us to ask whether the adaptation strategy is driven entirely by the target /m/'s phonetic duration as a consequence of

Proceedings of *AMP 2025*

the surrounding conditions, or whether listeners rely on the existence of surrounding conditions themselves. Therefore, the study tests how listeners combine cues from the immediate environment and word-level structure during online adaptation of English coda /m/ in Mandarin.

2 Method

2.1 *Participants* Nine Mandarin–English bilinguals (L1=Mandarin, L2=English) were recruited from international students studying in the United States (mean age 21 y/o, 5F, 4M).

2.2 *Stimuli* The stimulus set contained 105 English words, including 96 target items and 9 fillers. All target words contained a coda /m/ (e.g., lambda, bomb, impossible). Fillers were familiar loanwords in Mandarin without a coda /m/, and were used to screen for basic Mandarin fluency (e.g., California, email). Target items were chosen to provide coverage of factors discussed in related prior works, as shown below:

Table 1: Target stimuli by design factors

Category	**Details**
Syllable Counts	One-syllable: 11; Two-syllable: 63; Three-syllable: 16; Four-syllable: 6
Monomorphemic	Monomorphemic: 15; Non-Polymorphemic: 81
Resyllabification	Yes: 12; No: 84
Postnasal Consonant	
Labiality	[+labial]: 31; [-labial]: 59; No postnasal consonant: 6
Voicedness	[+voice]: 53; [-voice]: 36; No postnasal consonant: 7
Prenasal Vowel	
Height	High: 30; Mid: 40; Low: 26
Frontness	Front: 50; Mid: 30; Back: 16
Tenseness	Tense: 20; Lax: 76
Familiarity	Frequency counted by Baidu

These stimuli were designed to target factors proposed to influence coda /m/ adaptation, including prenasal vowel duration, prenasal vowel tenseness, prenasal vowel frontness and height, postnasal consonant voicing and labiality, whether /m/ and the following consonant belong to the same morpheme, whether they form a coda consonant cluster in the source syllable, and item familiarity as a proxy for orthographic knowledge.

Recordings were produced by a native speaker of American English in a sound-attenuated setting and were amplitude-normalized (70 dB).

2.3 *Procedure* The experiment was implemented in Gorilla (www.gorilla.sc) and consisted of a Typing task and a Multiple-choice task.

In the Typing task, participants heard an English stimulus and typed the perceived form in Pinyin (tone not required). In the Multiple-choice task, participants heard an English stimulus and selected which of two Mandarin repair options (one in the form predicted by VE and one in the form predicted by NS); option order was randomized, and participants responded via keyboard press.

All participants completed the typing task first and the multiple-choice task second. This ordering was chosen to avoid showing the VE and NS options in the multiple-choice task before the typing task, so that participants are not primed with only the two fixed strategies in case any other strategy appears.

Based on pilot feedback that the typing tasks were time-consuming, only 20% of the word list was assigned to the Typing task; the subset was selected to reflect the distribution of the design variables in the full set. For VE options in the Multiple-choice task, /u/ was used as the epenthetic vowel, in consistency with the dominant epenthetic vowel observed in pilot typing tasks.

2.4 *Coding and analysis* Responses were coded categorically as VE or NS (typing outputs were also coded into the same two categories).

Statistical tests were used to test how post-/m/ consonant properties, prenasal vowel properties, and word-level structure each predicted the probability of a VE response. Mixed-effects logistic regression was used to analyze the adaptation outcome. The dependent variable was the adaptation choice, coded as VE = 1 and NS = 0; only responses that were significant by a one-sided binomial test were retained, and non-significant trials were excluded from the regression analyses.

3 Results

3.1 *Identified Predictors* Each factor was first evaluated in isolation. For each candidate factor, the stimuli were divided into a feature-present and a feature-absent group, and their effect on two outcomes was evaluated: (i) adaptation strategy and (ii) target /m/ duration. For each item, a one-sided binomial test was used to determine whether it yields significant preferences of VE or NS among all corresponding responses. Table 2 summarizes all candidate factors considered in the evaluation stage, along with the direction of each factor's effect on adaptation choice and, where applicable, on /m/ duration. Factors that showed reliable effects in these univariate comparisons are boldfaced in the table and were carried forward to the mixed-effects logistic regression reported in the next section.

Table 2: Summary of tested predictors and their effects on adaptation strategies and target /m/ duration

Variable	**Cue type**	**Choice**	**Effect on /m/ duration**
/m/ duration (longer)	—	VE	—
Postnasal consonant [bilabial]	Local	VE*	Positive*
Postnasal consonant [+voice]	Local	VE*	Positive*
Prenasal vowel duration (longer)	Local	VE*	—
Prenasal vowel [+tense]	Local	VE*	Negative
Prenasal vowel [+front]	Local	VE	Negative
Syllable count (more)	Word-level	NS*	—
Morph. boundary [m+C]	Word-level	VE*	Positive*
Syllabic boundary [m. C]	Word-level	VE	Positive
Orthography	Word-level	—	—

Note: An asterisk marks a statistically significant effect. Boldface marks predictors included in the fitted model.

The labiality effect tested and later entered into the model turns out to be bilabiality, not a [+labial]/[−labial] distinction. Following prior work, we first tested postnasal consonant labiality and observed a clear tendency on the [−labial] side. Items that have [−labial] post-/m/ consonant did show a significant preference for VE, which is consistent with past paper. However, the complementary [+labial] group did not show a significant preference toward NS. This asymmetry led us to break down "labials" into bilabials and labiodentals. In this breakdown, labiodentals actually pattern with non-labials in adaptation strategy, as they also yield significant VE. In contrast, bilabiality is concluded to be the true dividing line. [+bilabial] postnasal consonants yield significant NS, while [−bilabial] postnasal consonants yield significant VE. A detailed breakdown is provided in Section 4.2.

As seen in the table, the factors carried to the modelling stage are bilabiality of the postnasal consonant, voicing of the postnasal consonant, prenasal vowel duration, prenasal vowel tenseness, syllable count within the word, and the presence of a morphological boundary between the target /m/ and the postnasal consonant. Target /m/ duration was included as an additional predictor in the model to test whether listeners' VE versus NS decisions are dominated by phonetics, as expected if the adaptation process is driven solely by phonetic detail.

3.2 *Modelling Results* To evaluate whether the adaptation strategy is primarily driven by a small set of cues or instead a more balanced integration of cues, three mixed-effects logistic regression models were fit. Model 1 included all predictors listed above. Model 2 added the interaction between prenasal vowel duration and vowel tenseness, as tense vowels are generally longer than their lax counterparts. Model 3 further added the interaction between postnasal consonant bilabiality and postnasal consonant voicing, building on Model 2. The three models were compared using leave-one-out cross-validation. On each fold, the model was fit to eight subjects and used to predict the adaptation strategy for the left-out subject, and the performance was summarized with log loss. Under this criterion, Model 3 achieved the lowest mean log loss compared to the other two. Accordingly, Model 3 was selected as the best-fitting model for reporting and interpretation in the Results that follow. The results of Model 3 are shown in Table 3:

Table 3: Selected model with fixed-effect estimates predicting the probability of VE

Predictor	Coef.	SE	z	p-value	95% CI
Intercept	-0.2712	0.1569	-1.728	8.400e-02	[-0.5788, 0.0364]
Duration of /m/	5.4382	1.7629	3.085	2.037e-03	[1.9829, 8.8936]
C is bilabial	-2.1729	0.1741	-12.477	9.926e-36	[-2.5142, -1.8315]
C is voiced	1.3597	0.1741	7.808	5.823e-15	[1.0184, 1.7010]
C_is_bilabial: C_is_voiced	0.7054	0.1741	4.051	5.110e-05	[0.3641, 1.0467]
Prenasal Vowel Tenseness	-0.3030	0.3616	-0.838	4.021e-01	[-1.0117, 0.4058]
Duration prenasal vowel (longer)	2.6066	1.5626	1.668	9.530e-02	[-0.4562, 5.6693]
Duration_prenasal_vowel: Prenasal_Tenseness	0.8726	0.4303	2.022	4.257e-02	[0.0292, 1.7160]
Syllable counts	0.6548	0.0777	8.424	3.629e-17	[0.5024, 0.8071]
C_diff_morph	0.3250	0.1741	1.866	6.200e-02	[-0.0163, 0.6663]

3.3 *Model Interpretation* The selected model above shows a good fit for the current dataset. In leave-one-subject-out cross-validation, it achieved a low mean log loss (0.214), meaning that the model usually assigns high probability to the strategy that participants actually chose. This is far better than a 50/50 predictor (log loss 0.693), indicating good predictive performance on held-out participants.

In the selected model, Choice was coded as VE = 1 and NS = 0, so positive coefficients indicate higher odds of a VE response. Target /m/ duration does indeed act as the strongest single predictor in the model ($\beta = 5.438$, $p = 0.002$), such that longer nasal duration was associated with substantially higher odds of choosing VE. This pattern suggests that listeners do rely heavily on a direct measure of nasal duration when selecting an adaptation strategy. At the same time, several effects remain substantial in size rather than being extremely small.

The next largest predictor is prenasal vowel duration ($\beta = 2.607$), such that longer prenasal vowels are associated with higher odds of VE, although this effect is below significance ($p = 0.095$). The vowel-duration-by-tenseness interaction is also reliable ($\beta = 0.873$, $p = 0.043$). Postnasal consonant bilabiality also shows a large effect ($\beta = -2.173$, $p < 1e^{-35}$), meaning that VE becomes less likely as the postnasal consonant is bilabial, while postnasal voicing shows a strong positive effect ($\beta = 1.360$, $p < 1e^{-14}$), meaning that VE becomes more likely when the postnasal consonant is voiced. The bilabiality-by-voicing interaction is significant as well ($\beta = 0.705$, $p = 5.11e^{-05}$), indicating that the voicing effect is not uniform and voicing increases VE more strongly when the postnasal consonant is bilabial than when it is non-bilabial. Syllable count contributes pretty significantly ($\beta = 0.655$, $p < 1e^{-16}$), with more syllables corresponding to higher odds of VE in this dataset. The remaining terms are smaller and less significant in this dataset. Vowel tenseness shows a smaller independent effect ($\beta = -0.303$, $p = 0.402$). Finally, the morphological-boundary factor shows a marginal effect as well($\beta = 0.325$, $p = 0.062$), suggesting that when the postnasal consonant is not in the same morpheme as the target /m/ (that is, when /m/+C spans a morpheme boundary), VE responses tend to be more likely, although the evidence is not strong enough here to treat it as a stable effect.

4 Discussion

4.1 *Local cues, word-level structure, and cue integration* Taken together, the results support a split between local cues adjacent to the target segment and broader word-level structure. The largest effects come from cues in the immediate /m/ environment, specifically, prenasal vowel duration, postnasal bilabiality and voicing; while word-level factors contribute more modestly, with factors such as syllable counts and morphological boundary information ranking lower. This supports the idea that listeners make adaptation decisions primarily from information that is available locally in the acoustic signal and its immediate segmental context, with broader structural pressures shaping preferences only secondarily. This ranking is shown in Figure 1.

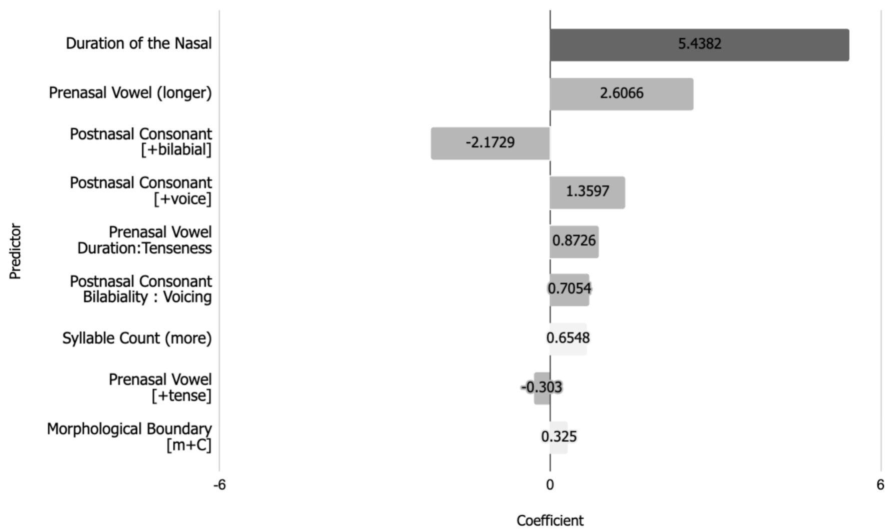

Figure 1: Coefficient magnitudes in the selected model

Previous work already suggests that English coda /m/ in Mandarin adaptation is not repaired in a uniform way. The present results contribute to the literature by clarifying how these sources of information line up when they are evaluated together. The fitted model does not point to a single decisive conditioning factor. Instead, it suggests a weighted system in which cues in the immediate /m/ environment carry the most explanatory load.

This pattern speaks to the debate over the phonetic versus phonological stance in loanword phonology. If repair were mainly driven by Mandarin well-formedness, then once coda /m/ is identified as illegal, the choice between VE and NS should mainly follow phonological pressures such as syllable-structure constraints and faithfulness to word shape, since Mandarin does not allow /m/ in a coda. Instead, in the fitted model, predictors tied to the local /m/ environment have larger coefficients than these word-level pressures.

At the same time, the pattern is not purely phonetic. While the target nasal duration does show the largest effect, it co-exists with other factors that also contribute to strategy choice. A closer look at the experimental results reveals multiple examples where the target /m/ of the shortest durations yields VE instead of NS. For example, the stimulus *hamster* has the fifth shortest target /m/ duration in the dataset (0.072s), which would lead a duration-only account to expect NS, but it is instead significantly adapted with VE. These kinds of exceptions show that a single perceptual stance cannot separate VE and NS choices cleanly. One way to capture this is to treat adaptation as a single decision process in which multiple cues are combined, rather than assuming that listeners first settle on a fixed perception and then apply repair rules separately. Cues in the local /VmC/ environment are the main basis for strategy choice, while broader, well-formedness pressures in Mandarin contribute less strongly.

4.2 *Postnasal consonant effects* In the current analysis, the postnasal consonant effect is specified in terms of bilabial place. With a wider range of following consonants, the dataset shows significantly that vowel epenthesis is more likely when the consonant after the target /m/ is non-bilabial, while nasal switch is more likely when the consonant after the target /m/ is bilabial. Earlier work described the relevant context with a [±labial] feature, which fits a /b/ vs. /d/ contrast in their stimuli but does not distinguish bilabials from other labials. In a feature-geometry view, an /m + bilabial C/ sequence can be represented as sharing the same LABIAL place across the two consonants. VE is likely disfavored in this context because inserting a vowel between /m/ and the following bilabial breaks the C-C adjacency that supports place sharing and forces the two consonants to be realized with separate place associations. The current grouping also shows that bilabiality, not [+labial], decides the pattern: under a LABIAL, bilabials and labiodentals should pattern together because both are LABIAL in standard feature geometry, but labiodentals in fact pattern with non-labials instead of with bilabials in the current dataset. A detailed breakdown is shown below:

Table 4: Comparison of adaptation strategies shared by different labial features

Feature	Adaptation Strategy	Binomial P-Value	Z-Test P-Value
Comparison of Labiality			
[+labial]	NS	0.4	0.0037**
[-labial]	VE	0.00001****	
Comparison Within [+labial] Category			
[+bilabial]	NS	0.00001****	0.00001****
[+labiodental]	VE	0.02*	
Labiodental vs. Non-Labial			
[+labiodental]	VE	0.02*	0.221
[-labial]	VE	0.00001****	
Bilabial vs. Non-Bilabial			
[+bilabial]	NS	0.00001****	0.00001****
[-bilabial]	VE	0.00001****	

This pattern suggests that listeners treat the place difference between a bilabial nasal (/m/) and a following labiodental consonant (/f, v/) as meaningful in this mapping, as the two group yield significantly different adaptation strategies, and the VE rates are significantly different (Z-Test P-Value= 0.003). If the process only used the native major place class LABIAL, then /f, v/ should behave like /p, b/, but they do not, and instead the VE rates of the labiodental group and the non-labial group are not significantly different. This means the conditioning factor needs a finer place contrast within LABIAL (bilabial vs. labiodental), not just [±labial], in this adaptation process.

4.3 *Prenasal vowel effects* In the current analysis, we can interpret the prenasal vowel tenseness effect in moraic terms. This aligns with Huang and Lin's idea that Mandarin loanword outputs tend to favor a bimoraic syllable. In their account, "bimoraic" means that the adapted syllable is usually heavy instead of light. The quality of the vowel can influence how much weight it adds (Huang & Lin, 2019). In this view, a lax prenasal vowel is more likely to be considered lighter. So, keeping a nasal in the rhyme through NS can help create a bimoraic syllable. On the other hand, a tense prenasal vowel is more likely seen as heavier. This makes VE more suitable because it moves /m/ out of the rhyme instead of adding more weight to the rhyme. What matters for this tenseness predictor in the current dataset is also its duration. Mandarin listeners often rely heavily on vowel duration to differentiate between English tense and lax contrasts (Wang & Munro, 1999). Therefore, "tense" in the present data can be seen as a representation of longer prenasal vowel duration. This leads us to consider prenasal vowel duration as the more direct variable. The longer the prenasal vowel is, the more likely the adaptation choice is to be VE, which is also consistent with Huang and Lin (2016). Add on the fact that Mandarin is a syllable-timed language (Mok, 2008) and that perception research shows listeners can track and compare timing patterns in speech (White, Mattys & Wiget, 2012; Smit & Rathcke, 2024), this timing pressure may make VE preferred because it creates an extra syllable slot, spreading the perceived duration over two syllables rather than concentrating a long interval in a single syllable.

4.4 *Word-level structure effects*

4.4.1 *Syllable counts* For the relationship between syllable count and the adaptation strategy, more syllable counts leads to less VE. Source words with more syllables have more potential spots where consonant clusters can cause Mandarin phonotactic violations; therefore, VE is often necessary at certain points in the adaptation. However, applying VE at every potential spot in a longer word would introduce too many extra syllables and thus induce repeated violations of faithfulness to the source word's syllabic count. The pattern that VE becomes less frequent as syllable count increases, despite longer words containing more potential consonant-cluster violation sites that could in principle be repaired by epenthesis, suggests that the intuitively available epenthesis repair is not applied at every potential site in longer words. This pattern suggests a kind of selective repair, in which VE potentially addresses the most marked phonotactic violations, while NS is used at other spots to minimize the number of syllables added. Limiting VE in longer words aligns with the Minimality Principle, which favors repairs that involve as few repair steps as possible (Paradis, 1996). In this way, VE involves more repair steps because it changes two things at a time: inserting a vowel and changing the nasal's syllabic position. While NS only changes the nasal place at the coda. Inserting vowels at every possible site would therefore require many repair steps, whereas mixing VE with NS keeps the overall amount of repair smaller.

4.4.2 *Linguistic boundaries* When we compared the types of boundaries observed for coda /m/ and the following consonant, a division was established. If /m/ and C belong to different morphemes, there is a significant tendency towards VE. In contrast, there is no such effect of syllable boundary, that is, whether /m/ and C belong to the same or different syllables does not increase the VE rate. One way to explain this is that VE can preserve the source word's morpheme boundary, because the inserted vowel acts as a clear divider between /m/ and the following consonant and keeps them separated in the output. Importantly, inserting a vowel does not change the fact that /m/ and the following C belong to different morphemes in the source word, so the morpheme boundary remains intact. Syllable-boundary knowledge is represented differently. In this case, it can be represented as the listeners' knowledge that /m/ is in coda position and the following C is in onset position, which suggests a syllable boundary between them. VE does not preserve this same cue, because inserting a vowel resyllabifies /m/ as an onset, so the original coda–onset relation that implied a boundary between /m/ and C is no longer maintained.

5 Conclusion

In this study, we aim to answer how Mandarin listeners use multiple sources of information simultaneously when adapting the English coda /m/ and what factors determine whether they use vowel epenthesis or nasal switch as an adaptation strategy. The results show that postnasal consonant properties are necessary to explain the adaptation choices, and that vowel properties are not sufficient on their own, because the model also shows different degrees of effect of other phonetic or word-level factors. More generally, the findings support a view of cue-integration in which some local cues exert stronger pressure on the decision than weaker non-local structure. This contrasts with a view of adaptation as driven by a single categorical repair rule. The probabilistic modelling we employed in this paper makes this combination explicit by estimating how each cue contributes to the likelihood of Vowel Epenthesis versus Nasal Switch.

6 Appendix

This appendix lists the stimuli 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				

References

Beguš, Gašper. (2019). Post-nasal devoicing and the Blurring Process. *Journal of Linguistics*, *55*, 689–753. https://doi.org/10.1017/S002222671800049X

Chen, Matthew (1972). *Nasals and nasalization in Chinese: Explorations in phonological universals* [Doctoral dissertation, University of California, Berkeley].

Chen, Matthew (1975). An areal study of nasalization in Chinese. *Journal of Chinese Linguistics*, *3*(1), 16–59.

Daland, Robert, Oh, Mira, & Kim, Syejeong. (2015). When in doubt, read the instructions: Orthographic effects in loanword adaptation. *Lingua*, *159*, 70–92. https://doi.org/10.1016/j.lingua.2015.03.0

Duanmu, San. (2000). *The Phonology of Standard Chinese*. Oxford University Press.

Dupoux, Emmanuel, Kazuhiko, Kakehi, Hirose, Yuki, Pallier, Christophe, & Mehler, Jacques (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

Hsieh, Feng-fan, Kenstowicz, Michael, & Mou, Xiaomin (2009). Mandarin adaptations of coda nasals in English loanwords. In A. Calabrese & W. L. Wetzels (Eds.), *Loan Phonology* (pp. 131–154). John Benjamins. https://doi.org/10.1075/cilt.307.05hsi

Huang, Ho-Hsin, & Lin, Yen-Hwei (2016). To epenthesize or not? Adaptations of English coda [m] in Standard Mandarin loanwords. *Proceedings of the Annual Meetings on Phonology*, *3*. https://doi.org/10.3765/amp.v3i0.3686

Huang, Ho-Hsin, & Lin, Yen-Hwei (2019). Vowel quality cues to variable nasal adaptation in Mandarin loanword phonology. *Proceedings of the Annual Meetings on Phonology*, *7*. https://doi.org/10.3765/amp.v7i0.4486

Hyman, Larry (1970). The role of borrowings in the justification of phonological grammars. *Studies in African Linguistics*, *1*(1), 1–48.

Jaroslavienė, J. (2019). A brief analysis of the Lithuanian and Latvian nasal sonorants and their velar allophones. *Proceedings of the III Round Table 'Current Trends in Phonetic Research'*, 33. Institute of the Lithuanian Language.

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

Kim, Soohee, & Curtis, Emily (2000). Phonetic duration of English /s/ and its borrowing into Korean. In N. M. Akatsuka, S. Strauss, & B. Comrie (Eds.), *Japanese/Korean Linguistics 10* (pp. 406–419). CSLI Publications.

Kim, Tae-Eun (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

Lehiste, Ilse., & Fox, Robert Allen (1992). Perception of prominence by Estonian and English listeners. *Language and Speech*, *35*(4), 419–434. https://doi.org/10.1177/002383099203500403

Li, Ya (2008). *Mandarin Speakers' Production of English and Mandarin Nasal Codas* [Master's thesis, University of Victoria].

Mok, Peggy Pik Ki (2008). On the syllable-timing of Cantonese and Beijing Mandarin. *Proceedings of the 8th Phonetics Conference of China (PCC 2008).*

Paradis, Carole (1996). The inadequacy of filters and faithfulness in loanword adaptation. In J. Durand & B. Laks (Eds.), *Current Trends in Phonology: Models and Methods*. University of Salford Publications.

Paradis, Carole, & Tremblay, Antoine (2009). Nondistinctive features in loanword adaptation. *Loan Phonology*, 211–224. https://doi.org/10.1075/cilt.307.09par

Anwyl-Irvine, Alexander L., Massonnié, Jessica, Flitton, Adam, Kirkham, N. Z., & Evershed, Jo K. (2020). *Gorilla in our midst: An online behavioural experiment builder*. Behavior Research Methods, 52, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Peperkamp, Sharon, & Dupoux, Emmanuel (2002). A typological study of stress 'deafness'. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 203–240). De Gruyter Mouton. https://doi.org/10.1515/9783110197105.1.203

Peperkamp, Sharon, & Dupoux, Emmanuel (2003). Reinterpreting loanword adaptations: the role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS 2003)*, 367–370.

Selkirk, Elizabeth O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press.

Shirai, Setsuko. (1999). *Gemination in loans from English to Japanese* [Master's thesis, University of Washington].

Shoji, Shinichi, & Shoji, Kazuko (2014). Vowel epenthesis and consonant deletion in Japanese loanwords from English. *Proceedings of the Annual Meetings on Phonology*. https://doi.org/10.3765/amp.v1i1.16

Silverman, Daniel (1992). Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology*, *9*(2), 289–328.

Smit, Eline A., & Rathcke, Tamara V. (2024). The role of native language and beat perception ability in the perception of speech rhythm. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-024-02513-4

Solé, Maria-Josep, Hyman, Larry M., & Monaka, Kemmonye C (2009). More on post-nasal devoicing: The case of Shekgalagari. *UC Berkeley Phonology Lab Annual Reports*, *5*. https://doi.org/10.5070/p70gk0g18d

Steriade, Donca (2001). The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology VI* (pp. 34–49). Cambridge University Press.

White, Laurence, Mattys, Sven L., & Wiget, Lukas (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*(4), 665–679.