Nuanced Phonological Mapping Influences the Speech Production of Bilingual Speakers:

A Case Study on Shanghai Dialect - Mandarin Bilinguals

by

Zhisheng Jin

Submitted in fulfillment of the requirements for the degree of
Bachelor of Arts in Chinese Language

Fudan University
May 2023

Abstract

To investigate the effects phonological mappings between Chinese dialects and Mandarin have on bilingual speakers, this study conducted a Chinese character naming experiment with balanced bilingual speakers of Shanghai Dialect and Mandarin. It's found that the naming reaction time was influenced by phonetic similarities and complexities of the mapping between Shanghai Dialect and Mandarin rimes. Whether the Mandarin rime is the most frequent match of the rime in the Shanghai Dialect phonetically and the phonetic similarity of the rime have significant interactive effects. The results show that phonetic similarity facilitates responses, which is consistent with previous research on cognate words. But more importantly, the complexity of the Shanghai Dialect's rime mapping interferes with speech production, leading to increased response time. When the Mandarin rime is the most frequent match of a Shanghai Dialect rime, the acceleration effect from the phonetic similarity disappeared. These results suggest that phoneme mapping between the Shanghai Dialect and Mandarin impacts bilingual speakers' representation activation during speech production, and the different realization of these phonemes causes interference during this rapid process.

Keywords: Bilingualism, Language production, Shanghai Dialect, Rhyme Correspondence, Cognate effects

Contents

List of Figures

List of Tables

List of Abbreviations

Corr	The complexity of the rime correspondence
diff	in the context of phonetic similarity, it means the two phoneme are phonetically different phonetically
FM	most Frequent Match of the vowel
H	in the context of HighestCorr, H is the most frequent match
HighestCorr	Whether the vowel is the most frequent match
MA	Morpheme Agreement
MD	Mandarin (only used if the space is not sufficient)
NH	in the context of HighestCorr, NH is not the most frequent match
NM	Null Model
PS	Phonetic Similarity
PU	Phonological Units
SD	Shanghai Dialect
sim	in the context of phonetic similarity, it means the two phoneme are phonetically similar phonetically
VU	Vowel Correspondence Uniqueness

1 Introduction

1.1 Research Questions and Motivations

In China, a significant number of individuals are proficient in both Mandarin and at least one regional dialect. Due to China's longstanding history of political unification of linguistically different regions, Mandarin, as the lingua franca, has had a profound influence on most of the dialects. More importantly, Chinese dialects are essentially linguistic offshoots that evolved from Zhou dynasty-era Chinese, sharing a common phylogenetic relationship (see Norman, 1988, page 19). While these dialects possess distinct linguistic systems in terms of phonology and lexicon, they share many morphemes due to their shared origin. However, their phonologies, including phoneme inventories, phonological rules, and prosodic strategies, vary greatly. Furthermore, the differences between dialects and Mandarin extend beyond phonology. From a sociolinguistics perspective, the language use case and language attitudes people have for regional dialects and Mandarin also differ greatly. Moreover, most dialects lack a written system, so there is a specific phenomenon called literary reading where for a single character there are different readings for Mandarin and dialects. As such, the relationship between Mandarin and Chinese dialects can be viewed as a bilingual one, which allows studies on Mandarin-dialect bilinguals to shed light on the internal mechanisms of bilingual speech production.

But beyond that, due to the shared historical origins of Mandarin and Chinese dialects, most morphemes across the two are cognates, which naturally results in phonological correspondences since the phonetic variation remains mostly systematic through generations. Many bilingual speakers possess an intuitive awareness of these correspondences. For instance, many elderly individuals who speak Shanghai dialect and learned Mandarin as a second language often mismatched phonemes and tones into their Mandarin pronunciation (e.g., pronouncing [mau] as [mɔ]). Conversely, younger speakers often "reverse-engineer" dialectal pronunciations based on Mandarin phonology. Additionally, dialectal strategies for literary reading often adhere to dialect-specific phonological rules.

What motivates this study is the fact that these correspondences are not always symmetrical, which leads to various mismatches between Mandarin and dialectal phonologies. For example, when we are mapping the onsets between Shanghai dialect (SD) and Mandarin, the sound [z] in SD may correspond to [ts], [s], [tʂ], [tʂʰ], [ʂ], or [ɹ] in Mandarin, while the sound [pʰ] corresponds exclusively to [pʰ] in Mandarin. Similarly, the SD rime [ua] can correspond to five candidates in Mandarin, namely [u], [a], [ua], [ai], and [uai], with varying probabilities. The rime [ɤ] in SD, although also present in Mandarin, nearly exclusively corresponds to Mandarin [ou], with an 84.75% correspondence rate (full table in the Appendix 9). These asymmetries can be utilized to examine how phonological representations are activated during speech production in Mandarin-dialect bilinguals.

This study exploits these linguistic features, specifically the correspondences between rimes in SD and Mandarin, as a case study that investigates the language selection mechanisms and phoneme activation processes during speech production in Mandarin-dialect bilinguals. By doing so, it seeks to provide evidence for the question on whether there are complex connections within the integrated phonological system of bilinguals. More specifically, the research question could be listed as follows:

- Do phonological correspondences between Mandarin and dialects affect how bilingual speakers produce speech?

- If so, do they interfere with the speech production or facilitate it?
- If not, what theory can explain why this prominent folk knowledge does not affect bilingual speech production.
- Is there any interactive effect between phonetic similarity and complexity of the mapping?

1.2 Contributions

This study contributes to psycholinguistics research in the following ways:

- It proposes a series of questions that have never been studied in the linguistic community in China based on linguistic features that are specific to Chinese languages.
- It constructed a experiment routine with a comprehensive set of stimuli with complex phonological mapping statuses and corresponding metrics that could be useful for future studies related to this topic.
- It provides evidence supporting the hypothesis that the phonological mappings between Mandarin and a dialect affect bilingual speakers' speech production, which could be useful for related bilingualism theories.

1.3 Variable Design

To explore the complex relationships in the mental phonological representations of Mandarin-SD bilinguals, we will control for three variables in this study. For clarity and ease in selecting corresponding stimuli, the potential states of these three variables have been categorized into seven word types, as shown in Table 1. These seven categories represent all practical combinations of the three variables. The definitions of the variables are described as follows:

Table 1: Examples of the combinations of the three variables

Stimulus	PS	VU	FM	SD Consonant	SD Rime	SD Tone	MD Consonant	MD Rime	MD Tone
龙	Similar	1.760078701	Yes	l	oŋ	23	l	oŋ	35
黄	Similar	34.07321152	Yes	h	uɑ̃	53	h	uɑŋ	35
美	Similar	82.33298683	No	m	e	53	m	ei	214
岛	Not similar	11.56416106	Yes	t	ɔ	34	t	au	214
冷	Not similar	65.83163059	Yes	l	ã	23	l	əŋ	214
歌	Not similar	82.11755476	No	k	u	53	k	ɤ	55
你	Non-existent word in SD			/	/	/	n	i	214

Phonetic Similarity (PS) Between Shanghai Dialect and Mandarin Rimes

Phonetic similarity is defined in terms of shared phonetic features. If two corresponding phonemes from Mandarin and SD share similar places and manners of articulation, they are categorized as similar. Specifically, pairs such as [e] and [ei], [yn] and [yən], as well as all anterior and posterior nasals except [ɑ], are treated as similar. This is based on both practical and objective considerations. Practically, differences in transcription standards between SD and Mandarin phonetics make it difficult to find sufficient stimuli. And since the transcriptions have the tendency to obfuscate things, these wiggle rooms help us to create a balanced dataset of stimuli. Objectively, SD-Mandarin bilinguals, who are the subjects of this study, tend to confuse these sounds in their perception and production, further supporting the classification.

The Complexity of Rime Correspondence (Corr) and High-Frequency Matches (Highest-Corr)

The uniqueness or complexity of rime correspondences and whether a rime is a high-frequency match are derived from statistical data based on the *Modern Chinese Character Frequency Table* from the (Academia Sinica Balanced Corpus of Modern Chinese (Accessible only with VPN)). Specifically, we first compiled a list of 30 SD rimes (excluding rimes with glottal stops for SD specific reasons) from a homophone table of SD in *The Urban Shanghai Dialect* (Xu, Tang, You, Qian, Shi, and Shen, 1988). Next, using the frequency table from the corpus, we assigned weights to each character and categorized them by rime, yielding total frequencies and frequencies per rime.

We quantified the complexity of the correspondences by calculating the information entropy of SD phonemes' potential correspondences in Mandarin. A low entropy value indicates a unique correspondences. In cases of the exceptions of the phonological correspondence, whether it is a high-frequency match is not discussed. On the other hand, a higher entropy value reflects non-unique, complex correspondences. Since the complexity is difficult to categorize, this variable will be treated as continuous, with a value for rime correspondence complexity included in the modeling process. The calculation for this complexity value is detailed later in this thesis.

For high-frequency correspondence, we identified the most frequently matched Mandarin rime for each SD rime and classified it as a high-frequency correspondence. All other matches were categorized as low-frequency. This binary classification serves two main purposes. First, given the limited amount of our data, a finer-grained classification would obscure general differences across levels in subsequent statistical analyses. Secondly, in many cases, the frequency of the highest match is substantially greater than that of other matches, making binary classification more effective in capturing this difference between rimes.

Control Group: Nonexistent Phonemes in Shanghai Dialect

For the control group, we used Mandarin characters containing morphemes not present in SD. As the focus of this study is the phonological structure of bilinguals, the control group is designed to minimize the activation of SD phonemes. Since these morphemes do not exist in Mandarin, they can limit the activation of SD phonology. This approach is analogous to Jared and Kroll (2001), which used non-cognate words as the control group. Given prior findings that phonological representations across languages can activate motor programs for each other, we acknowledge that the control group may still partially activate SD phonology, but still a lesser degree than non-control stimuli.

1.4 Research Hypotheses

This study examines the effects of the three variables defined above on participants' response times (RTs) in a Mandarin character-naming task. The following hypotheses are proposed:

- **Phonetic Similarity Hypothesis**: If two phonetically dissimilar phonemes produce a response pattern significantly different from the control group (nonexistent SD morphemes), this would indicate that SD phonology is activated, suggesting connections between distinct phonemes from different inventories. Conversely, if no significant difference is observed, it would suggest that parallel activation of Mandarin-SD phonemes occurs only for similar phonemes, aligning with findings in Indo-European language studies.
- **Rime Correspondence Complexity Hypothesis**: If the complexity of rime correspondence affects response time, two scenarios are possible:

    1. Slower Responses with Higher Complexity: This would indicate that activating SD phonemes interferes with production, likely due to the activation of different phonological representations or syllable motor programs.
    2. Faster Responses with Higher Complexity: This could result from the activation of multiple possibilities for the SD phoneme, leading to a dispersion effect that inhibits further activation of the motor programs.

  If there is no interaction between the variables, it would suggest that while SD phonemes activate Mandarin phonemes, subsequent processing of the activated SD phonemes is absent, necessitating new theoretical explanations.

- **High-Frequency Correspondence Hypothesis**: Due to the phonetic similarity of one-to-one correspondences, it is difficult to exclude the influence of phonetic similarity. Therefore, we will only consider one-to-many correspondences where the SD phoneme and Mandarin phoneme differ phonetically. If high-frequency correspondences significantly affect response times, there would be two possible outcomes:

    1. Faster Responses with Higher Frequency: This would indicate that SD phonemes not only activate Mandarin phoneme production but also do not elicit further activation of SD themselves.
    2. Slower Responses with Higher Frequency: Comparing these results to the control group could clarify whether inhibition effects are at play. If overall inhibition is observed, it would suggest that SD phonemes undergo strong further activation, but Mandarin production does not get the acceleration. Conversely, if high-frequency correspondences are merely slower than low-frequency ones, it may reflect mutual interference between SD and Mandarin phoneme activations.

  If no interaction between frequency and response time is observed, the variable should be analyzed in conjunction with complexity.

Phonological Environment Constraints

Historical sound changes in SD have imposed constraints on some correspondences. For example, the rime [e] corresponds to [uei] only when the onset is a dental, and [oŋ] corresponds to [əŋ] only when preceded by a labiodental fricative. In such cases, these correspondences become high-frequency matches. Since these constraints are often tied to phonological rules and pronunciation contexts, they are valuable for studying the processing of phonological rules. However, due to resource limitations, we only used the log-transformed frequency values to penalize such exceptions and to ensure a more balanced frequency calculation.

2 Literature review

2.1 Non-selective Activation in Bilingual Speech Production

In studies of bilingualism, a critical topic is the construction of the phonological systems in the minds of bilingual speakers. A major focus of this field of research is whether both language systems are simultaneously activated during linguistic tasks. Numerous studies support the simultaneous activation of both language systems (e.g., De Groot, Borgwaldt, Bos, and Van den Eijnden (2002); Green (1998); (Costa, Colomé, Gómez, and Sebastián-Gallés, 2003); Costa and Santesteban (2004); Costa, Heij, and Navarrete (2006);Kroll, Sumutka, and Schwartz (2005); Kroll, Bobb, Misra, and Guo (2008); Branzi, Martin, Abutalebi, and Costa (2014)). However, most of these studies emphasize lexical access rather than phonological representations, which is deeper and more intrinsic to the speakers.

In terms of the studies on the construction of bilingual phonological systems, the hypothesis of non-selective activation has also received extensive experimental support, and they have also explored the composition of phonological representations. Doctor and Klein (1992) found that English-Pashto bilinguals encountered significant difficulties in lexical judgment tasks involving homophones in English and Pashto (determining whether the stimulus was an English or Pashto word). These difficulties extended to non-words and pseudo-homophones, suggesting that the simultaneous activation of phonological representations in both languages led to interference. This study provided initial evidence supporting the non-selective activation hypothesis at phonological level. Subsequent studies have offered additional evidence. For example, Gollan, Forster, and Frost (1997) discovered that when the phonemes of a prime in the dominant language overlapped with the target stimulus in the non-dominant language, the priming effect was more pronounced. Similarly, Dijkstra, Grainger, and Van Heuven (1999) observed that Dutch-English bilinguals experienced interference effects from homophones of the two languages during lexical judgment tasks, resulting in slower reaction times compared to control conditions. Orthographic and semantic overlaps, however, facilitated faster processing. The authors interpreted these findings as competition between lexicons in both languages caused by the activation of identical phonological representations, resolved ultimately through inhibition. This further supports the notion that bilinguals' phonological systems possess a complex structure that activates downstream nodes.

Of particular relevance to this study is the research from Jared and Kroll (2001), who examined whether bilinguals activate phonological representations from both languages during speech production. They asked English-French bilinguals to perform English word-naming tasks with stimuli categorized as having either "French enemies" or "English enemies". A "French enemy" refers to an English word that follows English phonotactics but produces a different pronunciation in French. Conversely, an "English enemy" is an English word not only does not subject to French phonotactics but also violates English phonotactics. They found that for participants proficient in French, "French enemies" did not slow down the naming prior to activating French knowledge. However, once French knowledge was activated, naming speed slowed to levels comparable to "English enemies" words. Similarly, French-English bilinguals exhibited analogous patterns during the same task. The initial lack of influence from French might be due to stimuli being controlled only in English words, limiting the effect of French. Thus, in a follow-up study by Jared and Szucs (2002), they used hetero-phonic homographs, with French frequencies higher than English. This study revealed that hetero-phonic homographs caused naming difficulties regardless of whether participants were native French speakers or had brief exposure to French stimuli, further supporting the simultaneous activation of phonologies in both languages, albeit with activation levels influenced by native language or short-term exposure.

On the other hand, a series of studies have explored the nature of phonological representations themselves and the influence of other factors on these representations. In terms of whether phonological representations are shared between two language systems or interconnected, Roelofs (2003) used the form-preparation paradigm to test whether participants could prepare for phonemes even in different target languages. The form-preparation paradigm involves pairing stimuli where participants respond to the second word upon seeing the first. In monolingual experiments, this paradigm demonstrated that when the second word in a series shared an initial segment with others, participants experienced a preparation effect, which increased with the length of the shared segment. By manipulating whether the stimulus set consisted of words from a single language or mixed English and Dutch word pairs, they found that the preparation effect for words sharing the same initial segment was equally significant across languages. For example, in mixed-language stimuli where response words all began with /b/, the preparation effect was as robust as in monolingual conditions. This supports the hypothesis of shared phonological representations.

2.2 Cognate Effects in Bilingual Speakers

Since phonological representations are activated via lexical access, it is necessary to investigate whether cognate effects stem primarily from cross-linguistic lexical activation. Evidence that lexical activation occurs early in production comes from Hermans, Bongaerts, De Bot, and Schreuder (1998), who used a picture-naming interference paradigm with Dutch-English bilinguals. Participants named pictures in the target language (English) while experiencing phonological interference from either their native language (e.g., Dutch *mown* for "mountain") or target-language correspondence with native words (e.g., Dutch *berm* closely resembles Dutch *berg*, which means "mountain"). The study revealed that when interference preceded picture presentation, only target-language corresponding native words caused interference, while interference following picture presentation accelerated responses. Subsequent studies, such as Costa et al. (2003), found that cognates leads phonological acceleration effects. The WEAVER++ model (Roelofs, 2015) for bilingual speech production attributed these effects to parallel activation of morphemes rather than lexical correspondences. In model simulations, cognate effects persisted even when the activation of native-language morphemes was only 5% of the target language's. Removing lexical connections between target and native languages but linking their morphemes produced similar effects, suggesting that lexical activation might not be the answer to cognate effects. Moreover, the second model also replicated Hermans et al.'s finding that phonological interference effects peaked when native-language stimuli followed naming tasks. These simulations provided evidence that shows cogate effects could be a result from activation of deeper representations. Schwartz, Kroll, and Diaz (2007) conducted lexical naming tasks with English-Spanish bilinguals, manipulating orthographic and phonological similarity of cognates. They found that orthographically similar cognates slowed responses due to phonological differences, suggesting that orthographic representations activate phonological representations, causing interference. These findings indicate that lexical activation may not initially activate lexical items in non-target languages, but phonological representations in non-target languages are uncontrollably activated. All these previous findings lead to the question on where phonological representation could be one of the robust contributor to cognate effects, which is exactly what this paper tries to explore.

3 Methodology

3.1 Participants

The experiment recruited 18 Mandarin-Shanghai dialect bilingual participants, aged 30 to 50, as this age group typically comprises balanced bilinguals who began learning Mandarin in elementary school and have long-term mixed use of both languages. Participants completed a Bilingual Language Proficiency Test (Birdsong, Gertken, and Amengual, 2012) on the mobile survey platform of `Wen Juan Xing` before beginning the test. They were seated at a chair 50 cm far from a PC for the experiment running on `PsychoPy`. Due to technical difficulties, we had to manually initiated the second and third experimental sessions after the first session was completed. Since we do not need detailed acoustic information of the responses, the experiments were conducted in small spaces where ambient noise levels were below 40 dB. An Audio-Technica ATR2500 microphone with a 48kHz sampling rate in mono was used for recording. Upon completion of the experiment, participants received cash compensation of 30 RMB or an equivalent reward.

3.2 Stimuli

The design of the stimuli was crucial for this study. To quantify the correspondences in SD, we used the homophone dictionary from *The Dialect of Urban Shanghai* (Xu et al., 1988), extracting the Chinese characters associated with each rime and obtaining their corresponding Mandarin pronunciations. Using this raw data, a Python script was used to transform the homophone dictionary, each rime, and the characters containing each rime into class objects, enabling further quantification of the proportional correspondences between SD rimes and their Mandarin counterparts.

If all characters were weighted equally, rare characters would disproportionately influence the results, causing the calculated correspondences to deviate horrendously from the actual linguistic input of bilingual speakers. To address this issue, we matched the characters in the homophone dictionary with their frequencies in the Modern Chinese Character Frequency Table from the National Corpus of Modern Chinese, using these frequencies as weights for each character. By calculating the weighted probabilities of possible Mandarin rimes for each Shanghai dialect rime, based on character frequency, we derived quantified information about the correspondences between Shanghai dialect rimes and Mandarin phonemes (see Appendix 9).

As we've mentioned in the Introduction, linguistic materials are inherently complex, and true "unique correspondences" rarely exist. After quantifying the correspondences of Shanghai dialect phonemes, we found a strong negative correlation (r = -0.93) between the entropy of a rime's correspondence and the proportion of the most frequent Mandarin rime in the whole pool of candidates for that one SD rime. This indicates that as the correspondence becomes more complex, the most frequent Mandarin rime is less likely to dominate. For example, the entropy of [ioŋ] is as high as 1.997 bits, and its most frequent Mandarin correspondence accounts for only 45.8%, whereas [iaŋ] has an entropy of just 0.2682 bits, with its most frequent correspondence accounting for 96.42%. This aligns well with the variable design related to the uniqueness of correspondences.

However, relying solely on entropy to measure the uniqueness of correspondences risks overlooking cases where high entropy is driven by an abundance of exceptions. For instance, [ɨ] has an entropy of 0.9585 bits, but its most frequent Mandarin correspondence accounts for 81.49%, while [yn], with an similar entropy of 0.9437 bits, has only three possible correspondences, and its most frequent match accounts for just 65.89%. Such cases require further penalization. Therefore, we used the following formula to further assess the “uniqueness” or “complexity” of correspondences:

$$\text{Correspondence Complexity} = \frac{\text{Entropy} \times \log\left(\frac{1}{\text{Highest Correspondence Proportion} \times 100}\right)}{\log(\text{ of Candidates})} \tag{1}$$

In terms of our control variables, they are primarily introduced in the Introduction. We’ll discuss in detail about the quantification here. Word frequency is our primary control for its robust effect on speech production. Since the log-transformed word frequencies still exhibited unequal variances (p = 0.02), we conducted Welch and Brown-Forsythe ANOVA tests and they revealed that there were no significant differences in the log-transformed mean word frequencies between groups (Welch: 0.576, Brown-Forsythe: 0.394). Thus, word frequency was controlled between groups. Similarly, the log-transformed word frequencies of stimuli in the three experimental blocks (block1, block2, and block3) were tested, with Welch and Brown-Forsythe ANOVA showing no significant differences (Welch: 0.221, Brown-Forsythe: 0.120).

Ultimately, the experiment adopted three distinct variables: phonetic similarity, rime correspondence complexity, and whether the rime was the most frequent correspondence. We reprise the definition here for better clarity.

- Phonetic Similarity: We mostly refer to it as ”Similarity” in the tables. in the Values are sim (similar) or diff (dissimilar). This refers to whether the Mandarin rime of a Chinese character is **phonetically** similar to the actual pronunciation of the corresponding Shanghai dialect rime from an IPA perspective. The similarity of onset in the 70 actual stimuli was largely controlled, so any dissimilarity arises solely from rime differences.

- Rime Correspondence Uniqueness or Complexity: Written as ”Corr” in the tables. Values are positive real numbers calculated as described above, based on the complexity of the corresponding Shanghai dialect rime. A higher value indicates greater complexity in the correspondence.

- Whether the Rime Is the Most Frequent Match or Correspondent: Written as ”HighestCorr” in the tables. Values are H (yes) or NH (no). This is determined by whether the Mandarin rime of a stimulus **is** the most frequent correspondence for its Shanghai dialect rime. If it is, the value is H; otherwise, it is NH.

Last but not least, we also included the activation of Shanghai dialect knowledge as one extra variable in the experiment. Primarily in the form of the elicitation block in the middle of three blocks in one experiment, which would be discussed further in the following section.

3.3 Experiment

Each participant completed three experimental sessions, totaling 105 naming tasks, with each stimulus being a single-syllable Chinese character. Each session consisted of 35 lexical naming tasks based on the seven word types categorized in the Introduction (see Table 1), with 10 different lexical items per type, forming a total of 70 stimuli. In some cases, 2–3 stimuli may share the same final rime.

Before the experiment began, participants were verbally instructed on the experimental procedure and informed that their responses should prioritize speed over accuracy. The three sessions included both Mandarin and Shanghai dialect naming tasks. Prior to the Shanghai dialect naming task, participants were reminded to rely on their intuition when encountering characters difficult to name in Shanghai dialect.

In each trial, a cross appeared at the center of the screen to direct participants' focus for 2000 ms, after which it was replaced by a stimulus character displayed for 2000 ms. During the process, the microphone was activated for 1800 ms to record responses. This process was repeated for 35 trials to complete one session. The procedure of one trial is illustrated in Figure 1.

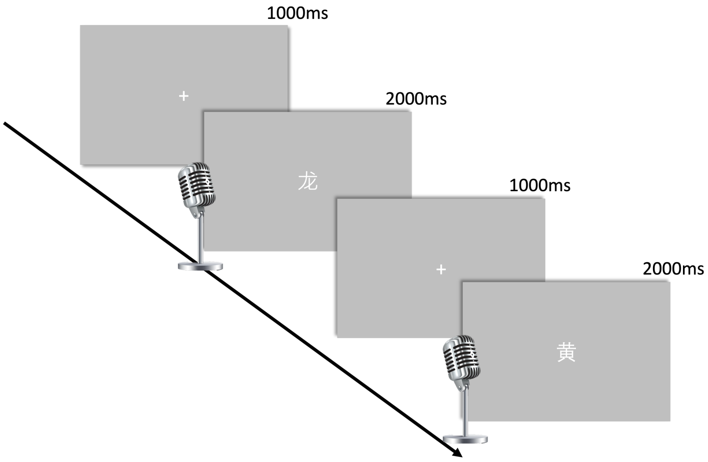


Figure 1: Illustration of the procedure of one trial

In the three sessions, the first session was a series of Mandarin naming tasks. The second session was a series of Shanghai dialect naming tasks for activating participants' Shanghai dialect knowledge. This session was designed to be a filler and no audio response was recorded; the 35 stimuli were Chinese characters outside the 70 actual stimuli. The third session went back to a series of Mandarin naming tasks. Due to constraints in PsychoPy's experimental design, the 70 actual stimuli were divided into two stimulus lists, referred to as A and B, and randomly assigned within the first and third sessions. To control for potential effects of the lists themselves, two sets of experiments were created. In the first set, List A was used in the first session, and List B was used in the third session. In the second set, List B was used in the first session, and List A was used in the third session. Among the valid participant data, 8 participants completed the first set, and 7 participants completed the second set, So the potential effects from the sets is randomized and thus controlled.

3.4 Data Pre-processing

Due to the limitations of the experimental setup, ambient noise affected the use of `Voicekey` module in PsychoPy, and `Voicekey` also lacked sufficient accuracy in detecting the onset of voiceless consonants. Moreover, using `Voicekey` and recording functions in PsychoPy on our PC caused unstable frame rates in the experiment, potentially leading to delays in microphone activation and invalidating responses. According to Roux, Armstrong, and Carreiras (2017), their automated speech onset detection software differed from manual annotation by less than 50 ms, which is acceptable for general statistical analyses. However, due to the relatively small sample size in this study and the significant impact of a 50 ms difference, so manual annotation was the optimal approach. Therefore, we manually annotated the audio files using Praat's `ProsodyPro tool` (Xu, 2013).

Our annotation guideline defines the onset of the release as the endpoint for response time measurement. This choice is based on the clear segmentation in the acoustic signal at the release burst compared to other phases, such as jaw-opening sounds (which could be more accurate but hard to define). Although many recordings contained distinct jaw-opening sounds, the delay between these sounds and the actual release burst was insignificant and unrelated to the following consonant articulation. Therefore, we do not consider the jaw-opening sound a reliable marker for response time.[1]

Additionally, the experimental program introduced a 10–40 ms fluctuation when calling the microphone. And we are certain that these fluctuations were caused by the microphone activation process. As a result, the calculation of the final response time requires adjustments using the audio file length. Specifically:

$$\text{Response Time} = 1800\,\text{ms} - \text{audio file length} + \text{duration from audio file start to the release} \tag{2}$$

This adjustment compensates for the reduction in response time caused by microphone activation delays.

---

[1] For a side note, many studies do not specify the specific boundary used for annotating the response time. And we believe that there could be methodological investigation on this issue in the future.

4 Results

A total of 21 participants took part in this experiment, including 10 males and 11 females. Due to technical issues, data from participants 2 through 7 were severely affected, leaving valid data for only 15 participants (7 males and 8 females). Based on the BLP (Bilingual Language Profile) assessment collected via `Wen Juan Xing`, all 15 participants were residents of central Shanghai, currently working there, and regularly using both Mandarin and Shanghai dialect. Their ages ranged from 35 to 50, with a mean age of 41.6 and a standard deviation of 5.28 years. In terms of the BLP scores, the mean score for Shanghai dialect was 174.14293 with a standard deviation of 9.61024, while the mean score for Mandarin was 156.3101 with a standard deviation of 18.228654. The mean scores for both languages are high according to the test, suggesting the participants have strong language skills in both languages. The ratio of the two mean scores was 1.1143018, which is close to 0, indicating that the participants were largely balanced bilinguals.

The 15 participants provided a total of 1050 responses. During annotation, the author was present in the room to note down the accuracy, and no incorrect responses were observed. This might be attributed to the relative simplicity of the character-naming task. The median response time was 624.35 ms, and the standard deviation was 134.56 ms. Among the responses, 37 were beyond the mean plus two standard deviations and were excluded, leaving 1,013 responses for further analysis.

In terms of random effects from participants, considerable variation was observed in their reaction times, as illustrated in Figure 2. The mean reaction time between the fastest and slowest participants differed by over 200 ms. This suggests that traditional ANOVA methods, which incorporate participant-level variability into the overall mean, could obscure the relationships between reaction time and independent variables. Furthermore, the unequal number of responses for each variable level excludes traditional ANOVA analysis from the options. Thus, this study employs a linear mixed-effects model to assess the impact of various variables on reaction time.

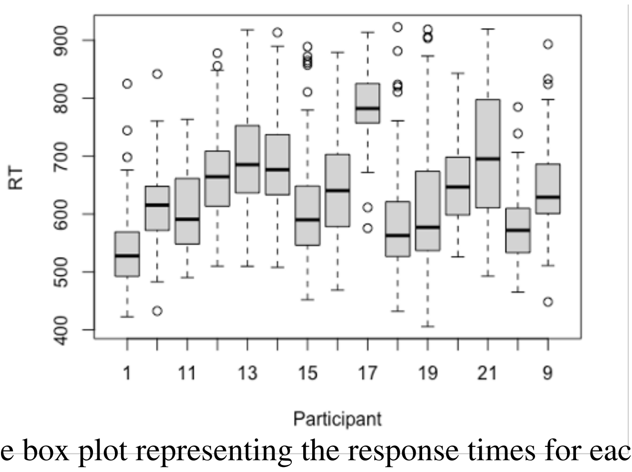


Figure 2: The box plot representing the response times for each participant

Also, the Q-Q plot of log-transformed reaction times show that the data were approximately normally distributed. However, the Shapiro-Wilk test for homo-scedasticity indicated hetero-scedasticity (p = 0.0001632), making the linear mixed-effects model the best choice for modeling the data.

4.1 Model Analysis

Using the `lmer4` package (Bates, Mächler, Bolker, and Walker, 2014), we constructed 8 models to determine whether the 3 variables had main effects on reaction time. Random intercepts were used to account for variability between participants and stimuli. The model components and likelihood ratio test results are shown in Table 2. Adding the variables Similarity (phonological similarity) and Corr (stimulus rime correspondence complexity) to the null model (`NM`) resulted in a significant improvement in Model `M2`, indicating that both variables exerted main effects on reaction time.

Table 2: Three basic models with their components and performances

Model	Factor I	Factor II	Factor III	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr (>Chisq)
NM	N/A	N/A	N/A	4	-3101.4	-3081.8	1554.7	-3109.4			
M1	Sim	N/A	N/A	5	-3103.8	-3079.2	1556.9	-3113.8	4.3475	1	**0.03706***
M2	Sim	Corr	N/A	6	-3107.4	-3077.9	1559.7	-3119.4	5.6619	1	**0.01734***
M3	Sim	Corr	HighestCorr	7	-3105.7	-3071.3	1559.9	-3119.7	0.2823	1	0.59522

To further investigate the interaction effects among the three variables, four additional models were constructed and compared to Model M3 (which lacks interactive effects) using likelihood ratio tests. The statistics are shown in Table 3. Model `M6`, which includes the interaction between Highest Correspondence and Similarity, demonstrated a significant interaction effect and had the lowest AIC. Residual analysis of `M6` (See Appendix 6b and 6a) suggests that residuals were randomly distributed with predicted values, where the Q-Q plot approximated a straight line, indicating that residuals met the assumptions. Therefore, `M6` was the best model to explain the results.

Table 3: Four models with interactive effects with their components and performances

Model	Factor IV	Factor V	Factor	Factor VII	npar	AIC	Chisq	Df	Pr (>Chisq)
M3	N/A	N/A	N/A	N/A	7	-3105.7	0.2823	1	0.595223
M4	Similarity Corr	N/A	N/A	N/A	8	-3104.6	0.9084	1	0.340543
M5	Similarity Corr	HighestCorr Corr	N/A	N/A	9	-3103.2	0.5538	1	0.456767
M6	Similarity Corr	HighestCorr Corr	HighestCorr Similarity	N/A	10	-3108.9	7.6819	1	**0.005578***
M7	Similarity Corr	HighestCorr Corr	HighestCorr Similarity	HighestCorr Similarity Corr	11	-3107.9	0.9947	1	0.318597

Statistical information for `M6` (See Appendix 7) showed that the standard deviation for random effects was larger for participants than for materials, indicating that inter-participant differences had a greater impact on reaction time than the materials themselves. When it comes to the fixed effects, Corr (the complexity of rime correspondence) increased by approximately 0.0003. Using the `lmerTest` package, the p-value for Similarity was found to be 0.279, while the p-value for Corr was 0.0003, indicating that Similarity lost its main effect when the interaction between HighestCorr and Similarity were taken into consideration. A scatter plot of Corr and reaction time, along with a linear regression (Figure 3), showed that reaction time increased with greater rime correspondence complexity, aligning with the statistical findings from `M2` and `M6`.

Continuing on the interactive effects. Plotting the two variables (Figure 4) revealed that for stimuli with phonetic similarity, being the highest correspondent increased reaction times. On the other hand, for stimuli without phonetic similarity, the highest correspondence resulted in shorter reaction times. When stimuli were the highest correspondents, phonetic similarity slowed down the reaction. Conversely, when stimuli were not the highest correspondent for that SD rime, phonetically similar stimuli had significantly faster reaction times. By splitting the data into two groups based on Similarity and further dividing them by Highest Correspondence, four linear mixed-effects models were constructed. Similarity had a main effect only when stimuli were not the highest correspondent, or the most frequent match, with higher similarity leading to faster reaction times,

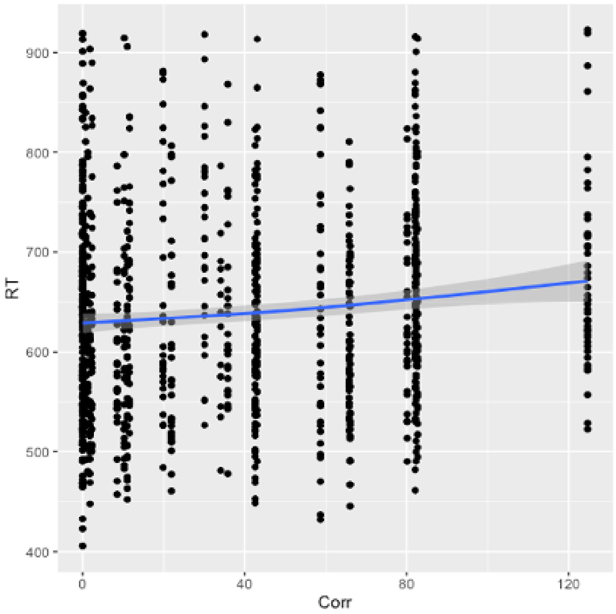


Figure 3: Mean RT vs. The complexity of the rime mapping of the stimuli (Corr)

while Similarity has no main effect on reaction time when stimuli are the highest correspondent. (Statistics of the four models are included in the Appendix 8)

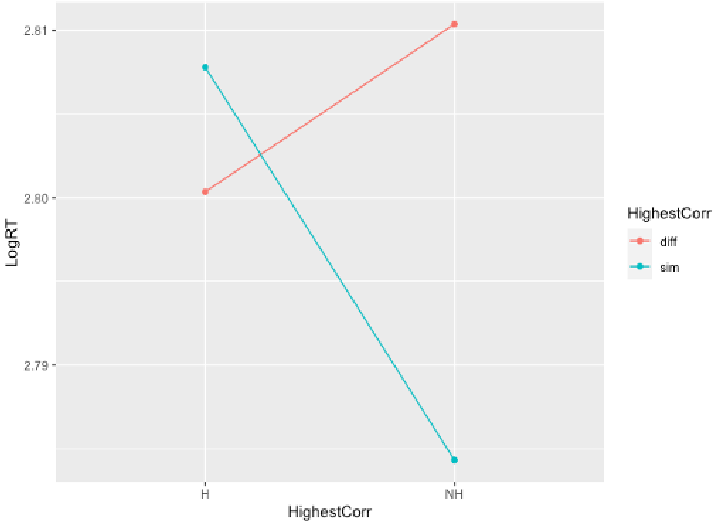


Figure 4: The interactive effects of phonetic similarity and whether the rime is the most frequent match on reaction time

Finally, the activation status of SD knowledge also plays a part in the effects on reaction time. Since Highest Correspondence did not significantly affect the model, six models were constructed to examine the impact of Shanghai dialect knowledge activation on reaction times and its interactions with other variables. Likelihood ratio test results indicated significant differences between Models `abM1` and `abM2` and the NM, suggesting that incorporating Shanghai dialect knowledge activation improved model performance. Model `abM2` also had the lowest AIC value among all models. However, the fitness of `abM2` was primarily due to the inclusion of the interaction between Highest Correspondence and Similarity. Adding interactions between Shanghai dialect knowledge activation and the other three variables did not result in any significant differences, suggesting no interaction effects. It is likely that participant fatigue after completing 70 trials influenced the changes in Shanghai dialect knowledge activation, or that Shanghai dialect knowledge exerted a global ef-

Table 4: Additional models evaluating the effect of dialect knowledge activation

Model	Factor IV	Factor V	Factor VI	Factor VII	Facotr VIII	Factor IX	npar	AIC	Chisq	Df	Pr (>Chisq)
abNM	N/A	N/A	N/A	N/A	N/A	N/A	7	-3105.7			
abM1	be4/after	N/A	N/A	N/A	N/A	N/A	8	-3110.3	6.6197	1	**0.010086***
abM2	be4/after	HighestCorr Similarityilarity	N/A	N/A	N/A	N/A	9	-3116.3	7.9431	1	**0.004827***
abM3	be4/after	HighestCorr Similarity	be4/after Similarity	N/A	N/A	N/A	10	-3116.3	2.0186	1	0.155382
abM4	be4/after	HighestCorr Similarity	be4/after Similarity	be4/after HighestCorr	N/A	N/A	11	-3115	0.6583	1	0.417151
abM5	be4/after	HighestCorr Similarity	be4/after Similarity	be4/after HighestCorr	be4/after Corr	N/A	12	-3113.1	0.0885	1	0.766117
abM6	be4/after	HighestCorr Similarity	be4/after Similarity	be4/after HighestCorr	be4/after Corr	HighestCorr Similarity be4/after	13	-3111.2	0.1861	1	0.666189

fect. As shown in Figure 5, the impact of conducting the experiment before or after Block 2 varied among participants, making it impossible to conclude that activating Shanghai dialect knowledge indeed influenced the production of phonetically similar rimes. This phenomenon may also be due to the limited activation of Shanghai dialect knowledge under the 35-item Shanghai dialect naming task.

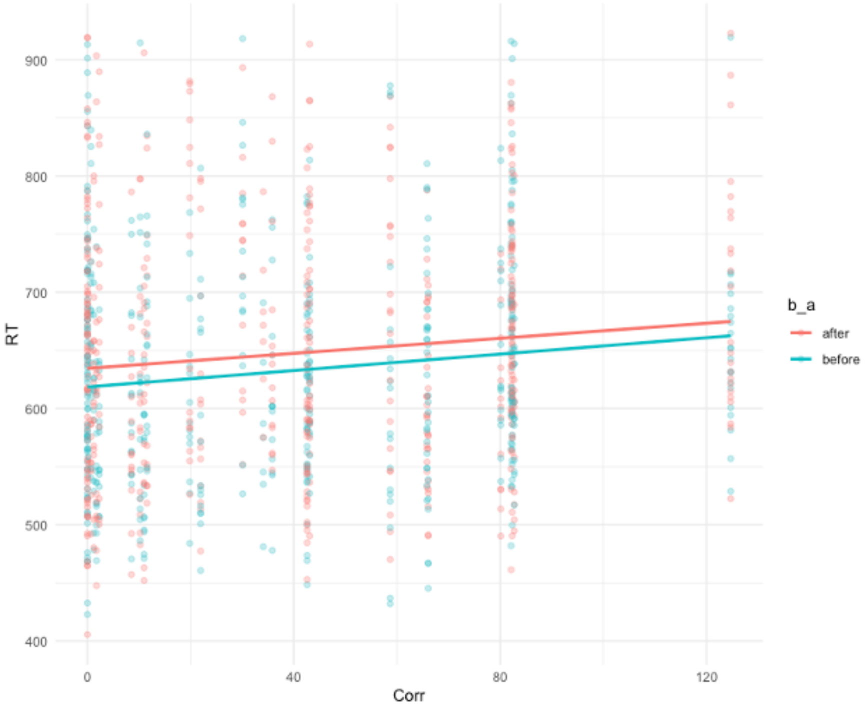


Figure 5: Linear regressions of Corr vs. RT under different knowledge activation levels

5 Discussion

This study investigated the reaction time of Chinese character naming tasks conducted on Shanghai dialect-Mandarin balanced bilinguals. In the experiments, three key variables were manipulated. Namely, they are Phonetic similarity between Shanghai dialect and Mandarin rimes, the complexity of the correspondence from a Shanghai dialect rime to many Mandarin rimes, and Whether the Mandarin rime is the most frequent correspondent of the Shanghai dialect rime. Additionally, we explored whether the activation of Shanghai dialect phonological knowledge influenced interactions between these variables by introducing a variable that accounted for whether stimuli were presented before or after the Shanghai dialect naming filler task.

The results from the experiment suggest three major findings:

- The complexity of the rime correspondence significantly affected reaction time, with more complex correspondences leading to slower response times. This is the first time such an effect has been observed in studies of cognates among bilinguals.
- The effect of phonetic similarity depended on whether the Mandarin rime was the highest-frequency correspondence for the rime in Shanghai dialect. Specifically, when the Mandarin rime was the highest-frequency correspondent, phonetic similarity had no significant impact on the reaction time, hence the disappearance of the classic cognate facilitation effect.
- When the Mandarin rime was not the highest-frequency correspondent, phonetic similarity significantly accelerated reaction times, and thus the cognate facilitation effect re-emerged.

5.1 Cross-language Activation and Interference Effects

The results suggest that Shanghai dialect phonological representations connect to downstream nodes corresponding to Mandarin motor programs, enabling cross-language activation. However, the interference caused by Shanghai dialect representations was weaker than the competition among Mandarin representations.

The notion of "correspondence" in this study refers to the connection between Shanghai dialect phonological representations and Mandarin motor programs (the realization of phonemes), rather than a direct link between two phonological representations. For example, the correspondence between Shanghai dialect [ɔ] and Mandarin [au] is grounded in motor program alignment rather than shared phonological representations. If the latter were true, we would expect high-frequency correspondences to show significant facilitation effects under conditions of phonetic similarity, which was not observed. Instead, the results indicated that high-frequency correspondences slowed down the reaction times, suggesting that interference originates from motor program competition rather than parallel phonological activation.

On the other hand, the interaction effects between phonological similarity and correspondence frequency points to a need for a more nuanced explanation. Namely, phonetic similarity only facilitated reaction times when the Mandarin rime was a non-high-frequency correspondent of the Shanghai dialect rime. High-frequency correspondents introduced interference, likely because multiple motor programs were activated, leading to competition.

This explanation is supported by the conditioned linear model (Appendix 8), where non-high-frequency correspondents with phonetic similarity had negative estimated values, suggesting facilitation. Conversely, high-frequency correspondences did not enhance reaction times, as shown in Figure 4.

5.2 Theoretical Implications

The findings challenge predictions made by existing models of bilingual language processing, such as the BIA+ model (Dijkstra and Van Heuven, 2002) and the WEAVER++ model (Roelofs, 2015). The BIA+ Model emphasizes cross-language integration of orthographic, phonological, and semantic representations but does not explicitly address cross-language phonological motor program interference. Our results suggest that while phonological similarity activates corresponding representations across languages, the observed interference cannot be fully explained by orthographic or semantic overlap. WEAVER++ Model posits shared phonological representations across languages and predicts that similar phonological representations should accelerate response times due to shared motor programs (phonetically similar). However, our findings contradict this prediction, as phonetic similarity under high-frequency correspondence conditions did not facilitate reaction times.

Overall, our results suggest that Shanghai dialect phonological representations activate their corresponding Mandarin motor programs, but Shanghai dialect motor programs provide limited interference. The absence of facilitation for high-frequency correspondences suggests that global inhibition of Shanghai dialect representations may have weakened their influence on Mandarin activation nodes. This global inhibition may also explain why facilitation effects are limited to non-high-frequency correspondences.

6 Conclusion

In conclusion, this study investigated the effects of phonetic similarity, correspondence complexity, and correspondence frequency on the speech production of Shanghai dialect-Mandarin balanced bilinguals. Through a systematic experiment, we identified three key findings that contribute to our understanding of bilingual phonological systems. Firstly, cognate facilitation effects are not uniform but depend on the correspondence frequency between Shanghai dialect and Mandarin phonological representations. Facilitation occurs when phonetic similarity is paired with non-high-frequency correspondences, while high-frequency correspondences only result in interference. Secondly, the complexity of phonological correspondences, quantified using an entropy-based formula, successfully predicts the change in reaction times. More complex correspondences lead to longer response times, suggesting the role of probabilistic phonological information in bilingual speech production. Last but not least, our findings support that cross-language phonological interference primarily arises at the motor program level, rather than at the level of abstract phonological representations. While Shanghai dialect phonological representations activate their corresponding Mandarin motor programs, the interference caused by Shanghai dialect motor programs is weaker than the competition among Mandarin motor programs, presumably stemmed from the global inhibition of Shanghai dialect.

These findings suggest that bilingual phonological systems are structured probabilistically and that phonological correspondences between languages play a significant role in speech production. By exploiting the unique phonological correspondences between Shanghai dialect and Mandarin, this study proposed a novel perspective on cross-language phonological interaction.

References

J. Norman. *Chinese*. ACLS Humanities E-Book. Cambridge University Press, 1988. ISBN 9780521296533. URL `https://books.google.com/books?id=wOPArZVCk-wC`.

Baohua Xu, Zhenzhu Tang, Rujie You, Nairong Qian, Rujie Shi, and Yaming Shen. *The Dialect of Urban Shanghai*. Shanghai Hope publishing, 1988.

Debra Jared and Judith F Kroll. Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of memory and language*, 44(1):2–31, 2001.

Annette MB De Groot, Susanne Borgwaldt, Mieke Bos, and Ellen Van den Eijnden. Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, 47(1):91–124, 2002.

David W Green. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, 1(2):67–81, 1998.

Albert Costa, Àngels Colomé, Olga Gómez, and Núria Sebastián-Gallés. Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, 6(3):167–179, December 2003. ISSN 1469-1841, 1366-7289. doi: 10.1017/S1366728903001111. URL `https://www.cambridge.org/core/journals/bilingualism-language-and-cognition/article/abs/another-look-at-crosslanguage-competition-in-bilingual-speech-production 868DD1A89FB228604AA6EEADBDF20418`.

Albert Costa and Mikel Santesteban. Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and l2 learners. *Journal of memory and Language*, 50(4):491–511, 2004.

Albert Costa, Wido La Heij, and Eduardo Navarrete. The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9(2):137–151, July 2006. ISSN 1469-1841, 1366-7289. doi: 10.1017/S1366728906002495. URL `https://www.cambridge.org/core/journals/bilingualism-language-and-cognition/article/abs/dynamics-of-bilingual-lexical-access/FE4EAB2DFBAA00B5324449096A1C43AA`.

Judith F Kroll, Bianca M Sumutka, and Ana I Schwartz. A cognitive view of the bilingual lexicon: Reading and speaking words in two languages. *International journal of bilingualism*, 9(1):27–48, 2005.

Judith F Kroll, Susan C Bobb, Maya Misra, and Taomei Guo. Language selection in bilingual speech: Evidence for inhibitory processes. *Acta psychologica*, 128(3):416–430, 2008.

Francesca M. Branzi, Clara D. Martin, Jubin Abutalebi, and Albert Costa. The after-effects of bilingual language production. *Neuropsychologia*, 52:102–116, January 2014. ISSN 0028-3932. doi: 10.1016/j.neuropsychologia.2013.09.022. URL `https://www.sciencedirect.com/science/article/pii/S0028393213003187`.

Estelle Ann Doctor and Denise Klein. Phonological processing in bilingual word recognition. In *Advances in psychology*, volume 83, pages 237–252. Elsevier, 1992.

Tamar H Gollan, Kenneth I Forster, and Ram Frost. Translation priming with different scripts: Masked priming with cognates and noncognates in hebrew–english bilinguals. *Journal of Experimental Psychology: Learning, memory, and cognition*, 23(5):1122, 1997.

Ton Dijkstra, Jonathan Grainger, and Walter JB Van Heuven. Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and language*, 41(4): 496–518, 1999.

Debra Jared and Carrie Szucs. Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition*, 5(3):225–239, 2002.

Ardi Roelofs. Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes*, 18(2):175–204, 2003.

Daan Hermans, Theo Bongaerts, Kees De Bot, and Robert Schreuder. Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: language and cognition*, 1(3):213–229, 1998.

Ardi Roelofs. Modeling of phonological encoding in spoken word production: From g ermanic languages to m andarin c hinese and j apanese. *Japanese Psychological Research*, 57(1):22–37, 2015.

Ana I Schwartz, Judith F Kroll, and Michele Diaz. Reading words in spanish and english: Mapping orthography to phonology in two languages. *Language and Cognitive processes*, 22(1):106–129, 2007.

David Birdsong, Libby M Gertken, and Mark Amengual. Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*, 2012.

Frédéric Roux, Blair C Armstrong, and Manuel Carreiras. Chronset: An automated tool for detecting speech onset. *Behavior research methods*, 49:1864–1881, 2017.

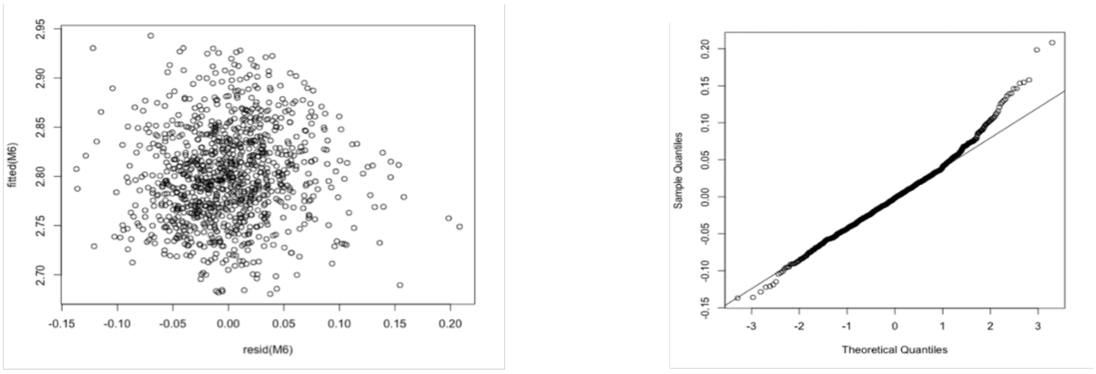
Yi Xu. Prosodypro—a tool for large-scale systematic prosody analysis. Laboratoire Parole et Langage, France, 2013.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting Linear Mixed-Effects Models using lme4, June 2014. URL `http://arxiv.org/abs/1406.5823`. arXiv:1406.5823.

Ton Dijkstra and Walter JB Van Heuven. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, 5(3):175–197, 2002.

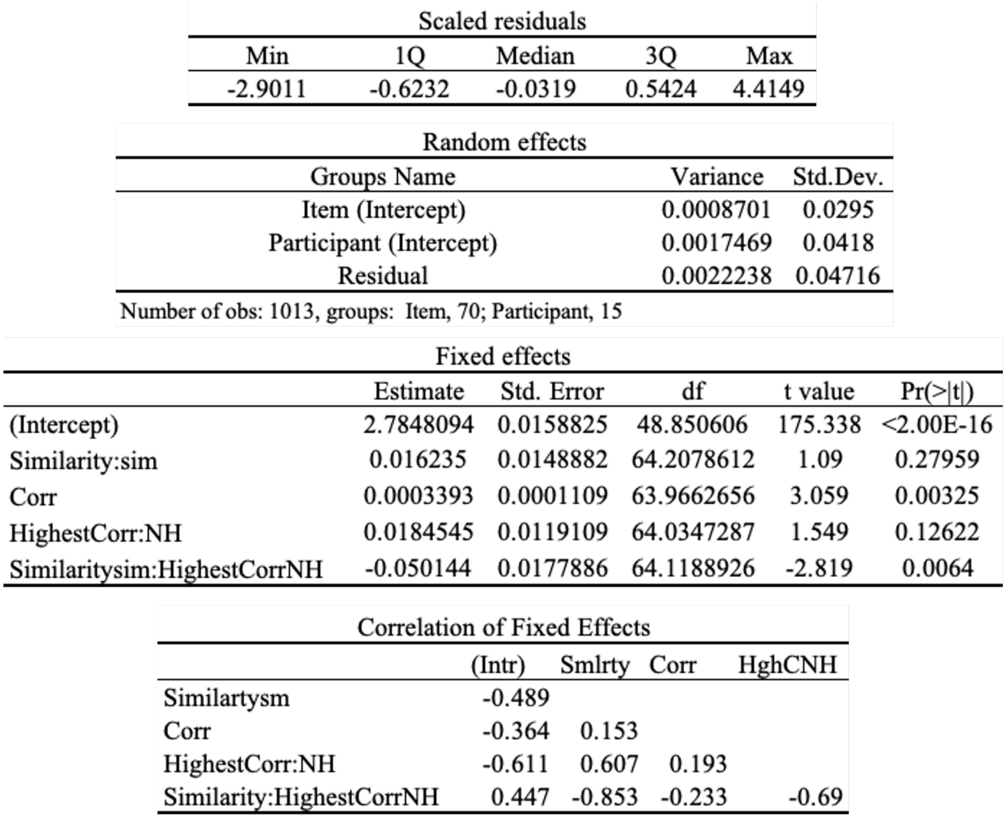
7 Appendix

7.1 Information on M6



(a) M6's residuals versus fitted values (b) The Q-Q plot of M6's residuals

Figure 6: Residual analysis of M6



Scaled residuals

Min	1Q	Median	3Q	Max
-2.9011	-0.6232	-0.0319	0.5424	4.4149

Random effects

Groups Name	Variance	Std.Dev.
Item (Intercept)	0.0008701	0.0295
Participant (Intercept)	0.0017469	0.0418
Residual	0.0022238	0.04716

Number of obs: 1013, groups: Item, 70; Participant, 15

Fixed effects

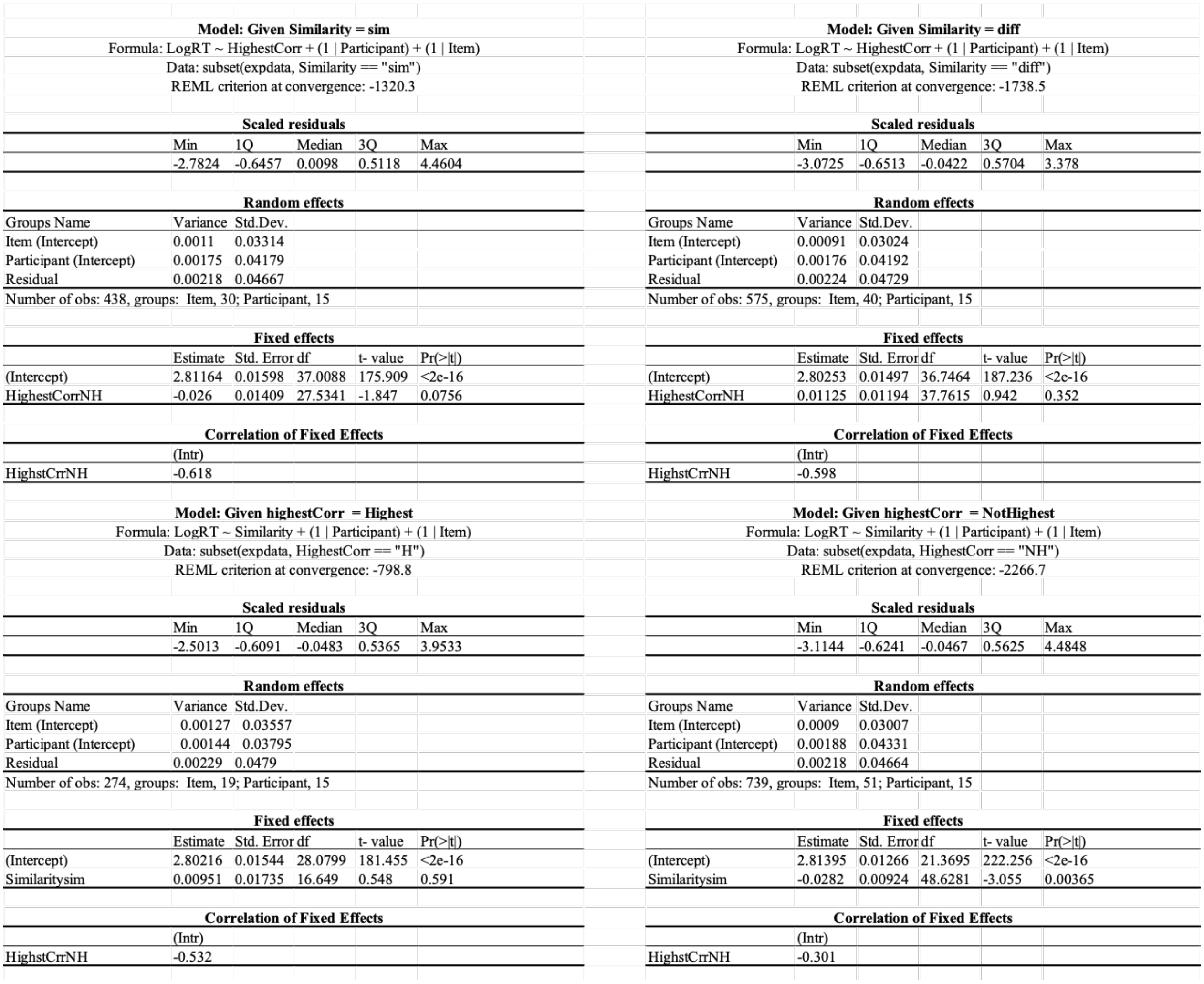
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.7848094	0.0158825	48.850606	175.338	<2.00E-16
Similarity:sim	0.016235	0.0148882	64.2078612	1.09	0.27959
Corr	0.0003393	0.0001109	63.9662656	3.059	0.00325
HighestCorr:NH	0.0184545	0.0119109	64.0347287	1.549	0.12622
Similaritysim:HighestCorrNH	-0.050144	0.0177886	64.1188926	-2.819	0.0064

Correlation of Fixed Effects

	(Intr)	Smlrty	Corr	HghCNH
Similartysm	-0.489			
Corr	-0.364	0.153		
HighestCorr:NH	-0.611	0.607	0.193	
Similarity:HighestCorrNH	0.447	-0.853	-0.233	-0.69

Figure 7: Statistics summary of M6

7.2 Conditioned linear models



Model: Given Similarity = sim

Formula: LogRT ~ HighestCorr + (1 | Participant) + (1 | Item)

Data: subset(expdata, Similarity == "sim")

REML criterion at convergence: -1320.3

Scaled residuals

Min	1Q	Median	3Q	Max
-2.7824	-0.6457	0.0098	0.5118	4.4604

Random effects

Groups Name	Variance	Std.Dev.
Item (Intercept)	0.0011	0.03314
Participant (Intercept)	0.00175	0.04179
Residual	0.00218	0.04667

Number of obs: 438, groups: Item, 30; Participant, 15

Fixed effects

	Estimate	Std. Error	df	t- value	Pr(>|t|)
(Intercept)	2.81164	0.01598	37.0088	175.909	<2e-16
HighestCorrNH	-0.026	0.01409	27.5341	-1.847	0.0756

Correlation of Fixed Effects

	(Intr)
HighstCrrNH	-0.618

Model: Given Similarity = diff

Formula: LogRT ~ HighestCorr + (1 | Participant) + (1 | Item)

Data: subset(expdata, Similarity == "diff")

REML criterion at convergence: -1738.5

Scaled residuals

Min	1Q	Median	3Q	Max
-3.0725	-0.6513	-0.0422	0.5704	3.378

Random effects

Groups Name	Variance	Std.Dev.
Item (Intercept)	0.00091	0.03024
Participant (Intercept)	0.00176	0.04192
Residual	0.00224	0.04729

Number of obs: 575, groups: Item, 40; Participant, 15

Fixed effects

	Estimate	Std. Error	df	t- value	Pr(>|t|)
(Intercept)	2.80253	0.01497	36.7464	187.236	<2e-16
HighestCorrNH	0.01125	0.01194	37.7615	0.942	0.352

Correlation of Fixed Effects

	(Intr)
HighstCrrNH	-0.598

Model: Given highestCorr = Highest

Formula: LogRT ~ Similarity + (1 | Participant) + (1 | Item)

Data: subset(expdata, HighestCorr == "H")

REML criterion at convergence: -798.8

Scaled residuals

Min	1Q	Median	3Q	Max
-2.5013	-0.6091	-0.0483	0.5365	3.9533

Random effects

Groups Name	Variance	Std.Dev.
Item (Intercept)	0.00127	0.03557
Participant (Intercept)	0.00144	0.03795
Residual	0.00229	0.0479

Number of obs: 274, groups: Item, 19; Participant, 15

Fixed effects

	Estimate	Std. Error	df	t- value	Pr(>|t|)
(Intercept)	2.80216	0.01544	28.0799	181.455	<2e-16
Similaritysim	0.00951	0.01735	16.649	0.548	0.591

Correlation of Fixed Effects

	(Intr)
HighstCrrNH	-0.532

Model: Given highestCorr = NotHighest

Formula: LogRT ~ Similarity + (1 | Participant) + (1 | Item)

Data: subset(expdata, HighestCorr == "NH")

REML criterion at convergence: -2266.7

Scaled residuals

Min	1Q	Median	3Q	Max
-3.1144	-0.6241	-0.0467	0.5625	4.4848

Random effects

Groups Name	Variance	Std.Dev.
Item (Intercept)	0.0009	0.03007
Participant (Intercept)	0.00188	0.04331
Residual	0.00218	0.04664

Number of obs: 739, groups: Item, 51; Participant, 15

Fixed effects

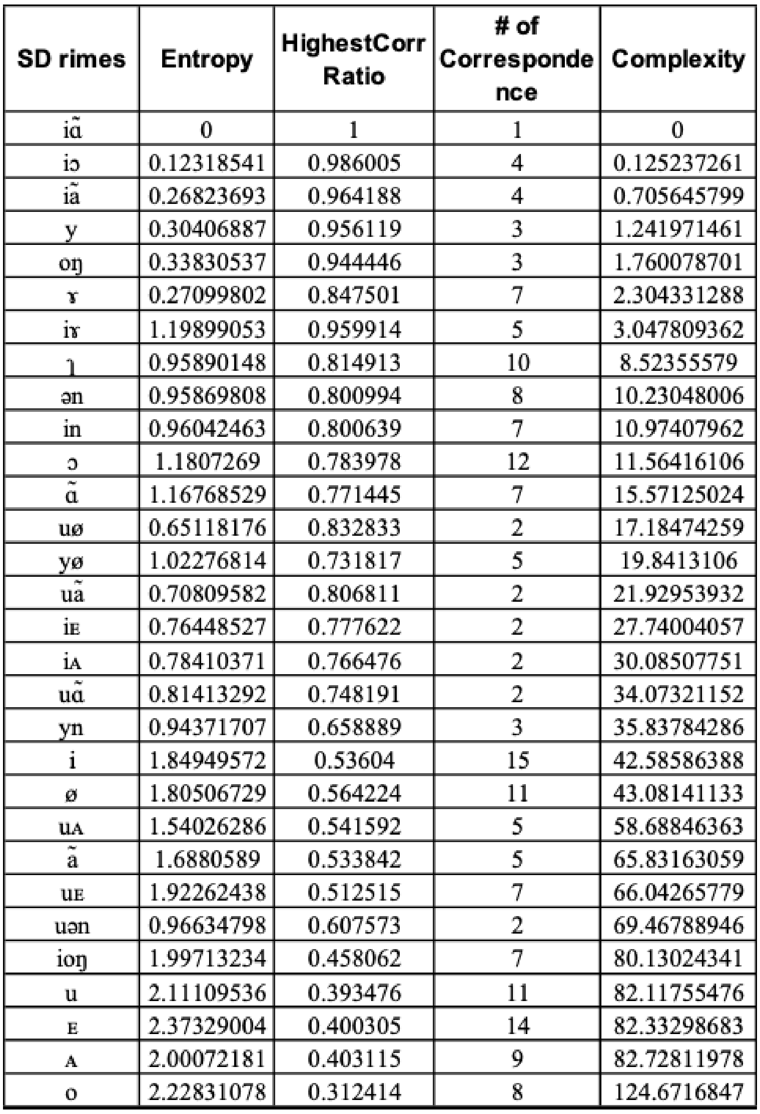
	Estimate	Std. Error	df	t- value	Pr(>|t|)
(Intercept)	2.81395	0.01266	21.3695	222.256	<2e-16
Similaritysim	-0.0282	0.00924	48.6281	-3.055	0.00365

Correlation of Fixed Effects

	(Intr)
HighstCrrNH	-0.301

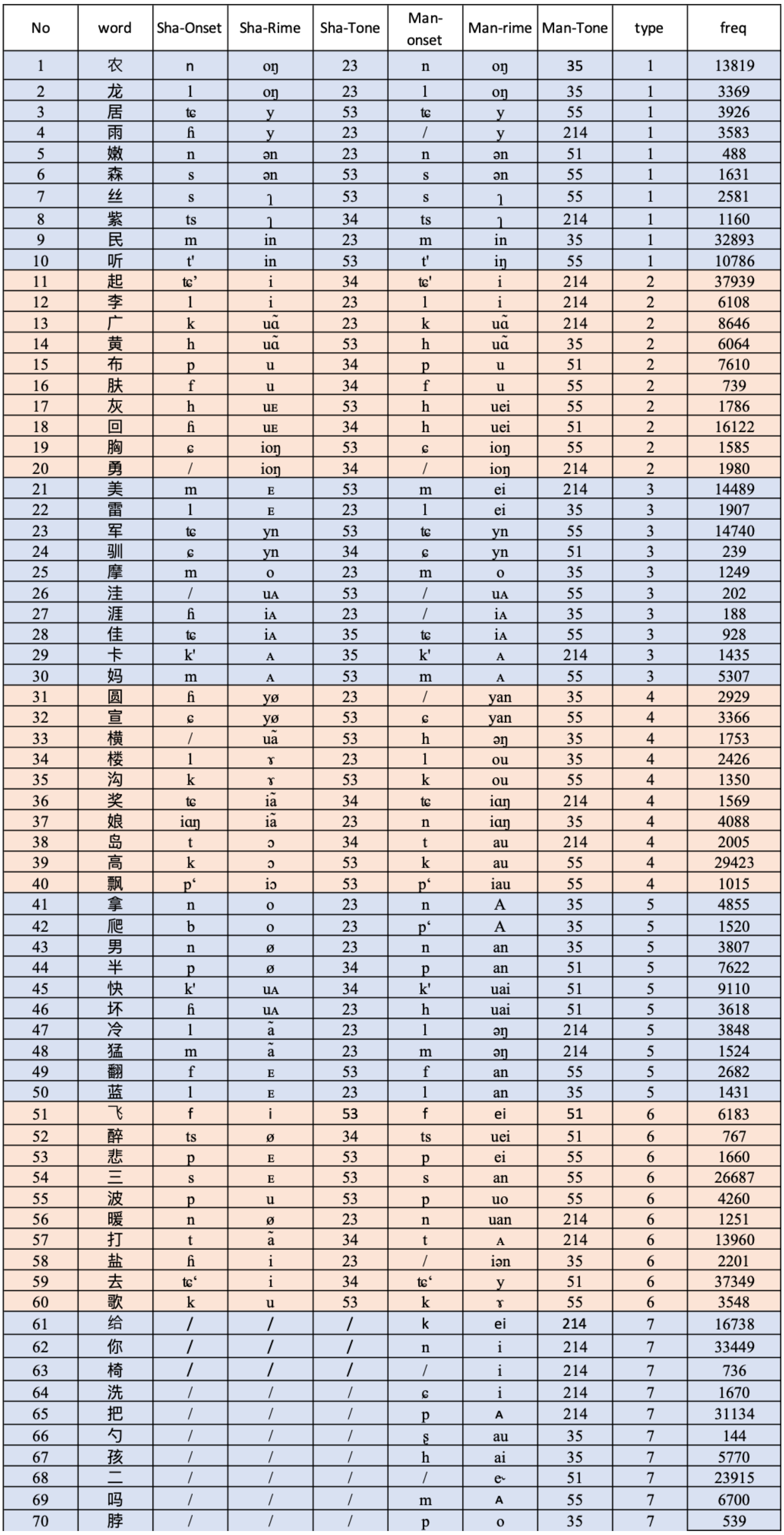
Figure 8: Statistics summary of the linear models conditioned on phonetic similarity or whether the rime is the most frequent match (highest correspondent)

7.3 Shanghai dialect and Mandarin correspondent analysis & List of Stimuli



SD rimes	Entropy	HighestCorr Ratio	# of Correspondence	Complexity
iɑ̃	0	1	1	0
iɔ	0.12318541	0.986005	4	0.125237261
iã	0.26823693	0.964188	4	0.705645799
y	0.30406887	0.956119	3	1.241971461
oŋ	0.33830537	0.944446	3	1.760078701
ɤ	0.27099802	0.847501	7	2.304331288
iɤ	1.19899053	0.959914	5	3.047809362
ɿ	0.95890148	0.814913	10	8.52355579
ən	0.95869808	0.800994	8	10.23048006
in	0.96042463	0.800639	7	10.97407962
ɔ	1.1807269	0.783978	12	11.56416106
ɑ̃	1.16768529	0.771445	7	15.57125024
uø	0.65118176	0.832833	2	17.18474259
yø	1.02276814	0.731817	5	19.8413106
uã	0.70809582	0.806811	2	21.92953932
iᴇ	0.76448527	0.777622	2	27.74004057
iᴀ	0.78410371	0.766476	2	30.08507751
uɑ̃	0.81413292	0.748191	2	34.07321152
yn	0.94371707	0.658889	3	35.83784286
i	1.84949572	0.53604	15	42.58586388
ø	1.80506729	0.564224	11	43.08141133
uᴀ	1.54026286	0.541592	5	58.68846363
ã	1.6880589	0.533842	5	65.83163059
uᴇ	1.92262438	0.512515	7	66.04265779
uən	0.96634798	0.607573	2	69.46788946
ioŋ	1.99713234	0.458062	7	80.13024341
u	2.11109536	0.393476	11	82.11755476
ᴇ	2.37329004	0.400305	14	82.33298683
ᴀ	2.00072181	0.403115	9	82.72811978
o	2.22831078	0.312414	8	124.6716847

Figure 9: Analysis on the rime correspondence between Shanghai dialect and Mandarin



No	word	Sha-Onset	Sha-Rime	Sha-Tone	Man-onset	Man-rime	Man-Tone	type	freq
1	农	n	oŋ	23	n	oŋ	35	1	13819
2	龙	l	oŋ	23	l	oŋ	35	1	3369
3	居	tɕ	y	53	tɕ	y	55	1	3926
4	雨	ɦ	y	23	/	y	214	1	3583
5	嫩	n	ən	23	n	ən	51	1	488
6	森	s	ən	53	s	ən	55	1	1631
7	丝	s	ɿ	53	s	ɿ	55	1	2581
8	紫	ts	ɿ	34	ts	ɿ	214	1	1160
9	民	m	in	23	m	in	35	1	32893
10	听	t'	in	53	t'	iŋ	55	1	10786
11	起	tɕ’	i	34	tɕ'	i	214	2	37939
12	李	l	i	23	l	i	214	2	6108
13	广	k	uɑ̃	23	k	uɑ̃	214	2	8646
14	黄	h	uɑ̃	53	h	uɑ̃	35	2	6064
15	布	p	u	34	p	u	51	2	7610
16	肤	f	u	34	f	u	55	2	739
17	灰	h	uᴇ	53	h	uei	55	2	1786
18	回	ɦ	uᴇ	34	h	uei	51	2	16122
19	胸	ɕ	ioŋ	53	ɕ	ioŋ	55	2	1585
20	勇	/	ioŋ	34	/	ioŋ	214	2	1980
21	美	m	ᴇ	53	m	ei	214	3	14489
22	雷	l	ᴇ	23	l	ei	35	3	1907
23	军	tɕ	yn	53	tɕ	yn	55	3	14740
24	驯	ɕ	yn	34	ɕ	yn	51	3	239
25	摩	m	o	23	m	o	35	3	1249
26	洼	/	uᴀ	53	/	uᴀ	55	3	202
27	涯	ɦ	iᴀ	23	/	iᴀ	35	3	188
28	佳	tɕ	iᴀ	35	tɕ	iᴀ	55	3	928
29	卡	k'	ᴀ	35	k'	ᴀ	214	3	1435
30	妈	m	ᴀ	53	m	ᴀ	55	3	5307
31	圆	ɦ	yø	23	/	yan	35	4	2929
32	宣	ɕ	yø	53	ɕ	yan	55	4	3366
33	横	/	uã	53	h	əŋ	35	4	1753
34	楼	l	ɤ	23	l	ou	35	4	2426
35	沟	k	ɤ	53	k	ou	55	4	1350
36	奖	tɕ	iã	34	tɕ	iɑŋ	214	4	1569
37	娘	iɑŋ	iã	23	n	iɑŋ	35	4	4088
38	岛	t	ɔ	34	t	au	214	4	2005
39	高	k	ɔ	53	k	au	55	4	29423
40	飘	p‘	iɔ	53	p‘	iau	55	4	1015
41	拿	n	o	23	n	A	35	5	4855
42	爬	b	o	23	p‘	A	35	5	1520
43	男	n	ø	23	n	an	35	5	3807
44	半	p	ø	34	p	an	51	5	7622
45	快	k'	uᴀ	34	k'	uai	51	5	9110
46	坏	ɦ	uᴀ	23	h	uai	51	5	3618
47	冷	l	ã	23	l	əŋ	214	5	3848
48	猛	m	ã	23	m	əŋ	214	5	1524
49	翻	f	ᴇ	53	f	an	55	5	2682
50	蓝	l	ᴇ	23	l	an	35	5	1431
51	飞	f	i	53	f	ei	51	6	6183
52	醉	ts	ø	34	ts	uei	51	6	767
53	悲	p	ᴇ	53	p	ei	55	6	1660
54	三	s	ᴇ	53	s	an	55	6	26687
55	波	p	u	53	p	uo	55	6	4260
56	暖	n	ø	23	n	uan	214	6	1251
57	打	t	ã	34	t	ᴀ	214	6	13960
58	盐	ɦ	i	23	/	iən	35	6	2201
59	去	tɕ‘	i	34	tɕ‘	y	51	6	37349
60	歌	k	u	53	k	ɤ	55	6	3548
61	给	/	/	/	k	ei	214	7	16738
62	你	/	/	/	n	i	214	7	33449
63	椅	/	/	/	/	i	214	7	736
64	洗	/	/	/	ɕ	i	214	7	1670
65	把	/	/	/	p	ᴀ	214	7	31134
66	勺	/	/	/	ʂ	au	35	7	144
67	孩	/	/	/	h	ai	35	7	5770
68	二	/	/	/	/	e˞	51	7	23915
69	吗	/	/	/	m	ᴀ	55	7	6700
70	脖	/	/	/	p	o	35	7	539

Figure 10: List of stimuli and their information