Phonological Mapping and Shared Abstract Phonological Representations: Cognate Effects Found in Shanghai dialect - Mandarin Bilinguals

19300110001 Zhisheng Jin
Department of Chinese Language and Literature
Fudan University

Abstract

Previous studies on bilingual speech production have exploited the cognate effects as an indicator of the connectionist structure of bilingual lexical and phonology storage. The current study uses cognates from Mandarin Chinese and the Shanghai dialect to test the WEAVER++ model and found that even when cognates have significant discrepancies in phonetics, the perfect phonological mapping relation between the Mandarin and Shanghai dialect representations ensures cognate facilitation effects, while this facilitation can be eliminated by the inhibitory effect stemming from the similarity of the phonetic representations of the cognate. The results are congruent with the prediction of the WEAVER++ model, which suggests that phonological representations are shared across languages and the phonetic encoding layer does not prevent activation from the unwanted language. Further, the finding also supports an exemplar-based model of phonology storage because the rule-based storage model is incapable of explaining the abstractness of the phonological representation.

Keywords: Bilingualism, speech production, phonological encoding, Chinese dialect

1. Introduction

1.1 Cognate Effects and Bilingual Speech Production Models

To construct an apt model for bilingual speech production, multiple theories have been put forwards to befit empirical evidence. There are two main consensuses among researchers - one is that lemmas compete for lexical selection, and the other is that the languages of one bilingual are both activated even in monolingual situations (Costa and Caramazza, 1999; Costa and Santesteban, 2004, 2006; Kroll et al., 2006; Branzi et al., 2014). One robust experimental effect has been detected and attested by an abundance of studies on bilingual language production (Costa et al., 2000) and has been considered a piece of evidence for parallel co-activation across languages.

Cognates are translation equivalents that share similar phonological representations and meanings (e.g. *guitar* in English and *guitare* in French). Cognate effects are mainly demonstrated by exploiting the tasks which mainly test the reaction time (RT). Being cognate may result in significantly different reaction times compared with the controlled conditions.

Cognate effects have an important role in constructing a functional structure for the bilingual lexicon and language production model not only because it is robust in all tasks deployed, but also for the reason that it is a kernel to investigate the status of the bilingual linguistic mechanism.

In the construction of models for bilingual lexicon, especially for speech production, most researchers have agreed on the shared concepts and separated representations between two languages, while the mediating process of lexical selection leaves much room for further hypotheses. There are several paradigms frequently used throughout this investigation, but not all of them should be considered proper for the purpose. First and foremost, to investigate the activation of the language not in use, researchers should strive to create a monolingual situation with undetectable interference for the better mimicking of the cognitive process (Costa et al., 2006). In this regard, the previously prevailing interference paradigm (Schriefers, Meyer, and Levelt, 1990; Starreveld and La Heij, 1996; Caramazza and Costa, 2000) and phoneme monitoring paradigm (Hermans, 2000; Colomé, 2001) both fall short in that they would introduce a certain degree of activation on the language not in use with their experimental conditions.

To be more specific, the interference paradigm attempted to disrupt the lexical process of the bilingual participants and considered their response latencies accompanied by the interference as a piece of evidence supporting parallel activation in the bilingual's lexicon (Costa & Caramazza, 1999; Hermans, Bongaerts, de Bot, and Schreuder, 1998). However, the results might not be the logical results of primitive parallel activation because the stimuli themselves induced in the cognitive process might exaggerate the activation of the language not in use. In terms of the phoneme monitoring tasks, participants are asked to decide whether a presented phoneme exists in the name of the corresponding picture. The given phonemes might arouse perplexity for processing when the name of the language not in use corresponding to the picture happened to contain them in the first place. Despite the removal of obvious hints for the language not in use, the given phonemes themselves can be processed and elevate the activation of the counterpart of the target lexicon.

For these concerns, cognates are thought to be the stimuli proper as a kind of linkage between two languages that can be easily manipulated with their various properties. That is to say, cognates are helpful for researchers to build a monolingual experiment with room for conditioning the variables to this the bilinguals' cognitive process.

In terms of theoretical constructions, cognate also plays a crucial role as the interface of the two lexica. As mentioned above, general agreements have already been made on the concept and the representation levels, while the process of lexical access is still contestable. The most controversial topic in this domain is the language-specific selection hypothesis versus the language-non-specific selection hypothesis. Namely, most of the theories on bilingual lexical access have supposed that the parallel activation of lexica of the two languages is on-line in nature (Costa et al., 2017). In this sense, the robust cognate effects can be construed as salient evidence for spread activation in that during the mapping from the concept to the lemmas, only parallel activation in the two lexica may contribute to the common facilitation effect found in the aforementioned experiments. However, if one constructs the model on a Hebbian basis, the cognates may win out their competitors by their relatively higher frequency due to the activation of the translation equivalents (Costa et al., 2017). In both cases, cognates are worthy of our investigation on a behavioral level for their unique place across the lexica.

On the other hand, cognate effects are not monotone in their nature, various studies have investigated how cognate effects might manifest themselves under task demands, language proficiency, and language used in the task. For example, Broersma et al. found that highly proficient English-dominated bilinguals exhibited a cognate inhibition effect in picture naming tasks in Welsh, whereas Welsh-dominated bilinguals show a facilitation effect in Welsh trials. They postulated that the lexicon competition within lemmas and the parallel activation of the representations are the underlying currents beneath the cognate effects. The inhibition effect might be resulted from lemma competition causing too much hindrance for English-dominated bilinguals in Welsh trials due to their lower frequencies in Welsh, while for Welsh-dominated bilinguals the parallel activation in the representational level outweighed the inhibition from lexicon competition. Furthermore, they found that cognate inhibition is a behavioral adaption effect, which entails that repression against unwanted language does occur to balance out the perplexity in the lexical selection process. These findings pointed out that cognate as the cross point of two lexica may further our understanding of the sophisticated mechanism in the cognitive process[1], especially in terms of the competition between representations.

Despite the astonishing findings in the previous bilingualism studies on cognate effects, there are still questions to be answered. First, cognate effects are insufficient for us to distinguish the parallel activation and the distributive account for the aggregated lexicon, which is crucial for understanding the structure of the bilingual mental lexicon. Secondly, dynamics in the cognate effects concerning various parameters have not been fully explored, while this endeavor may point to some interesting inference (Costa and Santesteban, 2004). More specifically, most researchers only consider the lemma selection while neglecting the phoneme retrieval process or seeing it as subordinate to the lemma selection. It is especially crucial to separate the two stages using the cognate paradigm in that for most of the cognates their meaning and phonetics are both similar. If the production is accelerated only because of their phonemic proximity or other variants that may result in the acceleration, it would be difficult to suggest that being a cognate is the sole reason for the cognate's idiosyncrasy. Thirdly, researches on the diversity of cognates, especially on manipulating various facets of cognates to incite different behavioral response are still scarce. Two astonishing pieces of research concerned with this innovative approach are Comesaña, Montserra, et al.(2015)'s investigation of the interplay of cognate facilitation effect and the orthographic overlap and Woumans et al.(2021)'s study on written word production. Thus, the current study aims

[1] In Broersma et al.'s study, a mixed-picture naming task is employed to create an asymmetric condition for L1 to L2 switch and vice versa. However this strategy might render their conclusion on cognitive inhibition as a behavioral adaptation effect questionable in that random language switch is not natural in everyday communication of the bilinguals, so the behavioral adaptation effect may be a result from the paradigm instead of the reflection of the nature of the bilingual lexicon selection process.

to introduce the diversity of the cognates and to see how it may lead to novel findings concerning the phonological encoding stage in bilingual language production.

1.2 Bilingualism in China and 'Chinese Cognates'

Most of the current studies on cognate effects are conducted among bilinguals with two Indo-European languages, for example, Spanish - English bilinguals or Dutch - English bilinguals. While the concept of cognate may be acquainted with those circumstances, on the other part of the continent, where 1.4 billion Chinese users reside, cognate has a somewhat different implication. Chinese is an all-encompassing name for various dialects scattered in the territory of China (Norman, 1988). Despite heated debates over the demarcations of those dialects, a consensus among Chinese linguists is that Chinese dialects can be classified into nine major groups as follows: Mandarin, Jin, Wu, Hui, Xiang, Gan, Kejia, Min, and Yue (Hou et al., 2002). Most of these dialects are independent language systems ramified from the Ancient Chinese used by the Zhou Dynasty on the North China Plain around the mid of the second millennium B.C. (Norman, 1988, p. 19). Their segregated usage distinct by different ethnic groups in various provinces of China contributed to their distinct status from the current koine, i.e. Mandarin Chinese.

Due to their confounding development and the interplay with the ever-existing koine approbated by the governments, the diversity among dialects is considerable, especially between the North and the South (Norman, 1988, p.213). In terms of the two languages in the current study, the Shanghai dialect has a distinct linguistic system compared with Mandarin, especially in its phonology and lexicon. Nevertheless, from a more specific view, the lexicon of the Shanghai dialect, just like all the other dialects of Chinese, is an amalgamation constituted of the Mandarin lexicon (usually new or abstract concepts) and colloquial expressions (usually ordinary objects and vivid descriptions). It is the former in which I can find parallel lexica with different phonological representations respective to the ones in Mandarin, the koine. For example, the phonological representation for computers in the Shanghai dialect is [di^{31} nɔ^{13}], while in Mandarin it is [tiən^{51} nau^{213}], while they share a biomorphic structure, their manners of articulation of the consonants also resemble each other, and even their rhymes are somewhat similar. Furthermore, if I take two parallel clusters of such 'cognates' into question, I can find a phonological mapping pattern between them:

Dialect	电	大	验	人	日	垠	爬	鼻	马
Mandarin	tiən˥˩	tʌ˥˩	iən˥˩	ɹən˧˥	ɹʅ˥˩	iən˧˥	p'a˥˩	pi˧˥	mʌ˨˩˦
Shanghai	di˧˩	du˩˧	ȵi˧˥	ȵin˩˧	ȵjəʔ˧	ȵin˩˧	bu˧˩	bjəʔ˧	mu˩˧

In this table, each group (indicated by the grey & white areas) consisting of three characters represents a distinct morpheme in both languages. The second character in the group shares the consonant with the first, and the third shares the rhyme. The mapping pattern in the pronunciation of the Shanghai dialect concurs with the Mandarin ones. The logic behind this phenomenon is their shared historical root in Ancient Chinese phonology. The rigid monosyllabic structure of Chinese phonology imposes this patent mapping rule onto their historical development, hence this kind of cognate can be more systematic paragons compared with the Indo-European ones and may entail a different kind of lexical processing approach in the mind of the bilinguals in question. Another good thing about this kind of cognate is that they consist of the majority of both of the lexica where most morphemes are identical, which precludes the doubts on whether cognates and other words are stored separately or whether it is the morphological encoding that contributes more to the cognate effects observed since these two parameters can be controlled.

The obvious question that follows is whether the mental lexicon of dialect-Mandarin bilinguals is also capable to demonstrate a similar cognate effect encountering the semi-cognates which share the same semantic concept and different phonological forms. And if the answer is yes, what model should we choose to explain the mechanism of the bilingualism in China which is so common across the country? Would it be different from our current findings in the bilingual experiments?

1.3 Bilingual Models on Phonological Encoding

One prevailing model in phonological encoding in bilinguals is the WEAVER++ model (e.g. Levelt, Roelofs, and Meyer, 1999; Roelofs, 2015). In this model, three stages are performed to output a phonological word after the lemma has been retrieved, which are the morphological encoding, the phonological encoding, and the phonetic encoding. The phonological encoding consists of segmental spell-out and prosodification, which is where the phonemic segments of the morphemes are retrieved, while the phonetic encoding is where the phonemic segments are substantialized into phonetic representations to articulate. On bilingual word production, the WEAVER++ model also implements the assumption that the phonemic segments are shared between different language systems in bilinguals' minds, while the phonetic representations which belong specifically to each language differ (Roelofs, 2003). While the nature of bilingual lexical selection is still under debate, in the current research I have managed to control the lexical selection thanks to the shared morpheme in the Shanghai dialect and Mandarin Chinese to accord with the WEAVER++ model, which hypothesizes that only lemma in targeted language will enter the further encoding process. Thus the effects observed should only result from the three stages mentioned above. On the other hand, the WEAVER++ model adopts condition-action rules which claim that the selection of phonemic and phonetic representations would follow the language tag when the language has been specified by the context (Roelofs and Verhoef, 2006). This proposition was backed up by monitoring the between-language facilitation effect while changing the SOA in a picture-naming paradigm in a second language with distractions from the first language (Hermans et al., 1998). The condition-action rules entail that there is no inhibition when the language tag has been sealed.

There are a series of research on whether the phonological representations and phonological activation are shared across two languages. Cognates, being phonologically related pairs in two languages, are widely utilized in those studies. Costa et al. (2000) employed picture-naming tasks to record the reaction time of the Catalan-Spanish bilinguals' word production and observed that words with translation equivalents that are phonologically related yielded faster responses than those that are not. Colomé (2001) came to a similar result by asking Catalan-Spanish bilinguals to conduct internal phonemic monitoring tasks with some of the stimuli having a translation equivalent to phonemic-related segments. These two studies both exhibited shared activation in phonological encoding.

To test whether phonological representations are shared, Roelofs (2003) used a form-preparation paradigm to test whether the phonemes were prepared even in different target languages and the results showed that while /p/ did not yield a preparation effect in English trials, /b/ yielded one across English and Dutch trials, which verified the assumption of shared phonological representations. However, research on the implicit priming effect found that the role of segments (elements prepared by the priming effect) in English and Dutch, was replaced by syllables and mora in Chinese and Japanese (Chen et al., 2002; O'Seaghdha et al., 2010; Kureta et al., 2006). These results imply that the nature of the metrical frame and the prosodification process is largely dependent on specific language properties and only when the metrical frame and the initials are shared can there be a phonological preparation before the articulation. Other researchers found that in Mandarin, atonal syllables have their psychological representations. With its restricted size of atonal syllable inventory (approximately 400) and morpheme-to-tonal syllable mapping, it seems reasonable for Mandarin speakers to put only the syllabic construction into function during the phonological encoding, but Wong and Chen's research on Cantonese discovered that sub-syllabic segments such as the onset and rime or the rime and coda can also bring about an implicit priming effect regardless of the tone in a picture-word interference paradigm (Wong & Chen, 2009). Further investigation of Cantonese speakers observed similar results in the implicit priming paradigm (Wong & Chen, 2012), indicating that the question of the functional phonological unit in various dialects of Chinese is still debatable. Thus, while shared phonological representations and activations have been recognized as a consensus, questions on the functional phonological units and the mechanism of the serial association of the segments to prosody in specific languages remain untapped. In terms of bilinguals, most of the investigations on this topic mainly consider bilingualism as a mediator to fathom the essence of the phonological representation, whilst in the bilingualism edifice, the depth of the phonological encoding process has not been gauged yet. Further research on this topic includes Li et.al. (2017) and Nakayama et al. (2016)'s research on Chinese and Japanese ESLs respectively which revealed that bilinguals prepare their phonological units differently with regard to the native monolinguals in English and the discrepancy diminishes alongside the raise of proficiency. In terms of backward transference, Verdonschot et al. (2013) and Ida et al. (2015) both found that the L2 of highly proficient bilinguals' L1 phonological planning would experience assimilation to their L2. In Wong & Chen's (2021) latest study, Chinese-English bilinguals whose residence in English-speaking regions was longer were found to be more syllable-focused in terms of phonological planning, which can be concluded as a backward transference effect. Both backward and forward transference effects can be nicely interpreted by a set of shared phonological representations, but there can be other explanations.

In a nutshell, following the WEAVER++ model, a spate of research has been conducted to test bilingualism's interplay in the phonological encoding process of bilinguals. Much has been discussed on the topic of PU (phonological units), but the nature of phonological representations and the parallel activation in the phonetic encoding process especially in Chinese is yet to be studied.

1.4 The Current Study

From our scope, no previous research has been conducted on Mandarin - Chinese dialect bilinguals' phonological encoding. This topic is worth being tapped upon in that China is a country with rich language resources. Even though the ever-changing koine reins the whole country and shapes every single dialect to at least some extent. If I view the diglossia in this country as a kind of bilingualism, I can not only transform the deadpan of the regime of Mandarin into a vociferous pool of language mutation and interaction, but fathom an extraneous king of bilingualism other than the Indo-European ones as well.

The current study sets out to investigate whether the phonological mapping rules bring an effect on the phonological encoding phase of Mandarin-dialect speakers. The question is intuitive for two reasons. First, almost all dialects share a similar morpheme set ramified from the Middle Chinese, the most obvious difference in them is putatively believed to be phonological. Secondly, most dialect speakers have at least some command of Mandarin, and youngsters are typically good at Mandarin for the mandatory inculcation in schools and the prevalence of mass entertainment. It is natural for dialect speakers to conclude a rudimentary phonological correspondence from their dialects to

Mandarin in their language experience. Whether this rustic phonological correspondence plays a role in their phonological planning is the key question in the current study.

In this study, I also adopt the WEAVER++ model (Roelofs, 1997) for its simplicity and for situating ourselves on the spectrum of the previous research on phonological encoding in bilinguals. There are a few assumptions to be made based on the WEAVER++ model. First of all, there is predicted to be a facilitation effect between close phonemic representations in the phonological encoding because they are shared between languages while competition in phonetic encoding from the activated phonemic representations due to their similarity and the cost of executive control. Thus, if the phonemic representations of one word are similar in two languages while the phonetic representations are not, the facilitation effect should be significant. However, the proximate units principle raised by O'Seaghdha et al. (2010) suggests contrarily that the Shanghai dialect may differ from Mandarin Chinese in its phonological units. Thus, little or no facilitation effect is expected due to the structural difference in their phonological units and the condition-action rules which implies that no cross-language spread activation from phonemic units to phonetic units would occur.

Finally and specifically, I used three variables to differentiate 5 kinds of stimuli: morpheme agreement (MA), phonological deductibility (PD), and phonetic similarity (PS). MA is straightforward as it evaluates whether all the morphemes in one disyllabic stimulus are the same. PD evaluates whether the phonetic segments of a stimulus in the Shanghai dialect can be deduced from its Mandarin counterpart. For instance, most of the words with the onset [ȵ] in the Shanghai dialect are pronounced as [ɹ] in the Shanghai dialect and most words with a rime [in] in the Shanghai dialect are pronounced as [ən] in Mandarin. So the word "人" with the pronunciation of [ȵin] in the Shanghai dialect has a solid PD when its Mandarin counterpart is pronounced as[ɹən]. PS evaluates the phonetic similarity regarding the natural class, and a detailed description of the criterion will be presented in the Experiment Material section.

2. Methods

2.1 Materials

From the book *Shanghai Urban Dialect* (Qian, Tang, 1988) I collected 111 words with 5 different phonological correspondence categories (referred as word types in the following), listed in the appendixes below. Our method to differentiate these categories is to compare each segment of their phonetic representations in Mandarin and the Shanghai dialect and to tell how different these phonetic representations are based on natural categories of these segments (i.e. places of articulation, vowel qualities, etc.). Due to the convoluting nature of the tonal sandhi in the Shanghai dialect, I did not intentionally control the tone of the representations.

The Type 1 words have the highest PS because their phonetic representations in Mandarin and Shanghai dialect only have at most one segment that is different in either voicing status or vowel duration, namely, these differences include '[o-uo] [ɔ-au] [ieʔ-i] [v-f] [b-p]'. The Type 2 and 3 words are the same in terms of whether the difference of the segments in the representations can be predicted or deduced following certain conspicuous laws of correspondence but different in terms of the position of the difference between the phonetic representations, i.e., the difference in representations of the Type 2 words are all located in the first morpheme while the Type 3 words the second. This variable is important in the sense that it is suggested in Wong & Chen's study that the position of the difference makes an impact on the RT of bilingual participants (2009). The Type 4 and Type 5 words are the same in the sense that they are the opposite of Type 2 and 3 words in that the difference of representations of the Type 4 and 5 words do not follow a predictable law of correspondence. In addition, the Type 4 and 5 words are different in that the stochastic difference in representations of the Type 4 words does not entail a difference in morpheme, i.e., although one is unable to predict the Shanghai dialect counterpart from the Mandarin representation, the representations themselves are not constructed by different morphemes. For example, in the word "围巾 (scarf)" the morpheme "围" pronounced as [ɦy] is only used in this single word, and the natives all write this word in this exact character, which implies that [ɦy] is just a variation of this morpheme "围", which most often is pronounced as [vei]. On the other hand, words in the fifth category are different in terms of morphemes between the two languages in question. For instance, the concept of "爆竹" in the Shanghai dialect is represented by "炮仗" [pau^{51} tʂu^{51}], which is another different word in Mandarin, though with the same meaning. A table summarizing the stimuli is listed below:

Word Type	Number	The differences in the first morpheme	The differences in the onset of the first morpheme
1	20	4	4
2	30	30	22
3	20	0	0
4	20	15	14
5	21	18	18

I gathered 111 black-white line-drawing picture stimuli to match the words from various sources and used Photoshop to crop them into 500*500 JPG files. To standardize the picture stimuli, I conducted an additional experiment on name agreement, familiarity, and imaginability. Due to the COVID-19 policy and the lockdown in Shanghai from March to June, I used a questionnaire to experiment with image agreement and concept familiarity. Detailed procedure and results will be discussed in the Experiment section. Sixteen participants who grew up in Shanghai with the

Shanghai dialect being his or her first language completed the questionnaire. After evaluating the score of each stimulus, 93 picture stimuli remained and were used in the experiment.

2.2 Participants

Due to COVID-19 policy and the lockdown in Shanghai from March to June, I was unable to recruit participants on campus or in scattered locations in Shanghai. In the end, 10 participants (5 males, 5 females) participated in Experiment Two (the main experiment). All participants grew up in Shanghai's urban districts (mainly Yangpu and Hongkou), whose ages range from 21 to 30 (mean = 27.4, SD=4.35). By implementing the BLP (Bilingual language proficiency) questionnaire (Birdsong et al., 2012), we evaluated their language proficiency and they yielded high scores in both languages. In terms of the participants of the questionnaire (experiment one), 3 of them are female, and their ages have a mean of 20.94 while the SD = 1.95. Their BLP results show that they are not balanced bilinguals but have a good command of the Shanghai dialect (mean = 130.67, SD=29.09). the language proficiency of the two groups of participants both follow normal distribution according to the Shapiro-Wilk test and have no significant difference according to the independent sample t-test (p=0.231) in terms of the Shanghai dialect. However, there was a significant difference in their Mandarin proficiency ($p<0.05$). Participants in Experiment Two are not as good at Mandarin as participants in Experiment One (mean = 178.92, SD = 16.40). This entails that their response to the stimuli would be more volatile and would yield a greater variance.

language	mean	SD
Shanghai dialect	143.45	19.13
Mandarin	161.53	13.93

2.3 Experiment Procedures

2.3.1 Experiment One Experiment One was for name agreement, familiarity, and imaginability of the stimuli. The procedure was a combination of a BLP questionnaire and 111 self-monitored pictures-naming tasks with self-rating of familiarity of the response word and the imaginability of the concept represented by the picture, basically following the norm established by Snodgrass & Vanderwart (1980). The questionnaire was run on an online questionnaire platform in China (www.wenjuanxing.com). To evaluate the name agreement of each stimulus, we used the H-index which put the number of responses for each picture stimulus into this formula:

$$\mathrm{H} = \sum_{i=1}^{k} p_i \log_2 p_i$$

p_i stands for the proportion of one response in all responses, and k represents the total number of responses. The lower the h-index is, the better the name agreement for the stimulus in question while naming failures are not counted in the calculation (Snodgrass & Vanderwart, 1980). Familiarity was rated on a 1-5 scale and imaginability on a 1-7 scale, following Snodgrass & Vanderwart's approach. The results are listed below:

Index	Mean	SD
H-index	0.62	0.74
Familiarity	4.84	0.16
Imaginability	6.76	0.23

The familiarity and imaginability both have high scores because the stimuli are very simple objects in daily life, so based on the h-index I removed stimuli whose h-index was 1 SD higher than the mean, leaving 93 stimuli to be incorporated into the experiment.

2.3.2 Experiment Two Experiment Two was the main experiment of this research. It was run on PsychoPy 2022.1.4 (Peirce, J. W. et al., 2019) on an LG-Gram 2017 13" laptop with a displace resolution of 1920*1080, while the audio responses being recorded by Audio-Technica ATR2500 at the sampling rate of 48000 Hz (the microphone module in PsychoPy could only run under this sampling rate). The experiments were conducted in a quiet room where the ambient noise is under -50 dBFS. Participants were seated 90 cm in front of the computer screen and were told to sit straight and still during the whole experiment.

The experiment consists of one training session with 10 stimuli that would not be used in the following trials, and 4 blocks of trials in which 5 types of stimuli were equally allocated and randomized. The first block has 24 stimuli in total, while the other has 23. Each trial starts with a cross fixated at the center of the screen. After 2500ms, the stimuli would take the place of the cross and the program would automatically start recording participants' responses at the same time. Participants should focus and name the object in the picture stimuli as fast as possible. After 2000ms, the trial would be over, and another cross would show up at the center of the screen, as the start of the next trial. Participants could not pause during each block of the experiment, but they were allowed to rest as long as they want during the break between every block. After completing the five sessions of the experiment, participants should fill in a questionnaire that contains a BLP test and familiarity rating test for correct responses corresponding to the stimulus used in the experiment. The familiarity rating test requires participants to rate their frequency of using the correct word responses on a 0-3 scale.

3. Results

The reaction times from the naming tasks were examined, accompanied by the familiarity rates of the respective stimuli. The start point of the recording is the same as the appearance of the stimulus, so we can obtain the RT directly from the onset of the articulation. The built-in VoiceKey in PsychoPy was tested to be implausible for consonant onsets, so I used Praat to manually annotate

930 collected audio responses and obtained the RTs by ProsodyPro (Xu, 2013). In the annotation, I labeled all invalid responses including errors, blanks, etc. I also found that most RTs that are over 1800ms result from participants' inability to recognize the stimuli in the first place. So I also ruled out all the responses over 1800ms to preclude their impact on the data analysis. Another reason for me to rule out these lagging responses is that variances between subjects and between stimuli are huge, and the RTs of word types 2 and 3 are lower than those of 1,4, and 5. So removing lagging responses would not skew the general results and would even help us to see a clearer picture. Six hundred thirty-six responses were collected after these procedures.

On the other hand, we used the familiarity test completed by the participants to rule out more results. The familiarity test shows that the mean familiarity = 2.59 (SD = 0.36), and the familiarities of 51 stimuli are higher than the mean. A large proportion of those stimuli whose familiarity is below the mean were found to have more than 4 participants giving error responses, so I removed them as well, which leaves 51 valid stimuli to be incorporated into the data analysis. In total, 510 responses from 10 participants have been incorporated into the data analysis.

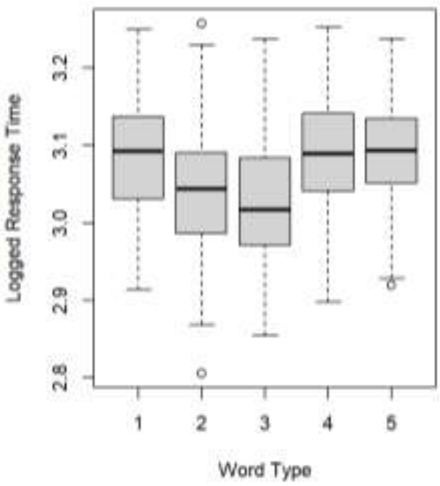

Fig 1. Logged reaction time for word types in the naming tasks of Experiment Two

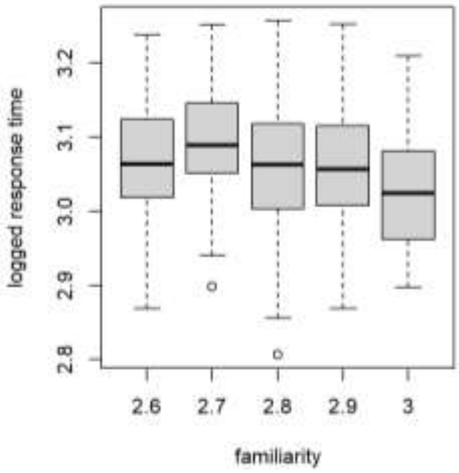

Fig 2. Logged reaction time for familiarity in the naming tasks of Experiment Two

The experiment was designed to be a 5*5 (word type and familiarity being within-subject variables) between 10 participants. The participants have different reaction capabilities and the stimuli with the same word types are fundamentally not the same in nature, so a series of Linear Mixed-Effects models were implemented using the lmerTest package (Kuznetsova, Brockhoff and Christensen, 2017) in the Rstudio, instead of MANOVA. In these models, participants and stimuli were set as random effects, and the word types and familiarity were set as fixed effects. Type 1 words were set to be the baseline of the control because the pronunciations of the lexical item in Mandarin and Shanghai dialect are generally the same, and we expected to see the maximum cognate facilitation effect from them.

Model 1: loggedRT ~ word_type + (1 | subject)
Model 2: loggedRT ~ word_type + (1 | subject) + (1 | stimulus)
Model 3: loggedRT ~ word_type + familiarity + (1 | subject) + (1 | stimulus)
Model 4: loggedRT ~ word_type + familiarity + word_type * familiarity + (1 | subject) + (1 | stimulus)

Model	#Df	LogLik	Df	Chisq	Pr(>Chisq)	
1	7	575.41				
2	8	606.79	1	62.7648	2.33E-15	***
3	9	605.84	1	1.9026	0.1678	
4	13	604.57	4	2.5478	0.6361	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1						

Table 1. Reaction Time likelihood ratio tests of Experiment Two: model comparison

To evaluate the LME models, likelihood ratio tests were also used to examine the effects of word type, familiarity, and their interaction. Table 1 shows the result of the likelihood ratio test and reveals that adding stimuli as a random effect significantly improved model A while adding familiarity and the interaction between familiarity and word type does not significantly improve the original model. Thus, we adopted model 2 as our linear mixed effect model, the summary is shown in Table 2.

Fixed effects						
	Estimate	Std.Error	df	t value	Pr(>|t|)	
(Intercept)	3.09E+00	1.77E-02	4.57E+01	175.071	<2.00E-16	***
word_type2	-4.55E-02	1.78E-02	4.60E+01	-2.557	0.01392	*
word_type3	-6.50E-02	2.11E-02	4.60E+01	-3.08	0.00349	**
word_type4	-8.63E-05	2.41E-02	4.60E+01	-0.004	0.99716	
word_type5	2.25E-03	1.96E-02	4.60E+01	0.115	0.90924	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

Random effects			
Groups	Name	Variance	Std.Dev.
stimulus	(Intercept)	0.0013589	0.03686
subject	(Intercept)	0.0008847	0.02974
Residual	0.004253	0.06522	
Number of obs: 510, groups: stimulus, 51; subject, 10			

Table 2. Fixed effects and random effects estimates result from Model 2

Model 2 reveals that there were significant effects of word Type 2 and word Type 3, indicating that the RT for the two kinds of stimuli was significantly different from the others, which matches the pattern in Fig 3, where the mean logged reaction times for Type 2 and Type 3 words are lower than the others.

Discussion

From the result, we can see a significant facilitation effect from the Type 2 and 3 words, but no facilitation for the baseline which is the type 1 words, which matches our prediction in the introduction. Moreover, there is no significant difference between model 2 and 3, which implies that familiarity does not play a role in affecting the reaction time, presumably because the stimuli we chose all yield high familiarity ratings from the participants. The type 1 words are the most typical cognates, as they resemble each other semantically, phonologically, and phonetically. The results suggest that type 1 words have no statistical difference from type 4 and type 5 words, while the type 5 words are supposed to be the confounding ones because the Mandarin and the Shanghai dialect morphemes are fundamentally different. The results countered our prediction and indicated that the Type 2 and Type 3 words are similar in nature while being different from the other types.

Type 2 and Type 3 words are typically the words with perfect correspondence across the two languages but some minute differences in the phonetic forms of the onset of the first syllables and the second syllables respectively. Compared with the Type 4 words which do not have perfect mappings between two languages, the facilitation effect could only be interpreted as the result of the phonological mapping between the two languages when all the variables were randomized and controlled. This finding is congruent with our prediction based on the WEAVER++ model, which states that phonological representations are shared between languages. To be more specific, if the phonological units of the two languages were not shared between the two languages, but strongly connected, the nodes of these separate phonological representations would have fired different prosodification rules and speech motor programs, then there would not be any facilitation effects for Type 2 and Type 3 words. The fact that facilitation only appears in Type 2 and Type 3 stimuli indicates that phonological representations are shared between languages.

More importantly, this finding supports the hypothesis that phonological mapping concluded from corpus-based synchronic phonology between two languages is psychologically real. The mapping rules purely concluded from historical phonology can be realized by the shared phonological representations in the mental storage and is highly abstract. Most of the stimuli from Type 2 and 3 words are drastically different in terms of phonetics, such as the morpheme “耳” in the word “耳朵(ear)”. The onset of it in Mandarin is [ɚ], while in the Shanghai dialect it is [n]. The result further implies that if the phonological representations are shared across languages in bilinguals’ minds, these differences were ignored at the phonological level where representations remain more abstract than expected. In terms of the storage of the phonological representations, the rule-based models or generative models are unable to explain this strong association across two languages in the sense that these merged phonological units contain information that cannot be abstracted into something that is related to the speech motor programs (because the mapping rules between different dialects and Mandarin are random from a synchronic perspective), but only to morphemes, which would lead to the redundancy of constructing the phonological level in the first place. The usage-based models like exemplar theories could explain the finding in Experiment Two contrarily. The merge or the link between two phonological units in two languages could be explained by frequent exposure and parallel activation (Guy, 2014).

The other interesting question is why the Type 1 words, the typical cognates, did not bring forth the usual cognate facilitation effects found in other research. Previous research has examined the inhibition effect in bilingual speech production. After the advance of the inhibition control model (Green,1998), an abundance of research supporting this model emerged, and the parallel activation of the two languages has been reckoned as a truism among researchers, the lack of facilitation effect of the type 1 words could be a result from this inhibitory control of the L2 language, Shanghai dialect. Namely, according to the ICM, during bilingual production, the lemmas are tagged with languages, so by activating the language in usage and inhibiting the output options from the language that is not in use, the inhibition is achieved (Van Assche, Duyck, and Gollan, 2013). And the short-lived, localized inhibition of a restricted set of words was supported by recent studies (e.g. Guo et al., 2011, Misra et al., 2012, etc.). The cognate facilitation effects could be offset by the inhibitory effect brought by the L2 language inhibition (the Shanghai dialect). Another possible explanation for this result is that the similarity between the cognates inhibited the output phonologically. Studies have found that the phonological representations are not separate (i.e. Costa, Santesteban and Cano, 2005, Roelofs, 2003) and the cognate facilitation effect is actually the result of the parallel activation of shared phonological representations, so only the prosodification and phonetic encoding part could offset the phonological boost. Thus, the finding supports the WEAVER++ model because even in L1 language production (the participants of Experiment 2 were better at Mandarin, while the AOA were similar), the prosodification rules of the language which was not in use and its speech motor programs were still activated. Either way, the inhibition could only happen after the phonological encoding.

In conclusion, the results from the experiments were analyzed by a series of linear mixed-effect models. The likelihood ratio test reveals that the familiarity of the stimuli does not affect the reaction performance, while the word type significantly influenced it. More specifically, the cognate stimuli with perfect phonological mappings between Mandarin and the Shanghai dialect and certain differences in phonetics yielded the largest cognate facilitation effects, and the position of the differences does not affect the RT significantly. The finding supports the hypothesis that the cognate stimuli with perfect phonological mapping and phonetic similarity encountered cognate inhibition effects stemming from inhibition control which inhibited the prosodification and phonetic encoding of unwanted language.

References

Acheson, D. J., Ganushchak, L. Y., Christoffels, I. K., & Hagoort, P. (2012). Conflict monitoring in speech production: Physiological evidence from bilingual picture naming. Brain and language, 123, 131-136.

Alexandra Kuznetsova, Per B. Brockhoff and Rune H. B. Christensen (2017) lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13), 1–26. doi:10.18637/jss.v082.i13

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. Neuropsychologia, 52, 102-116. doi:10.1016/j.neuropsychologia.2013.09.022

Broersma, M., Carter, D., & Acheson, D. J. (2016). Cognate costs in bilingual speech production: Evidence from language switching. FRONT PSYCHOL, 7, 1461-1461. doi:10.3389/fpsyg.2016.01461

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. Glot International 5:9/10, 341-345.

Broos, W. P. J., Bencivenni, A., Duyck, W., & Hartsuiker, R. J. (2021). Delayed picture naming in the first and second language. Bilingualism, 24(2), 389-400. doi:10.1017/S1366728920000620

Carolyn, Y., Fred, F., Ting, L., & Trey, H. (2004). A cross-culturally standardized set of pictures for younger and older adults: American and Chinese norms for name agreement, concept agreement, and familiarity. Behavior research methods, 36(4), 639.

Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., & García-Chico, T. (2015). Facilitative Effect of Cognate Words Vanishes When Reducing the Orthographic Overlap: The Role of Stimuli List Composition. J Exp Psychol Learn Mem Cogn, 41(3), 614-635. doi:10.1037/xlm0000065

COSTA, A., COLOMÉ, À, GÓMEZ, O., & SEBASTIÁN-GALLÉS, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. Bilingualism: Language and Cognition, 6(3), 167-179. doi:10.1017/S1366728903001111

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. Bilingualism: Language and Cognition, 2(3), 231-244. doi:10.1017/S1366728999000334

Costa, A., Heij, W. L., & Navarrete, E. (2006). The dynamics of bilingual lexical access. BILING-LANG COGN, 9(2), 137-151. doi:10.1017/S1366728906002495

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do Bilinguals Automatically Activate Their Native Language When They Are Not Using It? COGNITIVE SCI, 41(6), 1629-1644. doi:10.1111/cogs.12434

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects. Journal of memory and

language, 47(1), 91-124. doi:10.1006/jmla.2001.2840
Giezen, M. R., & Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. Bilingualism: Language and Cognition, 19(2), 264-276. doi:10.1017/S1366728915000097
Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism, 1(2), 67-81. doi:10.1017/S1366728998000133
Guy, G. R. (2014). Linking usage and grammar: Generative phonology, exemplar theory, and variable rules. Lingua, 142, 57-65. doi:10.1016/j.lingua.2012.07.007
Hou, J. Y. (2022). Introduction to modern Chinese dialects. Shanghai: Shanghai Education Pr.
Kang, C., Ma, F., & Guo, T. (2018). The plasticity of lexical selection mechanism in word production: ERP evidence from short-term language switching training in unbalanced Chinese–English bilinguals. Bilingualism, 21(2), 296-313. doi:10.1017/S1366728917000037
Kroll, J. F., Sumutka, B. M., & Schwartz, A. I. (2005). A cognitive view of the bilingual lexicon: Reading and speaking words in two languages. The international journal of bilingualism : cross-disciplinary, cross-linguistic studies of language behavior, 9(1), 27-48. doi:10.1177/13670069050090010301
Lee, M.-W., & Williams, J. N. (2001). Lexical access in spoken word production by bilinguals: evidence from the semantic competitor priming paradigm. Bilingualism: Language and Cognition, 4(3), 233-248. doi:10.1017/S1366728901000426
Leote, J., Castellote, J. M., Casanova-Molla, J., Navarro-Otano, J., Nunes, R. G., Ferreira, H. A., & Valls-Sole, J. (2018). Motor preparation in picture naming tasks. Brain Lang, 180-182, 24-30. doi:10.1016/j.bandl.2018.04.002
Li, P. (2009). Lexical Organization and Competition in First and Second Languages: Computational and Neural Mechanisms. 33(4), 629-664. doi:https://doi.org/10.1111/j.1551-6709.2009.01028.x
Norman, J. (1988). Chinese. Cambridge: Cambridge Univ. Pr.
Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. Behavior Research Methods. 10.3758/s13428-018-01193-y
Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese: Modeling of phonological encoding. Japanese psychological research, 57(1), 22-37. doi:10.1111/jpr.12050
Roelofs, A., & Verhoef, K. I. M. (2006). Modeling the control of phonological encoding in bilingual speakers. Bilingualism, 9(2), 167-176. doi:10.1017/S1366728906002513
Roelofs, A., & Verhoef, K. I. M. (2006). Modeling the control of phonological encoding in bilingual speakers. BILING-LANG COGN, 9(2), 167-176. doi:10.1017/S1366728906002513
Roelofs, A., & Verhoef, K. I. M. (2006). Modeling the control of phonological encoding in bilingual speakers. Bilingualism, 9(2), 167-176. doi:10.1017/S1366728906002513
Sadat, J., Martin, C. D., Magnuson, J. S., Alario, F. X., & Costa, A. (2016). Breaking Down the Bilingual Cost in Speech Production. Cogn Sci, 40(8), 1911-1940. doi:10.1111/cogs.12315
Traxler, M. J. (2012). Introduction to psycholinguistics: Understanding language science. Wiley-Blackwell.
Van Assche, E., Duyck, W., & Gollan, T. H. (2013). Whole-Language and Item-Specific Control in Bilingual Language Production. J Exp Psychol Learn Mem Cogn, 39(6), 1781-1792.

doi:10.1037/a0032859

Wang, J., Wong, A. W.-K., & Chen, H.-C. (2017). Time course of syllabic and sub-syllabic processing in Mandarin word production: Evidence from the picture-word interference paradigm. Psychon Bull Rev, 25(3), 1147-1152. doi:10.3758/s13423-017-1325-5

Wang, J., Wong, A. W.-K., & Chen, H.-C. (2021). Second language experience influences salience of phonological units in spoken word production in the first language. The international journal of bilingualism : cross-disciplinary, cross-linguistic studies of language behavior, 25(6), 1545-1559. doi:10.1177/13670069211031001

Wang, J., Wong, A. W.-K., Tsang, Y.-K., Wang, S., & Chen, H.-C. (2019). Behavioural evidence for segments as subordinate units in Chinese spoken word production: The form-preparation paradigm revisited. PLoS One, 14(11), e0225718-e0225718. doi:10.1371/journal.pone.0225718

Wong, A. W.-K., & Chen, H.-C. (2009). What are effective phonological units in Cantonese spoken word planning? Psychonomic Bulletin & Review, 16(5), 888-892. doi:10.3758/PBR.16.5.888

Woumans, E., Clauws, R., & Duyck, W. (2021). Hands Down: Cognate Effects Persist During Written Word Production. Frontiers in Psychology, 12, 647362-647362. doi:10.3389/fpsyg.2021.647362

Xu, Y. ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France. 7-10.