The Effect of Phonemic Similarity on Bilingual Word Superiority Effect

Liu Zhengyuan (MC364220), Chen Jielu (MC364353), Zhang Lizi (MC364456)

1. Introduction

The Word Superiority Effect (WSE) has been a well-documented phenomenon in cognitive psychology since its first systematic study in the 1970s. Suppose you are asked to find a specific letter in a sequence of letters, such as the letter "E". There are two scenarios: one where you are looking for the letter E within a word, like "apple", and another scenario where the letter E appears alone, such as in a sequence of letters without a clear meaning, like "p-d-f-e-l". You will find that when the letter E appears within a word, you will find it more quickly. This is called the word superiority effect, which illustrates that letters are recognized more accurately and quickly when presented within words as opposed to isolated or within non-word letter strings (Reicher, 1969; Wheeler, 1970). These foundational studies laid the groundwork for understanding the intricate processes of visual word recognition and the interplay of perceptual and cognitive factors influencing it.

Subsequent research has expanded on these initial findings by incorporating various linguistic and non-linguistic variables that modulate the WSE. Fort et al. (2010) suggested that visual information on phoneme identity contributes to the lexical activation process during word recognition. Chastain (1981) have found that the phonetic and spelling aspects of letter strings influence the recognition of constituent letters. Hildebrandt et al. (1995) identified lexical factors in the word superiority effect.

Despite extensive research on the WSE, and the identification of phonetic factors influencing it, few studies have focused on the impact of phonemic similarity on this very effect. Phonemic similarity refers to the degree of similarity in pronunciation between two words. Lukatela et al. (1990) found that phonemic similarity speeds up the recognition of words and pseudowords. Their research results also indicate that lexical processing units are typically activated by phonemic processing units. Their experiments found an intermediate layer of phonological units between lexical units and letter units during word processing, and the phonemic similarity effect in naming is based on the state of phonemic units, while the phonemic similarity effect in lexical decision-making is based on the state of word units. Their research demonstrates the influence of phonemic similarity on lexical recognition. Therefore, we hypothesize that phonological form is also part of the mental lexicon, and that phonemic similarity is a key variable of the mental lexicon.

Previous studies mainly concentrated on visual processing in word identification, to examine the effects of visual similarities on lexical recognition, which were measured by the reaction accuracy and speed towards memorization and identification tasks of separated letters in presenting words or separated lexical in context. Word Superiority Effect indicates the holistic recognition effect within the context by multitudes of literature. However, few investigated whether the phonological features affect lexical processing.

Bosma et al. (2023) summarized that the cross-linguistic transfer between two languages in bilinguals cannot be activated or shut down selectively, even in single-language contexts, providing an insight that the academically defined boundaries between languages do not correspond to the mechanism in which they are processed in the brain, which is engaged in conceptual chunks of vocabulary, the mental lexicon. The mental lexicon is thought to be a repository of lexical and conceptual representations, composed of organized networks of semantic, phonological, orthographic, morphological, and other types of information, which is "stored" in memories and mapped into the mechanism of word retrieval processes (Wulff et al., 2019). Cross-linguistic transfer, therefore, is represented by the retrieval speed when activating both languages in the cognition behaviorally, which could be examined through a word superiority paradigm. In sum, the more overlaps in the processing levels of a lexical stimulation, the faster and more accurate we expect to present in the experimental data from participants.

2. Experiment

2.1. Purpose

This experiment is to investigate if phonological form is a part of mental lexicon and to what extent that Chinese-English bilinguals' word superiority effect relies on semantic information. We hypothesize that as the phonological form is a part of mental lexicon, it would affect the cognitive performance on lexical identification due to the differences and similarities in the cross-linguistics.

2.2. Methods

Our experiment was conducted to examine that, based on our hypothesis, whether the overlaps of phonological symbols and forms could contribute to the lexical processing or not. The Word Superiority Effect (Teasdale et al., 1991) was utilized as the paradigm in the lexical processing procedure. A traditional WSE paradigm involves a presenting page of stimulus with words needed to be memorized, a masking page with four hashtags, and a response page with the letters to be recognized, as shown below.

As discussed above, the main effect of WSE underscores the holistic identification during processing lexical information. To limit the data sourcing, we manipulate the final letters in four-letter strings, considering it as fundamental operational differentiation during phonological, orthographic and semantic levels of lexical processing. We anticipate that the attributing overlaps of phonemic pronunciation and formations in the last position of letter strings could present a boosting effect in the modified word superiority paradigm. For instance, in the vocabulary "work", the final letter is "k", which pronunciation is defined as "/keɪ/" in International Phonetic Alphabet (IPA), and the whole pronunciation of "work" is defined as "/wɜːk/" in IPA, referring to a similarity or overlap of "/k/" for the ending "k" in articulation system. The matching conditions between real words and similarities constitute of the basic for later data analysis.

2.2.1. Materials

Four-letter word is the original stimulus in this experiment. Python was in the use of making up all the possible four-letter groupings in our experiment. These alphabetic groupings were classified into pseudo and reals words assessed by the Carnegie Mellon Pronouncing Dictionary, which is an open-source pronunciation dictionary, following IPA to return the articulation items of real words. We selected 120 words with proper frequency as stimuli, and their pairings of ending letters for recognition. The experiment was under a 2 (*Phonemic Similarity*: Overlapped, Not overlapped) × 2 (*Word Type*: Real Word, Pseudoword) within-subject design with 4 conditions in total, considering both the lexical meanings and the overlaps between ending sounds and ending letters.

To clarify, first, if we present the word "work" as a stimulation, and the last letter "k" for identification, the sound of "/k/" overlaps with the pronunciation of "/wɜːk/" in "work" at the ending, indicating a matching condition of real words and overlapped ending. Second, if we pair the word "work" with the letter "d" to identify, both "work" and "work" are real words, but the letter "d" does not match the ending sound in "work", which therefore is a real-word pair but not overlapped. Third, we present pseudo words, i.e. "loup", a four-letter string as well, which pronunciation is supposed to be "/lu:p/", overlapped with the ending letter "p" (/pɪ:/). That's a pseudo-word and overlapped pair. Fourth, in the last condition, we present pseudo words i.e. "nade", assumed to be sounded like "/neɪd/", in which the last letter "e" does not overlap with the articulation, therefore establishing a pseudo-word and not overlapped pair of stimulation. All the stimuli were presented in the experiment randomly.

2.2.2. Participants

We recruited 5 undergraduate students from University of Macau in this pilot study, including 4 females and 1 male. All participants were native Chinese speakers, whose second language is English. They had normal hearing and (corrected-normal) vision, with no reports of history of speech, language, neurological disorders, head damage, or mental illness. All participants achieved a score of over 425 on the College English Test Band 4 (CET-4) or an equivalent English proficiency level and volunteered to participate in the experiment.

2.2.3. Procedure

The experiment was conducted using PsychoPy (https://www.psychopy.org/, v2022.2.5). Participants were seated in front of a computer screen, instructed to minimize head movements and maintain fixation on the central point displayed on screen. A white fixation was presented at the center of the screen for 500ms, followed by the 300ms presentation of a priming word (e.g., "WORD"). Then, 1000-ms mask (i.e., "####") was shown to cover up the priming word with a 700 ms blank afterwards. Next, a target letter was presented to the subject, which was either the same (e.g., "###D") or different (e.g., "###K"). Subject was instructed to respond with the "j" key if the target letter was the same to the priming word's last letter, or the "f" key if it was not. If the subject responded or was failed to respond within 3000ms after the onset of the target letter, the program would automatically enter a random blank screen lasting up to 2000ms. Response accuracy and reaction times were recorded, with non-responses excluded from data analysis. All in all, participants were to run through a total of 240 trials costing around 20 to 30 minutes.

2.2.4. Data Analysis

Data selection was performed before analysis. For reaction time (RT) data, incorrectly responded trials were first excluded. Then, both unmatched and matched trials had to exist in pair for them to be retained, as the RT of unmatched trial would later be treated as baseline and be subtracted with the matched trial's RT representing the time saved. For accuracy (ACC) data, both incorrectly and correctly responded trials were retained with unmatched trials excluded.

Using the R statistical software (https://www.r-project.org/, version 4.2.2) and the belonging lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages, linear mixed-effects models (LMMs) were conducted on response time data, while generalized linear mixed-effects models (GLMMs) were conducted on accuracy data. The analysis of LMMs utilizes the Restricted Maximum Likelihood (REML) for parameter estimation, and its analyses of main effects and interaction effects are implemented using the anova function from the lmerTest package (Kenward & Roger, 1997). The main effects and interaction effects of the GLMMs are analyzed using the Anova function from the car package (https://cran.r-project.org/web/packages/car/index.html). Post-hoc tests for both models are conducted using the emmeans function in the emmeans package (https://cran.r-project.org/web/packages/emmeans/index.html). During the process of model analyses, models that exhibit singular fits, fail to converge, overfit, or possess a fixed effect matrix with rank deficiency will not be accepted. Among all these situations, overfitting means there is at least one pair of random slopes in the random effects model that have a correlation above 0.9 (Barr et al., 2013).

Table 1. Mean (standard deviation) for accuracy and reaction time

Word Type	Overlapped		Not Overlapped	
	ACC	RT	ACC	RT
Real Word	0.96 (0.24)	562 (236)	0.94 (0.24)	564 (235)
Pseudoword	0.94 (0.24)	567 (235)	0.94 (0.24)	566 (234)

Note: The unit of reaction time is millisecond.

Fixed factors include *phonemic similarity* and *word type*, where *phonemic similarity* is coded as overlapped (-1/2) and not overlapped (1/2), while *word type* is coded as pseudoword (-1/2) and real word (1/2). Subject and last letter pair are treated as random factors. Starting from the full model of the random effects, if the model fails to fit, the random slope that contributes the smallest explained variance would be removed until a successful fit is reached.

2.3. Results

The optimal models for accuracy and reaction time data in Experiment 1 are as follows:

ACC ~ phon_sim*word_type + (1+word_type|participant) + (1|letter_pair)

RT ~ phon_sim*word_type + (1|participant) + (1|letter_pair)

For the accuracy data, Wald chi-square test of the optimal model showed that there was no significant main effect for either *word type* ($\chi^2(1) = 2.219, p = 0.136$) or *phonemic similarity* ($\chi^2(1) = 0.3674, p = 0.544$), and no significant interaction between *word type* and *phonemic similarity* ($\chi^2(1) = 1.646, p = 0.199$).

For the reaction time data, F test of the optimal model showed the main effect of *phonemic similarity* was significant ($F(1, 17.19) = 4.477, p = 0.049$). Post-hoc analysis revealed that the reaction time for trisyllabic word was significantly higher than that for disyllabic word ($\beta = 57, SE = 26.9, df = 17.2, t = 2.116, p = 0.049$). The main effect of *word type* was not significant ($F(1, 484.74) = 0.109, p = 0.741$), and the interaction between *word type* and *phonemic similarity* was also not significant ($F(1, 484.93) = 1.514, p = 0.219$).

3. Discussion

This experiment investigated if phonological form is also part of metal lexicon, as we expected that Chinse-English bilinguals can recognize word-letter pair with higher phonemic similarity more quickly and accurately. The results of pilot studies showed that there was a significant main effect on phonemic similarity observed in reaction time data, with not overlapped word-letter pair taking longer time to process. The significance itself proves the phonological form should be a part of mental lexicon. Additionally, the post-hoc result meets our original expectations that higher level of phonemic similarity should reducing the difficulty for participants to identify the last letter in the word and the target letter to be the "same".

The other thing is, there was no significant main effect found on the word type, indicating that Chinese-English bilinguals recruited in this experiment either did not reach enough proficiency for generating a reasonable extent of semantic activation from either phonological or orthographical forms of words. That is to say, Chinese-English bilinguals likely made their judgements based on merely orthographic and phonological activation, rather than semantic activation. Therefore, if conducting ERP experiments in the future, we would not expect to see a significant N400 effect on Chinese-English bilingual participants.

3.1. Limitations and Future Perspectives

One of the areas for improvement in this research is the selection of vocabulary and the management of lexical repetition. Efforts should be made to balance each letter pair's occurrence, otherwise it is very likely to introduce practice effect (Bartels et al., 2010;

Donovan & Radosevich, 1999) to this study. It is also worth noting that the durational parameters we utilized in the experiment may not be in its best combination. Given that the participants we target at are all expected to be Chinese-English bilinguals, adjustments should be made to accommodate their proficiency so as to activate bilinguals' word superiority effect to the maximum extent possible.

By addressing these limitations and building upon our refined methodology, we aim to contribute to a more comprehensive understanding of bilingualism's effects on cognitive processes and to set the stage for innovative and impactful future research in the field of word superiority effect.

Reference

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. https://doi.org/https://doi.org/10.1016/j.jml.2012.11.001

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: A longitudinal study on frequent repetitive cognitive testing. *BMC Neuroscience*, *11*(1), 118. https://doi.org/10.1186/1471-2202-11-118

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48. https://doi.org/10.18637/jss.v067.i01

Bosma, E., Bakker, A., Zenger, L., & Blom, E. (2023). Supporting the development of the bilingual lexicon through translanguaging: a realist review integrating psycholinguistics with educational sciences. *European Journal of Psychology of Education*, *38*(1), 225-247.

Chastain, G. (1981). Phonological and orthographic factors in the word-superiority effect. *Memory & Cognition*, *9*(4), 389-397. https://doi.org/10.3758/BF03197564

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*(5), 795-805. https://doi.org/10.1037/0021-9010.84.5.795

Fort, M., Spinelli, E., Savariaux, C., & Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, *52*(6), 525-532. https://doi.org/https://doi.org/10.1016/j.specom.2010.02.005

Hildebrandt, N., Caplan, D., Sokol, S., & Torreano, L. (1995). Lexical factors in the word-superiority effect. *Memory & Cognition*, *23*(1), 23-33. https://doi.org/10.3758/BF03210554

Kenward, M. G., & Roger, J. H. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*, *53*(3), 983-997.

https://doi.org/10.2307/2533558
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13
Lukatela, G., Carello, C., & Turvey, M. T. (1990). Phonemic priming with words and pseudowords. *European Journal of Cognitive Psychology*, *2*(4), 375-394. https://doi.org/10.1080/09541449008406214
Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of experimental psychology*, *81*(2), 275.
Teasdale, N., Stelmach, G., Breunig, A., & Meeuwsen, H. (1991). Age differences in visual sensory integration. *Experimental brain research*, *85*, 691-696.
Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59-85.
Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New perspectives on the aging lexicon. *Trends in cognitive sciences*, *23*(8), 686-698.