Examining Language Interactivity Through Classifier Masked Primes in Chinese-English Bilinguals

Wen Jiang et al.

College of Liberal Arts and Sciences, University of Florida

Jorge Valdés Kroff, Department of Spanish and Portuguese Studies

Abstract

One universal characteristic of human languages is grammatical agreement dependencies. In Mandarin Chinese, classifiers, or measure words, are an obligatory grammatical class used to characterize a quantified noun. The classifier is inserted in between the numeral and the noun (i.e., 一張票 yī zhāng piào, one ticket). English lacks classifiers, so their use in Chinese may lead to cross-language influence for bilinguals and second language learners. Previous research has investigated language interactivity in bilinguals, commonly through the masked priming lexical decision task (LDT), a psycholinguistic task where participants decide if a sequence of characters form a real word or not. This study investigated Chinese-English bilinguals at the University of Florida (n= 24). The experimental group consisted of Chinese-dominant bilinguals who have been in an immersive English environment for at least one semester. The control group included monolingual or dominant English bilingual speakers with another language other than Chinese. Participants completed two LDTs in a within-subjects design. Part 1 featured Chinese (L1) classifier masked primes in an overt English (L2) LDT. Part 2 featured L1 translated masked primes sharing a common classifier with target words in the overt L2 LDT. The results indicated no interaction effect between group and congruency conditions in either part (p = 0.334; p = 0.889). These findings reveal no effect of classifier primes in Chinese-English bilinguals, suggesting minimal, or the absence of cross-language influence in this context.

Keywords: bilingual, Chinese, classifier, lexical decision, masked prime, measure word

Introduction

Today most of the world’s population speaks at least two languages, indicating that bilingualism or multilingualism is the norm (Dong et. al, 2021). How the brain processes and compartmentalizes languages is a question that has yet to have a clear answer. As a result, the field of bilingualism has grown in recent decades, with much of the research focused on language interactivity. One way to study this language interactivity is through the masked priming paradigm.

In the masked priming paradigm, participants complete a task related to a target word, which is preceded by a prime that appears for a moment (i.e., 50ms), so that the participant is not
consciously aware of seeing the prime. The prime is usually a word (or non-word) in the language(s) of interest. In regard to bilingual research, the use of translated primes is common. Translated masked priming involves the prime being in a language different from the target word. When the prime is in the speaker’s first language (L1) and the target word in the second language (L2), this is known as a forward translation. When this relationship is reversed, where the prime is in L2 and the target word in L1, it is called a backward translation (Chen et al., 2022). Studies have shown a stronger relationship in the forward rather than backward direction. This asymmetry is attributed to the differing levels of cognitive activation of each language (Keatley & de Gelder, 1994). This difference, in turn, is a result of varying levels of proficiency in unbalanced bilinguals (Xia & Andrews, 2015). The more proficient the speaker, the easier it is to access the language and the less mental effort required.

One model often adopted in language interaction research is the Bilingual Interactive Activation Plus (BIA+) model, proposed originally in 1998 (Dijkstra & van Heuven, 2022). The model proposes two subsystems in understanding bilingual language comprehension. One subsystem focuses on word identification and the other on task decision. Word identification involves the simultaneous activation of orthographic, phonological, and semantic representations of a word (Dong et al., 2021). For example, through the cognate facilitation effect, bilinguals can process cognates, words that share similar form and meaning across languages, faster than non-cognates because of these shared representations (i.e., English - attention; Spanish - atención). This advantage is not seen in the monolingual speaker, to whom the cognates do not matter (Poort & Rodd, 2017). The second subsystem, task decision, is based on executive function, which includes word inhibition, evaluation, and monitoring. The two subsystems interact in a top-down approach in which a stimulus (i.e., a word) activates word identification which in turn activates task decision.

Additionally, BIA+ suggests that the two languages of the bilingual are not actively used, but rather the target language is accessed only when needed. This is called resting level activation, which assumes a positively correlated relationship between language proficiency and level of activation (Dijkstra & van Heuven, 2022). For example, if a bilingual’s L1 is English and L2 Spanish, a higher frequency (commonly used) word in L2 can result in the same degree of activation at rest when compared to a lower frequency word in L1. This explains why many published studies found asymmetrical results in the priming effect.
Based on BIA+, McPhedran and Lupker compared the priming effect using translated masked primes in two tasks: lexical decision and semantic categorization (2020). The lexical decision task (LDT) is one of the most used tasks in bilingual language research. In this task, the participant must decide if a stimulus they see is a real word or not (McPhedran & Lupker, 2020). On the other hand, the semantic categorization task (SCT) requires the participant to indicate which words belong to the same semantic category.

The participants in McPhedran and Lupker’s study were students from the University of West Ontario who 1) spoke Mandarin Chinese and English and 2) had spent at least 5 years in Canada (2020). The study utilized both forward and backward masks. Participants completed a LDT followed by a SCT. The results revealed that the level of L2 proficiency influenced behavior in two ways. First, participants with higher reading and writing skills showed larger effects in the LDT. Second, participants with higher speaking and listening skills showed greater effects in the SCT. These results demonstrate that masked translation priming has different effects depending on not just language proficiency but also different components of proficiency.

When the individual’s L1 and L2 do not share a common orthography, such as Chinese and English, response times on LDTs can be longer (Dong et al., 2021). Orthography transparency, or how well the spelling of a word corresponds to its pronunciation, explains this (Borleffs et al., 2017). The more transparent a language is, the less ambiguous the relationship between orthography and phonology. Because Chinese is not based on an alphabetic writing system, orthography does not directly correlate to phonology. Even though pinyin, a romanization of the phonology of the characters, exists in Chinese it is only used in introductory courses to primary school aged children and second language learners. Therefore, research that investigates Chinese-English bilinguals are considered cross-scripts and often involve non-cognate primes.

In addition to differences in script, there are certain grammatical features found in Chinese but not in English. One example is the case of classifiers, more commonly referred to as measure words (Grüter et al., 2020). In Chinese, measure words are used in conjunction with numbers to characterize a class of nouns such as in examples (1) and (2). Measure words are required when counting nouns and can be further divided into subtypes, with sortal classifiers being the most common (Grüter et al., 2020). Sortal classifiers categorize nouns based on their physical features, such as size.
In example (1), the classifier 张 (zhāng) is used to count things that are flat while in example (2) the classifier 条 (tiáo) is used to count things that are long and thin.

(1) 一张纸
yi1 zhang2 zhi2
“a sheet of paper”

(2) 一条蛇
yi1 tiao1 she2
“a ∅ snake”

Different from Chinese, English lacks classifiers but there is a grammatical class that serves an analogous function: collective nouns. For example, in (3) and (4), the use of the word herd in its singular and plural form, respectively.

(3) a herd of cattle
[article] + [collective noun] + [preposition of]+ [noun].

(4) two herds of cattle
[article] + [collective noun] + [preposition of]+ [noun].

However, different from the optional use of collective nouns, measure words are obligatory to form grammatically correct sentences in Chinese. This allows the speaker to predict the upcoming word, by ruling out what cannot be followed (i.e., the classifier 张 zhāng precedes nouns that are categorized as flat and broad). This is known as predictive processing, or how a word is easier to process if preceded by a related word, such as belonging to the same grammatical class (Grüter et al., 2020). However, because English lacks classifiers, their use in Chinese can present certain difficulties to bilinguals and to second language learners of Chinese.
Authors Thierry and Wu (2007) found that Chinese-English bilinguals implicitly access Chinese (L1), when evaluating semantic similarity between pairs of nouns in English (L2). The results suggest that grammatical features from a speaker’s L1 can influence the L2, including those that are exclusively found in the L1. The question that arises is, if interactions exist in language processing, what happens when Chinese-English bilinguals, without realizing they see their L1, Chinese, in a L2, English, dominant environment? For example, if the primes are in Chinese but the task is exclusively in English.

Grüter et al. (2020) explored if classifiers allow speakers to anticipate upcoming nouns based on their classification into the grammatical class. The authors separated stimuli into two groups: semantic and grammatical. They found that L2 Chinese speakers, in contrast to L1 Chinese speakers, experienced more competition when viewing pairs of words that shared similar semantics even though the two words did not belong to the same grammatical class. Moreover, L1 speakers did not ignore semantic features, as previously thought, but rather they relied on semantic cues when the noun is already restricted to a certain grammatical class (i.e., when both words belong to the same classifier).

Even though these studies support the interaction between a bilingual’s L1 and L2, there has been little, if any, published research that has explored the topic of language interactivity in the context of measure words. Therefore, this study investigated this topic in Chinese-English bilinguals, using translated masked primes in a lexical decision task. The hypothesis was that Chinese-English bilinguals will have a shorter average reaction time when compared to controls, specifically in congruent conditions than incongruent conditions, where the classifier serves as a prime to predict the upcoming noun.

**Methodology**

**Participants**

The participants in this study were students attending the University of Florida (n = 24; ages 18-35). The experimental group consisted of Chinese-dominant bilinguals who have been in an immersive English-speaking environment for at least one semester. The control group consisted of monolingual or dominant English bilingual speakers with another language other than Chinese.
Stimuli selection

In both Experiments 1 and 2, participants saw a forward mask consisting of fourteen hashes (##############) for 500ms, chosen to match the orthographic length of the longest stimuli, followed by the prime for 50ms. The target stimuli then appeared for 2000ms or until a response was detected. All displayed text was in Consolas 18pt black font against a white background. Participants responded using keyboard presses, where “F” corresponded to yes and “J” corresponded to no. Prior to experimental trials, participants completed a practice session consisting of ten trials. Practice trials, lasting the same duration as experimental trials, were not primed but included the forward mask. No feedback was given on reaction times nor error rates in neither practice nor experimental sessions. Nonwords were selected from a nonword bank, differing from real words in one or two letters, and matched on orthographic length and spacing to the stimuli (i.e., copy paper and topy fover).

Classifiers were selected based on their order of introduction in college level beginning Chinese courses. The most common eight were chosen, along with sixteen nouns from each classifier group. All nouns were two Chinese characters in length, translated into English, and used as stimuli for both experiments (i.e., Chinese - 报纸 bàozhǐ; English - newspaper). The chosen nouns were verified for accuracy and familiarity with a native Mandarin Chinese speaker.

In Experiment 1, the classifier prime was either a correct (congruent) or incorrect (incongruent) classifier for the following stimuli. For example, in congruent conditions participants saw the prime 张 (zhāng) and stimuli 地图 (dìtú; map), while in incongruent conditions they saw the prime 条 (tiáo) and stimuli 地图 (dìtú; map). The prime was shown as a single simplified Chinese character. The stimuli were English lowercase letters. In Experiment 2, both primes and stimuli were in English. The primes, when translated into Chinese, either shared (congruent condition) or did not share (incongruent condition) a common classifier with the stimuli.

Procedure
Participants in both groups completed two LDTs in a within-subjects design. Experiment 1 featured Chinese (L1) classifier masked primes in an overt English (L2) LDT. Experiment 2 featured L1 translated masked primes that either shared or did not share a classifier with the stimuli in the overt L2 LDT. After completing both experiments, all participants then filled out a language background survey (adapted from Birdsong et al., 2012) and an English familiarity survey that asked questions about the stimuli used in the study. Participants in the bilingual group also completed two additional surveys: 1) Chinese familiarity survey with translated versions of the stimuli and 2) Chinese classifier test where participants filled in the blank with the correct classifier using the translated version of the stimuli.

At the conclusion, participants completed the AX-CPT on the computer and button pressed for yes or no. The AX-CPT is a cognitive task that measures both proactive control, prior to stimulus presentation, and reactive control, after stimulus presentation, through nonlinguistic means (Gonthier et al., 2016). In this study, instructions were presented in Courier New 18pt white font against a black background. The letters A and X were shown in red against a black background while the three letters in between were shown in white. To reduce mental fatigue, participants received a break halfway into the task.

**Results**

Repeated-measures ANOVA tests were performed in SPSS on participants’ mean reaction time (avgRT) on correct responses only across both experiments, while data on nonwords were not analyzed. Both between-subjects and within-subject analyses were done in a crossed 2 x 2 factorial design, including the variable group (bilingual, monolingual) and classifier congruency (congruent, incongruent). Main effects and their interaction are reported.

**Experiment 1 (Left aligned, Times New Roman 12, Bold)**

The avgRT for correct responses on real words was faster in controls (558.52ms, std.error = 27.44) than in bilinguals (718.59ms, std.error = 21.25, Table 1). Within the control group, the avgRT for correct responses in congruent conditions was 557.83ms and incongruent conditions 559.20ms.

**Table 1.** Mean RT, Std. Error, and the 95% CI Within Groups
Table 2. Experiment 1 Mean RT Within Group and Congruency Conditions

<table>
<thead>
<tr>
<th>Group</th>
<th>Congruency</th>
<th>Mean</th>
<th>SE</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Congruent</td>
<td>557.83</td>
<td>26.36</td>
<td>503.16</td>
<td>612.50</td>
</tr>
<tr>
<td></td>
<td>Incongruent</td>
<td>559.20</td>
<td>31.12</td>
<td>494.66</td>
<td>623.74</td>
</tr>
<tr>
<td>Bilingual</td>
<td>Congruent</td>
<td>706.82</td>
<td>20.42</td>
<td>664.47</td>
<td>749.17</td>
</tr>
<tr>
<td></td>
<td>Incongruent</td>
<td>730.36</td>
<td>24.11</td>
<td>680.37</td>
<td>780.35</td>
</tr>
</tbody>
</table>

Table 2 shows that in bilinguals the mean RT for congruent conditions was faster than for incongruent conditions (706.82ms, 730.36ms). Within-subjects tests revealed no significant effect of congruency conditions on RT ($F(1,22) = 1.229, p = 0.28$) nor an interaction effect of congruency and group ($F(1,22) = 0.97, p = 0.334$, Table 3). When comparing RT means between subjects, however, the difference in RT was significant ($F(1,22) = 21.27, p > 0.01$, Table 4).

Table 3. Repeated-Measures ANOVA Test for Congruency and Interaction Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Congruency</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruency</td>
<td>Linear</td>
<td>1</td>
<td>1745.16</td>
<td>1.30</td>
<td>0.28</td>
</tr>
<tr>
<td>Congruency*Group</td>
<td>Linear</td>
<td>1</td>
<td>1382.39</td>
<td>0.97</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Table 4. Repeated-Measures ANOVA Test with Each Group has Statistically Significant p-Value

<table>
<thead>
<tr>
<th>Group</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>18348707.88</td>
<td>1</td>
<td>18348707.88</td>
<td>1353.99</td>
<td>0.00</td>
</tr>
<tr>
<td>Group</td>
<td>288269.60</td>
<td>1</td>
<td>288269.60</td>
<td>21.272</td>
<td>0.00</td>
</tr>
<tr>
<td>Error</td>
<td>298133.00</td>
<td>22</td>
<td>13551.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The total number of incorrect responses was higher in bilinguals (5.00%) than controls (3.13%). However, in both groups, there was a slightly higher rate of incorrect responses in incongruent conditions (bilinguals = 6.00%, controls = 1.86%) than in congruent conditions (bilinguals = 5.31%, controls = 1.46%).

Experiment 2

When L2 primes were used, participants in the control group had a shorter avgRT for correct responses on real words compared to the bilingual group (523.56ms, 672.20ms, Table 5). The avgRT for correct responses in congruent conditions was longer than in incongruent conditions when examining across both groups (Table 6). When examining group effects, however, the difference in RT between subjects was significant (F (1,22) = 13.35, p < 0.01; Table 7). However, testing for interaction effects of congruency and group reveal statistically nonsignificant results (F (1,22) = 0.02, p = 0.889; Table 5). The results for within-subject effects of congruency likewise were statistically nonsignificant (F (1,22) = 0.16, p = 0.69, Table 5).

Table 5. Repeated-Measures ANOVA Test with Each Group has Statistically Significant p-Value

<table>
<thead>
<tr>
<th>Group</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>16112733.67</td>
<td>1</td>
<td>16112733.67</td>
<td>854.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Group</td>
<td>251895.69</td>
<td>1</td>
<td>251895.69</td>
<td>13.35</td>
<td>0.00</td>
</tr>
<tr>
<td>Error</td>
<td>415077.21</td>
<td>22</td>
<td>18867.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Mean RT, Std. Error and the 95% CI Within Each Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SE</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>523.56</td>
<td>32.38</td>
<td>456.42</td>
<td>590.71</td>
</tr>
<tr>
<td>Group</td>
<td>673.20</td>
<td>25.08</td>
<td>621.19</td>
<td>725.21</td>
</tr>
</tbody>
</table>

Table 7. Experiment 2 Mean RT Within Congruency Conditions

<table>
<thead>
<tr>
<th>Congruency</th>
<th>Mean</th>
<th>SE</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent</td>
<td>600.29</td>
<td>21.54</td>
<td>555.62</td>
<td>644.96</td>
</tr>
<tr>
<td>Incongruent</td>
<td>596.48</td>
<td>20.50</td>
<td>553.96</td>
<td>639.00</td>
</tr>
</tbody>
</table>

Within the control group, participants responded slower to stimuli in the congruent conditions than incongruent conditions (mean RT = 524.79ms, 522.33ms; std.error = 34.06ms, 32.42ms, Table 8). Similarly, bilinguals also responded slower to congruent conditions than incongruent conditions (mean RT = 675.78ms, 670.62ms; std.error = 26.380, 25.11, Table 8).

Table 8. Experiment 2 Mean RT within Group and Congruency Conditions

<table>
<thead>
<tr>
<th>Group</th>
<th>Congruency</th>
<th>Mean</th>
<th>SE</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Congruent</td>
<td>524.79</td>
<td>34.06</td>
<td>454.16</td>
<td>595.42</td>
</tr>
<tr>
<td></td>
<td>Incongruent</td>
<td>522.33</td>
<td>32.42</td>
<td>455.10</td>
<td>589.57</td>
</tr>
<tr>
<td>Bilingual</td>
<td>Congruent</td>
<td>675.78</td>
<td>26.38</td>
<td>621.07</td>
<td>730.49</td>
</tr>
<tr>
<td></td>
<td>Incongruent</td>
<td>670.62</td>
<td>25.11</td>
<td>618.54</td>
<td>722.70</td>
</tr>
</tbody>
</table>

The total number of incorrect responses was higher in bilinguals (5.00%) than controls (3.13%). This trend for controls to respond faster than bilinguals was also seen in Experiment 1. However, unlike Experiment 1, controls had a slightly higher rate of incorrect responses in congruent conditions (5.63%) than in incongruent conditions (4.38%) while controls had a higher
rate of incorrect responses in incongruent conditions (4.17%) than congruent conditions (2.08%). Analysis of language proficiency in the bilingual group revealed that participants’ self-rated proficiency is not correlated to mean RT in either experiment ($r^2 = 0.07$, $r^2 = 0.03$).

**Discussion**

The hypothesis that Chinese-English bilinguals will have a shorter RT than controls in the congruent conditions when compared to incongruent conditions was not supported. The results revealed no effect of classifier primes in Chinese-English bilinguals, suggesting minimal or the absence of cross-language influence in this context. In relation to the BIA+ model, simultaneous activation of orthographic and phonological representations was likely not seen at the level of word identification because of the absence of cognate facilitation between the Chinese classifier primes and English stimuli, while task decision may have been isolated to only the L2 context due to the speed at which L1 primes were shown. The lack of classifiers in English prevents bilinguals from utilizing cognitive shortcuts during the LDTs. On the contrary, if interaction between the two languages did occur, it may have required more mental effort. Therefore, the lack of predictive processing is explained in this context.

In Experiment 2, mean RTs decreased in congruent conditions for both groups, suggesting that the difference was not a direct result of the priming effect but rather attributed to other factors. However, this decrease was in the opposite direction when L1 primes were used. Across both groups, mean RT was faster in the incongruent condition, suggesting that when the prime and stimuli were in the same language, there is a possible effect of interference. Rather than using the congruent primes to predict the upcoming word, bilinguals may pause to think about the stimuli. Additionally, the faster mean RT within the control group in both experiments is as expected, given that controls were dominant English speakers who were born or raised in the U.S. While all the participants indicated high rates of familiarity with the English words used in the study, proficiency levels do influence the mean RT.

Many of the words used in the study shared semantic characteristics due to the nature of classifiers (i.e., *doctor* and *nurse*). In Chinese, most nouns are grouped into their specific classifier categories based on shared physical attributes. A possible future direction in examining the effect of classifier primes in LDTs is controlling for this confound. For example, choosing
words from the same classifier class that do not match on semantics but do match on grammatical properties (Grüter et al., 2020).

One major limitation of this study was the sample size (n=24). Recruiting Chinese-dominant bilinguals from an institution where English is the primary language of instruction was difficult because eligible participants must know how to read and write in simplified Chinese, ruling out many heritage speakers. A bigger sample size for analysis would lead to stronger statistical results.

Another area of improvement is greater homogeneity in the control group. Many of the participants in this group were Spanish-English bilinguals and some even Chinese-English bilinguals who did not qualify to be in the bilingual group because they could not read nor write Chinese. Future directions include recruiting a purely monolingual English control group, although this may prove challenging given the foreign language requirement in many public high schools. Another improvement in terms of participant homogeneity is within the bilingual group. Stricter screening procedures could rule out participants who are accustomed to traditional rather than simplified Chinese (i.e., Taiwan) and who use different classifiers due to regional dialects.

**Conclusion**

The goal of this study was to examine cross-language influence in Chinese-English bilinguals who speak Mandarin Chinese as their L1 and English as their L2. The study examined the effects of both L1 and L2 primes, separately, on mean RTs in a lexical decision task. Overall, there was no significant interaction between group and congruency conditions. This suggests the absence of predictive processing in the L1 to L2 direction with classifier masked primes, both translated and untranslated. Further research at the sentence level could further elucidate this phenomenon.
References


