

Running head: Morphological Congruency Effect

The activation of grammaticalized meaning in L2 processing:
toward an explanation of the morphological congruency effect

Authors:

Nan Jiang (njiang@umd.edu), 240-478-3778, The University of Maryland (correspondence author)

Guiling Hu (guilinghu@gmail.com), The University of Maryland

Anna Chrabaszcz (achrabaszcz@hse.ru), Neurolinguistics Laboratory, National Research
University Higher School of Economics

Lijuan Ye (lijuanstella@gmail.com), (717) 766-2511 ext. 2264, Messiah College

Abstract

Adult second language learners have been found to have great difficulty in learning inflectional morphemes that are not instantiated in their native languages. It is proposed that this difficulty is a result of L2 learners not being able to automatically activate the meaning expressed by the morpheme because of the design of their native language. Three sentence-picture matching experiments were done to test if the plural meaning was automatically activated in sentence processing by Chinese and Russian ESL speakers. The critical stimuli included sentences that were congruent or incongruent in number with the pictures (e.g., singular word *book* with multiple books). Consistent with the predictions of the proposal, Russian ESL speakers, like native English speakers, produced a number-mismatch effect, responding more slowly to sentence-picture pairs that were not matched in number. Chinese ESL speakers showed no such effect.

Key words: inflectional morpheme, sentence-picture matching, Russian, Chinese, plural morpheme

The activation of grammaticalized meaning in L2 processing:
toward an explanation of the morphological congruency effect

The acquisition of inflectional (or grammatical) morphemes such as English tense and plural markers by second language (L2) learners has been the focus of attention from the very early days of SLA research. In more recent years, it has been a test ground for a number of theoretical and empirical issues such as the critical period hypothesis, the role of Universal Grammar in adult L2 learning, the aspect hypothesis. One finding that emerged from this research is what Jiang, Novokshanova, Masuda, and Wang (2011) referred to as the morphological congruency effect. In their conception, an L2 morpheme is referred to as a congruent morpheme from a learner's perspective if it is also instantiated in a learner's first language (L1). An L2 morpheme whose meaning is not morphologically marked in a learner's L1 is considered an incongruent morpheme. The morphological congruency effect refers to the finding that an incongruent morpheme is much more difficult to acquire than a congruent one.

The evidence for such this morphological congruency effect can be found in many studies. Some of these studies examined the development of different morphemes by the same group of L2 learners. Some of the morphemes were congruent ones in that they had a counterpart in the learners' L1 but some other were incongruent morphemes as they were not instantiated in the learners' L1. For example, White (2003) examined the development of English third person singular, past tense, plural, and articles by a Turkish ESL speaker for whom articles were incongruent but plural and past were. Some other studies considered the development of the same L2 morpheme by learners whose L1 did or did not mark the related meaning morphologically, e.g., English past tense by German and Chinese ESL learners in Hawkins and

Liszka (2003). In both types of studies, L2 learners' development of grammatical morphemes was assessed by examining their accuracy rates in performing a linguistic task which varied from filling in the blank to spontaneous production. Participants in a large number of such studies reliably produced higher accuracy rates on congruent morphemes than on incongruent morphemes. In Hawkins and Liszka's (2003) study, for example, the accuracy rates for the congruent German ESL group, were above 95%, but the proficiency-matched incongruent Chinese ESL group showed much lower accuracy rates on both regular and irregular verbs (62.5% and 84.2%). Similar findings were obtained in studies reported by Bialystok (1997), Bialystok and Miller (1999), Dewaele and V'eronique (2001), Franceschina (2001, 2005), McDonald (2000), Parodi, Schwartz, and Clahsen (2004), and Sabourin, Stowe, and de Haan (2006).

The difficulty involved in learning an incongruent morpheme is also demonstrated in more recent studies that examined the processing of ungrammatical morphemes by adult L2 learners in online tasks. These studies showed a lack of native-like sensitivity to morphosyntactic errors involving incongruent morphemes on the part of adult L2 learners. For example, Guillelmon and Grosjean (2001) examined the development of French gender marking by English-speaking learners in a shadowing task. The participants were asked to repeat the last word in a noun phrase which was either consistent or inconsistent in gender with its modifier or determiner. The native speakers and young L2 learners of French (French onset age: 4:0) showed a delay in repeating (or shadowing) the target word when it disagreed with the preceding word in gender, but adult L2 learners showed no such sensitivity. Similarly, Scherag et al. (2004) showed that English native speakers who had resided in Germany for more than 15 years failed to show any sensitivity to gender marking in German in a priming study. In a series of studies that

examined English plural marking among Chinese, Japanese, and Russian ESL speakers in a self-paced reading task, Jiang and his collaborators (Jiang, 2004, 2007; Jiang, Novokshanova, Masuda, & Wang, 2011) demonstrated that Russian ESL speakers, for whom the plural marker was a congruent morpheme, were able to show a native-like sensitivity to plural marking errors, but Chinese and Japanese ESL speakers, for whom the plural marker was an incongruent morpheme, were not.

Explaining the morphological congruency effect

What makes an incongruent morpheme more difficult? Several scholars have attributed such difficulty to maturational constraints, sometimes within the UG framework (e.g., Franceschina, 2005; Hawkins & Liszka, 2003; Scherag et al., 2004). However, such proposals are often less than sufficiently detailed to be tested empirically. We hope to outline a more specific explanation for this effect and put this explanation to test in a series of experiments.

Our approach begins with the assumption that native-like morphological performance is dependent on the automatic activation of the meaning marked by a grammatical morpheme. In expressing the meaning of “I like the books”, the plural meaning has to be activated in the speaker’s mind in order to trigger the activation and articulation of the related plural marker in output. Similarly, in hearing this sentence, an individual has to automatically encode the plural marker –s and link it to the plural meaning in his or her mind in order to understand the speaker’s intended meaning of multiple books.

However, languages differ in what meanings are grammaticalized. While learning the native language, an individual learns, among other things, what meanings are grammaticalized in the language and thus should be automatically activated in receptive and productive language use. In learning a language which grammaticalizes singular and plural meanings such as English,

a child learns to automatically pay attention to and encode the number meaning as singular or plural. A child learning Arabic has to automatically make a distinction between singular, dual, and plural meanings as required by the language. In contrast, automatic encoding of the number meaning is not obligatory for children learning Chinese as their native language. This difference in what meanings are grammaticalized and thus have to be automatically activated in language processing across languages has a huge impact in the subsequent learning of a new language. Specifically, a learner faces two very different tasks while learning a congruent versus an incongruent L2 morpheme. In learning a congruent morpheme, the related meaning is already part of routinely activated meanings as a result of learning the L1, so the task is to link the meaning to a new morphological marker. However, in learning an incongruent morpheme, a learner has to first learn to automatically activate a meaning that is not automatically activated in his or her L1. We suggest that this is where the difficulty lies.

We suspect that the development of new grammaticalized meanings may be subject to maturational constraints. Note that the number of grammaticalized meanings in a particular language are usually limited and stable. A language is very unlikely to add a new grammaticalized meaning in one's life time. Thus, there is no need for an individual to develop new grammaticalized meanings after the establishment of one's first language. In this regard, one may compare the development of new grammaticalized meanings to the development of new phonological categories for speech perception. Grammaticalized meanings and phonological categories share two common properties: They are limited in number in a language and they remain stable in one's life time, the latter of which means that once they are developed, there is no need to learn new ones. If Kuhl (2000, 2004) is right in suggesting that the perceptual map of phonological categories is laid out very early in one's childhood and becomes very difficult to

alter later in life, the same can be true for grammaticalized meanings. The pattern of semantic activation involving grammaticalized meanings, such as number and tense, may be finalized in the process of learning one's L1 and becomes very difficult to change.¹

To sum up, we argue that the cause of an L2 learner's difficulty with incongruent morpheme lies in the difficulty in automatically activating a meaning that is not routinely activated in his or her L1.

Testing the explanation: the sentence-picture matching task

The present study was intended to test the assertion that the meaning related to a congruent morpheme is automatically activated in L2 processing but the meaning related to an incongruent morpheme is not. It was done through the use of a sentence-picture matching task. In this task, a picture of three objects was first presented on a computer monitor. It was then followed by a sentence that was presented above the picture. The participant was asked to decide whether the sentence correctly described the physical relationship among the three objects in the pictures. Half of the sentences provided a correct description and thus required a Yes response. The other half did not and should generate a No response.² The critical manipulation of the study was that each Yes item had two versions, a number-match (or n-match) version and a number-mismatch (or n-mismatch) version. The two versions shared the identical sentence but differed in the number of one of the objects. As shown in Figure 2, in the n-match version, there was only one basket in the picture, which matched the sentence-final word *basket* in number. In the n-mismatch version, there were more than one basket, so the picture and the sentence did not match in number.

(Figure 1 about here)

The use of the sentence-picture matching task in combination with the number-match/mismatch manipulation should allow us to determine whether singular/plural meanings were automatically activated and affected task performance. Note that the participants were asked to focus on the physical relationship among the three objects in performing the matching task. They were not asked to pay attention to the number of objects. If they followed the instructions correctly, they were expected to produce a positive response for all the Yes items, regardless of number agreement, as these sentences described the physical relationship among the three objects correctly. However, if the number meaning is automatically activated, a singular meaning would be activated and represented in association with the sentence-final word *basket*. A singular or plural meaning would be generated in association with the object of basket depending on whether there was one or more than one basket. Under this circumstance, an individual should (most likely subconsciously) notice a mismatch in number between the word *basket* in the sentence and multiple baskets in the picture while performing the matching task, which should result in a delay in producing a positive response. If no singular meaning is activated in processing the sentence, no n-mismatch would be present, and no delay in performing the task would be expected. Thus, we can compare the reaction time for the n-match and n-mismatch conditions to determine whether the plural meaning was activated in the matching process. A delay in the n-mismatch condition was taken as evidence for the automatic activation of the number meaning in sentence processing.

Three experiments were conducted. The first one was intended to confirm the usefulness of the method by testing native speakers of English and Chinese in their respective L1s. If the method was sensitive enough to capture the activation of the singularity/plurality meanings, only native speakers of English would show a n-mismatch effect. In the Chinese version of the

experiment, because of a lack of number marking in Chinese, there was no number mismatch between the sentence and the picture in the n-mismatch condition, so Chinese participants should show no difference between the n-match and n-mismatch conditions. The second experiment was the main experiment in which a group of L2-proficiency-matched advanced Chinese and Russian ESL speakers were tested in the same task in both their L1s and L2. The Russian and Chinese ESL participants represented L2 learners for whom the singular/plural distinction was a congruent and an incongruent one, respectively. If we were correct in suggesting that only meanings grammaticalized in L1 get automatically activated in L2 processing, only Russian ESL speakers would produce a n-mismatch effect. Chinese ESL speakers should show no difference. If both ESL groups showed a n-mismatch effect, or if neither group did, then the proposal should be rejected. Experiment 3 was done to reinforce our interpretation of the results of Experiment 2 by demonstrating that the method was powerful enough to detect the activation of lexical meanings among Chinese ESL speakers.

Experiment 1: Assessing the effectiveness of the method

The purpose of the first experiment was to assess the effectiveness of the sentence-picture matching task for examining the activation of the singular/plural meanings in real-time language processing. The experiment consisted of an English part and a Chinese part. The test materials in the two parts were identical except for the language of the sentences. If the method was effective, only English-speaking participants would produce a n-mismatch effect, that is, responding to the number-match items faster than the number-mismatch items.

Method

Participants. A total of seventy-three college students served as participants in the study. Five of them were excluded due to high error rates. The remaining sixty-eight participants were

32 native English speakers who were taking an undergraduate psychology course at an American university and 36 native Chinese speakers studying in a Chinese university. They participated in the English and Chinese part of the experiment, respectively.

Stimuli. The test materials, first constructed in English, included 80 test items. Each item consisted of a picture and a sentence that described the physical relationship of the objects in the picture, as illustrated in Figure 1. The picture contained three easily recognizable objects whose names were familiar to the participants. The objects were placed side by side horizontally or positioned diagonally with one object placed right above or below one of the other two objects. The size of the pictures was approximately 4" by 4". Most sentences described the spatial relationship between the objects (e.g., *the banana is above the red apple*). Other sentences referred to the color or shape of the objects to reduce monotony of the task (e.g., *The shirt is of the same color as the tall tree; the arrow is pointing to the blue towel*). All of the sentences ended with a noun phrase which consisted of the definite article, a modifier, and a singular noun, as shown in the above examples. Half of the sentences described the pictures correctly and, thus, were match items; the other half did not, and thus were non-match items. The former served as critical stimuli and the latter as fillers.

For each of the 40 critical match items, two versions were created. The n-match version showed a single object referred to by the last singular noun of the corresponding sentence, and the n-mismatch version showed multiple objects labeled with the same singular noun (see Figure 1). The two versions were otherwise identical.

Two presentation lists were constructed. Each list consisted of 40 Yes-response items and 40 No-response items. Half of the Yes items were in the n-match condition and the other half in the n-mismatch condition. These Yes items were counterbalanced across the two presentation

lists such that the two versions of the same item never appeared on the same list. If the n-match version appeared in List A, its n-mismatch version appeared in List B, and vice versa. Both lists shared the same 40 No-response filler items.

To construct the test materials for the Chinese part of the study, one of the present authors translated the English sentences into Chinese, and the translation was checked by two additional Chinese-English bilingual speakers for its accuracy and idiomaticity. For example, the English sentence *The bike is above the tall basket* was translated as 自行车在高的篮子上方 (*zixingche zai gao de lanzi shangfang*). Even though the same set of pictures were used from the n-match and n-mismatch conditions of the English part, there was no difference in the two conditions in number matching in the Chinese stimuli because of a lack of number marking in these Chinese sentences,

Procedures. The participants were tested individually and were assigned to one of the two presentation lists randomly in their respective language. Once they were seated in the lab, they were asked to read and sign a consent form and then read the instructions for the test, which explained that they would match sentences to pictures on the basis of the spatial or other physical relationship of the objects in the pictures. The instructions emphasized the need for both accuracy and speed. The participants also had twelve practice items before the test items. All test items began with the appearance of a picture at the center of a computer monitor, which was followed 1330 milliseconds (ms) later by a sentence two lines above the picture. The stimulus onset asymmetry (SOA) of 1330 ms was determined on the basis of a pretest to ensure that the participants had enough time to view the pictures. Both the picture and the sentence remained on the monitor until a response was given. The participants responded by pressing two buttons, one for “Yes”, and the other for “No.” The data included the participants’ reaction time (RT), which

was the interval between the onset of the sentence and the execution of their response, and their error rate (ER). The task was self-paced in that the participant could choose to take a break between items or immediately go to the next item.

Results and Discussion

Participant and item means were first computed for both RT and ER for further statistical analysis. RT of incorrect responses and outliers were excluded from the computation of these means. The outliers were any RT that fell two standard deviations outside of the same subject mean or outside the low and high cutoffs set at 1200 millisecond (ms) and 6000 ms, respectively. 7.3% of data were affected by this data treatment. The mean RT and ER for the n-match, n-mismatch, and No-response conditions are shown in Table 1.

(Table 1 about here)

As is clear from the ER data in Table 1, both groups of participants performed the sentence-picture matching task with high accuracy, which meant that they were able to focus on the physical relationship between the objects while performing the matching task. As our primary purpose was to examine whether the two groups of participants were sensitive to number match between sentences and their corresponding pictures, the statistical analysis was focused on the comparison of the number-match and mismatch conditions in each group.

The English-speaking group was 113 ms faster in responding to the n-match items than they were in responding to the mismatch items. The difference was significant in both participant (t_1) and item (t_2) analyses, $t_1=2.33$, $p<.05$; $t_2=2.45$, $p<.05$. There was also a 3.9% difference in ER, and the difference was significant in participant analysis only, $t_1=2.61$, $p<.05$; $t_2=1.71$, $p=.096$.

The Chinese participants showed no significant difference in RT between the two conditions. They responded to the mismatch items 37 ms faster than the match items, but the difference was not significant, $t_1=0.97$, $p>.05$; $t_2=0.56$, $p>.05$. Neither was the difference in ER, $t_1=1.57$, $p>.05$; $t_2=0.60$, $p>.05$.

The results demonstrated a clear difference in the performance of English and Chinese speakers. English speakers took longer to respond to items that did not match in number between the sentences and pictures, but Chinese speakers did not. Could the n-mismatch effect found in English speakers be attributed to the fact that the mismatch pictures had more objects in them than the match pictures and, thus, should take longer to recognize? The answer is negative for two reasons. First, the pictures were shown before the sentences, and the 1330 ms SOA should allow the participants enough time to recognize all objects based on our estimate from the pretest. Second, if the perception of a single vs. multiple tokens of objects had led to the difference in RT among English speakers, the same should have happened among Chinese speakers, but the number of tokens of objects had no effect on the latter group.

The observation of the n-mismatch effect in RT among English-speaking participants indicated that the method was sensitive enough to capture the activation of the number meanings. We also consider the activation of such number meanings observed in this task an automatic process as it was not required by or relevant to the task. The lack of this n-mismatch effect among Chinese speakers helped strengthen our interpretation of the English results. If it was something other than the activation of the number meanings that contributed to the mismatch effect found among English speakers, we would have observed the same effect among Chinese speakers. The results from both groups of participants taken together suggest singular and plural meanings are automatically activated in language processing only among those whose L1 mark

these meanings morphologically and this activation can be successfully detected in the sentence-picture matching task.

Experiment 2: The activation of grammaticalized number meanings in L2 processing

Experiment 2 is the main experiment in which we examined the main assertion of the psycholinguistic explanation outlined in the introduction. Specifically, it was intended to examine the proposal that the meaning correspondent to a congruent L2 morpheme is automatically activated, but the meaning related to an incongruent L2 morpheme is not.

Advanced Russian and Chinese ESL speakers were tested in the same sentence-picture matching task in both their L1 and L2. Based on the results of Experiment 1, Russian ESL were expected to show a number-mismatch effect in RT while performing the matching task in Russian, as the singular/plural meanings are morphologically marked in Russian. Based on our proposal, we also expected the Russian ESL speakers to show a number-mismatch effect in English, because they met the two conditions for the observation of this effect: the automatic activation of the singular and plural meaning (as a result of learning L1 Russian) and a strong link between these meanings and the right morphological markers (at their high level of English proficiency). In contrast, we expected Chinese ESL speakers to show no n-mismatch effect in either their L1 or L2. In addition to replicating the L1 Chinese results found in Experiment 1, the prediction of a lack of n-mismatch effect in L2 followed directly from our view that adult L2 learners are unable to grammaticalize and automatically encode meanings in L2 processing that are not grammaticalized in their L1.

Method

Participants. Twenty-six Russian ESL speakers and 28 Chinese ESL speakers participated in the experiment. Twenty Russian ESL speakers and all 28 Chinese ESL speakers

were graduate students studying in an American university at the time of testing, and the other six Russian ESL speakers were mostly English teachers and professional translators who were living and tested in Russia. A cloze test taken from Bachman (1982) was used to assess their English proficiency. Three participants in each group did not finish the cloze test. The average score for the remaining 50 Russian and Chinese participants was 20.8 and 21.5, respectively, out of 30 possible points. The difference was not significant, $F(1,48)=.15$, $p=.70$. In comparison to the average score of 16.8 obtained from 418 university students who participated in Bachman's study, these participants could be best considered as advanced ESL speakers.

Materials. The test materials were the same as those in Experiment 1 except that a Russian version of the experiment was created for Russian ESL participants. The creation of the Russian version followed the same procedure as the Chinese version in Experiment 1.

Procedure. All participants were tested in English first and in their native language second. This is to prevent any potential influence of the native language on their performance in the second language. Otherwise, the procedure was the same as that in Experiment 1.

Results

The same data treatment procedure used in Experiment 1 was adopted in this experiment, except that a higher high cutoffs of 7000 ms was adopted for the L2 data than that of 6000 ms for the L1 data for both groups. Data treatment affected 8.3% of the data for Russian ESL speakers and 6.9% of the data for Chinese ESL speakers. Data from two Russian ESL speakers and on two items in the Chinese part of the experiment had to be discarded due to high ERs (20% or higher). The mean RTs and ERs after data treatment for the two groups of participants were shown in Table 2.

(Table 2 about here)

Two sets of analyses were done on RT and ER means for each group of participants, first a two-way ANOVA examining the main effect of language (L1 and L2), the main effect of number-match (match and mismatch), and the interaction between the two variables, followed by t tests comparing the number-match and mismatch conditions for each language separately. In all analyses, both participant means (F_1 , t_1) and item means (F_2 , t_2) were used. We report the results of these analyses below.

Russian ESL speakers. These participants showed a main effect of both language and congruency. Their performance in Russian was 895 ms faster than that in English and the difference was significant, $F_1(1,23)=65.7$, $p<.05$; $F_2(1,39)=208.9$, $p<.05$. They also responded to the n-match items 130 ms faster than to the n-mismatch items and the difference was also significant, $F_1(1,23)=10.1$, $p<.05$; $F_2(1,39)=8.0$, $p<.05$. There was no significant interaction between the two in RT, both $F_s<1$. The ER data showed no main effect of language, both $F_s<1$, or number match, both $F_s<1$. Neither was significant interaction, $F_1(1,23)=1.5$, $p=.24$; $F_2(1,39)<1$. The paired-samples t tests of the RT data showed that the number-mismatch effect in L1 approached significance in participant analysis, $t_1=1.90$, $p=.07$, and significant in item analysis, $t_2=2.40$, $p<.05$. The number-mismatch effect in L2 English was significant in both participant and item analyses, $t_1=2.79$, $p<.05$, $t_2=1.97$, $p=.05$. No significant number-mismatch effect was found in ER in either language (L1, $t_1=1.65$, $p=.113$, $t_2=1.16$, $p=.25$; L2, both $t_s<1$).

Chinese ESL speakers. There was a significant main effect of language in both RT ($F_1(1,27)=41.8$, $p<.05$, $F_2(1,37)=422$, $p<.05$) and ER ($F_1(1,27)=16.4$, $p<.05$; $F_2(1,37)=8.6$, $p<.05$). No significant main effect of congruency was found in RT (both $F_s<1$) or ER (both $F_s<1$), neither was any significant interaction between the two variables in RT (both $F_s<1$) or ER (both $F_s<1$). Paired-samples t tests produced no reliable number-mismatch effect in either L1 (both

$t_s < 1$) or L2 ($t_1 = 0.73$, $p = .47$, $t_2 = 1.10$, $p = .28$) in RT. No reliable number-mismatch effect was found in either L1 or L2 in ER, either (L1: $t_1 = 1.29$, $p = .21$, $t_2 = 0.98$, $p = .33$; L2: both $t_s < 1$).

The results from the two groups of participants first replicated those of Experiment 1. While performing the task in their respective L1s, Chinese and Russian speakers showed a different pattern of results. Russian speakers showed a reliable n-mismatch effect, but Chinese speakers did not. These results reinforced our contention that a n-mismatch effect in sentence-picture matching would only occur when the singular and plural meanings are morphologically marked in a language and automatically activated in language processing.

More importantly, the two groups produced different results while performing the task in L2, too. Russian ESL speakers showed a mismatch effect in performing the task in English, but Chinese ESL speakers showed no such an effect. These L2 results suggested that the singular/plural meanings are automatically activated and played a role in the matching process for Russian ESL speakers, but not for Chinese ESL speakers. These results are consistent with the view that L2 learners can routinely activate grammaticalized meanings in L2 processing only when such meanings are also grammaticalized in L1.

To provide a more detailed account of the L2 results, let's consider the example in Figure 1. When Russian ESL speakers processed the picture containing three single objects in the number-matched condition (Version a in Figure 1), a singular meaning was automatically activated in association with the three objects, including the critical object basket. A singular meaning was also activated in processing the word *basket* in the sentence. No number mismatch was present in the messages constructed for comparison. A positive Yes response was not adversely affected under this circumstance. In processing a n-mismatch item, however, a plural meaning was generated in association with the basket in the picture because of the presence of

multiple baskets in the picture, but a singular meaning with the word *basket* in the sentence. This mismatch caused a delay in reaching a Yes response to the n-mismatch items.

It was a very different scenario for Chinese ESL speakers, however. Note that due to a lack of plural marking in Chinese, there was no mismatch in number between the picture and sentence in the Chinese part of the experiment. Both items in both the singular and plural objectsA Chinese speaker may generate a plural meaning in picture processing for the mismatch condition, but and thus no delay in processing the “mismatch” items. Assuming that the plural meaning was not part of the routinely activated meaning among these Chinese ESL speakers, the mismatch in number between the English sentences and the pictures did not matter to them linguistically and psycholinguistically. This explains the lack of the n-mismatch effect in their results.

One may argue that, for some reason, the activation of the number meaning is harder to detect among Chinese ESL speakers in the task adopted in the study. We explored this possibility in Experiment 3 by demonstrating a lexical n-mismatch effect among Chinese ESL speakers using the same task. We suggested earlier that the number meanings are lexicalized, not grammaticalized, for Chinese speakers, which means that such meanings are linked to words such as *yige* (one), *jige* (several), *xuduo* (many) in Chinese and to words such as *one*, *several*, and *many* in English. A lexical n-mismatch effect in this context refers to a delay in responding to a n-mismatch item where a plural noun such as *baskets* is preceded by a quantifier such as *several* is presented with a picture with only one basket. Empirical demonstration of a lexical n-mismatch effect among Chinese ESL speakers could serve two purposes: showing that the task is sensitive enough to capture the automatic activation of the number meaning among Chinese ESL speakers and confirming that such meanings are readily activated in the presence of a lexical

item. Such a finding would help strengthen our proposal that what is lacking for Chinese ESL speakers is the presence of a grammaticalized (not lexicalized) number meaning in English processing.

Experiment 3: Lexical number-mismatch effect among Chinese ESL speakers

To demonstrate that lexicalized singular/plural meanings are automatically activated in language processing for Chinese ESL speakers and the task is powerful enough to capture such activation, we modified the test materials used in the first two experiments to create critical trials such as the one illustrated in Figure 2. These test items were different from those in the first two experiments in mainly two ways. First, a word such as *several* was used to indicate the plural meaning lexically in these sentences. Second, the pictures with multiple tokens of the critical object were in the n-match condition and the pictures with a single object were in the n-mismatch condition because of the plural nature of the sentence, as illustrated in Figure 2. If Chinese ESL speakers do activate and represent singular and plural meaning lexically in English and picture processing, we would expect them to produce a lexical number-mismatch effect in that they would show a delay in responding to the n-mismatch items.³

(Figure 2 about here)

Method

Participants. Twenty-six Chinese ESL speakers participated in the experiment. They were all graduate students studying in an American university at the time of testing.

Materials and procedure. A total of eighty items were constructed for this experiment, all in English. Forty of them had sentences and pictures matched in physical relationship and served as critical stimuli and the other forty were nonmatching filler items. Each of the forty matching

items had two versions. The sentences for both versions were identical. All sentences ended with a NP that had a plural quantifier such as *several*, *two*, *many*, *a few*, and a plural head noun. The picture had multiple tokens of the referent of the head noun for the n-match condition, and only a single object of the referent for the n-mismatch condition. The construction of the trial lists and the procedure were the same as those in Experiment 2.

Results

The data were treated in the same way they were in the first two experiments except a higher high cutoff of 1800 ms was adopted in light of the longer RTs in this data set. Data treatment affected 6.1% of the data. The participants' mean RTs and ERs for the n-match and mismatch conditions are shown in Table 3. Paired samples t tests showed that the 311-ms lexical n-mismatch effect was significant in both participant and item analyses, $t_1=3.90$, $p<.05$; $t_2=2.61$, $p<.05$. So was the difference in ER, $t_1=6.17$, $p<.05$; $t_2=3.38$, $p<.05$. Many participants reported noticing the mismatch between the plural NP in the sentence such as *several paper bags* and a single bag in the picture and thus providing a negative response, incorrectly, even though a sentence described the physical relationship among the three objects correctly and a positive response was expected. This contributed to the high ERs for the mismatch condition.

(Table 3 about here)

Two things are clear from the results. First, this sentence-picture matching task is adequate in assessing whether a meaning is activated in language processing. Given the right manipulation of the test materials and the task, the activation of a meaning can be detected by means of measuring the participants' RT under different conditions. In the context of this particular experiment, the activation of a meaning revealed itself through a longer RT for the n-mismatch condition. Second, lexical meanings related to number, such as *several*, *two*, or *many*,

get automatically activated in language processing among Chinese ESL speakers, which lead them to notice the number mismatch between the sentence and the picture and respond more slowly to the mismatch items, or make more errors on these items. The finding of a lexical number-mismatch effect in this experiment drew a sharp contrast to the results of Experiment 2 in which Chinese ESL speakers showed no number-mismatch effect when the number meaning was expressed through the means of morphology. Taken together, the results from the Chinese ESL speakers in Experiments 2 and 3 showed that lexicalized L2 meanings were automatically activated but not grammaticalized L2 meanings that were not grammaticalized in L1.

General Discussion

We suggested in the introduction that the meaning that is not grammaticalized or morphologically marked in a learner's L1 is not automatically activated in L2 processing unless lexical means are involved in expressing the meaning and this is the cause of the difficulty L2 learners have in acquiring incongruent morphemes in L2. Three sentence-picture matching experiments were conducted to test the assertion. Three findings emerged from these experiments:

- a. Speakers of a language that marks the number meaning morphologically (English and Russian speakers in Experiments 1 and 2) showed a number-mismatch effect in L1 processing in that they took longer in responding to test items whose critical noun in a sentence and its referent in the picture did not match in number than to items that did not involve such mismatch, but speakers of a language that does not mark number morphologically (Chinese speakers in Experiments 1 and 2) did not show this effect.
- b. When ESL speakers performed the task in their L2 English, Russian ESL speakers showed a number-mismatch effect, but not Chinese ESL speakers (Experiment 2).

c. Chinese ESL speakers were able to show a lexical number-mismatch effect when the mismatch occurred between a number word and an object (Experiment 3).

These findings suggest that the n-mismatch effect observed in this task provides a tangible and reliable means to determine whether the number meanings are automatically activated or not in the process of language processing. More importantly, the findings provided empirical confirmation for the proposition that a meaning that is morphologically marked in L2 but not in L1 is not automatically activated (without lexical means) in L2 sentence processing. These findings strengthen the plausibility of our approach to the explanation of the morphological congruency effect. That is, the difficulty in learning an incongruent L2 morpheme lies in the inability to automatically activate a meaning that is marked by the incongruent morpheme.

Two issues related to the explanation of the morphological congruency effect deserve more discussion. One is how maturational constraints affect morpheme acquisition in particular and the other is how such a meaning-based model can be applied to the acquisition of L2 morphemes that are less semantic in nature such as gender marking.

The role of maturational constraints, or a critical period, in adult L2 learning is well recognized. For example, in explaining why adult L2 learners are less capable of relying on implicit learning in developing native-like proficiency in an L2, Ellis (2011) commented that “the L2 learner’s neocortex has already been tuned to the L1, incremental learning has slowly committed it to a particular configuration, and it has reached a point of entrenchment where the L2 is perceived through mechanisms optimized for the L1” (p. 40). This neural commitment view is also adopted by Kuhl in the discussion of phonological development who suggests that “learning results in a commitment of the brain’s neural networks to the patterns of variation that

describe a particular language. This learning promotes further learning of patterns that conform to these initially learned, while interfering with the learning of patterns that do not conform to those initially learned” (Kuhl, 2004, p. 832).

Several researchers have also explained L2 learners’ difficulty with incongruent morphemes in terms of maturational constraints (e.g., Franceschina, 2005; Hawkins & Liszka, 2003; Scherag et al., 2004), though in a less than specific manner. Within the meaning activation approach taken in the present study, maturational constraints can be said to affect what meanings are routinely activated in language processing. Lexical meanings are activated by lexical input in comprehension and by intention and what is available in the lexicon in production. The activation of lexical meanings is more fluid in the sense that it is completely controlled by situation-specific input and thoughts. In contrast, the activation of grammaticalized meanings is more patterned in the sense that certain grammaticalized meanings have to be activated in association with other meanings due to the design of the language and regardless of specific situations. For example, a number meaning (singular or plural) has to be activated in association with a countable noun in English. This patterned activation of grammaticalized meanings is language-specific and determined by the grammaticalization patterns of a language. Learning a language in childhood comprises learning this particular pattern of semantic activations, that is, what grammaticalized meanings should go with what lexicalized meanings. Because the number of grammaticalized meanings is limited and stable, it is reasonable to think that the patterned activation of grammaticalized meanings is finalized at a young age in the process of learning one’s L1, just like the finalization of perceptual categories for speech. It is this particular pattern of activation of grammaticalized meanings that is constrained by maturation. Thus, maturational constraints affect the acquisition of L2 inflectional morphemes by putting a particular pattern of

semantic activation in place in childhood L1 learning that is difficult to alter. Whether the resistance to alteration is the result of neural commitment, as Ellis (2011) and Kuhl (2004) suggested, or merely “because of stylistic differences in learning at different times in life,” as Bialystok (1997, p. 132) argued, is debatable.

The second issue has to do with how a meaning-activation approach applies to grammatical morphemes that are less semantic than plural and tense markers. Good examples of the former are the third person singular marker in English and gender markers in Spanish and French. These markers are clearly not as semantic as plural and tense markers in English as they do not usually express a meaning a speaker wants to convey. Instead, they are more syntax-motivated and have no semantic meaning associated with them. How such morphemes differ from more semantic morphemes in acquisition is yet to be determined. Before more evidence becomes available, there are two good reasons to examine the development of these morphemes in the current model as well. First, this model can serve as an initial point of departure if one defines meaning broadly to include what is marked by grammatical person, gender, or case. Similar to English plural and tense markers, the third person singular meaning, for example, though more syntactic than semantic, has to be activated in the process of speech production or comprehension. Thus, the idea of a finalized pattern of meaning activation in childhood L1 acquisition discussed earlier should apply to these morphemes as well. Second, published studies point to the existence of a morphological congruency effect in the acquisition of gender marking by adult L2 learners (e.g., Franceschina, 2005; Guillelmon & Grosjean, 2001; Sabourin et al., 2006; Scherag, et al., 2004). Given the presence of some conflicting findings (Foote, 2011; Keating, 2009; Sagarra & Herschensohn, 2010, 2011),⁴ further research is certainly necessary to

examine the morphological congruency effect in the acquisition of gender markers in particular, and in the acquisition many other inflectional morphemes in general.⁵

Conclusion

The difficulty associated with inflectional morphemes in L2 learning has been widely recognized. However, we know much less about the exact nature of this difficulty. The present study represents our effort toward a better understanding of what underlies this difficulty from a cognitive perspective. The meaning-activation model of L2 morpheme acquisition proposed in this study provides a conceptual framework for exploring L2 morpheme acquisition in general and for understanding the difficulty in the learning of incongruent morphemes in specific.

We also hope to consider this study as part of a more challenging but worthwhile endeavor of identifying what is not acquirable in adult L2 acquisition. Both common sense and research evidence have shown that adult L2 learners are very unlikely to develop native-like competence in all aspects of language. It is surprising in this regard that there has been little orchestrated, deliberate, and conscious effort in SLA research in identifying what is not acquirable. We believe that the identification of unacquirable morphosyntactic structures should be a central topic in SLA research. Progress on this topic is prerequisite for understanding what is unique about adult L2 acquisition.

Aknoweldgement

We would like to thank Zhang Weimin and Fang Yanhua for their assistance in collecting data from the Chinese speakers in Experiment 1, and Cao Yongheng for assistance in data collection for Experiment 3. Part of the results were presented at the 2008 annual meeting of the American Association for Applied Linguistics held in Washington, D.C. from March 27 to April 1, 2008.

References

- Bachman, L. F. (1982). The Trait Structure of Cloze Test Scores. *TESOL Quarterly*, 16(1), 61-70.
- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13, 116-137.
- Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, 2, 127-145.
- Dewaele, J.-M., & V'eronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, 4, 275-296.
- Ellis, N. C. (2011). Implicit and explicit SLA and their interface. In Sanz, C., & Leow, R. (Eds.). *Implicit and explicit language learning: conditions, processes, and knowledge in SLA and bilingualism* (pp. 35-47). Washington, DC: Georgetown University Press.
- Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics*, 32, 187-220.
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213-247.
- Franceschina, F. (2005). *Fossilized second language grammars: the acquisition of grammatical gender*. Amsterdam: John Benjamins.
- Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition*, 29, 503-511.
- Hawkins, R. & Liszka, S. (2003). Locating the source of defective past tense marking in

- advanced L2 English speakers. In R. van Hout, H. Aafke, F. Kuiken, and R. Towell (Eds.). *The interface between syntax and lexicon in second language acquisition* (pp. 21-44). Amsterdam: John Benjamins.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25, 603-634.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57, 1-33.
- Jiang, N., Novokshanova, E., Masuda, K., & Wang, Xin. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, 61, 940-967.
- Keating, G. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, 59, 503-535.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences USA*, 97, 11850-11857.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews/Neuroscience*, 5, 831-843.
- Lucy, J. (1992). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
- McDonald, J. L. (2000). Grammaticality Judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21, 395-423.
- Parodi, T. Schwartz, B. D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, 42, 669-705.
- Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22, 1-29.

- Scherag, A., Demuth, L., Rosler, F., Neville, H. J., & Roder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, 93(3), B97-B108.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge UK: Cambridge University Press.
- White, L. (2003). Fossilization in steady state l2 grammars: persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, 6, 129-141.
- Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L2 and L2 Spanish. *Lingua*, 120, 2022-2039.
- Sagarra, N., & Herschensohn, J. (2011). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning*, 61(1), 80-116.

End Notes

1. This is certainly related to the linguistic relativity issue which is not reviewed here due to space constraints. Interested readers may refer to Lucy (1992 and Slobin (1996).
2. Several mental processes were conceivably involved in performing the task. First, one had to recognize the objects, note their physical relationship (as directed by the instructions), and encode and store this information in the mind. Then, the sentence had to be processed and its message stored. Finally, the two messages were compared to see if they matched. If they did, a positive response was in order. Otherwise, a negative response was generated.
3. The same idea can also be tested by simply changing the original sentences used in Experiment 2 by replacing the definite article with the word *one*, e.g., *The bike is above one tall basket*. When this sentence is paired with a picture with one tall basket and another picture with multiple tall baskets, the same lexical mismatch effect should be found among Chinese ESL speakers.
4. Native-like performance in at least some of these studies may have a methodological explanation. In the two studies by Sagarra and Herschensohn, for example, the instructions emphasized accuracy rather than speed which to some extent defeats the purpose of using the self-paced reading task (see Jiang, 2004). The RT data were consequently much longer in these studies, often in the 800s ms and 900s ms range, rather than a typical 300s ms and 400s ms (see Foote, 2011, Jiang, 2004, 2007 for comparison). Participants in Foote (2011) were mostly Spanish instructors, an atypical group of participants who were more likely to pay attention to grammatical accuracy even in performing a self-paced reading task.
5. My colleague Robert DeKeyser brought to my attention a difference between gender marking in Spanish and plural marking in English. English nouns and verbs have an uninflected form as a

base form to which the plural or tense markers are then added. Spanish nouns and adjectives usually do not have an uninflected form. They are usually either feminine or masculine. This difference is likely to affect how gender marking is learned. Its exact effect is yet to be explored.

Table 1

English and Chinese speakers' mean RTs, mean ERs, and their standard deviations (in parentheses) for the number-match (n-match), number-mismatch (n-mismatch), and No-response conditions.

		English Speakers N=32		Chinese Speakers N=36	
		Reaction Time (ms)	Error Rate (%)	Reaction Time (ms)	Error Rate (%)
Yes-response items	N-match	2888 (707)	6.4 (5.3)	2831 (589)	11.6 (5.6)
	N-mismatch	3001 (769)	10.3 (7.8)	2794 (583)	13.5 (6.6)
	Difference	113	3.9	-37	2.9
No-response items		2787 (639)	5.7 (6.0)	2629 (556)	9.7 (5.5)

Table 2.

Russian and Chinese ESL speakers' mean RTs (in ms) and ERs (in percentage) for the n-match and n-mismatch conditions while performing the task in L1 and L2.

		Russian ESL (n=24)		Chinese ESL (n=28)	
		RT	ER	RT	ER
L1	N-match	2484 (451)	7.2 (9.4)	2615 (458)	6.1 (7.1)
	N-mismatch	2581 (529)	9.3 (12.3)	2604 (538)	4.4 (5.5)
	Differences	97*	2.1	-11	-1.7
L2	N-match	3344 (829)	8.1 (9.2)	3702 (811)	11.5 (9.4)
	N-mismatch	3506 (757)	7.6 (7.3)	3750 (788)	10.5 (10.6)
	Differences	162**	-0.5	48	-1

* significant in participant or item analysis

** significant in both participant and item analysis

Table 3. Chinese ESL participants RT and ER in responding to congruent and incongruent items in Experiment 3.

	RT	ER
Congruent	3840 (911)	7.5 (9.3)
Incongruent	4151 (1033)	20.1 (11.3)
Differences	311**	12.6

** significant in both participant and item analysis

Figure Captions

Figure 1

A test item in the sentence picture matching task. Each item had a set of three objects accompanied by a sentence that described the physical relationship of the objects. Each item had two versions, a number-match version with a single tall basket (a) and a number-mismatch version with two baskets (b).

Figure 2

N-match (a) and mismatch (b) items in Experiment 3

The bike is above the tall basket.



(a)

The bike is above the tall basket.



(b)

- (a) The notebook is to the left of several paper bags.



- (b) The notebook is to the left of several paper bags.

