BUCLD 39 Proceedings
To be published in 2015 by Cascadilla Press
Rights forms signed by all authors

Real-time Processing of Classifier Information by L2 Speakers of Chinese

Elaine Lau and Theres Grüter

Native (L1) speakers take advantage of prenominal cues, such as gender-marked articles and classifiers, to identify an upcoming noun during online processing (e.g., Lew-Williams & Fernald, 2007; Huettig et al., 2010). The extent to which non-native (L2) speakers are able to do so remains a topic of on-going investigation. Findings from learners of gender-marking languages have not been entirely consistent, and point to the influence of a number of language- and learner-specific factors, as discussed in more detail below. No previous findings from L2 learners of classifier languages are available, as far as we know. The goal of the present study is to extend research on the facilitatory effect of prenominal cues in the online processing of an L2 by looking at classifiers in Chinese, which are both similar and different along potentially relevant dimensions from gender-marked articles in Indo-European languages. We report the findings from a visual-world eye-tracking experiment with L1 and L2 speakers of Chinese, closely following the procedures and design of Lew-Williams and Fernald's (2007, 2010) work on the processing of gender-marked articles in L1 and L2 Spanish.

1. Classifiers in Chinese

Classifiers are morphemes marking the noun class of the following noun. Their presence is obligatory when the noun phrase includes a demonstrative, as illustrated in (1).

(1) na *(tiao) maozin
that CL towel
'that towel'

There is a large inventory of classifiers in Chinese, ranging from the commonly agreed number of about 75 (Erbaugh, 2004) to an exhaustive list of about 902 from the *Hanyu Liangci Cidian* (A Dictionary of Chinese Classifiers; Chen et al.,

---

* Both authors are at the University of Hawaiʻi at Mānoa; Elaine Lau: elau@hawaii.edu; Theres Grüter: theres@hawaii.edu. We are grateful to Casey Lew-Williams for sharing his materials with us, and to Yongsi Huang for assistance with participant recruitment and testing. This study was supported by an Elizabeth Carr Holmes Scholarship to Elaine Lau from the Department of Second Language Studies at UH.

1988). The specific type of classifier relevant here is the type known as sortal (Lyons, 1977) or qualifying classifiers (Hu, 1993), which categorize nouns according to their inherent properties, such as shape, animacy, natural kind or function, of the noun (Chao, 1968; Matthews & Yip, 2011). The number of nouns associated with a particular classifier varies greatly among different classifiers, and membership within a classifier category is not necessarily homogeneous in nature, as the association of the nouns to a classifier class can be based on different dimensions. For example, the classifier *zhi* is used to denote birds and other small animals such as rabbits and cats. Tigers and lions also fall into this class by extension from cats, as do small objects such as ears, hands, shoes and socks due to their small size and manipulability (Erbaugh, 2002). Thus unlike noun classes in Indo-European languages based on grammatical gender, where membership, for the most part, is not defined by semantic criteria, there is a significant, though not exhaustive, semantic component to the noun class system in classifier languages such as Chinese.

2. Processing of classifiers in L1 Chinese

Previous studies on the processing of classifiers by L1 Chinese speakers have demonstrated that cues from the classifier facilitate reference resolution during online processing. Using a visual world paradigm, Huettig et al. (2010) showed that speakers of Chinese efficiently utilized the information from the classifier in the speech stream to locate the target object in a visual scene. The contrast in the looking patterns between conditions in which the classifier was present or absent suggested that classifiers served as a predictive cue for listeners to identify the upcoming noun. Moreover, when the object being named was not contained in the visual scene, fixations were briefly attracted to objects depicting nouns from the same classifier class.

Using a similar visual world paradigm setup as Huettig et al. (2010), Klein et al. (2012) included, in a single visual display, (i) the target object (e.g., *men* ‘door’, which takes the classifier *shan*, a classifier for flat vertical things), (ii) a classifier-consistent competitor (e.g., *chuanghu* ‘window’), which uses the same classifier as the target, (iii) a phonological competitor (e.g., *menpiao* ‘ticket’), in which the first syllable of the noun is homophonous with that in the target noun, but the noun does not use the same classifier as the target, and (iv) a classifier competitor (e.g. *shan* ‘fan’), which is homophonous with the classifier itself. When the auditory stimuli included the specific classifier (*shan*), looks to (iii) and (iv) dropped rapidly, while those to (ii) continued to rise until well after target noun onset, indicating that on hearing the classifier, participants used this information proactively to narrow down the set of possible referents to those consistent with the classifier.

Tsang and Chambers (2011) investigated the differential contributions of semantic versus class-membership information encoded in classifiers during the online processing of Cantonese, a Chinese language with a similar classifier system as Mandarin Chinese. Their findings indicate that semantic information did not play a strong role when the target object was a prototypical member of the classifier class (i.e., displaying all its defining semantic features, e.g., long, narrow and flexible for the classifier *tiu4*). In this case, no competition was found between a target object (e.g., *geng2gan1*, *tiu4*, 'scarf') and a non-classifier-consistent competitor that shared the same prototypical semantic features (e.g., *kei4*, *zi1*, 'flag'), but solely from a classifier-consistent competitor. A slightly stronger effect was observed with classifier-consistent competitors that shared the prototypical semantic features (e.g., *tiu3sing2, tiu4,* 'jump rope') than with those that did not (e.g., *so2si4, tiu4,* 'key'). However, when the target object itself was not a prototypical member of the classifier class, participants attended more to competitors which displayed the prototypical semantic features, including competitors not from the same classifier class. The authors concluded that (shape) classifiers influence predictive processing "primarily through their grammatical constraints" (p. 1065), with classifier semantics acting as a secondary cue that becomes apparent only in certain circumstances, such as with non-prototypical members of a class.

In sum, native speakers of Chinese efficiently utilize information encoded by the classifier during reference resolution in online processing. They appear to use both class-membership and semantic cues to do so, with the strength of these cues potentially varying depending on the prototypicality of the nouns involved.

3. Acquisition and processing of noun class information in an L2

To date, no previous study that we are aware of has explored the online processing of classifiers by L2 speakers. Offline studies have shown that classifiers are often a source of difficulties for L2 learners of Chinese. Although learners appear to become aware of the obligatory presence of classifiers in required contexts, they often fail to select the appropriate classifier (Gao, 2009); instead, they sometimes select a non-matching classifier, or they often use the general classifier *ge* (Liang, 2008; Polio, 1994), a phenomenon also observed in L1 Chinese-speaking children (Fang, 1985; Hu, 1993) and L1 adult casual speech (Erbaugh, 2002). L2 learners thus appear to use the general classifier, which is acceptable in many but not all contexts, as a syntactic place holder to fill the classifier position, complying with structural but not semantic constraints pertaining to classifiers.

While classifier processing has not been investigated in L2 yet, another prenominal cue, namely gender-marking on determiners, has been extensively examined. Gender is a noun classification system in which the categories are based on biological sex (feminine, masculine, and in some cases neuter/inanimate). Although gender systems show some relations with biological sex, most nouns are arbitrarily assigned to gender classes, without reference to any inherent properties of the associated object. Therefore, within each gender class, membership is highly heterogeneous (Aikhenvald, 2003). In visual world studies similar to those described above for Chinese, L1 speakers of gender-marking languages were consistently faster in identifying the target after hearing a gender-marked determiner when objects on the screen were from different gender classes than when they shared the same gender (e.g. Dussias et al., 2013; Hopp, 2013; Lew-Williams & Fernald, 2007, 2010). In contrast, adult L2 learners of gender-marking languages do not consistently succeed in exploiting grammatical gender as a predictive cue in the same experimental contexts. Some studies have found facilitatory effects of gender-marked articles with L2 learners, particularly for learners whose L1 is also a gender-marking language (Dussias et al., 2013), for learners who independently demonstrated fast lexical access speed in the L2 (Hopp, 2013), and for learners who were implicitly taught novel nouns and their gender class through repeated exposure at the beginning of the experiment (Grüter et al., 2012). Yet the original paradigm used by Lew-Williams and Fernald (2007) with L1 speakers of Spanish has consistently failed to reveal facilitatory effects for English-speaking learners of Spanish, both at intermediate (Lew-Williams & Fernald, 2010) and near-native (Grüter et al., 2012) levels of proficiency. Note that the stimuli in this paradigm consist of inanimate nouns, such as *la pelota* ('the-FEM ball'), for which there is no semantic basis for noun class assignment. Importantly for our purposes, in a related experiment, Lew-Williams and Fernald (2009, Experiment 3) included animate, gendered nouns, such as *la niña* ('the-FEM girl'), and observed a clear facilitatory effect for both L1 and (intermediate proficiency) L2 speakers of Spanish. Thus it appears that L2 learners were able to make use of the *semantic* information encoded by the gender-marked article, but they were not able to exploit the more abstract cue signaling noun-class membership during online processing.

If this interpretation is on the right track and L2 learners are more successful at exploiting semantically informative cues, we should expect L2 learners of Chinese to be able to take advantage of (semantically informative) classifiers to identify an upcoming noun in online processing. In other words, we would expect them to pattern more like the L2 learners of Spanish in the '*la niña* study' than those in the '*la pelota* study'. The goal of our study is to explore this prediction.

4. Method

4.1 Participants

L1 participants were 19 native Chinese-speaking adults living in Honolulu, Hawai'i (age: 20.6 - 36.7 years, *M* = 25.6, *SD* = 4.7). Three additional L1 participants were excluded due to poor calibration or early exposure to English. All of the remaining 19 L1 participants had moved to the U.S. during adolescence or later (age of arrival: 14 - 36 years, *M* = 20.6, *SD* = 6.2). L2 participants were 16 native English-speaking adults from the University of Hawai'i community (age: 19.3 - 70.7 years, *M* = 32.6, *SD* = 15.2). Four additional L2 participants were excluded due to poor calibration or early exposure to Chinese as a heritage language. None of the remaining 16 L2 participants were exposed to Chinese during early childhood, or had spent extensive amounts of time in a Chinese-speaking environment. Table 1 provides a summary of relevant background and self-rating information obtained from a questionnaire completed prior to the experimental session.

Table 1. Language background and self-reported proficiency.

	L1	L2
	M (SD)	*M (SD)*
Age of first exposure to Chinese	.8 (2)	18 (4.2)
Self-rated proficiency (on a 10-point scale):		
Speaking	9.4 (1.1)	4.8 (1.8)
Understanding	9.7 (0.6)	6 (1.9)
Overall proficiency	9.4 (0.9)	5.4 (1.9)

As an additional measure of proficiency in Chinese, all participants completed a written cloze test, designed by the first author, as well as a *classifier knowledge test* designed to assess whether participants knew the correct pairing of classifiers with the target nouns (see 4.2.2 for further detail). Results from these measures are reported in section 5.

4.2 Materials

4.2.1 Eye-tracking experiment

Materials were created closely following the design of Lew-Williams and Fernald (2007). In place of the two gender-marked articles, the classifiers *tiao* (long, string-like objects) and *zhang* (flat surfaced objects) were chosen for this experiment because they are frequently used in modern Chinese (Academia Sinica, 2013; Ministry of Education of the People's Republic of China, 2013), and were included in the vocabulary taught in elementary level Chinese courses designed for English-speaking adults (Liu et al., 2009; Ning & Montanaro, 2011). These two classifiers have well-defined and salient perceptual features, and occur with a variety of objects that are visually distinct and common in ordinary life. Table 2 lists the nouns chosen for each classifier in the experiment, which were also chosen from the same introductory textbooks as the classifiers.

Table 2. Classifier-noun pairings.

tiao (long, string-like objects)	*zhang* (flat surfaced objects)
chuan 'boat'	*chuang* 'bed'
yu 'fish'	*zhuozi* 'table'
maozin 'towel'	*ditu* 'map'
kuzi 'trousers'	*xinyongka* 'credit card'

Each trial consisted of a speech stimulus and a visual stimulus. The speech stimulus consisted of a simple Chinese sentence with the target classifier noun phrase embedded in a carrier sentence: *kandao* ('see')/ *zhaodao* ('find') *na* ('that')

CLASSIFIER NOUN *ma* (question particle)? The classifier noun phrase was always headed by the distal demonstrative *na* (‘that’) followed by the classifier, and referred to one of two objects on the screen. All speech stimuli were recorded by a female native speaker of Chinese, and were normalized to control for the length of the classifier region (normalized duration = 242ms), and the pre-classifier region, i.e., the sentence initial verb + demonstrative (normalized duration = 627ms).

Visual stimuli consisted of two pictures shown on a black screen. The pictures were coloured illustrations of animals and objects on a grey background. An example is given in Figure 1. One picture served as the target, matching the noun in the speech stimulus, and the other as the distractor. Side of target presentation was counter-balanced across trials. To enhance the variety of visual stimuli, two tokens were created for each noun, as shown by the two different tokens of 'towel' in Figure 1.

(a) (b)

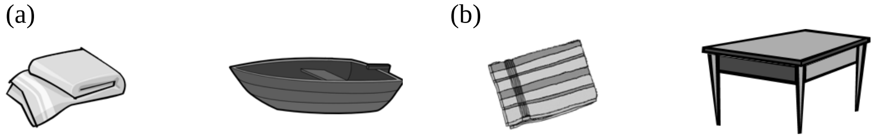

Figure 1. Sample visual stimuli: (a) SAME classifier condition; (b) DIFFERENT classifier condition.

The experiment consisted of two experimental conditions: the SAME classifier condition and the DIFFERENT classifier condition. In the SAME condition (Figure 1a), the objects depicted by the two images shared the same classifier, such as ‘towel’ and ‘boat’ which both take the classifier *tiao* (CL$_{\text{STRING}}$). On the other hand, in the DIFFERENT classifier condition (Figure 1b), the objects take different classifiers, such as ‘towel’ and ‘table’, which take the classifiers *tiao* (CL$_{\text{STRING}}$) and *zhang* (CL$_{\text{FLAT OBJ}}$) respectively. Thus the classifier is informative with regard to the identification of the upcoming noun only in the DIFFERENT classifier condition. Hence if participants take advantage of information encoded by the classifier, they should identify the target earlier in the DIFFERENT than in the SAME classifier condition.

Each noun served as the target on four trials and as the distractor on four trials, for a total of 32 experimental items, 16 in each of the two experimental conditions. 24 filler items were created and interspersed between the target items. Four lists in pseudo-randomized orders were created, and counter-balanced across participants.

4.2.2 Classifier knowledge test

The purpose of the classifier knowledge test was to assess whether participants knew the correct pairings of the nouns and classifiers used in the main experiment. The test consisted of 40 phrases containing a blank to be filled. The target items were the eight nouns used as the test items in the experiment; they were interspersed among 32 filler items. The target items took the form of a simple noun phrase, headed by either a demonstrative or a numeral, a context in which a classifier is required. An example is given in (2). The expected answer was the specific classifier associated with the noun. Participants could provide the answer in Chinese characters or pinyin romanization.

(2) 一 _____ 小魚 Expected answer: 條 *tiao*
a small fish CL_{LONG, STRING-LIKE OBJ}

All the vocabulary tested and used in the questions was selected from the textbook used in the elementary level Chinese course at the University of Hawai‘i (Liu et al., 2009; Ning & Montanaro, 2011). This provided some assurance that the test was designed at an appropriate level for learners.

4.3 Procedure

All participants were asked to complete the background questionnaire and the classifier knowledge test at least 7 days prior to the experimental session in order to minimize any effects of priming from the target items in the knowledge test on the visual world experiment. During the experimental session, a vocabulary check was performed with the participants immediately prior to the eye-tracking experiment, as in Lew-Williams and Fernald (2007). Participants were provided with the images that they were going to see in the experiment together with the Chinese nouns referring to these objects, and they were asked to provide the English translation of the noun, and to indicate their familiarity with the Chinese word. Classifiers were avoided during the labelling to avoid priming the particular noun-classifier pairings. The purpose of this check was to ensure that participants had the correct label for the objects depicted in the images used in the experiment, and to strengthen the association between the lexical items and these images. All participants correctly translated the nouns into English, and indicated being familiar or very familiar with the Chinese word in almost all cases.

Participants were then seated comfortably in front of a computer screen for the eye-tracking experiment, which was conducted on an SMI RED250 system, tracking at 250Hz. During the experiment, participants were told only to look and listen, and no responses would be required. Each trial started with a 2000ms presentation of the visual stimulus prior to the presentation of the speech stimulus. An 800ms interval was placed between trials. The experiment lasted approximately 5 minutes. Immediately after the eye-tracking experiment, participants completed the written cloze test as an independent measure of proficiency in Chinese.

5. Results

Unsurprisingly, the L1 group performed significantly better than the L2 group on the cloze test (L1: M = 95%, SD = 3; L2: M = 57%, SD = 27; $t(34) = 6.8$, $p < .01$) and on the classifier knowledge test (L1: M = 88%, SD = 23; L2: M = 60%, SD = 33; $t(34) = 3.4$, $p < .01$). On the classifier knowledge test, the most common type of non-target response in both groups was the use of the general classifiers (*ge,* or the plural form *xie*) in place of the more specific target classifier (M = 4.9% for L1ers and M = 22.9% for L2ers). Such responses do not sound completely natural to native speakers, and were thus scored as incorrect, as the objective of the test was to assess if participants knew the specific classifier-noun pairing. Substitution of an incorrect specific classifier (e.g., using *zuo*, the classifier for things in block, e.g., clock and building, instead of *zhang* for 'bed') was relatively rare in both groups (M = 0.6% for L1ers and M = 9% for L2ers). The substantial use of general classifiers by the L2 group makes it difficult to determine with certainty whether those participants knew the specific classifier for each noun, as their use of the general classifier may well reflect a general strategy. (In retrospect, we believe that a forced-choice task could have provided a better assessment of knowledge of classifier-noun pairings.) For this reason, and in order to retain the same number of data points from all participants, we did not eliminate data from the eye-tracking experiment based on participants' performance on the classifier knowledge test, although given a more successful measure of classifier-noun knowledge, it would perhaps be desirable to do so in future studies of this kind.

Figure 2 illustrates the time course of proportion of fixations towards the target in the SAME and DIFFERENT classifier conditions, starting from the acoustic onset of the classifier. As the length of the classifier was approximately 240ms, four 240ms windows for analysis were created from the acoustic onset of the classifier. The four windows are superimposed as w1-w4 in Figure 2. W1 corresponds to the classifier, w2 and w3 align approximately with the first and second syllable of the noun, and w4 aligns roughly with the question particle. Visual inspection of Figure 2 indicates the L1 group oriented faster to the target in the DIFFERENT than the SAME classifier condition, with the effect emerging in w3, while the noun was unfolding. The pattern in the L2 group is less clear. A (smaller) advantage for the DIFFERENT classifier condition appears to emerge late in w3 and continue into w4.

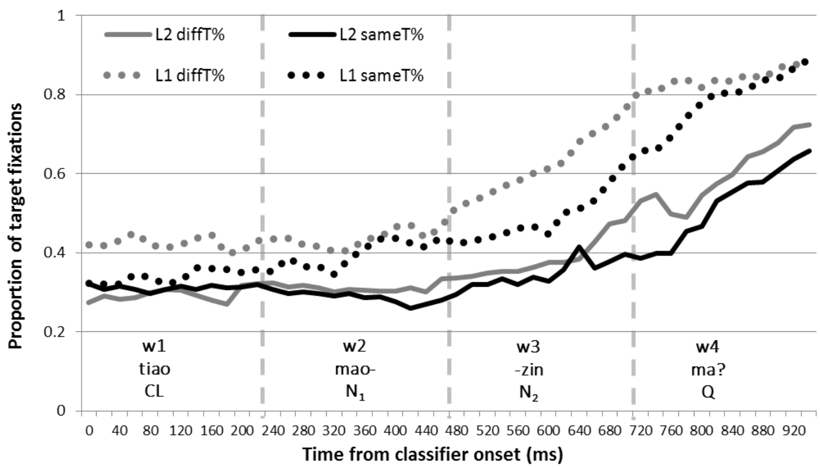

Figure 2. Proportion of looks to the target (out of fixations to any area of the screen): L1 vs. L2.

A 2 (condition) × 4 (window) × 2 (group) mixed ANOVA was performed on the mean proportion of target fixations. This analysis yielded a significant main effect for condition ($F(1, 33) = 10.0$, $p = .003$), indicating more looks to the target in the DIFFERENT than in the SAME classifier condition overall. A significant main effect was also obtained for window ($F(3, 99) = 69.84$, $p < .001$); the interaction between window and condition was not significant ($F(3, 99) = 1.49$, $p = .23$). The main effect for group was also significant ($F(1, 33) = 11.52$, p = .002); the interaction between group and condition was not significant ($F(1, 33) = 1.89$, $p = .18$), while the interaction between group and window was ($F(3, 99) = 4.07$, $p = .03$). The three-way interaction did not reach significance ($F(3, 99) = 1.91$, $p = .14$). Given the limited sample size in this study, power to detect potentially relevant effects in this omnibus test was limited, and results must be interpreted with great caution. In order to address our primary research question – can L2 learners of Chinese take advantage of an informative classifier? – we thus also conducted simple pairwise comparisons between conditions for each group in the windows of interest. For the L1 group, a clear advantage for the DIFFERENT classifier condition emerged in w3 ($t(18) = 3.98$, $p = .001$, $d = .91$). For the L2 group, there was no evidence of facilitation in w3 ($t(15) = 1.05$, $p = .31$, $d = .18$), but the comparison approached significance in w4 ($t(15) = 2.04$, $p = .059$, $d = .37$). Again, these results must be interpreted with caution, given multiple comparisons and small sample size.

Some further support that the small-to-medium effect observed in w4 for the L2 group is indicative of facilitation due to an informative classifier comes from a closer look at fixations to both target and distractor images, as shown in Figure 3. L2 participants' looks towards the target and the distractor begin to diverge early in w3 in the DIFFERENT classifier condition, whereas the split for the

SAME classifier condition appears about 200 milliseconds later, at the end of w3. The earlier decrease in looks to the distractor in the DIFFERENT classifier condition provides further indication that the L2 group seems to have derived at least some benefit from an informative classifier.

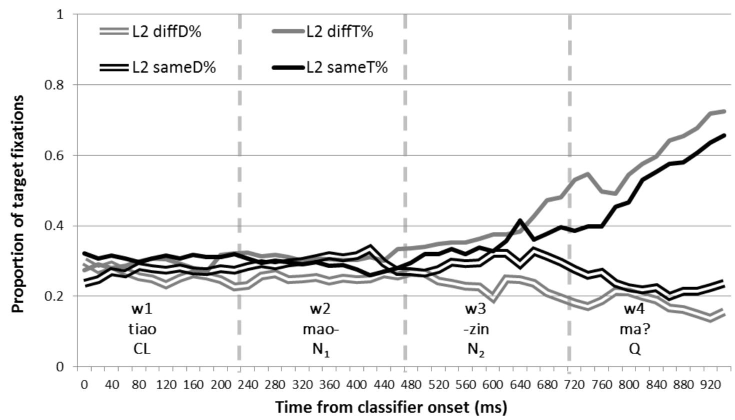

Figure 3. Target vs. distractor fixations in the L2 group.

Due to small sample size, the effect of L2 proficiency could only be explored descriptively in this study. For this exploratory purpose, we split the L2 group by performance on the written cloze test. As shown in Figure 4, the pattern in the higher proficiency group (n=8) seems to approximate that in the L1 group more closely, suggesting that proficiency most likely does play a role here. Participants in the lower proficiency group (n=8) could not identify the target until after the offset of the noun, and there are no clear differences between the two conditions for this group.

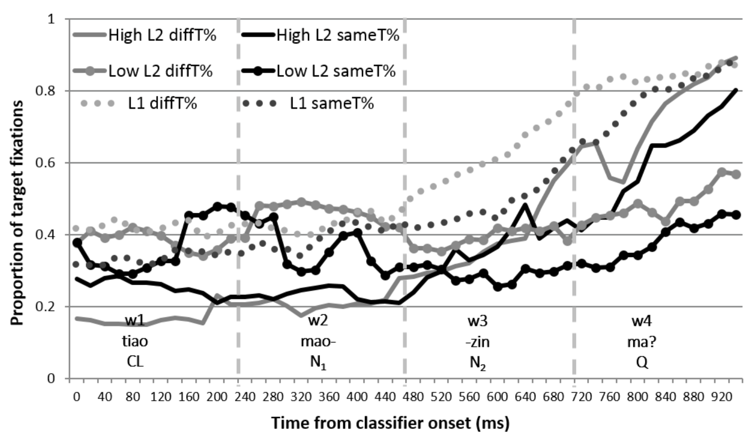

Figure 4. Proportion of looks to the target: split by L2 proficiency.

6. Discussion and Conclusion

In this study, we have replicated findings from previous work (Huettig et al., 2010; Klein et al., 2012; Tsang & Chambers, 2011) showing that Chinese native speakers use information encoded by a prenominal classifier to predict an upcoming noun during online comprehension. In addition, we have presented results from L2 learners of Chinese. While our L2 sample was small, and we were not able to determine conclusively to what extent participants in this group had full knowledge of the classifier-noun pairings used in the main experiment, the eye-movement results from this L2 group were nevertheless indicative of at least some facilitation due to information encoded on a classifier. Notably, this finding stands in contrast to the complete absence of facilitation observed for L2 learners of Spanish in the paradigm we followed as closely as possible (Lew-Williams & Fernald, 2007, 2010).

Why might classifiers provide better cues for non-native speakers than grammatical gender? There are at least two factors that appear relevant. First, classifiers are *semantically* informative, similar to semantic gender marking. Indeed, our results seem to align more closely with those from Lew-Williams and Fernald's (2009) '*la niña* study' than with those from their 2007 '*la pelota* study', in line with the hypothesis that L2 learners may be more successful at taking up semantically informative cues than cues based on statistical co-occurrence alone. Pursuing this hypothesis further in the context of Chinese classifiers will be particularly informative in light of Tsang and Chambers' (2011) observation that semantic properties of classifiers seem to play a *smaller* role for native speakers of Chinese during online processing.

A second factor that could help explain why Chinese classifiers provide a more informative cue than gender-marked articles in Spanish is the fact that there is a much larger number of classifiers in the Chinese classifier system than there are genders in any gender marking language. Most gender systems involve only two or three classes, whereas Chinese has at least 75 classifier classes (Erbaugh, 2004). As a result, a classifier is a more informative cue than a gender-marked determiner in that it narrows down the set of potential nouns that can follow to a considerably larger extent. Again, the Chinese classifier system presents an ideal context for pursuing this question further as classifier classes differ greatly among each other in terms of membership size. If membership size relates (inversely) to informativity, we would expect, all else being equal, that a classifier whose class has few members should be a better cue than one which is associated with a larger number of nouns. We therefore hope that the experiment we have presented here may serve as a starting point for a more extensive exploration of how native and non-native speakers of Chinese make use of information encoded on classifiers of various types during real-time comprehension.

References

Academia Sinica. (2013). *Academia Sinica Balanced Corpus of Modern Chinese.* Retrieved from http://app.sinica.edu.tw/kiwi/mkiwi/

Aikhenvald, Alexandra Y. (2003). *Classifiers: a typology of noun categorization devices (Oxford studies in typology and linguistic theory)*. New York, NY: Oxford University Press.

Chao, Yuen Ren. (1968). *A grammar of spoken Chinese.* Berkeley, CA: University of California Press.

Chen, Baocun, Che, Guicheng, Chen, Hao, & Zhang, Zaizhan. (1988). *Hanyu liangci cidian* [A dictionary of Chinese numeral classifiers]. Fuzhou, China: Fujian Renmin Chubanshe [Fujian People's Publishing House].

Dussias, Paola E., Valdes Kroff, Jorge R., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition, 35,* 353-387.

Erbaugh, Mary S. (2002). Classifiers are for specification: Complementary functions for sortal and general classifiers in Cantonese and Mandarin. *Cahiers de Linguistique-Asie Orientate, 31,* 33-69.

Erbaugh, Mary S. (2004). Chinese classifiers: Their use and acquisition. In Ping Li, Li Hai Tan, Elizabeth Bates, & Ovid J. L. Tzeng (Eds.), *Handbook of East Asian psycholinguistics: Vol. 1. Chinese* (pp. 39-51). Cambridge, UK: Cambridge University Press.

Fang, Fuxi. (1985). An experiment on the use of classifiers by 4 to 6 year olds. *Acta Psychological Sinica, 17,* 384-392.

Gao, Helena Hong. (2009). *Cognitive Barriers to the Learning of Chinese Noun Classifiers by Native Swedish Speakers.* Paper presented at Second Conference of the Swedish Association for Language and Cognition, Stockholm University.

Grüter, Theres, Lew-Williams, Casey, & Fernald, Anne. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*(2), 191-215.

Hopp, Holger. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*(1), 33-56.

Hu, Qian. (1993). Overextension of animacy in Chinese classifier acquisition. In Eve Clark (Ed.), *The proceedings of the twenty-fifth annual child language research forum* (pp. 127-136). Stanford, CA: Center for the Study of Language and Information.

Huettig, Falk, Chen, Jidong, Bowerman, Melissa, & Majid, Asifa. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture, 10,* 39-58.

Klein, Natalie M., Carlson Greg N., Li, Renjie, Jaeger T. Florian, & Tanenhaus, Michael K. (2012). Classifying and massifying incrementally in Chinese language comprehension. In Diane Massam (Ed.), *Count and mass across languages* (pp. 261-282). Oxford, UK: Oxford University Press.

Lew-Williams, Casey, & Fernald, Anne. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18,* 193-198.

Lew-Williams, Casey, & Fernald, Anne. (2009). Fluency in using morphosyntactic cues to establish reference: How do native and non-native speakers differ? In Jane Chandlee, Michelle Franchini, Sandy Lord & Gudrun-Marion Rheiner (Eds.), *Proceedings of the*

33rd Annual Boston University Conference on Language Development (pp. 290-301). Somerville, MA: Cascadilla Press.

Lew-Williams, Casey, & Fernald, Anne. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*, 447-464.

Liang, Neal Szu-Yen. (2008). The acquisition of Chinese shape classifiers by L2 adult learners. In Majorie K. M. Chan & Hana Kang (Eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics* (NACCL-20) (pp. 309-326). Columbus, OH: Ohio State University.

Liu, Yuehua, Yao, Tao-chung, Shi, Yaohua, Bi, Nyan-Ping, & Ge, Liangyan. (2009). *Integrated Chinese: Level 2, part 1 (simplified and traditional character) textbook (Chinese edition)*. Boston, MA: Cheng & Tsui.

Lyons, John. (1977). *Semantics*. Cambridge, UK: Cambridge University Press.

Matthews, Stephen, & Yip, Virginia. (2011). Cantonese: A comprehensive grammar (2nd ed.). London, UK: Routledge.

Ministry of Education of the People's Republic of China. (2013). *通用规范汉字表* [Table of General Standard Chinese Characters].

Ning, Cynthia Y., & Montanaro, John S. (2011). *Encounters: Chinese language and culture, Student Book 1*. Yale, NJ: Yale University Press.

Polio, Charlene. (1994). Non-native speakers' use of nominal classifiers in Mandarin Chinese. *Journal of the Chinese Language Teachers Association*, *29*(3), 51-66.

Tsang, Cara, & Chambers, Craig G. (2011). Appearances aren't everything: Shape classifiers and referential processing in Cantonese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1065-1080.