# *How Does Adult Learners' L1 Interact With Word Frequency in the Error Rates and Patterns of L2 Classifier Use: A Cross-Linguistic Comparison*

Kun Yu, The Hong Kong University of Science and Technology, Hong Kong SAR
Yin-To Chui, The Hong Kong University of Science and Technology, Hong Kong SAR

The Asian Conference on Education 2023
Official Conference Proceedings

**Abstract**
Previous literature has reported mixed findings on the effect of L1 (classifier languages vs non-classifier languages) on the acquisition of L2 classifiers. This study aims to investigate whether any effect of L1 may be modulated by the word frequency of the target classifiers. Chinese classifiers were categorized into pre-established frequency bands A to C in descending order of word frequency, and learner data was extracted from the HSK Dynamic Composition Corpus containing sentences with classifiers written by L1-Korean and L1-English learners, with a total of 5248 sentences analyzed. Two native Mandarin speakers reported the error rates, and errors were also further categorized into four different types (i.e. misuse, omission, overuse, misorder) to investigate any differences in error patterns across different L1 groups. Results showed a significant interaction between L1 and classifier word frequency, with comparably low error rates for both L1 groups on high-frequency classifiers, but a significantly higher error rate for L1-Korean learners on low-frequency classifiers. Aligned with exemplar theory, more frequently encountered classifiers may have a relatively robust representation in the mental lexicon for both L1 groups, as opposed to the less stable representation for less-frequently encountered classifiers which are more prone to transfer effects from L1 to L2.

Keywords: Classifier, Word Frequency, L2 Acquisition, Cross-Linguistic Comparison

## Introduction

Previous literature has reported mixed findings on the effect of L1 (classifier languages vs non-classifier languages) on the acquisition of L2 classifiers. This study aims to investigate whether any effect of L1 may be modulated by the word frequency of the target classifiers.

Classifiers can be defined as "morphemes that classify and quantify nouns according to semantic criteria" (Senft, 2000). They are common in East Asian languages such as Mandarin Chinese, Japanese, and Korean, but are absent in most European languages. One example of a Mandarin classifier is *tiao2*, a mandatory grammatical unit between a number/determiner and a head noun that describes the noun as being a thin one-dimensional object (e.g. *yi4 tiao2 xian4 [a-CL-line]*). Another noteworthy classifier in Mandarin Chinese is *ge4*, a generic classifier that can be used with 40% of common nouns and which is often overused by both L1 and L2 speakers of Chinese (e.g. Zhang, 2007).

## L1 Effects on the Acquisition of L2 Classifiers

An early hypothesis was that L1-speakers of a classifier language (e.g. Japanese, Korean) may learn L2 classifiers better than those of a non-classifier language (e.g. most European languages) by means of positive transfer. However, studies yielded mixed findings (e.g. Polio, 1994; Paul & Grüter, 2016). For example, Polio (1994) tested the use of Chinese classifiers by L1-English and L1-Japanese intermediate learners of Chinese. In their study, production data was elicited where participants narrated a story after watching the Pear Film. Results showed that although both groups of learners did not avoid using classifiers generally, no L1 effect was found.

Later studies have found that the existence of an L1 effect might depend on additional factors, such as learner proficiency of the target language and the actual linguistic task used. In terms of proficiency, Liang (2008) tested L1-English vs L1-Korean learners' use of eight Chinese shape classifiers on an object-classifier mapping task (e.g. *tiao2* for a one-dimensional object, *zhang1* for a two-dimensional object) across three proficiency levels (i.e. advanced, intermediate, beginner), and found that both Korean and English groups performed similarly at the beginner and advanced level, and a group difference was found only for the intermediate level, suggesting that any effect of L1 might be sensitive to where the learners are on their developmental path of L2 acquisition. In terms of task, Zhang & Gnevsheva (2022) compared L1-Japanese learners (classifier language) against L1-English and L1-Arabic learners (non-classifier languages) at the intermediate level, using three written tasks based around a description of the same picture – short composition, free-cloze, and multiple-choice cloze. The short composition was aimed at capturing naturalistic use of classifiers, while the two cloze tests captured compulsory application of classifiers in a more constrained manner. Results showed that L1-Arabic learners performed significantly worse than L1-Japanese learners only in the free-cloze test but not the composition task, suggesting that the effect of L1 might be sensitive to the actual task used for testing, and that learners might have used test-taking strategies (e.g. avoiding use of difficult classifiers) in certain tasks (e.g. composition) that led to an inflation of accuracy score. Both of the above results advanced our understanding that the effect of L1 (classifier vs non-classifier language) might be modulated by previously under-examined factors.

Importantly, Zhang & Gnevsheva (2022) also tested a hypothesis as to whether the types of classifiers themselves might modulate the effect of L1. Based on a typology by Gao and Malt

(2009), classifiers were categorized based on association with their head nouns – shape, animate, inanimate, and concept. For example, animate classifiers categorize animate nouns according to size and shape, and inanimate classifiers categorize inanimate nouns according to their function. Although results did not reach significance, a numerical trend was found where L1-English and L1-Arabic speakers performed disproportionately worse on concept classifiers than other types of classifiers.

**Frequency Effects, and an Exemplar Account of Lexical Acquisition**

As far as classifier type – as well as its potential interaction with L1 – is concerned, we put forward the hypothesis that the *word frequency* of the classifiers could be a potential interacting factor instead. At least two pieces of evidence may substantiate this claim. First is the observation that the typology of classifiers (based on association with head nouns) might be confounded with word frequency. For example, concept classifiers such as *cheng2* (associated with the concept of probability; e.g. *yi4 cheng2 sheng4 suan4 [1-CL-probability of winning]* often occupy the low frequency bands of Chinese word usage, in a classification of Chinese word/character frequency data officially published by the Ministry of Education of the People's Republic of China (Hanban, 2001).

Second, the construct of word frequency is closely tied to the exemplar theory of language processing that has become increasingly influential over the past two decades (Ambridge, 2020). The exemplar theory was proposed in response to the inadequacies of traditional abstractionist models that dominated the field in the 20th century, and has received support from fields encompassing phonological, lexical, and syntactic processing. It posits that instead of abstracting basic phonological/lexical/phonological units, learners store all instances of linguistic exposure in their episodic memory which subsequently alter their own usage of different sounds/words/grammatic structures. For example, listeners may not just abstract stable phonological word forms in the mental lexicon because recently heard exemplars will alter their own production and usage (Pierrehumbert, 2001). Crucially, in lexical acquisition, it has been found that novel words that are phonological neighbors of high-frequency words (i.e. more exemplars) were accessed faster than novel words that are neighbors of low-frequency words (Vitevitch et al., 2014). One possibility proposed by the authors of that study was that neighbors of high-frequency words have more exemplars to "latch on" to in memory. It is thus reasonable to assume that high-frequency words have a more stable representation in the lexicon *by means of* all the exemplars stored in memory.

Thus, for our current study, we suspect that L1 effects on the acquisition of Chinese classifiers might be similarly modulated by the word frequency of target classifiers, as high-frequency classifiers have a more stable representation across learners of different L1s and may have a low error rate that is comparable across L1 groups, while the representations for low-frequency classifiers are more transient and an L1 group difference may emerge.

**Methods**

Chinese classifiers were categorized into pre-established frequency bands A to C in descending order of word frequency, following the official statistics published under the Ministry of Education of the People's Republic of China (Hanban, 2001). This resulted in a total of 49 Band A classifiers, 42 Band B classifiers, and 20 Band C classifiers. Example classifiers for each Band are listed in Table 1 below.

| Band A | Band B | Band C |
|---|---|---|
| *ben3*[本] | *zhu1* [株] | *juan3* [卷] |
| *feng1*[封] | *shou3* [首] | *ding3* [顶] |
| *tiao2*[条] | *fen4* [份] | *zhan3* [盏] |
| *zhang1*[张] | *duo3* [朵] | *xiang1* [箱] |
| *zhi1*[支] | *bu4* [部] | *cheng2* [成] |

Table 1: Example classifiers from each frequency band

Learner data was extracted from the HSK Dynamic Composition Corpus which contains over 10,000 HSK composition papers of foreign exam takers. We extracted sentences with classifiers written by L1-Korean and L1-English (aggregate data from US, UK, Australia, Canada) learners. Because of the special status of *ge4* as a generic classifier that can be associated with over 40% common nouns and tends to be overused even by L1-Chinese speakers, we opted to omit *ge4* in our search.

Two native Mandarin speakers reported the error rates. Accuracy was binary-coded (1 for correct, 0 for incorrect). A sentence was considered to have a classifier error if it conformed to one of four conditions: 1. Misuse, where an incorrect classifier was substituted for the correct one; 2. Omission, where a classifier was omitted when it was required in the sentence; 3. Overuse, where a classifier was added when it should not; 4. Misorder, where a correct classifier was used but placed in the wrong position in the phrase. Results yielded an 85% inter-rater reliability. Only target sentences for which the raters agreed on the error status were included in subsequent analysis. This yielded a total of 5248 target sentences.

**Results and Discussion**

Figure 1 shows the overall group results. General observations reveal comparable error rates for L1-English and L1-Korean learners on high-frequency classifiers, but perhaps unexpectedly, a *higher* error rate for L1-Korean learners than L1-English learners on low-frequency classifiers.

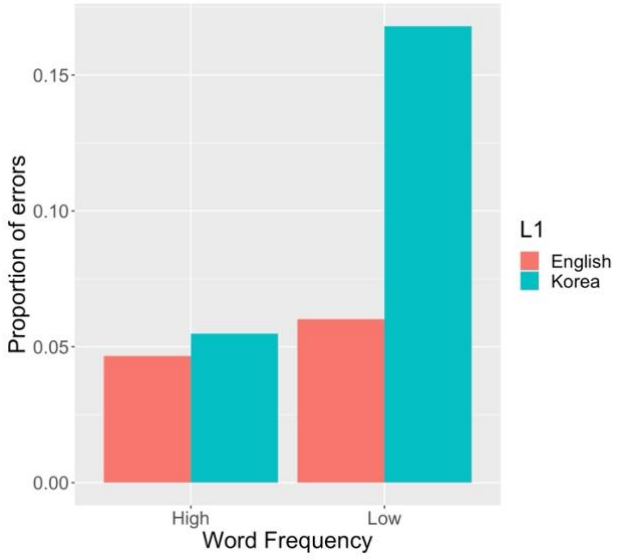


Figure 1: Proportion of overall classifier errors by L1-English and L1-Korean learners

A 2 (L1: Korean vs English) x 2 (Frequency: High vs Low) ANOVA conducted on the error rates confirmed this trend. We found a significant main effect of L1 [$F(1,5244) = 6.62$, $p = .01$], with L1-Korean learners making more errors than L1-English learners. Expectedly, we also found a significant main effect of frequency [$F(1,5244) = 98.12$, $p < .001$], with more errors on low-frequency classifiers than high-frequency classifiers. Crucially, the interaction between L1 and Frequency was significant [$F(1,5244) = 14.66$, $p < .001$], suggesting that the L1 effect was different for high-frequency vs low-frequency classifiers. Separate analyses showed an L1 effect for low-frequency classifiers only [$t(285.1) = -4.24$, $p < .001$] and not for high-frequency classifiers [$t(871.5) = -0.89$, $p = .375$], with L1-Korean learners performing *worse* than L1-English learners on low-frequency classifiers.

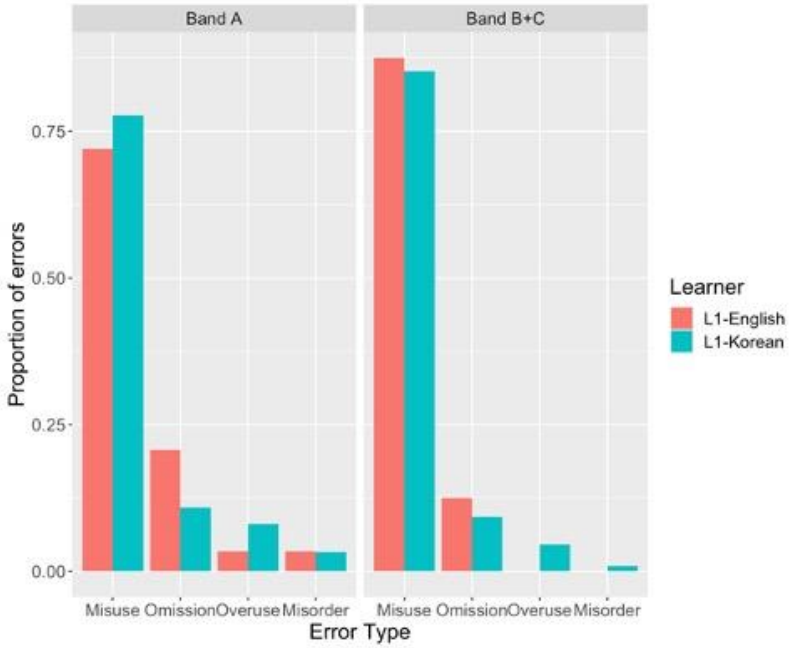


Figure 2: Classifier error patterns for L1-English and L1-Korean speakers

Figure 2 further broke down the errors into the four error types (i.e. misuse, omission, overuse, misorder). Results seem to reveal a slight tendency to *omit* classifiers for L1-English learners, as opposed to a tendency to *overuse* classifiers for L1-Korean learners.

Overall, the results confirm our prediction that word frequency of classifiers plays a role in modulating the effect of L1 (classifier vs non-classifier language) in the acquisition of L2 classifiers. In particular, both L1 groups maintained a relatively low and comparable error rate for high-frequency classifiers, suggesting that the high-frequency classifiers may have occupied a more stable representation in the mental lexicon of both L1 groups and are more resilient to errors. For low-frequency classifiers, somewhat unexpectedly, L1-Korean learners performed worse than L1-English learners. This means that the interaction could *not* result from a positive transfer for L1-Korean learners and they did not manage to leverage existing knowledge about classifiers to acquire L2 classifiers that are less frequently encountered and thus may have a less robust representation. One possibility is that perhaps due to a lack of classifiers in their L1 inventory, L1-English learners might be more cautious in applying them in written sentences (indeed, we found a smaller number of sentences in L1-English data than L1-Korean data), while L1-Korean learners may be free from this constraint given the familiarity of classifiers in their L1 system, using them even erroneously. A second possibility, as revealed in Figure 2, is that the transfer from L1-Korean classifier system to

L1-Chinese classifier system may not be always positive. Since Figure 2 shows a tendency of L1-Korean learners to overuse classifiers, it may be the case that the L1-Korean learners have erroneously used classifiers that are present in Korean but not in Chinese.

## Conclusion

The present study showed that an L1 effect (classifier language vs non-classifier language) exists for the acquisition of L2 classifiers, but is modulated by the word frequency of the target classifiers, with an L1 group difference emerging for low-frequency classifiers only. While the present study proposed explanations for the counter-intuitive trend of L1-Korean learners performing worse than L1-English speakers in terms of constraints of usage, or specific error patterns like overuse of L2 classifiers, further empirical research is needed to compare different classifier-L1s (e.g. L1-Japanese vs L1-Korean) in the acquisition of L2-classifiers across different word frequencies in order to confirm whether the transfer effects are universal to classifier languages or specific to Korean.

**References**


Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language, 40*(5-6), 509-559.

Gao, M. Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes, 24*(7-8), 1124-1179.

Liang, N. S. Y. (2008). The acquisition of Chinese shape classifiers by L2 adult learners. In *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)* (Vol. 1, pp. 309-326).

Paul, Jing Z. & Theres Grüter. (2016). Blocking effects in the learning of Chinese classifiers. *Language Learning 66*(4). 972–999.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. *Typological studies in language, 45*, 137-158.

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

Senft, G. (Ed.). (2000). *Systems of nominal classification* (Vol. 4). Cambridge University Press.

Vitevitch, M. S., Storkel, H. L., Francisco, A. C., Evans, K. J., & Goldstein, R. (2014). The influence of known-word frequency on the acquisition of new neighbours in adults: Evidence for exemplar representations in word learning. *Language, cognition and neuroscience, 29*(10), 1311-1316.

Zhang, H. (2007). Numeral classifiers in mandarin Chinese. *Journal of East Asian Linguistics, 16*(1), 43-59.

Zhang, J., & Gnevsheva, K. (2022). The effects of L1, task, and classifier type in Chinese-L2 learners' use of classifiers. *Chinese as a Second Language Research, 11*(1), 33-59.

国家汉语水平考试委员会办公室考试中心. (2001). *汉语水平词汇与汉字等级大纲.*