Reading and Writing
https://doi.org/10.1007/s11145-020-10058-w

Check for updates

Word-meaning inference in L2 Chinese: an interactive effect of learners' linguistic knowledge and words' semantic transparency

Tianxu Chen[1] · Keiko Koda[2] · Seth Wiener[2]

Abstract

Recent second language (L2) reading studies have examined how characteristics of an individual learner and word-specific linguistic properties jointly predict successful L2 word-meaning inference, i.e., a learner's ability to guess the meaning of an unknown word. Semantic transparency is one such word-specific property. Each constituent morpheme in a compound word contributes differently to the word's overall meaning. As a result, words vary in the degree to which they are transparent or opaque. The ability to infer the meanings of these compounds is associated with broad linguistic knowledge that varies greatly across learners. This study explores how L2 linguistic knowledge as an individual learner factor and semantic transparency as a word-specific property interact to affect L2 learners' word-meaning inference in Chinese. Two hundred and twelve adult L2 learners of Chinese in China completed six paper–pencil tasks, which measured participants' short-term memory, linguistic knowledge, and meaning inference of compound Chinese words that varied in their semantic transparency. The results showed that linguistic knowledge interacted with semantic transparency in word-meaning inference: morpheme knowledge and grammatical knowledge were significant contributors of word-meaning inference only for transparent words when L2 learners' short-term memory was controlled. These findings underscore how reading in an L2 involves the interplay between the learners' characteristics (e.g., linguistic knowledge) and language-specific properties (e.g., semantic transparency).

Keywords Linguistic knowledge · L2 Chinese · Semantic transparency · Word-meaning inference

Word knowledge in a second language (L2) plays an important role in L2 reading development (e.g., August, Carlo, Dressler, & Snow, 2005) and successful adult

✉ Tianxu Chen
tianxuc@muc.edu.cn

Extended author information available on the last page of the article

Published online: 20 May 2020

L2 acquisition (e.g., Nassaji, 2003). The reciprocity between word knowledge and reading comprehension has also been presumed in recent studies (e.g., Zhang & Koda, 2018). Yet, word knowledge is often developed incidentally through word-meaning inference in which learners make appropriate predictions of an unknown word's meaning using morphemic knowledge and morphological analysis (e.g., Koda, 2005; Nagy & Anderson, 1984; Wesche & Paribakht, 2010). This inference ability requires readers to integrate prior knowledge outside the text (e.g., pragmatic and sociocultural information) with linguistic information encoded in print (i.e., the function and forms of specific linguistic units at morpheme, word, and sentence levels; Koda, 2015). Thus, L2 learners' linguistic knowledge plays a critical role in word-meaning inference: learners with less linguistic knowledge assign greater cognitive attention to morphemes within unknown words, which limits learners' cognitive capacity to analyze additional information, such as morphological structures (e.g., Hamada, 2014; Koda & Miller, 2018; Reder, Liu, Keinath, & Popov, 2016; Zhang, Lin, Wei, & Anderson, 2014).

In addition to learners' linguistic knowledge, word-specific factors (e.g., lexical frequency, morphological family size, and semantic transparency) also affect the way that learners acquire new L2 words: some words are easier to acquire and infer the meanings of, whereas other words are more difficult (Balota, 1994; Dronjic, 2011; Frantzen, 2003). Among these word-specific factors, semantic transparency—i.e., the degree to which the meaning of each constituent morpheme in a compound word contributes to the word's overall meaning; Wang, Lin, & Gao, (2010)—is important when it comes to the activation of meanings of constituent morphemes (e.g., Libben, Gibson, Yoon, & Sandra, 2003). Importantly, the notion of semantic transparency is not dichotomous but varies along a continuum from semantically transparent to opaque within the same processing system. Whereas "classroom" is more transparent because both "class" and "room" contribute to the meaning of "classroom," "ragtime" is more opaque because neither the meaning of "rag" nor "time" contributes to the meaning of "ragtime" (see Mok, 2009).

Thus, when a reader encounters an unknown L2 word, at least two factors are at play: learner-specific linguistic knowledge of the target language and word-specific properties of the target word. However, little is known about how word-specific properties interact with individual linguistic abilities to facilitate word-meaning inference. This study uses Mandarin Chinese (hereafter "Chinese") to explore how these two factors interact to affect L2 reading skills.

Chinese serves as an ideal language to manipulate semantic transparency. Unlike most European languages, Chinese forms words primarily by compounding two monosyllabic morphemes (Kuo & Anderson, 2006; Taylor & Taylor, 2014). Packard (2000) categorizes Chinese morphemes into four types based on their boundedness (free or bound) and linguistic role (function or content): free-function ('function word'), free-content ('root word'), bound-function ('affix'), bound-content ('bound root'). This yields four types of complex words in Chinese: a compound word containing two root words (e.g., *lan2tian1*蓝天,[1] "blue" + "sky" → "blue sky"), a bound

[1] This manuscript uses Pinyin romanization, which indicates the spoken syllable and lexical tone.

root word containing one or two bound roots (e.g., *wei3ren2* 伟人, "great" + "person" → "great man"; *yu3jing4* 语境, "language" + "condition" → "context"), a derived word containing a bound root or root word plus a word-forming affix (e.g., *bi3zhe3* 笔者, "pen" + "a word-forming affix" → "author"), and a grammatical word containing a root word plus a grammatical affix (e.g., *wo3men2*我们, "I" + "a grammatical affix" → "we").

The formation of Chinese "compounding" words (i.e., compound words and bound root words) typically includes five constructions: coordinate (e.g., *zi3nü3* 子女, "son" + "daughter" → "offspring"), verb-object (e.g., *dai4ye4* 待业, "wait for" + "job" → "wait for employment"), verb-complementary (e.g., *jian3shao3* 减少, "reduce" + "few" → "reduce-[to loss]"), subject-predicate (e.g., *di4zhen4* 地震, "earth" + "shake" → "earth-quake"), and subordinate (e.g., *dian4nao3* 电脑, "electricity" + "brain" → "computer"). Subordinate constructions are the most common in modern Chinese (Yuan & Huang, 1998) and typically contain a modifier-head morphemic combination. As these previous examples have highlighted, there is a gradient range in which each morpheme contributes to a Chinese compound word's meaning.

In this study we take advantage of this naturally occurring range of transparency among Chinese compound words. We test a large sample of adult L2 Chinese learners engaged in a study abroad experience in China using a variety of tasks, including measures of short-term memory, linguistic knowledge, and word-meaning inference. We test how individual L2 learner factors and word-specific semantic transparency interact to affect L2 learners' word-meaning inference in Chinese. This allows us to explore whether learners with greater linguistic knowledge better infer the meaning of all unknown L2 words or only those unknown words that are more transparent.

Literature review

Learner-specific linguistic knowledge and word-meaning inference

Not surprisingly, learners' L2 linguistic knowledge affects word-meaning inference (Jeon & Yamashita, 2014): learners with greater linguistic knowledge of units (e.g., morphemes, words and sentences) are better able to successfully infer the meanings of unknown words (e.g., Chen, 2018, 2019b; Koda & Miller, 2018). There are at least three reasons for this finding.

First, morpheme knowledge plays an important role in unlocking the meaning of unfamiliar morphologically complex words. Logically, the ability to make use of prior morphemic information stored in long-term memory allows L2 learners to efficiently extract partial information from familiar morphemes and construct the meaning of a newly encountered word (Nagy & Scott, 2000). This finding is supported by extensive L2 English research showing the role of morpheme-related knowledge in word-meaning inference (e.g., De Bot, Paribakht, & Wesche, 1997; Parel, 2004; Paribakht & Wesche, 1999; Wesche & Paribakht, 2010). Crucial to the present study, this type of linguistic knowledge varies greatly across learners. For example, Parel (2004) tested 302 high school beginning-level ESL learners in Canada on a variety of reading tasks. Parel found that morpheme knowledge facilitated learners' word-meaning inference. Moreover, morpheme knowledge helped some learners overcome their limited vocabulary knowledge and implement appropriate lexical inferencing strategies.

Second, vocabulary knowledge also plays an important role in L2 word-meaning inference (e.g., Hamada, 2009; Ke & Koda, 2017; Nassaji, 2004; Prior, Goldina, Shany, Geva, & Katzir, 2014; Zhang & Koda, 2018). For example, Nassaji (2004) investigated 21 ESL learners with different language backgrounds, and found that L2 learners with greater vocabulary knowledge used more strategies in word-meaning inference than learners with reduced vocabulary knowledge. Nassaji also found that learners' vocabulary knowledge contributed to effective word-meaning inference. Prior et al. (2014) examined the predictive role of vocabulary knowledge in L2 word-meaning inference among 53 L1 Russian-speaking high school students in Israel. The authors' results indicated that Hebrew vocabulary knowledge was strongly related to L2 word-meaning inference ability, which in turn affected L2 reading comprehension. With respect to our target language, Chinese, previous research from both English-speaking learners of L2 Chinese (Ke & Koda, 2017) and American learners who speak Chinese as a heritage language (Zhang & Koda, 2018) have demonstrated that vocabulary knowledge as a significant contributor of L2 word-meaning inference. In other words, L2 learners with greater vocabulary knowledge better inferred the meanings of unknown words than those learners with reduced vocabulary knowledge.

Third, grammatical knowledge also plays a role in L2 word-meaning inference (e.g., Haastrup, 1991; Kaivanpanah & Alavi, 2008; Paribakht & Wesche, 1999). Kaivanpanah and Alavi (2008) found that grammatical knowledge affected L2 word-meaning inference in both simple and complex texts based on 102 EFL learners' reading tests and nine EFL learners' introspective think-aloud data. Kaivanpanah and Alavi claimed that L2 learners with greater grammatical knowledge were able to integrate information from different sources by making extensive use of L2 linguistic knowledge sources. In contrast, L2 learners with limited grammatical knowledge mostly relied on word-to-word translation during text-meaning construction.

In sum, an L2 learner's linguistic knowledge, which involves sub-lexical, vocabulary, and grammatical knowledge, affects how the meaning of an unknown L2 word is inferred. This linguistic knowledge is highly varied given each learner's individual experience with the language. In this study, we operationalize linguistic knowledge as morpheme knowledge, vocabulary knowledge, and grammatical knowledge. We use this measure to examine a wide range of learners immersed in their L2.

Word-specific semantic transparency and word-meaning inference

Semantic transparency directly influences how people recognize morphologically complex words because it determines the extent to which a compound word can be processed through decomposition (Libben et al., 2003). Previous studies on L2 Chinese readers have shown how semantic transparency affects a learner's word-meaning inference ability (e.g., Xu, 2014; Zhang & Zeng, 2010). For example, Xu (2014) tested 20 L1 English-L2 Chinese learners and found the learners better inferred the meaning for transparent words than opaque words with or without contextual information.

Two factors jointly determine semantic transparency: semantic extension and semantic ambiguity (Cacciari & Glucksberg, 1994). Semantic extension refers to a sense of a constituent morpheme that is extended from its literal and concrete meaning. By definition, extended meanings are more abstract and, thus, more figurative than concrete meanings. In this study, semantic extension is indexed by classifying compound words into four categories: (1) fully transparent (concrete modifier and concrete head; e.g., *che1piao4* 车票 vehicle + ticket → bus ticket), (2) semi-transparent (extended modifier and concrete head; e.g., *hong2ren2* 红人 red + person → popular people), (3) semi-opaque (concrete modifier and extended head; e.g., *min2feng1* 民风 folk + wind → folk custom), and (4) fully opaque (extended modifier and extend head; e.g., *leng3men2* 冷门 cold + door → unpopular) words (see Chen, 2019b). Note that the head of each compound is the stem that determines the semantic category of that word. As a result, the difference between semi-transparent and semi-opaque words involves whether the morpheme as a morphological head conveys an extended meaning.

Semantic ambiguity refers to the number of semantically different senses a morpheme carries (including concrete meanings and extended meanings). For instance, the morpheme *yan3* 眼 (eye) typically has three senses for native Chinese speakers. It represents "eye" in *yan3jing4* 眼镜 eye + glass → eyeglass, "hole" in *qiang1yan3* 枪眼 gun + hole → bullet hole, and as a measure word in *yi4yan3jing3* 一眼井 one + [a measure word] + well → a well. It should be noted that whether the morpheme has a single meaning or multiple meanings in this study is determined based on the morpheme knowledge of participants. For example, as mentioned above, *yan3* 眼 (eye) is a polysemantic morpheme for native Chinese speakers because this morpheme represents multiple meanings. However, *yan3* 眼 (eye) in this study is treated as a monosemantic morpheme because the participants had learned only one meaning of the morpheme used in two words *yan3jing4* 眼镜 (eyeglass) and *yan3jing0* 眼睛 (eyes) at the time of data collection. Similarly, a polysemantic morpheme is operationalized as a morpheme that has multiple meanings within the instructional context. Therefore, semantic ambiguity is also indexed in this study by categorizing compound words into four groups: monosemantic (single meaning modifier + single meaning head), semi-monosemantic (multiple meaning modifier + single meaning head), semi-polysemantic (single meaning modifier + multiple meaning head), and polysemantic (multiple meaning modifier + multiple meaning head). The difference between semi-monosemantic and semi-polysemantic words is whether the morpheme as a morphological head conveys a single meaning.

The present study

Linguistic knowledge and semantic transparency thus affect the extent and the way in which learners successfully infer meanings of unknown words. To date, however, it remains unclear whether these properties interact such that learners with varying levels of linguistic knowledge differ in their ability to infer the meanings of a wide range of semantically transparent L2 words. To explore this issue, the present study addressed two questions: (1) Does linguistic knowledge affect meaning inference of words having varying semantic transparency? (2) If yes, how do higher and lower proficiency learners differ in the way they infer the meaning of semantically transparent and opaque words?

Method

Participants

A total of 212 adults studying Chinese in China (taken from Chen, 2019c) took part in the study. Self-reported participant background information revealed that participants' L1 included 133 L1 English, 26 L1 Russian, 19 L1 Spanish, 12 L1 French, eight L1 Italian, six L1 German, five L1 Portuguese, two L1 Danish, and one L1 Dutch. See Table 1 for participant information. None of the participants were heritage language learners and none had previously studied Japanese, a language involving character-based writing components. At the time of data collection, all participants had studied Chinese in China for at least one semester but no more than four semesters. During each semester, participants received 20–24 h of classroom instruction per week. We estimate that at the time of testing our participants could recognize at least 1000 simplified Chinese characters and 1600 words based on the vocabulary lists from the participants' textbooks.

Measures

All participants completed six paper–pencil tasks, including one short-term memory task, three linguistic knowledge tasks (i.e., morpheme knowledge, vocabulary knowledge, and grammatical knowledge), and two word-meaning inference tasks (i.e., multiple choice and translation). The latter five tasks' scores were used to estimate a participant's corresponding L2 Chinese reading abilities.

Short-term memory task

Short-term memory (e.g., Baddeley, 1992) was measured using a digit span task (e.g., Conway et al., 2005). Participants were shown 14 numbers that ranged from three to nine digits. Each number appeared on a computer screen for 1 s then disappeared. Participants were then asked to write down on paper the number they saw. Numbers were separated by an 8-second interval. Each correct response was calculated as one point.

Linguistic knowledge tasks

Morpheme knowledge was measured using a select checklist (e.g., Chen, 2019a; Anderson & Freebody, 1983; Ku & Anderson, 2003; Zhang, 2015). Participants

Table 1 Participant information

Mean age	Gender	Mean study abroad length in China	Mean outside language practice with L1 Chinese speakers
23.07 years old (*SD* = 4.50)	104 males 108 females	11.43 months (*SD* = 6.56)	14.72 h/week (*SD* = 6.56)

were shown a printed Chinese character and asked if they knew the character's meaning or not by indicating on paper. Participants were shown 75 real characters, which were selected from a character corpus of L2 Chinese learners (Xing, 2018) in terms of their frequency and 25 nonce characters, which were taken from Chen (2019a). The real characters included 50 free characters (i.e., characters can stand alone as words), and 25 bound characters (i.e., characters cannot stand alone as words). Zhang's (2015) scoring formula was adopted: score $=(h-f)/(1-f)$ ($h=$ real hits/total real characters; $f=$ false alarm/total nonce characters). Real characters selected as known were coded as "real hits" and nonce characters selected as known were coded as "false alarms".

Vocabulary knowledge was measured using a select checklist format (Chen, 2019a). Participants were shown 100 disyllabic words (75 real words and 25 pseudowords), in which 75 real words were randomly selected from the HSK—a standardized test measuring Chinese proficiency—vocabulary list (Hanban, 2010), including five words at the 1st level (the lowest difficulty level), five words at the 2nd level, 25 words at the 3rd level, 25 words at the 4th level, 10 words at the 5th level, and five words at the 6th level (the highest difficulty level). The difficulty levels were based on the frequency of word usage in day-to-day activities of L1 Chinese speakers. Participants were asked whether they knew the word's meaning by indicating on paper. Participants were explicitly told to choose "YES" if they knew the meaning of the entire word as opposed to only one or two characters within the word. Zhang's (2015) scoring formula was used: score $=(h-f)/(1-f)$.[2]

Grammatical knowledge was measured using two parts of the HSK test (Office of Chinese Language Council International, 2007). In the first part, participants were shown sentences with four blanks and asked where a grammatical construction was most appropriate. In the second part, participants were shown a sentence with one blank and asked to choose the most appropriate grammatical construction out of four options. Both parts involved 15 sentences for a total of 30 items. Each correct response was calculated as one point. The reliability of the task across all 30 items was acceptable (Cronbach's $\alpha=.73$).

Word-meaning inference tasks

A series of pilot tests were conducted to control the effects of word-specific factors on the word-meaning inference task. These word-specific variables included syllable length, word property, word construction, morpheme combinations, character complexity (i.e., numbers of character strokes), character frequency, word frequency, and semantic transparency. Specific item selection followed several steps. First, 1009 characters that appeared in the participants' textbooks *Road to*

[2] Cronbach's α was not applied to the morpheme knowledge and vocabulary knowledge stimuli due to the nonce characters and pseudowords involved, but the reliability of the checklist format has been established by previous studies (see Anderson & Freebody, 1983; Chen, 2018, 2019a). Additionally, this study used Zhang's (2015) scoring formula rather than d-prime because the stimuli did not contain an equal number of real and nonce/pseudoword targets.

Success and *Chinese in 10 Days* were selected. It was assumed that the participants would have learned these characters by the time of data collection. Second, 9762 low-frequency (less than five per million characters) two-character real words were identified in Cai and Brysbaert (2010). These words did not appear in the participants' textbooks but consisted of the 1009 learned characters; that is, learners were assumed to be familiar with the individual characters but not the compound words. From these possible target words, 106 nouns were selected based on their semantic extension and ambiguity. Next, 34 native Chinese speakers were asked to evaluate the semantic transparency of these 106 words using a four-point scale (1 = completely opaque, 4 = completely transparent) for each morpheme. A Pearson correlation coefficient showed that L1 learners' evaluation and semantic transparency (extension + ambiguity) were significantly and positively correlated with each other ($r = .61$, $p < .001$). Therefore, the semantic transparency scores (i.e., the composite scores of semantic extension and semantic ambiguity ranging from 3 to 8, see scoring details in "Appendix 1") were used to represent words' semantic transparency in this study. A follow-up cluster analysis was conducted to categorize the semantic transparency into two groups based on the semantic transparency scores. The Average Silhouette measure of cohesion and separation was 0.7, indicating a good cluster quality. The "transparent" group had 43 items (40.6%) and the "opaque" group included 63 items (59.4%). In the end, 20 items were selected as transparent words (e.g., *yu3yi1*雨衣, rain + coat → raincoat) and 20 as opaque words items (e.g., *xin1ji1* 心机, heart + machine → thinking and scheming) (see full stimuli and their characters in "Appendix 2"). All 40 items were two-character "N–N" nouns with a subordinate construction because this type of noun is the most common in Chinese (Yuan & Huang, 1998). The target items had moderate character complexity (Mean of stroke numbers = 7.32, $SD = 1.26$), high character frequency (Mean = 588.05/million characters, $SD = 433.04$; Cai & Brysbaert, 2010), but low word frequency (Mean = 0.78/million characters, $SD = 0.92$; Cai & Brysbaert, 2010). A series of two-sample t tests between the transparent and opaque word groups showed that there were no significant differences in word frequency $t\ (38) = -1.08$, $p = .29$ and character complexity, $t\ (38) = 1.59$, $p = .12$. The difference in character frequency, however, approached significance, $t\ (38) = -1.95$, $p = .06$. Importantly, the semantic transparency scores differed between the transparent words (mean = 7.50; $SD = 0.53$) and opaque words (mean = 4.10; $SD = 0.88$), $t\ (38) = 11.33$, $p < .001$. Thus, the only difference between items in the two groups was their level of semantic transparency.

These 40 words were used in two inference tasks. In the first task, a multiple-choice task, a two-character word (e.g., *chuan2piao4* 船票boat ticket) was shown along with four potential meanings in English. The participants were required to choose the correct translation from the options. In addition to the correct meaning, the options included one related to the meaning of the first morpheme (e.g., a ship's captain), one related to the meaning of the second morpheme (e.g., a receipt), and one unrelated distractor (e.g., a machine). The order of answers was randomized for each trial. Ten additional distractors trials (i.e., two-character verbs with different word constructions, including coordinate, verb-object, and verb-complementary) were included but not scored. Each correct response was calculated as one point, and the maximum score in this task was 40. The reliability of the task was good (Cronbach's $\alpha = .85$).

In the second task, a translation task, participants were shown a two-character word and asked to infer its meaning by writing down the corresponding English translation. Ten transparent items and 10 opaque items were randomly selected from target items in the multiple-choice task. Five distractors verbs were also included for a total of 25 items. Following Chen (2019c), the answers were evaluated by two L1 Chinese raters. If participants provided a correct answer or a translation that was synonymous with the correct answer, they received 2 points. If they provided multiple answers, but one of them was correct, they received 1 point (e.g., "boat ticket" and "vehicle" for 'boat ticket'). If participants provided a partially correct answer, they also received 1 point (e.g., "a kind of ticket" for 'boat ticket'). If a morpheme-related but word meaning-unrelated answer was provided, they received 0 points (e.g., "vehicle" for 'boat ticket'). Unrelated or missing answers received 0 points. The maximum score in this task was 40. The reliability of the task was acceptable (Cronbach's $\alpha = .77$).

Procedure

All measures were administered to the participants in group-sessions, with 10–20 participants per session. Participants first received IRB-approved informed consent forms and agreed to voluntarily participate in this study. Testing took place in a quiet classroom after participants' daily classes. The participants were allowed to take a short break during the test, if needed. In total, the testing session lasted approximately 60 min. Participants were first asked to complete the translation task and then finish the remaining tasks in a counterbalanced order. At the end of the test, the multiple-choice version of the word-meaning inference task was distributed to all participants. Upon completion of the whole test, the participants were paid for their time.

Results

A total of 203 participants were included in the statistical analysis because nine participants' results for at least one task were either above or below 2.5 *SD* from the overall mean. Table 2 lists the descriptive data, including numbers of items, mean accuracy, standard deviations (*SD*s), minimums (Min) and maximums (Max). A Pearson product-moment correlation was first run to assess the intercorrelations among all variables (Table 3).

The main results in Table 3 showed that (1) morpheme knowledge, vocabulary knowledge, and grammatical knowledge were significantly correlated with each other ($ps < .001$); (2) morpheme knowledge, vocabulary knowledge, and grammatical knowledge were significantly correlated with the meaning inference of transparent words in the multiple choice and translation tasks ($ps < .05$); (3) morpheme

Table 2 Summary of participants' results (N = 203)

Task	Numbers of Items	Mean	*SD*	Min	Max
Short-term memory	14	.67	.11	.43	1.00
Linguistic knowledge					
Morpheme knowledge	100	.80	.11	.52	.98
Vocabulary knowledge	100	.60	.21	.11	.92
Grammatical knowledge	30	.53	.14	.20	.90
Word-meaning inference					
Multiple choice	18 (transparent words)	.71	.15	.33	1.00
	19 (opaque words)	.34	.15	.05	.68
Translation	10 (transparent words)	.56	.18	.15	.94
	10 (opaque words)	.26	.19	.00	.72

Means, *SD*s, Minimum and Maximum pertain to the proportion of correct item selections in all the tasks except for the morpheme knowledge and vocabulary knowledge tasks, which followed Zhang's (2015) scoring formula. In the word-meaning inference task (multiple choice), two items in the transparent word group and one item in the opaque word group were excluded due to non-normality

Table 3 Correlations among variables in all tasks (N = 203)

	1.	2.	3.	4.	5.	6.	7.	8.
1. Short-term memory	–	.11	.04	−.01	.08	.06	.10	.03
2. Morpheme knowledge		–	.59***	.38***	.32***	.10	.19**	.06
3. Vocabulary knowledge			–	.53***	.35***	−.12	.16*	−.04
4. Grammatical knowledge				–	.33***	.08	.19**	.08
5. Multiple-choice Transparent					–	.16*	.31***	.19**
6. Multiple-choice Opaque						–	.11	.28***
7. Translation Transparent							–	.28***
8. Translation Opaque								–

*** < .001, ** < .01, * < .05

knowledge, vocabulary knowledge, and grammatical knowledge were not correlated with the meaning inference of opaque words in the multiple choice and translation tasks (*p*s > .05); and (4) there were significant correlations between learners' performance on certain word-meaning inference tasks for both transparent words ($p < .001$) and opaque words ($p < .001$). The results indicate that each facet of learners' linguistic knowledge and their word-meaning inferencing for transparent and opaque words had similar correlations.

Next, a hierarchical multiple regression was performed to determine the relative contributions of morpheme knowledge, vocabulary knowledge, and grammatical knowledge to the meaning inference of transparent words. The composite scores of both word-meaning inference tasks (multiple choice and translation represented as a mean percentage) were used to represent learners' word-meaning inferencing abilities. It should be noted that a null model for the meaning inference of opaque words

Table 4 Hierarchical regression analysis with the meaning inference of transparent words as the dependent variable (N = 203)

	R	R^2	R^2 Change	B	SE	$\hat{\beta}$	t	Sig.
Step 1	.109	.012						
Short-term memory				.128	.082	.109	1.552	.122
Step 2	.306	.094	.082***					
Short-term memory				.090	.080	.077	1.131	.259
Morpheme knowledge (Week 4)				.349	.082	.288	4.256	.001
Step 3	.321	.103	.009					
Short-term memory				.094	.079	.080	1.189	.236
Morpheme knowledge				.262	.102	.216	2.577	.011
Vocabulary knowledge				.078	.055	.120	1.438	.152
Step 4	.375	.141	.038**					
Short-term memory				.104	.078	.089	1.338	.183
Morpheme knowledge				.230	.100	.190	2.293	.023
Vocabulary knowledge				.010	.058	.015	.167	.867
Grammatical knowledge				.214	.073	.230	2.948	.004

*** < .001, ** < .01

confirmed the correlation findings: no significant correlation existed among any aspect of linguistic knowledge and the meaning inference of opaque words.

For transparent words, a four-stage hierarchical multiple regression was conducted with short-term memory as the control variable, morpheme knowledge, vocabulary knowledge, and grammatical knowledge as predictors, and word-meaning inference as the dependent variable. The Model entered the variables in the order ranking from relatively lower-level linguistic knowledge (i.e., morphemic units) to relatively higher-level linguistic knowledge (i.e., syntactic units). Specifically, short-term memory was entered during the first stage of the regression as a control variable. Morpheme knowledge was entered during the second stage. Vocabulary knowledge was entered during the third stage, and grammatical knowledge was entered during the fourth stage. The results are provided in Table 4. Morpheme knowledge was a significant contributor when L2 learners' short-term memory was controlled, F (2, 200) = 10.36, $p < .001$. Morpheme knowledge explained 8.2% of the variance. Morpheme knowledge and grammatical knowledge were also significant contributors (ps < .05) when three aspects of linguistic knowledge were involved in the model, F (4, 198) = 8.12, $p < .001$. Grammatical knowledge explained an additional 3.8% of the variance. However, vocabulary knowledge was not a significant contributor no matter whether morpheme knowledge was involved in the model ($p = .15$) or morpheme knowledge and grammatical knowledge were jointly involved in the model ($p = .87$).

To further explore the interaction between linguistic knowledge (as a learner property) and semantic transparency (as a word property), participants were grouped into higher (top 25%, $N = 51$) and lower (bottom 25%, $N = 51$) proficiency

Table 5 Scores of all tasks in lower- (N = 51) and higher-linguistic knowledge (N = 51) groups

Tasks	Linguistic knowledge groups	Mean	*SD*
Short-term memory	Lower	.67	.12
	Higher	.69	.10
Linguistic knowledge			
Morpheme knowledge	Lower	.67	.09
	Higher	.89	.04
Vocabulary knowledge	Lower	.35	.14
	Higher	.82	.05
Grammatical knowledge	Lower	.40	.09
	Higher	.69	.09
The meaning inference of transparent words			
Multiple choice	Lower	.63	.16
	Higher	.78	.09
Translation	Lower	.51	.17
	Higher	.63	.15
The meaning inference of opaque words			
Multiple choice	Lower	.33	.17
	Higher	.36	.15
Translation	Lower	.26	.21
	Higher	.26	.17

learners based on their composite morpheme knowledge, vocabulary knowledge, and grammatical knowledge scores. The means and *SDs* of the two groups are listed in Table 5. Two-sample *t* tests showed that (1) morpheme knowledge, vocabulary knowledge, and grammatical knowledge in the higher and lower groups significantly differed ($ps < .001$); (2) the higher and lower groups' word-meaning inference of transparent words in the multiple-choice task and the translation task significantly differed ($p < .001$ and $< .01$, respectively); and (3) the higher and lower groups' word-meaning inference of opaque words did not significantly differ ($p = .25$ and $= .94$, respectively).

Finally, a two-way ANCOVA was conducted using semantic transparency (two levels: transparent and opaque) as a within-subjects variable, linguistic knowledge (two levels: higher and lower) as a between-subjects variable, and short-term memory as a covariate on word-meaning inference. The results showed a significant effect of semantic transparency, F (1,99) = 23.35, $p < .001$, $\eta_p^2 = 0.19$, a significant effect of linguistic knowledge, F (1,99) = 9.31, $p < .01$, $\eta_p^2 = 0.09$, and a significant two-way interaction between semantic transparency and linguistic knowledge, F (1,99) = 15.77, $p < .001$, $\eta_p^2 = 0.14$ (see Fig. 1). Specifically, learners with higher linguistic knowledge were better able to infer the meanings of semantically transparent unknown words than learners with lower linguistic knowledge. In contrast, all learners, regardless of their linguistic knowledge, performed similarly when inferring the meanings of semantically opaque unknown words. Simple main effects analysis confirmed that a significant difference existed between the higher and lower linguistic

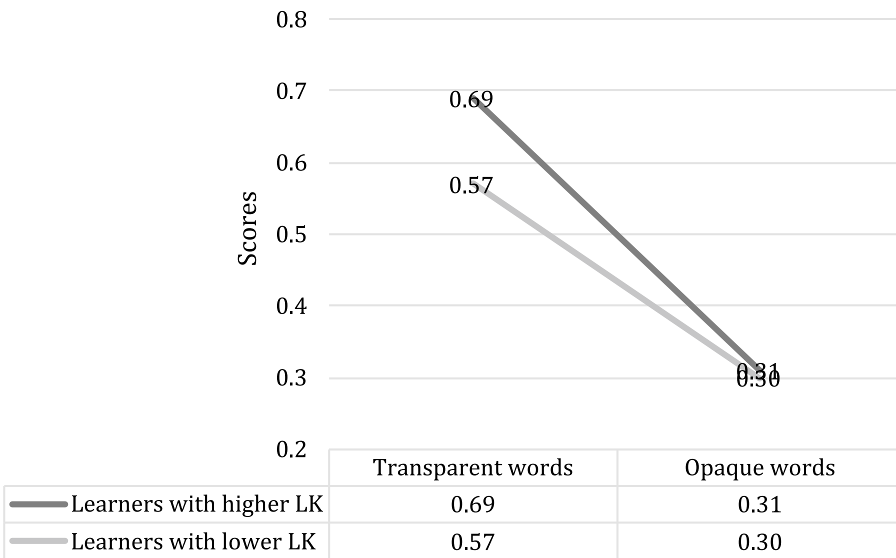

Fig. 1 The meaning inference of transparent and opaque words between higher and lower linguistic knowledge learners

knowledge learners when inferring the meanings of transparent words ($p < .001$), while there was no significant difference between the two groups when inferring the meanings of opaque words ($p = .60$).

Discussion

This study probed the relationship among semantic transparency, linguistic knowledge, and L2 learners' word-meaning inference abilities. Research Question One concerned whether linguistic knowledge affects meaning inference of words with varying semantic transparency. Research Question Two concerned how higher and lower proficiency learners differ in the way they infer the meaning of semantically transparent and opaque words. The results from over 200 L2 learners of Chinese indicated that learners more accurately inferred the meanings of transparent words than opaque words. Additionally, linguistic knowledge and semantic transparency interacted to affect L2 word-meaning inference. Specifically, L2 learners with greater linguistic knowledge more accurately inferred the meanings of semantically transparent words compared to learners with reduced linguistic knowledge. In contrast, all leaners, irrespective of their linguistic knowledge, struggled to accurately infer the meanings of semantically opaque words.

The finding that semantic transparency affects L2 word-meaning inference is consistent with previous L2 Chinese studies (e.g., Xu, 2014; Zhang & Zeng, 2010). After all, transparent words can be morphologically analyzed, while opaque words cannot. The meaning combinations of morphemes in transparent words are closely related to the whole word meaning. L2 learners can take advantage of transparent words' internal morphological information and the relationship between these morphemes (Chen, 2019b). Table 5 shows that even learners with relatively lower linguistic knowledge could, to some extent, infer the meanings of transparent words (63% and 51% accuracy in the multiple-choice and the translation tasks, respectively).

In contrast, the meaning combinations of morphemes in opaque words tend to be unanalyzable and less likely to be segmented. L2 learners cannot take advantage of internal morphological information or the functional relationships between these morphemes (e.g., Gan, 2008; Zhang & Zeng, 2010). Table 5 shows that L2 learners' mean accuracies for opaque words were only approximately 30% in both meaning-inference tasks. This finding is also in line with previous L1 studies that assert that semantically opaque words hamper word-internal morphological processing (e.g., Libben, 1998; Libben et al., 2003; Wang & Peng, 1999, 2000). Interestingly, learners with higher and lower linguistic knowledge demonstrated different patterns when inferring the meanings of constituent morphemes in opaque words. For example, when presented with the opaque word *xin1ji1* 心机 "heart" + "machine" → "thinking and scheming", learners with higher and lower linguistic knowledge had the same accuracy (both 17.6%) in the translation task. Yet, roughly half of the 42 participants in the higher linguistic knowledge group provided answers similar to the morpheme-by-morpheme meaning "a heart machine." In contrast, only a quarter of the 42 participants in the lower linguistic knowledge group responded with "heart" and "machine". This example seems to indicate that learners with higher linguistic knowledge are more likely to retrieve the meanings of the individual morphemes within unknown words. More qualitative evidence is needed in the future to further evaluate this claim.

The finding that linguistic knowledge affects the extent to which L2 learners benefit from transparent words is also consistent with previous L2 studies (e.g., Prior et al., 2014; Zhang & Koda, 2012). More importantly, the present results not only confirmed the role of linguistic knowledge in L2 Chinese learners (e.g., Ke & Koda, 2017; Zhang & Koda, 2018), but also demonstrated the interactive effect of linguistic knowledge and semantic transparency on L2 word-meaning inference (c.f. Tse et al., 2017). This finding extends previous research that has documented how individual sources of linguistic knowledge such as morpheme knowledge and grammatical knowledge play significant roles in the meaning inference L2 words.

Linguistic knowledge is therefore not a "panacea" that functions in word-meaning inference under any circumstance. Rather, linguistic knowledge facilitates learners' inferencing abilities under certain circumstances (e.g., when faced with an unknown transparent word as in the present study). These kinds of interactive effects between a learner's characteristics (e.g., linguistic knowledge, metalinguistic awareness, and learning strategies) and language-specific properties (e.g., semantic transparency, word frequency, and morpheme combination) are assumed to be universal in L2 learning. Here we provided evidence of this interaction using Mandarin Chinese, a language that allowed for a systematic examination of semantic transparency (see Packard, 2000; Taylor & Taylor, 2014). Our results are compatible with the multilevel interactive-activation framework (see details in Taft, Zhu, & Peng, 1999) in which activation units are organized hierarchically in levels corresponding to submorphemic components, morphemes, and words/multi-morphemes. Thus, the familiarity of morphemes and their relations can reduce a learner's memory load (e.g., Reder et al., 2016), and facilitate word-meaning inference of morphologically complex words.

Implications

Although this study targeted L2 Chinese, the findings are generalizable to word learning in all languages since the constructs of linguistic knowledge, semantic transparency, and word-meaning inference exist in all languages. First, L2 learners should focus on developing knowledge of linguistic units (e.g., morphemes and words) particularly in the beginning levels; that is, curriculum and instruction should support word development in all languages. Second, language instructors and textbook editors should note that different levels of semantic transparency affect the way learners infer meanings and the extent to which learners can use their prior knowledge of morphemes' meanings and relations. L2 learners could benefit from learning how to use different strategies when inferring the meanings of words with different levels of semantic transparency. For transparent words, learners could reach the correct meaning of an unknown word containing familiar morphemes once they have a sense of how morphemes combine with each other within the transparent word and are willing to guess. For opaque words, learners need to extend more effort in word memorization and make use of contextual information.

Limitations and conclusions

Admittedly, there are several limitations to this study. First, the present study did not find that vocabulary knowledge plays an independent role in word-meaning inference. This result, however, does not necessarily weaken the importance of vocabulary knowledge in L2 word-meaning inference, but rather highlights the role of contextual information. Unlike previous studies in which L2 learners infer unknown words within sentences (e.g., Ke & Koda, 2017; Zhang & Koda, 2018), this study did not provide contextual information for the word-meaning inference tasks. In natural learning environments, learners rarely encounter an unknown word without contextual information. Thus, future studies should investigate how L2 learners combine word-internal and word-external information to infer meanings of unknown words.

Second, different measures of word-meaning inference should be included in future studies. Because tasks may affect L2 learners' performances, researchers should aim to use a variety of tasks. Although L2 learners seemed to perform better, on average, in the multiple-choice task than the translation task (see mean scores in Tables 2, 5), the results in both tasks were consistent. We acknowledge that L2 learners presumably used more testing strategies (e.g., guessing) in the multiple-choice task (e.g., Budescu & Bar-Hillel, 1993) than the translation task. This assumption, however, should be thoroughly tested in future studies.

Third, the issue of morphemes' semantic ambiguity should be more carefully examined in future studies. For example, the morpheme *ying3* 影 within the word *hei1ying3* 黑影 ("black" + "shadow" → "twilight") has different high-frequency meanings, such as "shadow" in *ying3zi0* 影子 and "influence" in *ying3xiang3* 影响. If a learner cannot distinguish which meaning of the morpheme *ying3* 影 matches better with the morpheme *hei1* 黑 (black), the learner may report an incorrect (but semantically plausible) answer such "a bad effect." To what degree morpheme-specific frequency effects influenced the results remains an open question.

Finally, participants' L1 backgrounds should be better controlled. Testing participants with the same L1 could minimize the possible effects of cross-language transfer (see Koda, 2005) on L2 word-meaning inference.

In summary, the present study investigated the relationships between linguistic knowledge and semantic transparency in word-meaning inference among college-level adult learners of L2 Chinese. The main findings showed that (1) learners' linguistic knowledge and words' semantic transparency interact to affect word-meaning inference in L2 Chinese; and (2) for transparent words, linguistic knowledge (i.e., morpheme and grammatical knowledge) facilitates word-meaning inference in L2 Chinese. For opaque words, linguistic knowledge does not affect learners' inferencing abilities.

Funding Funding was provided by Key Research Project of "Double First-Class" Initiative by Minzu University of China (Grant No. 20SYL011).

Appendix 1

See Table 6.

Table 6 The result of semantic transparency

Semantic transparency	Transparency scores	Numbers of items	L1 raters' mean scores	L1 raters' mean scores	Semantic extension	Extension scores	Semantic ambiguity	Ambiguity scores
Most transparent ↓ Least transparent	**8**	21 items	3.63	3.63	Concrete + concrete on figurative	4	Mono + Mono	4
	7	11 items	2.92	3.36	Concrete + concrete	4	Poly + Mono	3
		11 items		2.47	Extended + concrete Concrete	3	Mono + Mono	4
	6	10 items	2.64	3.40	Concrete + concrete	4	Mono + Poly	2
		7 items		2.15	Extended + concrete	3	Poly + Mono	3
		7 items		2.38	Concrete + extended	2	Mono + Mono	4
	5	10 items	2.51	3.12	Concrete + concrete	4	Poly + Poly	1
		5 items		2.28	Extended + concrete	3	Mono + Poly	2
		8 items		2.13	Concrete + extended	2	Poly + Mono	3
	4	3 items	1.87	1.57	Extended + concrete	3	Poly + Poly	1
		9 items		2.25	Concrete + extended	2	Mono + Poly	2
	3	4 items		1.78	Concrete + extended	2	Poly + Poly	1

Concrete = a concrete morpheme; Extended = an extended morpheme; Mono = a monosemantic morpheme; Poly = a polysemantic morpheme

Appendix 2

See Table 7.

Table 7 Stimuli in word-meaning inference tasks and their characteristics

Stimuli	Semantic Transparency	Semantic transparency scores	Word frequency (per million characters)	Character frequency (per million characters)	Character complexity (stroke numbers)
笔试	Transparent	8	0.3	347.27	9
冰柜	Transparent	8	1.34	86.68	7
病床	Transparent	8	2.24	285.41	8.5
菜刀	Transparent	8	0.57	96.80	6.5
饭盒	Transparent	8	1.19	136.03	9
画室	Transparent	8	0.42	287.00	8.5
球台	Transparent	8	0.36	403.01	8
舞场	Transparent	8	0.57	614.61	10
羊群	Transparent	8	1.19	76.48	9.5
雨衣	Transparent	8	1.64	193.53	7
班车	Transparent	7	1.97	712.06	7
宝座	Transparent	7	1.70	262.71	9
报馆	Transparent	7	0.12	308.85	9
春装	Transparent	7	0.09	210.21	10.5
花期	Transparent	7	0.03	379.18	9.5
机房	Transparent	7	0.51	791.49	7
家史	Transparent	7	0.18	1285.08	7.5
年货	Transparent	7	0.06	648.51	7
市价	Transparent	7	0.63	205.16	5.5
月票	Transparent	7	0.36	313.44	7.5
病号	Opaque	6	0.83	422.74	7.5
带鱼	Opaque	6	0.06	560.48	8.5
底盘	Opaque	6	1.22	261.62	9.5
风车	Opaque	6	2.44	735.34	4
笔名	Opaque	5	0.83	522.15	8
黑市	Opaque	5	3.82	252.70	8.5
火速	Opaque	5	0.78	264.23	7
机床	Opaque	5	0.06	614.93	6.5
蓝图	Opaque	5	3.01	154.51	10.5
美工	Opaque	5	0.45	961.86	6
体院	Opaque	5	0.06	428.62	8
外商	Opaque	5	0.03	412.83	8
外线	Opaque	5	0.83	500.81	6.5
白条	Opaque	4	0.12	641.76	6
行情	Opaque	4	1.19	1519.91	8.5
泪花	Opaque	4	0.18	199.05	7.5
音响	Opaque	4	4.38	258.15	9
收成	Opaque	3	1.37	865.27	6
心机	Opaque	3	0.48	1412.44	5
中游	Opaque	3	0.36	876.54	7.5

References

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Huston (Ed.), *Advances in reading research* (pp. 231–256). Greenwich, CT: JAI.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50–57.

Baddeley, A. (1992). Working memory. *Science, 255,* 556–559.

Balota, D. A. (1994). Visual word recognition: The journey from features to meaning. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 303–358). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Budescu, D., & Bar-Hillel, M. (1993). To guess or not to guess: A decision-theoretic view of formula scoring. *Journal of Educational Measurement, 30*(4), 277–291.

Cacciari, C., & Glucksberg, S. (1994). Understanding figurative langauge. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 447–477). San Diego, CA: Academic Press.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), e10729.

Chen, T. (2018). The contribution of morphological awareness to lexical inferencing in L2 Chinese: Comparing more-skilled and less-skilled learners. *Foreign Language Annals*, *51*(4), 816–830.

Chen, T. (2019a). Joint contribution of reading subskills to character meaning retention in L2 Chinese. *Journal of Psycholinguistic Research*, *48*(1), 129–143.

Chen, T. (2019b). The role of morphological awareness in L2 Chinese lexical inference: From a perspective of word semantic transparency. *Reading and Writing: An Interdisciplinary Journal*, *32*, 1275–1293.

Chen, T. (2019c). The contribution of morphological awareness to word-meaning inference in L2 Chinese: From perspectives of language-related and learner-related factors. Unpublished doctoral dissertation.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A review and a user's guide. *Psychonomic Bulletin & Review, 12,* 769–786.

De Bot, K., Paribakht, T., & Wesche, M. (1997). Towards a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. *Studies in Second Language Acquisition, 19,* 309–329.

Dronjic, V. (2011). Mandarin Chinese compounds, their representation, and processing in the visual modality. *Writing Systems Research, 3*(1), 5–21.

Frantzen, D. (2003). Factors affecting how second language Spanish students drive meaning form context. *The Modern Language Journal, 87*(2), 168–199.

Gan, H. (2008). 语义透明度对中级汉语阅读中词汇学习的影响 [The effects of semantic transparency on vocabulary learning in intermediate Chinese reading]. 语言文字应用 *[Applied Linguistics], 1,* 82–90.

Haastrup, K. (1991). *Lexical inferencing procedures or talking about words: Receptive procedures in foreign language learning with special reference to English*. Tubingen: Gunter Narr.

Hamada, M. (2009). Development of L2 word-meaning inference while reading. *System, 37,* 447–460.

Hamada, M. (2014). The role of morphological and contextual information in L2 lexical inference. *The Modern Language Journal, 98*(4), 992–1005.

Hanban (Confucius Institute Headquarters). (2010). 新汉语水平考试大纲 *[Chinese Proficiency Test Syllabus]*. Beijing: The Commercial Press.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160–212.

Kaivanpanah, S., & Alavi, S. M. (2008). The role of linguistic knowledge in word-meaning inferencing. *System, 36,* 172–195.

Ke, S., & Koda, K. (2017). Contributions of morphological awareness to adult L2 Chinese word meaning inferencing. *The Modern Language Journal, 101*(4), 742–755.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York, NY: Cambridge University Press.

Koda, K. (2015). Development of word recognition in a second language. In X. Chen, V. Dronjic, & R. Helms-Park (Eds.), *Second language reading: Cognitive and psycholinguistic perspectives* (pp. 70–98). New York: Routledge.

Koda, K., & Miller, R. (2018). Cross-linguistic interaction in L2 word meaning inference in English as a foreign language. In H. Pae (Ed.), *Writing systems, reading processes, and cross-linguistic influences: Reflections from the Chinese, Japanese and Korean languages* (pp. 293–312). Philadelphia: John Benjamins.

Ku, Y. M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing: An Interdisciplinary Journal, 16,* 399–422.

Kuo, L. J., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist, 41*(3), 161–180.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language, 61,* 30–44.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language, 84,* 50–64.

Mok, L. W. (2009). Word-superiority effect as a function of semantic transparency of Chinese bimorphemic compound words. *Language and Cognitive Process, 24*(7/8), 1039–1081.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304–330.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 269–284). Mahwah, NJ: Lawrence Erlbaum Associates.

Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal, 87*(2), 261–276.

Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Canadian Modern Language Review, 61*(1), 107–134.

Office of Chinese Language Council International. (2007). 汉语水平考试 *[HSK tests]*. Beijing: Yuwen Press.

Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.

Parel, R. (2004). The impact of lexical inferencing strategies on second language reading proficiency. *Reading and Writing: An Interdisciplinary Journal, 17,* 847–873.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition, 21*(2), 195–224.

Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading comprehension. *Reading and Writing: An interdisciplinary journal, 27,* 1467–1484.

Reder, L., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review, 23,* 271–277.

Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language, 40*(4), 498–519.

Taylor, I., & Taylor, M. (2014). *Writing and literacy in Chinese, Korean and Japanese*. Amsterdam: John Benjamins Publishing Company.

Tse, C., Yap, M., Chan, Y., Sze, W., Shaoul, C., & Lin, D. (2017). The Chinese lexicon project: A megastudy of lexical decision performance for 25000+traditional Chinese two-character compound words. *Behavior of Research Methods, 49,* 1503–1519.

Wang, M., Lin, C. Y., & Gao, W. (2010). Bilingual compound processing: The effects of constituent frequency and semantic transparency. *Writing Systems Research, 2*(2), 117–137.

Wang, C., & Peng, D. (1999). 合成词加工中的词频、词素频率及语义透明度 [The roles of surface frequencies, cumulative morpheme frequencies, and semantic transparencies in the processing of compound words]. 心理学报 *[Acta Psychologica Sinica], 31*(3), 266–273.

Wang, C., & Peng, D. (2000). 多词素词的通达表征: 分解还是整体 [The access representation of poly morphemic words decomposed or whole]. 心理科学 *[Journal of Psychological Science], 23*(4), 395–398.

Wesche, M. B., & Paribakht, T. S. (2010). *Lexical inferencing in a first and second language: Cross-linguistic dimensions*. Bristol: Multilingual Matters.

Xing, H. (2018). *25亿字语料汉字字频表 [A corpus of character frequency based on 2.5 billion characters]*. Retrieved from http://faculty.blcu.edu.cn/xinghb/zh_CN/article/167473/content/1437.htm#article.

Xu, Y. (2014). 面向汉语二语教学的常用复合词语义透明度研究 *[Semantic transparency research on frequency used Chinese compound words for second language teaching](Unpublished doctoral dissertation)*. Beijing: Beijing Normal University.

Yuan, C., & Huang, C. (1998). 基于语素数据库的汉语语素及构词研究 [A study on Chinese morphemes and word formation based on a Chinese morpheme data bank]. 世界汉语教学*[Chinese Teaching in the World], 12*(2), 7–12.

Zhang, H. (2015). Morphological awareness in vocabulary acquisition among Chinese-speaking children: Testing partial mediation via lexical inference ability. *Reading Research Quarterly, 50*(1), 129–142.

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing: An Interdisciplinary Journal, 25,* 1195–1216.

Zhang, H., & Koda, K. (2018). Vocabulary knowledge and morphological awareness in Chinese as a heritage language (CHL) reading comprehension ability. *Reading and Writing: An Interdisciplinary Journal, 31,* 53–74.

Zhang, J., Lin, T. J., Wei, J., & Anderson, R. C. (2014). Morphological awareness and learning to read Chinese and English. In X. Chen, Q. Wang, & Y. Luo (Eds.), *Reading development of monolingual and bilingual Chinese children* (pp. 3–22). Dordrecht: Springer Science.

Zhang, J., & Zeng, Y. (2010). 影响中级水平留学生汉语新造词语理解的三个因素 [Three factors of the understanding of Chinese coinage for intermediate international students]. 语言文字应用 *[Applied Linguistics], 19*(2), 118–126.

Affiliations

Tianxu Chen[1] · **Keiko Koda**[2] · **Seth Wiener**[2]

Keiko Koda
kkoda@andrew.cmu.edu

Seth Wiener
sethw1@andrew.cmu.edu

[1] College of International Education, Minzu University of China, 27 Zhongguancun South Ave, Haidian District, Beijing 100081, China

[2] Department of Modern Languages, Carnegie Mellon University, Baker Hall 160, Pittsburgh, PA 15213, USA