Chinese idioms as constructions

Frequency, semantic transparency and their processing

Te-hsin Liu and Lily I-Wen Su
National Taiwan University

Chinese Quadrisyllabic Idiomatic Expressions (henceforth QIEs) are highly productive in the modern language. They can be used to understand the cognitive processing of structure and meaning during reading comprehension, as in the patterning of [*qian-A-wan-B*] '1k-A-10k-B' (e.g. one-thousand army ten-thousand horse). However, little is known about the underlying mechanisms of QIEs during reading comprehension. Adopting the framework of Construction Grammar, in the present study, we aimed to study the convergence and divergence between native speakers and L2 learners in the processing of Chinese idiomatic constructions.

In the present study, twenty-three native university-level Mandarin speakers and twenty-three L2 learners of intermediate and advanced levels of Mandarin, all speakers of the non Sinosphere, participated in the experiment, and were instructed to make a semantic congruency judgment during the presentation of a QIE. Our results showed that, for both native speakers and L2 learners, semantically transparent idiomatic constructions elicited much shorter RTs than semantically opaque idiomatic constructions. Our behavioral results also showed that native speakers processed low frequency QIEs faster than high frequency ones, implying semantic satiation to impede the interpretation of high frequency idioms. For L2 learners, it was semantic transparency, rather than frequency, that played a more prominent role in idiom processing.

Keywords: Chinese idiomatic expressions, semantic transparency, frequency, construction grammar, second language acquisition

1. Introduction

Idiomatic expressions are multiword conventional expressions whose meaning cannot be inferred directly from their constituent parts. Idiomaticity has been of

https://doi.org/10.1075/lali.00091.liu
Language and Linguistics 22:4 (2021), pp. 558–592. ISSN 1606-822X | E-ISSN 2309-5067

enduring interest to theoretical linguists, and much research attention has been devoted to explore the respective weights of semantic transparency (Gibbs et al. 1989) and frequency (Cacciari & Tabossi 1988; Tabossi et al. 2009). However, the role of construction and the interplay among semantic transparency, frequency, and construction deserve further investigation. The present study addressed these issues by examining the processing of Chinese Quadrisyllabic Idiomatic Expressions (QIEs).

Chinese QIEs mostly possess syntactic, phonological and semantic regularity, but are usually considered as semantically unpredictable. Even though many studies have been devoted to QIEs, what yet to be probed is their multifaceted nature and their linguistic productivity, i.e. their interrelated syntactic, semantic, and pragmatic features when a constructional view is adopted. To this end, this study examines Chinese QIEs with [*qian-A-wan-B*] '1k-A-10k-B' patterning, which is highly productive in the modern Chinese language. The high productivity of [*qian-A-wan-B*] '1k-A-10k-B' implies that the construction is semantically decomposable, thereby enabling native speakers to create novel idiomatic expressions based on this construction. The role played by frequency during idiom processing is therefore of interest. The construction [*qian-A-wan-B*] '1k-A-10k-B' has a high-type frequency. As Bybee & Thompson (2000) noted, this high-type frequency ensures that a construction is frequently used, thereby strengthening its representational schema and enhancing its accessibility for further use with new items. How would native speakers therefore react to a low-token-frequency idiom of a high-type frequency? What role does construction play in the association of form and meaning?

To probe the above topics, we used a semantic judgment task to investigate the underlying mechanisms of QIEs during reading comprehension. Twenty-three native university-level Mandarin speakers and twenty-three L2 learners of intermediate and advanced levels of Mandarin, all speakers of non Sinosphere, participated in the experiment. The first goal of the present study was to investigate the respective weights of semantic transparency and frequency during idiom processing given that previous studies have provided inconsistent results. Second, we examined the relationship between frequency and construction. Due to the high-type frequency of [*qian-A-wan-B*] '1k-A-10k-B' and its association between form and meaning, we hypothesized that token frequency is not important for native speakers because they can infer the meaning of a low-frequency idiom with the same patterning. By contrast, L2 learners, who lack exposure to idiomatic constructions, have no access to the meaning of the construction and must rely on the semantic transparency of QIEs to decode their semantic content.

2. Literature review

2.1 Critical factors contributing to idiom processing

Over the past 30 years, many psychological studies on idiom processing have been conducted, with a focus on compositionality, semantic transparency, and frequency. Swinney & Cutler (1979) were the first to propose the Lexical Representation Hypothesis (LRH), positing that idioms are *noncompositional*. Lacking semantic compositionality, idioms are therefore represented mentally as morphologically complex words and are recognized using the same retrieval processes applied during word recognition. This hypothesis predicts that, for native speakers, the processing time of idioms (e.g. *break the ice*) is faster than that of nonidioms (i.e. *break the glass*) with the same syntactic structure. However, Swinney & Cutler (1979) did not classify idioms according to their semantic transparency and thus failed to explain why the processing of *pop the question* and that of *kick the bucket* are different. Semantic transparency may be linked to either meaning predictability (Plag 2003: 46) or analyzability (Zwitserlood 1994). Plag, for example, noted that words are semantically transparent if "… their meaning is predictable on the basis of the word-formation rule according to which they have been formed." Zwitserlood suggested that "[t]he meaning of a fully transparent compound is synchronically related to the meaning of its composite words …"

To address this problem, Gibbs et al. (1989) proposed the Idiom Decomposition Hypothesis (IDH), which posits that idioms are processed differently depending on whether they are decomposable or nondecomposable.

People read sentences containing decomposable idioms faster than sentences containing nondecomposable idioms. Semantic compositionality refers to the fact that the constituents of some idioms "carry identifiable parts of the idiomatic meaning" (Nunberg et al. 1994: 496). For example, in *pop the question*, there is a clear correspondence between *pop* and *question* and the relevant parts of the figurative meaning "propose marriage". By contrast, in *kick the bucket*, the correspondence between *kick* and *bucket* and the meaning of "die suddenly" is much less clear. Thus, *pop the question* is semantically decomposable, whereas *kick the bucket* is nondecomposable. Gibbs et al. (1989) also found that the syntactic processing of an idiom is closely related to its semantic compositionality; for decomposable idioms, the same mechanisms of lexical retrieval and syntactic parsing occur during the comprehension of literal expressions, whereas the processing mechanisms of nondecomposable idioms are similar to those operating in the recognition of individual words.

The contribution of Gibbs et al. (1989) is twofold. First, they classified idioms from the perspective of semantic compositionality and found a positive corre-

lation between semantic compositionality and semantic transparency. Second, they related semantic compositionality to syntactic flexibility, arguing that the more transparent an English idiom is, the more flexible is its syntactic structure. Thus, native speakers process a semantically transparent idiom syntactically. This explains why *pop the question* can be changed to *the question was popped*. By contrast, the more opaque an English idiom is, the less flexible is its syntactic structure. Therefore, *kick the bucket* cannot be changed to **the bucket was kicked*, as the processing of this idiom is the same as that for the recognition of an individual word.

Cacciari & Tabossi (1988) took a different view when they proposed their Configuration Hypothesis (CH). They claimed that nothing is inherent in idiomatic expressions that makes them easy to recognize. In addition to idiomaticity, the major difference between idiomatic and literal expressions is that idiomatic expressions are known to speakers and listeners, whereas literal expressions may be entirely novel. Tabossi et al. (2009) tested predictions deriving from the three main theories of idiom recognition: the Lexical Representation Hypothesis (LRH), the Idiom Decomposition Hypothesis (IDH), and the Configuration Hypothesis (CH). Using a semantic judgment paradigm, where people were asked to decide whether a string was meaningful, they observed that participants were faster at judging decomposable idioms, nondecomposable idioms, and clichés than matched controls. No significant difference was observed between decomposable and nondecomposable idioms, implying that semantic transparency plays no role in idiom processing. Based on these results, Tabossi et al. (2009) claimed that, in line with the CH, it is frequency, rather than idiomaticity, that explains the rapid recognition of idioms. However, Tabossi et al.'s results were based on a lexical decision task undertaken by twelve native Italian speakers; therefore, we cannot ascertain whether the lack of statistical significance was attributable to the characteristics of the participants.

Notably, the frequency examined in the aforementioned experiment refers to *token frequency*, rather than the *type frequency* of an idiom. *Token frequency* counts how often a particular form appears in the input whereas *type frequency* refers to the number of distinct lexical items that can be substituted in a given slot in a construction, whether it is a word-level construction for inflection or a syntactic construction specifying the relations between words (Ellis 2013). Bybee & Hopper (2001) noted that the productivity of phonological, morphological, and syntactic patterns is a function of type, rather than token frequency. A question arises as to how speakers would process a low-token-frequency idiom of high-type frequency. Meanwhile, the role of construction deserves further investigation. In the following section, we review some basic claims made within cognitive linguistics to see how this approach can explain the processing of Chinese QIEs.

2.2 Cognitive linguistics and Construction Grammar (CG)

Nunberg et al. (1994: 492–493) argued that one of the central characteristics of idioms is *conventionality*, in that "their meaning or use can't be predicted, or at least entirely predicted, on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another" (Nunberg et al. 1994: 492). Thus, if idioms are conventional, as in the traditional componential model, then they must be stored in the mind of a speaker, and their interpretation should be predicted by the general and linking rules for syntactic and semantic components. However, this is not the case, as many idioms are semantically opaque, such as *kick the bucket* and *pull a fast one*. The interpretation rules cannot be applied to these idioms because parts of the syntactic phrase do not correspond to parts of the semantic phrase. CG has therefore been developed to address the conventionality of idioms. As Croft & Cruse (2004: 225) noted, "it is not an exaggeration to say that CG grew out of a concern to find a place for idiomatic expressions in the speaker's knowledge of a grammar of their knowledge."

Fillmore et al. (1988) argued that a construction is a schematic idiom. Some elements of the construction are *lexically open* and therefore cannot be listed as phrasal lexical items. Consequently, the syntactic, semantic, and pragmatic properties of schematic idioms cannot be predicted from the general rules for the syntactic and semantic components of the language. Instead, these properties are directly associated with the construction *per se*. Goldberg (1995) defined "construction" as "the pairing of form and meaning," which is the essential foundation of language construction. She illustrated that there are many grammatical phenomena in English that cannot be derived from the nature of vocabulary nor from the application of grammar rules; instead, they are shaped by specific constructions. CG focuses on the importance of "use," as the pairing of form and meaning is shaped by repeated use. Several empirical studies have provided evidence for the constructional approach. For example, the ditransitive construction in English is directly associated with the meaning of "transfer" (Pinker 1989; Goldberg 1992; Goldberg et al. 2005; Hovav & Levin 2008). Ahrens (1995) showed that when people were asked what the nonsense verb *moop* means in {*She mooped him something.*}, 60% of the subjects responded that it meant 'give', and the rest offered meanings that preserved the meaning of literal or metaphorical transfer (e.g. 'tell'). Kaschak & Glenberg (2000) also demonstrated that subjects relied on constructional meaning when they encountered nouns used as verbs in novel ways (e.g. {*She crutched him the ball.*}).

Although the study of English idioms within the framework of CG began three decades ago, the study of Chinese idiomatic expressions from a construc-

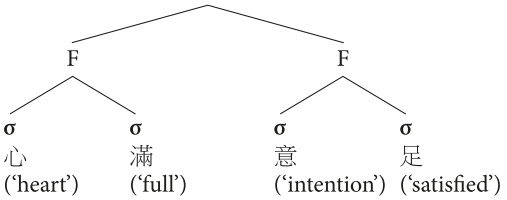
tional perspective is still in the nascent stage. The next section provides some background on the characteristics of Chinese idioms and shows how CG can be used to explain their regularity and idiomaticity.

2.3 The definition of Chinese idioms and their linguistic features

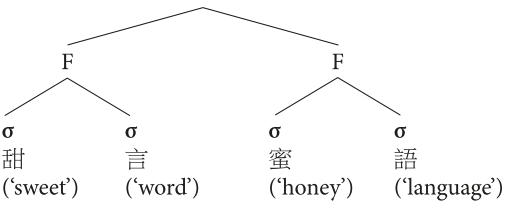
More than 90% of Chinese idiomatic expressions exhibit four-character patterns (Xu 2006: 108). Most Chinese idioms are derived from literary classics, where the meaning of an idiom is usually greater than the sum of the meanings conveyed by the four characters. Liu (1990) stated that the most important feature of many Chinese QIEs is their double planes of meaning: people must therefore see beyond the literal meaning to understand the true meaning of a Chinese QIE.[1] Such QIEs are frequently used in ordinary Chinese speech and writing, as they make a point using only four characters (Wang 1990). The grammatical construction of idioms in European languages usually resembles regular phrase constructions and can be subjected to syntactic operations such as passive voicing, topicalization, quantification, tense marking, and insertion (Cacciari & Glucksberg 1994; Moon 1998; Cacciari 2014). By contrast, many Chinese QIEs do not follow the usual grammatical structure and syntax of modern spoken Chinese language and are instead highly compact and synthetic (Tsou 2012). Tsou therefore described that Chinese QIEs comprise three characteristics: (a) four syllables or logographs; (b) relatively fixed structure and patterns; and (c) figurative meaning and semantic opacity.

Tsou (2012) suggested, moreover, that there are finite possibilities for the internal morphological and syntactic structures of QIEs and noted that, in terms of syntax, 35% of QIEs have a coordinative structure, and that 21.5% of QIEs have an attributive structure. Furthermore, the subject-predicate structure accounts for 17.5% of QIEs, and the predicate-object structure accounts for 15%. The coordinate form can be divided into a two-plus-two syllable structure, in which the first foot and the second foot are syntactically and semantically parallel, as shown in Figure 1:

1. For example, the idiom 刻舟求劍 [cut-boat-seek-sword] is based on a historical account of a man, during the Spring and Autumn period, who dropped his sword into the river and etched a mark onto the side of his boat to indicate the place where this occurred. The figurative meaning derived from this idiom is that of an action made pointless by changed circumstances.



a. 心滿意足 [*xin man yi zu*] 'perfectly contented'



b. 甜言蜜語 [*tian yan mi yu*] 'smooth-tongued'

Figure 1. Illustration of the internal morphology of two QIEs

Figure (1a) displays two subject-predicate structures in parallel, in which both the subjects and predicates are semantically equivalent. The same morphological mechanism applies to Figure (1b), which is composed of two synonymous adjective-noun structures.

Drawing upon the CG assumption, Su (2002) investigated three productive Chinese idiomatic constructions and suggested that the existence of these constructions denotes a format for mental representation, namely the conceptual structure. For example, in X-*lai*(come)-X/Y-*qu*(go), the X/Y slots in these constructions are typically filled by monosyllabic verbs of activity, as X and Y are near synonyms. The expression therefore denotes a repetitive action of a certain duration. The repetitive, iterative, and durative properties in turn lead to specific semantic constraints placed on the verb selected for X/Y. For example, **si*(die)-*lai*(come)-*si*(die)-*qu*(go), 死來死去, 'to die again and again,' is not a possible construction because the act of dying cannot be repeated. Chen (2012) focused on the internal structures of QIEs but referred to these as "quasi-fixed structures." They identified five basic structures, two of which are the pattern of second/fourth syllables to be filled (e.g. *ban-X-ban-Y* 'half-X-half-Y') and the pattern of first/third syllables to be filled (e.g. *X-lai-Y-qu* 'X-come-Y-go'). Liu et al. (2017) examined Chinese idioms with the patterning of [yi-X-#-Y] 'one-X-number-Y' by looking into instances that fall under three specific prefabs:[2]

2. A lexically-filled prefab is taken as a subtype of construction (Bybee 2006).

[yi-X-yi-Y] 'one-X-one-Y' (e.g. 一心一意), [yi-X-er-Y] 'one-X-two-Y' (e.g. 一石二鳥), and [yi-X-qian-Y] 'one-X-thousand-Y' (e.g. 一字千金). Their results suggest that the three prefabs of [yi-X-#-Y] display a negative correlation in terms of the numeral in the third syllable (i.e. one/two/thousand): the higher the number, the fewer the number of meaning. The productivity of the QIE in question may actually come from the interaction among several sources: the polysemy of the numeral *yi* 'one', the contrast between the numerals (one, two, and thousand), and the many meanings subsumed under the construction discussed, as shown in Table 1:

Table 1. Meaning and examples of the construction [yi-X-#-Y]

***Construction: Yi*-X-#-Y**					
***Prefab 1: Yi*-X-*yi*-Y**		***Prefab 2: Yi*-X-*er*-Y**		***Prefab 3: Yi*-X-*qian*-Y**	
Meaning of prefab	**Example**	**Meaning of prefab**	**Example**	**Meaning of prefab**	**Example**
wholeness	一心一意	emphasis	一乾二淨	contrast	一字千金
repetition	一來一往	repetition	一來二去		
each and every	一言一行	comment	一石二鳥		
comment	一字一板				
contrast	一龍一豬				

They concluded, on the one hand, that the polysemous construction [yi-X-#-Y] is motivated by the interaction between syntactic parallelism and semantic contrast. On the other hand, the compositionality of idiomatic expressions can be considered as a multifaceted continuum rather than as a binary distinction. The above studies were the first to investigate Chinese QIEs under the framework of Construction Grammar; however, what remains to be probed is the underlying mechanisms of idiom processing during reading comprehension.

Two recent studies have explored the comprehension of Chinese QIEs among native speakers from a neurolinguistic perspective. Zhang et al. (2013) used the Event-Related Potentials (ERPs) to explore the effects of semantic compositionality on Chinese idiom processing. Primed by their literal meanings, 146 QIEs with different degrees of compositionality and non-idiomatic phrases were visually presented to 18 subjects for a semantic judgment task. Their behavioral results demonstrated that semantic compositionality (i.e. semantic transparency) exerts a great influence on Chinese idiom processing. Meanwhile, subjects responded significantly more slowly to non-idiomatic literal phrases than to all Chinese idioms, thus validating the CH (Cacciari & Tabossi 1988). Yang et al. (2016) investigated the role of the right hemisphere in Chinese idiom processing, and

found that both the left hemisphere and the right hemisphere are involved in idiom processing. More importantly, idioms elicited more activation than do non-idioms in the right parietal cortex, and the activation strength decreased as a function of semantic transparency of idioms.[3] In sum, both semantic transparency and frequency were found to have an effect on Chinese idiom processing. What remains unclear is how construction interacts with frequency and semantic transparency. Meanwhile, the convergence and divergence between native speakers and L2 learners when processing idioms merits further investigation. In the next section, we shall review how non-native speakers learn idioms in order to better examine the difference between the two groups of speakers.

2.4 L2 acquisition of idioms

Ellis (2002) argued that L2 acquisition is different from L1 acquisition in terms of conceptual development, language input, and transfer from L1. Moreover, although knowledge of the world and knowledge of language develop simultaneously during L1 acquisition, L2 adult learners build on preexisting conceptual knowledge. Regarding the role of language input, a typical L1 pattern of acquisition results from naturalistic exposure, whereas L2 acquisition mainly occurs in classroom environments; such acquisition can distort patterns of exposure, function, and social interaction (Ellis & Laporte 1997). Concerning the role of L1 transfer, L2 learners build on preexisting L1 knowledge to learn L2, which means that speakers of different languages prioritize different aspects of events in narrative discourse (Berman & Slobin 1994). Moreover, Ellis (2006) observed that L2 learners referred to their native tongue to survey their use of L2, so their inductions were often affected by transfer, with L1-tuned expectations and selective attention. This observation was supported by Hu & Fong (2010), who explored how cross-cultural differences hinder adequate idiom interpretation. For example, in Chinese, *xin* (heart) is associated with emotional and intellectual domains, whereas dualities of *mind* vs. *body* are found in English. Hu & Fong's analysis of 40 Chinese learners' interpretation of English idioms showed that L1 negative transfer occurred during this interpretation.

Research has also shown that, similar to children below the age of 7 years, L2 learners tend to engage in a literal reading of idioms in their target language (Cieślicka 2006) and process multiword expressions analytically, dividing them into separate words (Wray 2002: 206–210; Fitzpatrick & Wray 2006). Consequently, in the L2 acquisition of Chinese, *semantic mistakes* were found to be

3. It should be noted that the three types of stimuli presented to subjects in Yang et al.'s (2016) study were matched for frequency.

one of the major errors made by non-native speakers (Wang 2001; Liu 2005). For example, the QIE 一日千里 [one-day-thousand-miles] 'one day, a thousand miles' has connotations of 'making progress at a tremendous pace.' It is therefore not uncommon to observe inappropriate sentences such as "*我的男朋友病了，我**一日千里**的坐火車去巴黎看他" (Liu 2005), literally meaning, 'My boyfriend was sick, so I traveled **a thousand miles in a day** to see him in Paris.' The above phenomenon points to the importance of semantic transparency in the acquisition of idioms by L2 learners. Skoufaki (2008) also suggested that transparent idioms were easier to comprehend even if they were unfamiliar to L2 learners. Skoufaki had advanced learners of English guess at the meaning of unknown English idioms varying along the dimension of transparency, and found that transparent idioms had significantly more correct guesses than opaque idioms (see also Boers & Demecheleer 2001).

2.5 The current study

Previous studies have indicated that research on idiom comprehension should consider several important variables, including semantic transparency, frequency, the role of construction, and how idioms are processed by non-native speakers. The current study addressed these challenges by exploring how Chinese idiomatic constructions are processed. As noted in § 2.3, Chinese QIEs mostly possess syntactic, phonological, and semantic regularity, but are usually considered to be semantically unpredictable. Even though many studies have been devoted to QIEs, what yet to be explored is their multifaceted nature and their linguistic productivity when a constructional view is adopted. The study of Chinese idioms, from a constructional perspective, is of great importance, given their syntactic/semantic parallelism and productivity, but no psycholinguistic research on this topic has been conducted to date.

In the present study, we selected the construction [*qian-A-wan-B*] '1k-A-10k-B' for examination for two reasons. First, it is composed of 50 idiomatic sequences, has a high-type frequency, and is very productive. Moreover, *qian* '1,000' and *wan* '10,000' are numerals; such concreteness should be relatively easy to understand for L2 learners and facilitates a comparison between native speakers and L2 learners. From a syntactic perspective, the A and B of [*qian-A-wan-B*] '1k-A-10k-B' can be nouns of near synonyms, as in [*qian-cun-wan-luo*] (千村萬落: [1k-village-10k-hamlet] 'thousands of villages and hamlets'). They can also be semi-bound morphemes that form a disyllabic verb as in [*qian-hu-wan-huan*] (千呼萬喚: [1k-call-10k-shout] 'after repeated calls') and [*qian-xin-wan-ku*] (千辛萬苦: [1k-hard-10k-painstaking] 'extremely painstaking.'). The construction [*qian-A-wan-B*] '1k-A-10k-B' is symmetrical in that *qian* '1,000' and *wan* '10,000' are

numerals and A and B generally possess the same syntactic function. From a semantic perspective, *qian* '1,000' and *wan* '10,000' both mean 'large in amount'. When they are combined with the aforementioned synonyms, the construction intensifies the degree to which an action is executed or the amount of a noun. From a contextual perspective, because the construction involves the pairing of form and meaning, the comprehension of idiomatic sequences with the patterning of [*qian-A-wan-B*] '1k-A-10k-B' does not rely on the context in which they occur; instead, the construction *per se* provides the "context" upon which the native speaker relies to decode the semantic content.

Based on their morphological and syntactic similarities, exemplars composed of [*qian-A-wan-B*] '1k-A-10k-B' are judged by native speakers as belonging to the same category. The semantics of the construction emerges from every idiomatic sequence with the same construction. This means that each token of [*qian-A-wan-B*] '1k-A-10k-B' has an effect on representation, and each occurrence of the construction maps onto the exemplar cloud for the construction. Figure 2 schematizes this exemplar cloud:

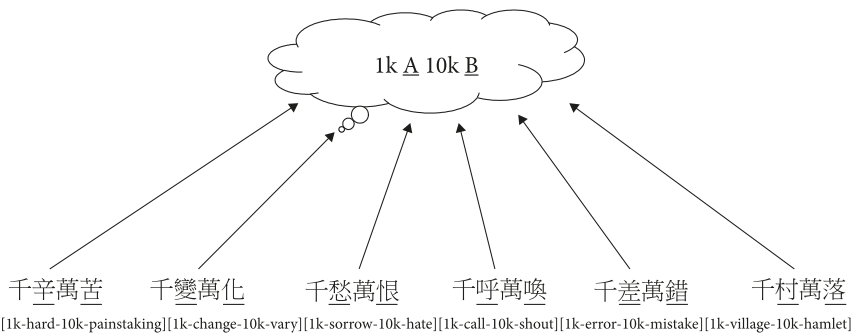


Figure 2. Illustration of the exemplar cloud of [*qian-A-wan-B*] construction

Studies on the respective weights of semantic transparency (Gibbs et al. 1989) and frequency (Cacciari & Tabossi 1988; Tabossi et al. 2009) have provided inconsistent results. Following the IDH, the high productivity of [*qian-A-wan-B*] '1k-A-10k-B' implies semantic decomposability and high semantic transparency, thereby enabling native speakers to create novel idiomatic expressions based on this construction. A question arises as to whether idioms composed of [*qian-A-wan-B*] '1k-A-10k-B' are transparent in the same way. If not, what role does semantic transparency play during idiom processing? The IDH posits that transparent expressions are faster than opaque idioms. Meanwhile, how would native speakers comprehend invented idioms which obey the morphological and semantic constraints of the construction? The CH posits that all expressions, if they are

well known, are easier to process and hence will be processed faster than novel expressions, and semantic transparency is not relevant in determining processing time. The present paper probes these topics.

Furthermore, the frequency examined in the CA refers to token frequency, rather than type frequency. The construction [*qian-A-wan-B*] '1k-A-10k-B' has a high-type frequency, which, according to Bybee & Thompson (2000), ensures that it is used frequently, thereby strengthening its representational schema and enhancing its accessibility for use with new items. Given the high-type frequency and the association between form and meaning, we therefore predicted that token frequency is not important for native speakers because they can infer the meaning of a low-frequency idiom with the same patterning. By contrast, L2 learners, who lack exposure to idiomatic constructions, have no access to the meaning of the construction and must rely on the semantic transparency of QIEs to decode their semantic content.

To verify the aforementioned assumptions, we designed three types of stimuli. The first consisted of idiomatic constructions. Forty-eight [*qian-A-wan-B*] '1k-A-10k-B' constructional idioms were selected, half of which had a high token frequency and half had a low token frequency. Among the high-frequency idioms, half were semantically transparent, and half were semantically opaque. The same grouping applied to low-frequency idioms. The second type of stimuli consisted of quasi-idiomatic constructions. These are invented idioms that do not exist in the modern Chinese language, but they nevertheless obey the morphological and semantic rules of construction; for example, the second and the fourth syllable are occupied by synonyms. The third type of stimuli consisted of pseudo-idiomatic constructions. These are invented idioms that violate the morphological rule of the construction. Additionally, 16 QIEs not belonging to the [*qian-A-wan-B*] '1k-A-10k-B' construction were included in the study to serve as controls.

3. Materials and methods

3.1 Participants

Twenty-three native university-level Mandarin speakers (age range 20–28 years, mean age 24.1, SD 3.0) and twenty-three L2 learners of intermediate and advanced levels of Mandarin (age range 20–32 years, mean age 24.8, SD 3.6), all speakers of non Sinosphere, participated in the experiment and were paid for their services. All L2 learners had passed a placement test and were assigned to Mandarin courses at an intermediate or advanced level at the authors' institution. The revised Peabody Picture Vocabulary Test (PPVT, Lu & Liu 1998[1994])

was also administered for L2 learners to estimate their knowledge of the Mandarin lexicon. As in the English version, the subject hears a word and selects the corresponding picture from a set of four options. The L2 learners obtained age-equivalent scores of 11.4 years (mean raw score = 83.45). All subjects had normal or corrected-to-normal vision and were unaware of the research purposes. Data of three L2 learners were discarded in the statistical analysis due to excessive high error rates.

3.2 Materials

The first questionnaire focused on rating the semantic distance, and it contained 48 existing idioms of [*qian-A-wan-B*] '1k-A-10k-B,' along with the most frequent 300 QIEs found in Chinese teaching materials. Non-[*qian-A-wan-B*] '1k-A-10k-B' idioms were included in the semantic transparency rating task for two reasons. First, if we only tested the transparency rating of [*qian-A-wan-B*] '1k-A-10k-B' idioms, the participants would be rating the same construction repeatedly, which might influence the reliability of the ratings. By rating [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions as well as other idioms, we were able to compare the semantic transparency of these idiomatic constructions to that of idioms that are not construction-based. To obtain an objective measure of the semantic transparency of each stimulus, we asked a separate group of 60 native speakers to rate the semantic distance between the figurative and literal meaning on a 1–7 Likert scale, where 1 indicates *least similar* and 7 indicates *most similar*. To assess Inter-Rater Reliability (IRR), a two-way mixed, consistency, average-measures Intra-Class Correlation (ICC) (McGraw & Wong 1996) was calculated to determine the degree to which raters were consistent in their ratings of the semantic distance. An ICC of 0.91, which is in the excellent range (Cicchetti 1994), indicated that there was a high degree of agreement between raters. Semantic distance ratings were therefore deemed to be suitable for testing our hypotheses.

The result of the semantic transparency rating showed that the mean transparency value of non-[*qian-A-wan-B*] '1k-A-10k-B' idioms was 4.02 (*SD* = 1.56). The most transparent idiom was 簡單明瞭 (*jian-dan-ming-liao*; meaning: 'easy to understand'), whereas the least transparent idiom was 別開生面 (*bie-kai-sheng-mian*; meaning: 'to break a new path'). The 48 [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions had a mean value of 4.54 (*SD* = 1.58), two-thirds of which had a value higher than 4.02. This result demonstrated that, owing to the pairing of form and meaning and the productivity of the construction, idiomatic constructions were generally more transparent than non-constructional ones. Table 2 displays the result of the semantic distance rating.

Table 2. Result of semantic distance rating

Idiom types	Example	Semantic distance	Meaning
Non-constructional idioms			
Transparent	簡單明瞭 *(Jian-dan-ming-liao)*	6.97 (0.18)	easy to understand
Opaque	別開生面 *(Bie-kai-sheng-mian)*	1.0 (0)	to break a new path
Mean value	4.02 (1.56)		
Constructional idioms			
Transparent	千叮萬囑 *(Qian-ding-wan-zhu)*	6.33 (0.87)	repeatedly urging
Opaque	千頭萬緒 *(Qian-tou-wan-xu)*	1.03 (0.18)	very complicated
Mean value	4.54 (1.58)		

For the frequency of idioms, we originally drew upon the Center for Chinese Linguistics Chinese Corpus (CCL Chinese Corpus, developed by Peking University, Zhan et al. (2003)), which contains 4.77 hundred million characters. However, 7 of the 48 existing constructions, compiled in current Chinese dictionaries, could not be found in the corpus. We then used the Google Search Engine to obtain the latest frequency information for our stimuli. As noted by Meyer et al. (2003), frequency information generated by search engines must be interpreted with caution. Nevertheless, such information is "suggestive" of the frequency with which certain words and grammatical constructions occur. We therefore individually entered the 48 [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions into the Google Search Engine to obtain up-to-date frequencies.

A separate group of 580 native speakers of Mandarin who did not participate in the experiment or the semantic transparency test rated familiarity with the idiom on a 1–7 Likert scale (1 = totally unfamiliar; 7 = very familiar. ICC = 0.95). The result of the familiarity test indicated that the 48 [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions had a mean familiarity value of 4.12 ($SD = 1.90$). The mean value of quasi-idiomatic constructions was 1.92 ($SD = 1.57$), and the mean value of pseudo-idiomatic constructions was 1.49 ($SD = 1.02$). The result of the ratings demonstrated a clear graded effect across the three types of idiomatic constructions [$F\ (2,77) = 27.74$, $p < 0.001$]. A post hoc Scheffé test showed that the ratings of idiomatic constructions were significantly higher than those of quasi-idiomatic and pseudo-idiomatic constructions (both $ps < 0.001$), but no difference was observed between quasi-idiomatic and pseudo-idiomatic constructions ($p = 0.698$). The 16 existing idioms that were not construction-based had a mean familiarity value of 6.55 ($SD = 0.88$).

To determine whether a correlation exists between semantic transparency and frequency, a correlation analysis was conducted. The result indicated no correlation ($r = -0.12$, $p = 0.458$). Additionally, a correlation analysis was also con-

ducted to determine whether a correlation exists between familiarity and frequency. The results showed a high correlation between these two factors ($r = 0.69$, $p < 0.001$). This means that high-frequency idioms are generally familiar to native speakers of Mandarin. No correlation was found between familiarity and semantic transparency ($r = -0.12$, $p = 0.83$).

The final set of experimental stimuli included (1) 48 [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions halved on the basis of semantic transparency and frequency. The mean values for semantically transparent and opaque idioms were 5.48 and 3.63, respectively. (2) 16 quasi-idiomatic constructions, where A and B are either synonyms or semantically related. This tested speakers' comprehension of newly created idioms, given that [*qian-A-wan-B*] '1k-A-10k-B' is highly productive. (3) 16 pseudo [*qian-A-wan-B*] '1k-A-10k-B' trials that violated the semantic restriction of the construction and (4) 16 existing idioms that were not construction-based, such as 頭頭是道 *(tou-tou-shi-dao*; meaning: 'clear and logical') and 九牛一毛 (*jiu-niu-yi-mao*; meaning: 'a drop in the ocean').

A summary of the experimental stimuli is presented in Table 3. It is important to note that quasi-idiomatic constructions and pseudo-idiomatic constructions are invented; therefore, only word-by-word meanings can be provided. The full list of linguistic materials is displayed in the Appendix.

Table 3. Examples of experimental conditions

Condition	Example	Meaning	No. of items
Idiomatic constructions			
a. High Frequency/ High Transparency	千變萬化 *(Qian-bian-wan-hua)*	Countless changes	12
b. High Frequency/ Low Transparency	千頭萬緒 *(Qian-tou-wan-xu)*	Very complicated	12
c. Low Frequency/ High Transparency	千思萬慮 *(Qian-si-wan-lu)*	Think repeatedly	12
d. Low Frequency/ Low Transparency	千倉萬箱 *(Qian-cang-wan-xiang)*	Massive storage of food	12
Quasi-idiomatic constructions			
	千德萬賢 *(Qian-te-wan-xian)*	[1k-ethic-10k-virtue]	16
	千花萬香 *(Qian-hua-wan-xiang)*	[1k-bloom-10k-perfume]	
Pseudo-idiomatic constructions			
	千桶萬庭 *(Qian-tong-wan-ting)*	[1k-bucket-10k-court]	16
	千跳萬手 *(Qian-tiao-wan-shou)*	[1k-jump-10k-hand]	
Not construction-based idioms			
	頭頭是道 *(Tou-tou-shi-dao)*	Clear and logical	16
	九牛一毛 (*Jiu-niu-yi-mao)*	A drop in the ocean	

3.3 Experimental task and procedure

The task was programmed using E-prime software (Schneider et al. 2002). The experiment was conducted in a sound-attenuated chamber where participants were seated on a comfortable chair in front of a 17-inch computer screen. They were then provided instructions before performing the experimental task. During the task, they were required to read all the presented trials silently and then decide whether the meaning offered corresponds to the preceding stimulus by pressing corresponding buttons. The percentages of correct "Yes" and "No" responses were 50%.

To compare differences between native speakers and L2 learners in comprehending quasi-idiomatic and pseudo-idiomatic constructions, we used a semantic judgment task, rather than a non-semantic task (for example, a font judgment task). Each trial began with the presentation of a hashtag (#) displayed at the center of the screen for 1000 ms. Each stimulus was then displayed for 3000 ms, followed by its meaning. The participants were given a maximum of 7 seconds to decide whether the meaning offered was correct or incorrect. The inter-trial-interval varied between 0 and 1800 ms (mean = 900 ms). Ten practice trials were provided so that the participants could familiarize themselves with the task. The total duration of the experiment was approximately 15 minutes. The experimental procedure is depicted in Figure 3.

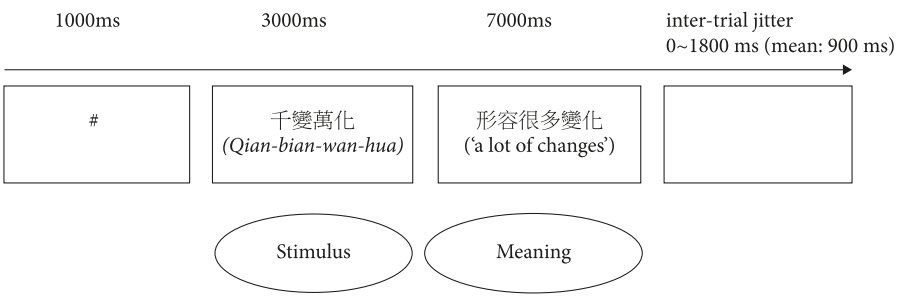


Figure 3. Illustration of experimental procedure

The length of the meaning offered was carefully controlled (average length = 6.9 characters, range = 6–8 characters). Given that native speakers and L2 learners were reading the same stimuli, the meanings that were offered contained words that were easy for L2 learners to understand. Note that the results of three L2 learners were excluded from analysis due to high error rates.

3.4 Results

3.4.1 *Accuracy rates*

The accuracy rates for native speakers and L2 learners are presented in Figure 4. To compare the performance of native speakers and L2 learners, we conducted a two-way repeated measures ANOVA on accuracy rates, in which language (Mandarin speakers vs. L2 learners) served as the between-subject variable and construction type as the within-subject variable. Mauchly's test of sphericity indicated that sphericity could not be assumed [$\chi^2(5) = 32.007$, $p < .001$, $\varepsilon = .447$], therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .739$). The effect of construction type was significant [$F(2.2, 90.8) = 24.9$, $p < 0.001$, $\eta^2 = .38$]. In addition, the ANOVA also showed a significant main effect for language background, native speakers obtained higher accuracy rates (76.65% vs. 62.27%) and pairwise comparisons showed a reliable difference between native speakers and L2 learners [$F(1, 41) = 63.3$, $p < 0.001$, $\eta^2 = .61$]. The interaction between construction type and language was also significant [$F(2.2, 90.8) = 3.82$, $p = 0.02$, $\eta^2 = .085$]. The results indicate that different types of construction yield significant behavioral outcomes, with native speakers obtaining better results on all types of idioms than L2 learners (see Figure 4).

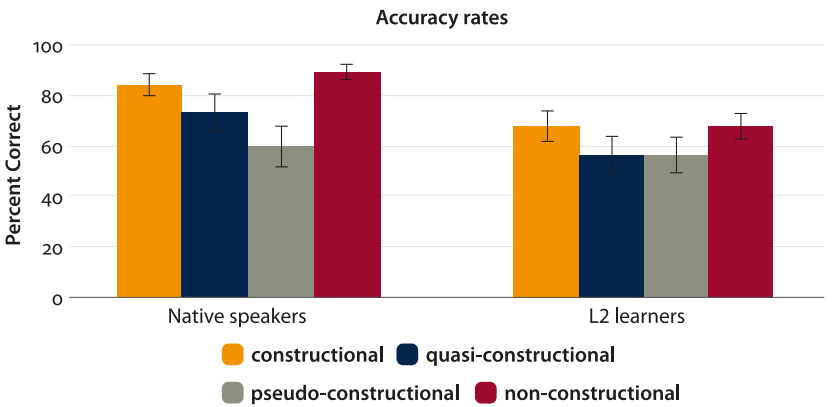


Figure 4. Accuracy rates for native speakers and L2 learners (the error bars represent standard deviations)

Notably, native speakers exhibited a regressive tendency in accuracy rates from constructional to pseudo-constructional idioms. On the other hand, for L2 learners, the accuracy rates between quasi-constructional and pseudo-constructional idioms were the same.

Post hoc Scheffé tests on construction type showed that the difference in accuracy rates among constructional idioms, quasi-constructional idioms, and pseudo-constructional idioms was significant. The difference in accuracy rates between constructional idioms and non-constructional idioms was not significant. This suggested that quasi-constructional idioms, which obey the morphological restrictions of idiomatic constructions, were understood less well by our participants compared with constructional idioms, presumably due to their novelty. This result is not particularly surprising, if one considers the importance that familiarity has in the processing of idioms (Schweigert 1991; Cronk et al. 1993; Giora & Fein 1999). Meanwhile, participants comprehended significantly better quasi-constructional idioms than pseudo-constructional idioms.

Table 4. Result of post hoc tests on construction type

	Type 1	**Type 2**	**Type 3**	**Type 4**
Type 1	–	***	***	n.s.
Type 2	***	–	*	***
Type 3	***	*	–	***
Type 4	n.s.	***	***	–

(Type 1: Constructional; Type 2: Quasi-constructional; Type 3: Pseudo-constructional; Type 4: Non-constructional.)
* < .05 ** < .01 *** $p < 0.001$

In order to examine whether the effect of construction is different for each group of subjects, we conducted a pair of post hoc Scheffé tests on construction type. The result is displayed in Table 5.

Table 5. Result of post hoc tests on construction type for native speakers and L2 learners

	Native speakers				**L2 learners**			
	Type 1	**Type 2**	**Type 3**	**Type 4**	**Type 1**	**Type 2**	**Type 3**	**Type 4**
Type 1	–	*	***	*	–	*	*	n.s.
Type 2	*	–	***	***	*	–	n.s.	***
Type 3	***	***	–	***	*	n.s.	–	**
Type 4	*	***	***	–	n.s.	***	**	–

(Type 1: Constructional; Type 2: Quasi-constructional; Type 3: Pseudo-constructional; Type 4: Non-constructional.)
* < .05 ** < .01 *** $p < 0.001$

It could be observed that, for native speakers, the difference among the four types of idioms was salient. The accuracy rate for quasi-constructional idioms, obeying semantic and morphological constraints of Chinese idiomatic construction, was lower than constructional idioms. Despite this fact, quasi-constructional idioms obtained a significantly higher accuracy rate than pseudo-constructional idioms, indicating that idioms violating the underlying semantic restriction of the construction were not well comprehended by native speakers. In light of these results, it is reasonable to assume that native speakers were aware of the underlying semantic and morphological constraints of idiomatic constructions. On the contrary, this awareness was absent among L2 learners, given that there was no difference in accuracy rate between quasi-constructional and pseudo-constructional idioms.

To investigate how semantic transparency, frequency, and language background interact, the accuracy rates for Type I trials (i.e. constructional idioms) were analyzed using a three-way repeated measures ANOVA, in which language (Mandarin speakers vs. L2 learners) served as the between-subject variable and semantic transparency and frequency as the within-subject variables. Given that there were only two levels of repeated measures, there was no need to conduct the Mauchly's test of sphericity. The assumption of Mauchly's sphericity was assumed.

The effect of semantic transparency, the frequency x language interaction, as well as the semantic transparency x frequency interaction were significant [semantic transparency: $F\ (1,41) = 140.8$, $p < 0.001$, $\eta^2 = 0.775$, frequency x language: $F\ (1,41) = 6.16$, $p = 0.017$, $\eta^2 = 0.131$, semantic transparency x frequency: $F\ (1,41) = 7.17$, $p = 0.011$, $\eta^2 = 0.15$]. In addition, the ANOVA also showed a significant main effect for language background, native speakers obtained higher accuracy rates (79.66% vs. 63.72%) and pairwise comparisons showed a reliable difference between native speakers and L2 learners [$F\ (1,41) = 23.89$, $p < 0.001$, $\eta^2 = .368$]. However, the effect of frequency, the semantic transparency x language interaction, as well as the three-way interaction were not significant [frequency: $F\ (1,41) = 0.368$, $p = 0.547$; semantic transparency x language: $F\ (1,41) = 0.009$, $p = 0.925$; language x semantic transparency x frequency: $F\ (1,41) = 0.192$, $p = 0.664$].

Because the interaction between frequency and language background was significant, we conducted a pair of two-way repeated measures ANOVA on each group of subjects, in which semantic transparency and frequency served as the within-subject variables. The accuracy rates for native speakers and L2 learners are displayed in Figure 5.

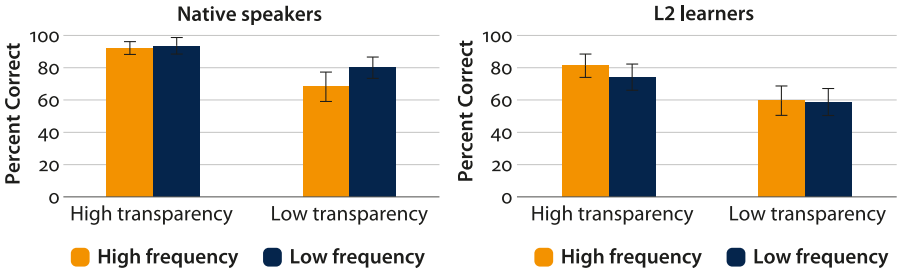


Figure 5. Accuracy rates for native speakers and L2 learners as a function of semantic transparency and frequency (the error bars represent standard deviations)

These results showed that semantically transparent constructions were comprehended more easily than semantically opaque constructions for both native speakers and L2 learners [native speakers: $F\,(1,22)=96.03$, $p<0.001$, $\eta^2=0.814$; L2 learners: $F\,(1,19)=52.64$, $p<0.001$, $\eta^2=0.735$]. More importantly, the effect of frequency was significant for native speakers, with low-frequency idiomatic expressions better comprehended than high-frequency ones in both semantically transparent and opaque conditions [native speakers: $F\,(1,22)=9.276$, $p=0.006$, $\eta^2=0.297$; L2 learners: $F\,(1,19)=1.082$, $p=0.311$]. The semantic transparency x frequency interaction was significant for native speakers, but not for L2 learners [native speakers: $F\,(1,22)=5.249$, $p=0.032$, $\eta^2=0.193$; L2 learners: $F\,(1,19)=2.33$, $p=0.143$].

The above results indicated that, for both group of speakers, semantically transparent idioms were comprehended more easily than semantically opaque idioms. The main effect of frequency as well as the semantic transparency x frequency interaction were significant only for native speakers, such that low-frequency idiomatic constructions yielded a higher accuracy rate than high-frequency idiomatic constructions in low transparency condition. This indicated that, when interpreting idiomatic constructions, the effect of frequency was different for native speakers and L2 learners. The question therefore arises as to why, for native speakers, low-frequency idiomatic constructions did not impede semantic judgment. We explore this problem further in the following section.

3.4.2 *Reaction times*

The RTs of correct responses for both groups were analyzed using a two-way repeated measures ANOVA, in which language (Mandarin speakers vs. L2 learners) served as the between-subject variable and construction type as the within-subject variable. Native speakers displayed a progressive increase in reaction time from constructional to pseudo-constructional idioms, as can be seen in Figure 6.

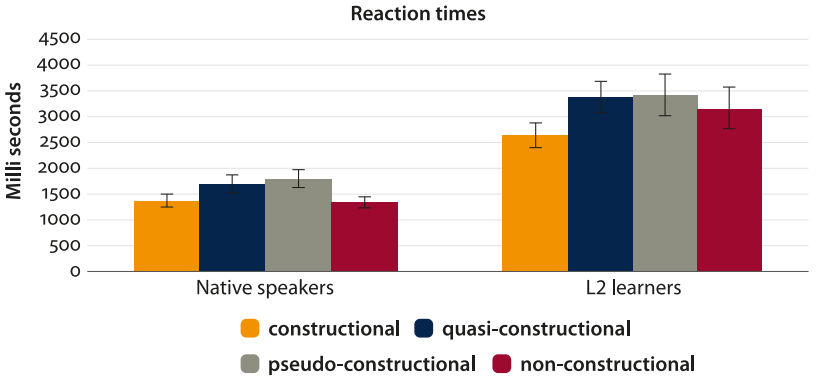


Figure 6. Reaction times for native speakers and L2 learners by construction type (the error bars represent standard deviations)

The assumption of sphericity had been violated [$\chi^2(5) = 20.721$, $p = .001$, $\varepsilon = .594$], therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .726$). The effect of construction type was significant [$F(2.18, 89.28) = 50.69$, $p < 0.001$, $\eta^2 = .553$]. In addition, the ANOVA also showed a significant main effect for language background, L2 learners took longer to respond than native speakers (1519 ms. vs. 3118 ms.) and pairwise comparisons showed a reliable difference between native speakers and L2 learners [$F(1, 41) = 124.25$, $p < 0.001$, $\eta^2 = .752$]. The construction type x language interaction was also significant. [$F(2.18, 89.28) = 9.96$, $p < 0.001$, $\eta^2 = .195$].

Post hoc Scheffé tests showed that the RTs of constructional idioms were significantly different compared with other types of idioms; however, the difference in RTs between quasi-constructional and pseudo-constructional idioms was not significant (see Table 6).

Table 6. Post hoc Scheffé comparisons of correct responses by type of construction

	Type 1	Type 2	Type 3	Type 4
Type 1	–	***	***	***
Type 2	***	–	n.s.	***
Type 3	***	n.s.	–	***
Type 4	***	***	***	–

(Type 1: Constructional; Type 2: Quasi-constructional; Type 3: Pseudo-constructional; Type 4: Non-constructional.)
* < .05 ** < .01 *** $p < 0.001$

To examine whether the effect of construction is different for each group of subjects, we conducted a pair of post hoc Scheffé tests on construction type. The result is displayed in Table 7.

Table 7. Post hoc Scheffé comparisons of correct responses by type of construction for native speakers and L2 learners

	Native speakers				L2 learners			
	Type 1	**Type 2**	**Type 3**	**Type 4**	**Type 1**	**Type 2**	**Type 3**	**Type 4**
Type 1	–	***	***	n.s.	–	***	***	***
Type 2	***	–	n.s.	***	***	–	n.s.	**
Type 3	***	n.s.	–	***	***	n.s.	–	**
Type 4	n.s.	***	***	–	***	**	**	–

(Type 1: Constructional; Type 2: Quasi-constructional; Type 3: Pseudo-constructional; Type 4: Non-constructional.)

* < .05 ** < .01 *** $p < 0.001$

The above result indicated that, for native speakers, the difference in RTs among constructional, quasi-constructional, and pseudo-constructional idioms was significant. The fact that quasi-constructional idioms obtained longer RTs than constructional idioms could be attributed to unfamiliarity with these invented idioms. However, quasi-constructional idioms still obtained significantly shorter RTs than pseudo-constructional idioms. No difference in RTs was observed between constructional and non-constructional idioms for native speakers. For L2 learners, the difference in RTs was not significant between quasi-constructional and pseudo-constructional idioms. However, the difference in RTs between constructional and non-constructional idioms was significant, with the latter obtaining longer RTs. This result will be further explored in § 4.

To investigate how semantic transparency, frequency, and language background interact, the RTs of Type I trials (i.e., [*qian-A-wan-B*] '1k-A-10k-B' idiomatic constructions) were analyzed using a three-way repeated measure ANOVA, in which language (Mandarin speakers vs. L2 learners) served as the between-subject variable and semantic transparency and frequency as the within-subject variables. Given that there were only 2 levels of repeated measures, the assumption of Mauchly's sphericity was assumed. The effect of semantic transparency and the frequency x language interaction were significant [semantic transparency: $F\ (1,41) = 27.73$, $p < 0.001$, $\eta^2 = 0.403$, frequency x language: $F\ (1,41) = 4.731$, $p = 0.035$, $\eta^2 = 0.103$]. However, the effect of frequency, the semantic transparency x language interaction, the semantic transparency x frequency inter-

action, as well as the three-way interaction were not significant [frequency: $F(1,41)=0.724$, $p=0.4$; semantic transparency x language: $F(1,41)=0.009$, $p=0.927$; semantic transparency x frequency: $F(1,41)=3.092$, $p=0.086$; language x semantic transparency x frequency: $F(1,41)=0.838$, $p=0.365$]. Pairwise comparisons showed a reliable difference between native speakers and L2 learners [$F(1,41)=103.05$, $p<0.001$, $\eta^2=.715$]. The result is displayed in Figure 7.

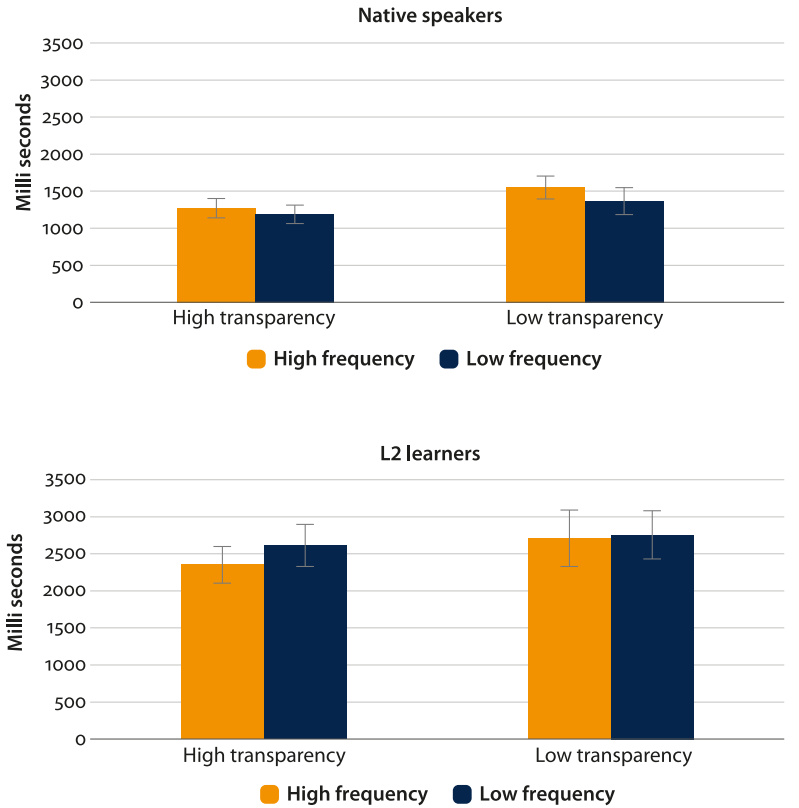


Figure 7. Mean RTs (in milliseconds) for native speakers and L2 learners as a function of semantic transparency and frequency (the error bars represent standard deviations)

Given that the frequency x language interaction was significant, we conducted a pair of two-way repeated measures ANOVA on each group of subjects, in which semantic transparency and frequency served as the within-subject variables. The results demonstrated that, for both native speakers and L2 learners, semantically transparent idiomatic constructions elicited a much shorter RT than semantically opaque idiomatic constructions [native speakers: $F(1,22)=38.26$, $p<0.001$, $\eta^2=0.635$; L2 learners: $F(1,19)=7.79$, $p=0.012$, $\eta^2=0.291$]. The effect

of frequency was significant for native speakers, with low-frequency idiomatic constructions eliciting a shorter RT than high-frequency idioms. This outcome echoed the result for accuracy rates, where low-frequency idioms yielded a higher accuracy rate. The effect of frequency was not significant for L2 learners [native speakers: $F\ (1,22)=9.77$, $p=0.03$, $\eta^2=0.346$; L2 learners: $F\ (1,19)=2.291$, $p=0.147$]. The semantic transparency x frequency interaction was not significant for both group of speakers [native speakers: $F\ (1,22)=0.917$, $p=0.349$; L2 learners: $F\ (1,19)=1.99$, $p=0.174$].

4. Discussion

This section first discusses how native speakers and L2 learners processed different types of constructions, and then clarifies the interaction among semantic transparency, frequency, and construction. The result from this study showed that native speakers exhibited a regressive tendency in accuracy rates from constructional to pseudo-constructional idioms, and displayed a progressive increase in reaction time from constructional to pseudo-constructional idioms. For native speakers, the difference in accuracy rates and that in RTs among these three types of constructions reached significance. Quasi-constructional idioms, obeying morphological and semantic constraints of the construction but unfamiliar to native speakers, received a lower accuracy rate and longer RTs compared with constructional idioms. The results provided support for the CH, which capitalizes on the role of frequency. Meanwhile, compared with pseudo-constructional idioms, quasi-constructional idioms received a higher accuracy rate and shorter RTs, indicating that native speakers were aware of the underlying semantic restriction of the construction. On the contrary, this awareness was absent among L2 learners, such that the difference in accuracy rates and that in RTs between quasi-constructional and pseudo-constructional idioms were not significant. This phenomenon seems to suggest that advanced L2 learners lack knowledge on the semantics as well as the morphological constraints associated with the construction. This may pose a challenge for the acquisition of idioms, given that idiomatic constructions are highly productive in Mandarin. Lacking the ability to make analogies means that L2 learners must recite every idiom they encounter, which is inefficient.

Native speakers processed constructional idioms faster than quasi-constructional and pseudo-constructional idioms due to familiarity with the former. However, a different picture emerged when we took semantic transparency into consideration. Regarding RTs, for both native speakers and L2 learners, semantically transparent idiomatic constructions elicited much shorter RTs than

semantically opaque idiomatic constructions. Thus, although idiomatic constructions were relatively transparent compared with non-constructional idioms, the result of the behavioral experiment showed they are not transparent in the same way, in that opaque idiomatic constructions yielded lower accuracy rates and longer RTs among both groups. Our result thus supports the IDH (Gibbs et al. 1989), positing that idioms are processed differently depending on their compositionality and semantic transparency.

Apart from the importance of semantic transparency, it can be observed that the effect of frequency is different for native speakers and L2 learners: whether in semantically transparent or opaque conditions, native speakers obtained higher accuracy rates on low-frequency idiomatic constructions; for L2 learners, high-frequency idiomatic constructions were better comprehended than low-frequency ones in semantically transparent condition. The result of RTs echoed the result for accuracy rates, with semantically transparent idioms obtaining shorter RTs for both groups, and low-frequency idiomatic constructions eliciting shorter RTs than high-frequency idioms for native speakers.

In sum, native speakers exhibited a higher accuracy rate and shorter RTs for low-frequency idiomatic constructions, therefore confirming our prediction regarding the interaction between frequency and construction. This result can be explained in terms of the framework of CG and by semantic satiation (Jakobovits 1962). Semantic satiation is a phenomenon in which a word loses its meaning following continuous repetition (Severance & Washburn 1907; Bassett & Warne 1919, among others). Several experiments show the semantic satiation effect occurring in various cognitive tasks such as rating words and figures presented repeatedly in a short time and in tasks involving verbally repeating words then grouping them into categories. For example, the continuous repetition of "7" should increase the time needed to add 7 and 4 (Jakobovits & Lambert 1962). Galmar (2012) reported a similar phenomenon in Mandarin, showing that the locus of satiation is semantic rather than orthographic. Such language-independent phenomena can be explained by the fact that, in the cortex, verbal repetition repeatedly arouses a specific neural pattern that corresponds to the meaning of the word. Rapid repetition causes repeated peripheral sensorimotor and central neural activation, inducing reactive inhibition (Jakobovits 1962).

In the current study, it was observed that, in both semantically transparent and opaque contexts, low-frequency idiomatic constructions were comprehended better and processed faster by native speakers than high-frequency ones. We therefore suggest that, because the construction of idioms involves the pairing of form and meaning, native speakers can interpret a low-token-frequency idiom based on the meaning conveyed by the construction. This means that, when processing low-frequency idioms, intervention of a construction with direct asso-

ciation between form and meaning facilitates semantic decoding. However, the recognition of high-frequency idioms was negatively affected by semantic satiation, which hindered native speakers' interpretation. For L2 learners, the RT results demonstrated that semantic transparency plays a more salient role than frequency. In the current study, although [*qian-A-wan-B*] '1k-A-10k-B' can be analyzed literally as "a large amount of" plus nouns or "repeated actions," some idioms bearing the same construction still have a figurative meaning that is semantically distinct from its literal meaning, thereby inhibiting L2 learners' comprehension. Some examples of semantically opaque idiomatic constructions that are particularly difficult for L2 learners are shown in Table 8:

Table 8. Examples of semantically opaque idiomatic constructions

Idiomatic constructions	Literal meaning	Figurative meaning	Frequency
a. 千門萬戶 *(Qian-men-wan-hu)*	[1k-door-10k-window]	A big house or lots of inhabitants	high
b. 千倉萬箱 *(Qian-cang-wan-xiang)*	[1k-warehouse-10k-box]	Massive storage of food	low
c. 千乘萬騎 *(Qian-cheng-wan-ji)*	[1k-ride-10k-bestride]	Lots of carriages and cavalry	high
d. 千匯萬狀 *(Qian-hui-wan-zhuang)*	[1k-flow-10k-state]	Lots of genres and types	low

Previous research has demonstrated the importance of semantic transparency for L2 learners in the acquisition of idiomatic expressions (Boers & Demecheleer 2001; Skoufaki 2008). Most Chinese idioms are derived from literary classics, where the meaning of an idiom is usually greater than the sum of the meanings conveyed by the four characters. With their musical rhythm, structural symmetry, and double planes of meaning, Chinese QIEs are frequently used in ordinary Chinese speech and writing. Consequently, acquiring competence in figurative language is particularly important. The result of RTs showed that L2 learners processed constructional idioms significantly faster than non-constructional idioms (see Figure 6 and Table 7). We therefore suggest that teaching Chinese QIEs to L2 learners need not be reduced to memorization, but that it can be enhanced by raising students' awareness of the semantics associated with a construction. This method may help L2 learners make analogies with new idioms containing the same construction.

Our understanding of idiomatic constructions in Mandarin can be increased through future research in the following two areas. On the one hand, given that Mandarin involves an ideographic writing system, the meaning of a character can be derived without knowing the pronunciation. The role of pictograms in idiom processing also merits further research. Meanwhile, Yang et al. (2016) examined the neural mechanisms involved in Chinese idiom comprehension and found that although both brain hemispheres are involved in idiom processing, they play different roles. It would be instructive to investigate the neural mechanisms involved in idiom processing from a constructional perspective to compare how the comprehension of idiomatic constructions differs from other types of idioms. Last but not least, being a pioneering study probing the processing and acquisition of Chinese idiomatic expressions under the framework of Construction Grammar, the data presented in the current research provide new insights, but is also limited in terms of the sample size. Therefore, the results and interpretations reported here will need to be validated in future research endeavors.

Acknowledgements

This research is part of the project granted by the Ministry of Science and Technology, Taiwan (Project number: 105-2420-H-002-008-MY2. Project name: The Classification of Chinese Four-Character Idioms and their Underlying Cognitive Mechanism–A Neural Consideration), and a continuation of the project "Brain Lateralization as a Benchmark in Learning Chinese" of the Cutting-Edge Steering Research Project granted by National Taiwan University (project members: Tai-li Chou, Shu-kai Hsieh, Chia-lin Lee, Chiarung Lu, and I-ni Tsai). The authors would like to thank all the research assistants who made the project possible.

Abbreviations

ANOVA	Analysis of Variance
CCL	The Center for Chinese Linguistics
CG	Construction Grammar
CH	Configuration Hypothesis
ERPs	Event-Related Potentials
ICC	Intra-Class Correlation
IDH	Idiom Decomposition Hypothesis
IRR	Inter-Rater Reliability
LRH	Lexical Representation Hypothesis
PPVT	Peabody Picture Vocabulary Test
QIEs	Quadrisyllabic Idiomatic Expressions
RT	Reaction Time

Appendix. Linguistic materials used for targets

1. Idiomatic constructions:

High frequency idiomatic constructions		**Low frequency idiomatic constructions**	
Semantically transparent		**Semantically transparent**	
千變萬化 *Qian-bian-wan-hua*	Constant permutations	千仇萬恨 *Qian-chou-wan-hen*	Deep hatred
千辛萬苦 *Qian-xin-wan-ku*	Much hardship	千歡萬喜 *Qian-huan-wan-xi*	Extremely happy
千年萬載 *Qian-nian-wan-zai*	A very long time	千刀萬剁 *Qian-dao-wan-duo*	A thousand cuts and myriad pieces
千軍萬馬 *Qian-jun-wan-ma*	A huge army	千變萬狀 *Qian-bian-wan-zhuang*	Have much variety
千山萬水 *Qian-shan-wan-shui*	A long and arduous journey	千端萬緒 *Qian-duan-wan-xu*	With many thoughts in mind
千叮萬囑 *Qian-ding-wan-zhu*	Exhort repeatedly	千轉萬變 *Qian-zhuan-wan-bian*	Constant permutations
千差萬別 *Qian-cha-wan-bie*	Completely different	千嬌萬態 *Qian-jiao-wan-tai*	Beautiful appearance and figure
千家萬戶 *Qian-jia-wan-hu*	Every family	千村萬落 *Qian-cun-wan-luo*	Many villages
千難萬險 *Qian-nan-wan-xian*	Many hazards and difficulties	千依萬順 *Qian-yi-wan-shun*	Always obedient
千真萬確 *Qian-zhen-wan-que*	Absolutely true	千變萬態 *Qian-bian-wan-tai*	Constant permutations
千恩萬謝 *Qian-en-wan-xie*	Thank again and again	千思萬慮 *Qian-si-wan-lü*	Think or consider repeatedly
千態萬狀 *Qian-tai-wan-zhuang*	Have much variety	千支萬派 *Qian-zhi-wan-pai*	Many (philosophical, martial, etc.) sects

High frequency idiomatic constructions		Low frequency idiomatic constructions	
Semantically opaque		**Semantically opaque**	
千峰萬壑 *Qian-feng-wan-huo*	Many mountains and valleys	千兵萬馬 *Qian-bing-wan-ma*	A huge army
千秋萬世 *Qian-qiu-wan-shi*	A long, long time	千回萬轉 *Qian-hui-wan-zhuan*	Go through ups and downs
千門萬戶 *Qian-men-wan-hu*	A big house or lots of inhabitants	千章萬句 *Qian-zhang-wan-ju*	Many phrases and articles
千呼萬喚 *Qian-hu-wan-huan*	Call repeatedly	千倉萬箱 *Qian-cang-wan-xiang*	Massive storage of food
千山萬壑 *Qian-shan-wan-huo*	Many mountains and valleys	千推萬阻 *Qian-tui-wan-zu*	Do everything to decline
千條萬縷 *Qian-tiao-wan-lü*	Many threads	千緒萬端 *Qian-xu-wan-duan*	Tangled thoughts
千乘萬騎 *Qian-sheng-wan-ji*	Lots of carriages and cavalry	千生萬死 *Qian-sheng-wan-si*	Very dangerous situation
千言萬語 *Qian-yan-wan-yu*	Have many words to say	千齡萬代 *Qian-ling-wan-dai*	Generation after generation
千刀萬剮 *Qian-dao-wan-gua*	A thousand cuts and myriad pieces	千匯萬狀 *Qian-hui-wan-zhuang*	Lots of genres and types
千頭萬緒 *Qian-tou-wan-xu*	So many thoughts in one's mind	千言萬說 *Qian-yan-wan-shuo*	Have many words to say
千絲萬縷 *Qian-tiao-wan-lü*	Tangled connections	千妥萬當 *Qian-tuo-wang-dang*	Very appropriate
千秋萬歲 *Qian-qiu-wan-sui*	A long, long time	千岩萬壑 *Qian-yan-wan-gu*	A group of mountain ranges

2. Quasi-idiomatic constructions

千思萬計	*Qian-si-wan-ji*	[1k-think-10k-plan]
千呼萬唱	*Qian-hu-wan-chang*	[1k-shout-10k-sing]
千笑萬鬧	*Qian-xiao-wan-nao*	[1k-laugh-10k-mischief]
千花萬香	*Qian-hua-wan-xiang*	[1k-bloom-10k-perfume]
千寶萬玉	*Qian-bao-wan-yu*	[1k-treasure-10k-jade]
千海萬江	*Qian-hai-wan-jiang*	[1k-sea-10k-river]
千德萬賢	*Qian-de-wan-xian*	[1k-ethic-10k-virtue]
千朝萬暮	*Qian-zhao-wan-mu*	[1k-morning-10k-night]
千仙萬妖	*Qian-xian-wan-yao*	[1k-fairy-10k-goblin]
千糟萬亂	*Qian-zao-wan-luan*	[1k-terrible-10k-messy]
千兵萬將	*Qian-bing-wan-jiang*	[1k-soldier-10k-general]
千快萬慢	*Qian-kuai-wan-man*	[1k-fast-10k-slow]
千載萬日	*Qian-zai-wan-ri*	[1k-year-10k-day]
千蟹萬蝦	*Qian-xie-wan-xia*	[1k-crab-10k-shrimp]
千聖萬言	*Qian-sheng-wan-yan*	[1k-saint-10k-speech]
千文萬墨	*Qian-wen-wan-mo*	[1k-essay-10k-ink]

3. Pseudo-idiomatic constructions

千舞萬遊	*Qian-wu-wan-you*	[1k-dance-10k-game]
千跳萬手	*Qian-tiao-wan-shou*	[1k-jump-10k-hand]
千硯萬山	*Qian-yan-wan-shan*	[1k-inkstone-10k-mountain]
千髮萬畫	*Qian-fa-wan-hua*	[1k-hair-10k-painting]
千石萬星	*Qian-shi-wan-xing*	[1k-stone-10k-star]
千竹萬酒	*Qian-zhu-wan-jiu*	[1k-bamboo-10k-wine]
千目萬杯	*Qian-mu-wan-pei*	[1k-eye-10k-cup]
千米萬葉	*Qian-mi-wan-ye*	[1k-rice-10k-leaf]
千時萬磨	*Qian-shi-wan-mo*	[1k-time-10k-grind]
千桌萬蝦	*Qian-zhuo-wan-xia*	[1k-table-10k-shrimp]
千犬萬里	*Qian-quan-wan-li*	[1k-dog-10k-mile]
千桶萬庭	*Qian-tong-wan-ting*	[1k-bucket-10k-court]
千求萬睡	*Qian-qio-wan-shui*	[1k-beg-10k-sleep]
千夫萬指	*Qian-fu-wan-zhi*	[1k-man-10k-point]
千斤萬笑	*Qian-jing-wan-ziao*	[1k-kilogram-10k-smile]
千魚萬木	*Qian-yu-wan-mu*	[1k-fish-10k-wood]

4. Not construction-based idioms

平分秋色	*Ping-fen-qiu-se*	To share on a fifty-fifty basis
九牛一毛	*Jiu-niu-yi-mao*	An iota from a vast quantity
三綱五常	*San-gang-wu-chang*	Three principles and five virtues
三人成虎	*San-ren-cheng-hu*	Repeated rumor becomes a fact
土崩瓦解	*Tu-beng-wa-jie*	Completely collapse
大器晚成	*Da-qi-wan-cheng*	Great minds mature slowly
翻雲覆雨	*Fan-yun-fu-yu*	As changeable as clouds and rain
郎才女貌	*Lang-cai-nü-mao*	An ideal couple
有教無類	*You-jiao-wu-lei*	Teach equally regardless of background
牛頭馬面	*Niu-tou-ma-mian*	Goblins living in hell
滄海桑田	*Cang-hai-sang-tian*	Time brings great changes to the world
龍飛鳳舞	*Long-fei-feng-wu*	Lively and vigorous flourishes in calligraphy
目眩神迷	*Mu-xuan-shen-mi*	Be dazzled and stunned
見賢思齊	*Jian-xian-si-qi*	To emulate those better than oneself
虎頭蛇尾	*Hu-tou-she-wei*	Fine start and poor finish
頭頭是道	*Tou-tou-shi-dao*	Clear and logical

References

Ahrens, Kathleen. 1995. *The mental representation of verbs*. San Diego: University of California at San Diego. (Doctoral dissertation.)

Bassett, M. F. & Warne, C. J. 1919. On the lapse of verbal meaning with repetition. *The American Journal of Psychology* 30(4). 415–418. https://doi.org/10.2307/1413679

Berman, Ruth A. & Slobin, Dan I. (eds.). 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale: Lawrence Erlbaum Associates.

Boers, F. & Demecheleer, M. 2001. Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal* 55(3). 255–262. https://doi.org/10.1093/elt/55.3.255

Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733. https://doi.org/10.1353/lan.2006.0186

Bybee, Joan L. & Hopper, Paul J. (eds.). 2001. *Frequency and the emergence of linguistic structure* (Typological Studies in Language 45). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45

Bybee, Joan L. & Thompson, Sandra. 2000. Three frequency effects in syntax. In Juge, Matthew L. & Moxley, Jeri L. (eds.), *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society, 14–17 February 1997: General session and parasession on pragmatics and grammatical structure (BLS 23)*, 378–388. Berkeley: Berkeley Linguistics Society.

Cacciari, Cristina. 2014. Processing multiword idiomatic strings: Many words in one? *The Mental Lexicon* 9(2). 267–293. https://doi.org/10.1075/ml.9.2.05cac

Cacciari, Cristina & Glucksberg, Sam. 1994. Understanding figurative language. In Gernsbacher, Morton Ann (ed.), *Handbook of psycholinguistics*, 447–477. San Diego: Academic Press.

Cacciari, Cristina & Tabossi, Patrizia. 1988. The comprehension of idioms. *Journal of Memory and Language* 27(6). 668–683. https://doi.org/10.1016/0749-596X(88)90014-9

Chen, Changlai. 2012. *Xiandai Hanyu leiguding duanyu yanjiu* [Quasi-fixed structures in Modern Mandarin]. Shanghai: Xuelin Press.

Cicchetti, Domenic V. 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 6(4). 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cieślicka, Anna. 2006. Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research* 22(2). 115–144. https://doi.org/10.1191/0267658306sr263oa

Croft, William & Cruse, D. Alan. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803864

Cronk, Brian C. & Lima, Susan D. & Schweigert, Wendy A. 1993. Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research* 22(1). 59–82. https://doi.org/10.1007/BF01068157

Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24(2). 143–188. https://doi.org/10.1017/S0272263102002024

Ellis, Nick C. 2006. Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27(2). 164–194. https://doi.org/10.1093/applin/aml015

Ellis, Nick C. 2013. Construction grammar and second language acquisition. In Hoffmann, Thomas & Trousdale, Graeme (eds.), *The Oxford handbook of construction grammar*, 365–378. New York: Oxford University Press.

Ellis, Nick C. & Laporte, Nadine. 1997. Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In de Groot, Annette M.B. & Kroll, Judith F. (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 53–83. Mahwah: Lawrence Erlbaum Associates.

Fillmore, Charles J. & Kay, Paul & O' Conner, Mary Catherine. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538. https://doi.org/10.2307/414531

Fitzpatrick, Tess & Wray, Alison. 2006. Breaking up is not so hard to do: Individual differences in L2 memorization. *The Canadian Modern Language Review* 63(1). 35–56. https://doi.org/10.3138/cmlr.63.1.35

Galmar, Bruno. 2012. *An investigation of the nature of verbal satiation in Chinese*. Tainan: National Cheng Kung University. (Doctoral dissertation.)

Gibbs, Raymond W., Jr. & Nayak, Nandini P. & Cutting, Cooper. 1989. How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28(5). 576–593. https://doi.org/10.1016/0749-596X(89)90014-4

Giora, Rachel & Fein, Ofer. 1999. On understanding familiar and less-familiar figurative language. *Journal of Pragmatics* 31(12). 1601–1618. https://doi.org/10.1016/S0378-2166(99)00006-5

Goldberg, Adele E. 1992. The inherent semantics of argument structure: The case of the English ditransitive construction. *Cognitive Linguistics* 3(1). 37–74. https://doi.org/10.1515/cogl.1992.3.1.37

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. & Casenhiser, Devin M. & Sethuraman, Nitya. 2005. The role of prediction in construction-learning. *Journal of Child Language* 32(2). 407–426. https://doi.org/10.1017/S0305000904006798

Hovav, Malka Rappaport & Levin, Beth. 2008. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics* 44(1). 129–167. https://doi.org/10.1017/S0022226707004975

Hu, Ying-Hsueh & Fong, Yu-Ying. 2010. Obstacles to CM-guided L2 idiom interpretation. In De Knop, Sabine & Boers, Frank & De Rycker, Antoon (eds.), *Fostering language teaching efficiency through cognitive linguistics* (Applications of Cognitive Linguistics (ACL) 17), 293–316. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110245837.293

Jakobovits, Leon A. 1962. *Effects of repeated stimulation on cognitive aspects of behavior: Some experiments on the phenomenon of semantic satiation*. Montreal: McGill University. (Doctoral dissertation.)

Jakobovits, Leon A. & Lambert, Wallace E. 1962. Semantic satiation in an addition task. *Canadian Journal of Experimental Psychology* 16(2). 112–119. https://doi.org/10.1037/h0083238

Kaschak, Michael P. & Glenberg, Arthur M. 2000. Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language* 43(3). 508–529. https://doi.org/10.1006/jmla.2000.2705

Liu, Shuxin. 1990. *Hanyu miaoxie cihuixue* [Chinese descriptive lexicology]. Beijing: The Commercial Press.

Liu, Te-hsin & Lu, Chiarung & Tsai, I-ni & Su, I-wen. 2017. Idiomaticity and regularity of Chinese four-character idioms: The case of [*yi*-X-#-Y]. (Paper presented at the 29th North American Conference on Chinese Linguistics (NACCL-29), Rutgers University, New Brunswick, 16–18 June 2017.)

Liu, Yuehua. 2005. Chengyu yu duiwai Hanyu jiaoxue [Idiomatic expressions and teaching Chinese as a foreign language]. In Teng, Shou-hsin (ed.), *Ershiyi shiji Huayu jigou yingyun celue yu jiaoxue guoji yantaohui lunwenji* [Proceedings for the 2005 International Symposium on Operational Strategies and Pedagogy for Chinese Language Programs in the 21st Century], 298–303. Taipei: Mandarin Training Center, National Taiwan Normal University.

Lu, L. & Liu, H.-H. 1998[1994]. *Xiuding Bi-Pao-De tuhua chihui ceyan* [Peabody picture vocabulary test – revised]. Taipei: Psychological Publishing. (Mandarin Chinese version, an earlier version is published in 1994.)

McGraw, Kenneth O. & Wong, S. P. 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1(1). 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Meyer, Charles F. & Grabowski, Roger & Han, Hung-Yul & Mantzouranis, Konstantin & Moses, Stephanie. 2003. The world wide web as linguistic corpus. In Leistyna, Pepi & Meyer, Charles F. (eds.), *Corpus analysis: Language structure and language use* (Language and Computers: Studies in Practical Linguistics 46), 241–254. Amsterdam: Rodopi. https://doi.org/10.1163/9789004334410_014

Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press.

Nunberg, Geoffrey & Sag, Ivan A. & Wasow, Thomas. 1994. Idioms. *Language* 70(3). 491–538. https://doi.org/10.1353/lan.1994.0007

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge: The MIT Press.

Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841323

Schneider, Walter & Eschman, Amy & Zuccolotto, Anthony. 2002. *E-prime user's guide*. Pittsburgh: Psychology Software Tools.

Schweigert, Wendy A. 1991. The muddy waters of idiom comprehension. *Journal of Psycholinguistic Research* 20(4). 305–314. https://doi.org/10.1007/BF01074283

Severance, Elizabeth & Washburn, Margaret Floy. 1907. The loss of associative power in words after long fixation. *The American Journal of Psychology* 18(2). 182–186. https://doi.org/10.2307/1412411

Skoufaki, Sophia. 2008. Investigating the source of idiom transparency intuitions. *Metaphor and Symbol* 24(1). 20–41. https://doi.org/10.1080/10926480802568448

Su, Lily I-Wen. 2002. Why a construction – That is the question! *Concentric: Studies in Linguistics* 28(2). 27–42.

Swinney, David A. & Cutler, Anne. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior* 18(5). 523–534. https://doi.org/10.1016/S0022-5371(79)90284-6

Tabossi, Patrizia & Fanari, Rachele & Wolf, Kinou. 2009. Why are idioms recognized fast? *Memory & Cognition* 37(4). 529–540. https://doi.org/10.3758/MC.37.4.529

Tsou, Benjamin K. 2012. Idiomaticity and classical traditions in some East Asian languages. In Manurung, Ruli & Bond, Francis (eds.), *Proceedings of the 26th Pacific Asia Conference on Language, Information and Computation (PACLIC 26)*, 39–55. Depok: Faculty of Computer Science, Universitas Indonesia.

Wang, Li. 1990. Idiom and classical illusion. In Wang, Li (ed.), *Wang Li wenji* [Collection of Wang Li], vol. 11, 653–672. Jinan: Shandong Educational Press.

Wang, Ruojiang. 2001. Liuxuesheng chengyu pian wu youyin fenxi-cidian pian [Analysis of idiom errors made by foreign students]. *Jinan daxue Huawen xueyuan xuebao* [TCSOL Studies: Journal of College of Chinese Language and Culture of Jinan University] 2001(3). 28–35.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Xu, Sheng-huan. 2006. Proximity and complementation: Studies of the formation mechanism of idioms from cognitive point of view (II). *Jinan daxue Huawen xueyuan xuebao* [TCSOL Studies: Journal of College of Chinese Language and Culture of Jinan University] 2016(3). 33–41.

Yang, Jie & Li, Ping & Fang, Xiaoping & Shu, Hua & Liu, Youyi & Chen, Lang. 2016. Hemispheric involvement in the processing of Chinese idioms: An fMRI study. *Neuropsychologia* 87. 12–24. https://doi.org/10.1016/j.neuropsychologia.2016.04.029

Zhang, Hui & Yang, Yiming & Gu, Jiexin & Ji, Feng. 2013. ERP correlates of compositionality in Chinese idiom comprehension. *Journal of Neurolinguistics* 26(1). 89–112. https://doi.org/10.1016/j.jneuroling.2012.05.002

Zhan, Weidong & Guo, Rui & Chen, Yirong. 2003. *CCL yuliaoku jiansuo xitong* [The CCL Corpus of Chinese Texts: 700 million Chinese Characters, the 11th Century B.C.–present]. Beijing: Center for Chinese Linguistics, Peking University. (http://ccl.pku.edu.cn:8080/ccl_corpus) (Accessed 2021-03-15.)

Zwitserlood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9(3). 341–368. (*Language and Cognitive Processes* is currently known as *Language, Cognition and Neuroscience*.) https://doi.org/10.1080/01690969408402123

Authors' addresses

Te-hsin Liu (corresponding author)
Graduate Program of Teaching Chinese as a Second Language
National Taiwan University
No. 1, Sec. 4, Roosevelt Road
Taipei 10617
Taiwan

tehsinliu@ntu.edu.tw

Publication history

Date received: 13 April 2018
Date accepted: 9 May 2019