華語文教學研究 12.3:1-30, 2015

The Acquisition of Semi-fixed Idiomatic Expressions in Chinese*

Li LIU
Faculty of English Education,
Guangdong University of Foreign Studies

Hin Tat CHEUNG
Department of Linguistics and Modern Languages,
Hong Kong Institute of Education

Abstract

Semi-fixed templates with open slots such as *bu-A-bu-B* (not-A-not-B), *qian-A-wan-B* (hundred-A-thousand-B), *A-yan-B-yu* (A-words-B-utterances) abound in Chinese idioms and are used productively in discourse (Ching 1964; Zhang 2002; Ji 2007). Yet the productivity of these schematic idioms is not unrestrictive since possible A, B morphemes that fill in the open slots must conform to certain structural and semantic constraints. How children come to learn the constraints on their productivity raises an interesting issue in language acquisition. This pilot study investigates young and adult native speakers' competence on the partial productivity of Chinese schematic idioms in terms of three related constructions -- XAXB, XAYB and AXBY (in which X and Y are the fixed templates), and explores the factors that influence their acceptability judgments. Specifically, this study tested the effect of input frequency, structural complexity, internal semantic relation and chunk effect of open morphemes with 4th graders and adults in a graded Grammaticality Judgment task. Results showed even 4th graders began to form abstract generalizations on the legitimate words that can enter the open slots of schematic expressions. The results provided empirical support for the psychological reality of constructional claims for idiomatic expressions under Chinese context.

Keywords: semi-fixed expressions, schematic idioms, partial productivity

* This work is in part supported by MOE (Ministry of Education in China) Youth Project of Humanities and Social Sciences in 2015 given to the first author.

1. Introduction

Idioms and idiomaticity are universal phenomenon in human languages (Moon 1998). Idiomatic expressions have long been held as lexically fixed and their meaning as non-compositional. They are viewed as idiosyncratic and peripheral in language system, and thus restricted to the mental lexicon (Chomsky 1981; Katz and Postal 1963). Yet the existence of semi-fixed idiomatic expressions, termed as formal or schematic idioms (Fillmore et al. 1988), poses serious challenges for this lexicon-grammar separation. Schematic idioms are not fully lexicalized and have open slots for variants in actual use, as in *the X-er, the Y-er*; *No X, no Y*. Apparently they cannot simply be assigned as fixed lexical items due to their productive behavior.

Such schematic expressions are particularly abundant in Chinese idioms (Ching 1964; Ji 2007; Shi 1979; Zhang 2002). More than 270 schematic patterns are identified in Jiang (2000), which are capable in creating novel phrases, as shown in the following examples:

a. *bu-A-bu-B* (not-A-not-B):
e.g. bu-da-bu-xiao (not-big-not small); *bu-nan-bu-nv* (not-male-not-female)
b. *qian-A-wan-B* (hundred-A-thousand-B):
e.g. qian-xin-wan-ku (hundred-difficulty-thousand-hardship)
c. *A-yan-B-yu* (A-words-B-utterances):
e.g. hua-yan-qiao-yu (flowery-words-artful-utterances)

The productivity of these schematic expressions, however, is not unlimited. Mature speakers tend to have intuitions on the legitimate morphemes that can enter the open slots and they are capable of figuring out the ultimate interpretation of an expression despite variations of the filled morphemes. This raises interesting learnability issues in language acquisition: How do children come to learn the restrictive productivity? How do they figure out the interpretation of new variations? Yet few studies so far have addressed the issue of productivity in Chinese schematic idioms, let alone its learning process in language acquisition.

2. Construction grammar and idiomatic expressions

As mention above, the existence of schematic idioms poses serious challenges for the componential model in generative tradition (Croft and Cruse 2004). These expressions have grammatical structures which are fully or partially productive (as in *let alone*; *what's X doing in Y*). Yet the meaning may not be canonically mapped onto the components in surface structure. That is, their meaning and structure are not predictable from general syntactic and linking rules. Should these expressions be represented in the lexicon or as syntactic principles? The assignment of these expressions into the lexicon as in the generative grammar does not account for their productivity properly. In addition, mature speakers tend to have intuitions on which are legitimate selections in the open slots and which are not. A proper grammar should include this knowledge and provide an equal account for the productive behaviors of these schematic idioms as with other regular grammatical structures.

2.1 Schematic idioms as constructions

Construction Grammar zooms in on this issue and argues that such schematic expressions cannot simply be listed in the lexicon. They should be viewed as form-meaning pairings on their own terms (Croft and Cruse 2004; Hilper 2014; Taylor 2002; Wulff 2008). Construction Grammar defines the basic units in a linguistic system as constructions, a symbolic unit of a form-meaning pairing (Goldberg 1995, 2006; Langacker 1991). "Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist" (Goldberg 2006). Thus defined, idioms are cited as a typical example of constructions due to its unpredictability in meaning. Analysis on expressions such as *let alone, what's X doing Y* demonstrated that those traditionally thought irregular expressions can actually be quite productive, and rule-based within the framework of its own construction (Fillmore et al. 1988; Kay and Fillmore 1999). The syntactic and semantic properties of schematic idioms, therefore, must be directly associated with the construction.

The constructional approach also provides a better account for the

learnability issue arising from idiom variations. That is, how could speakers know which lexical items are legitimate in the open slots as instances of the schema and which are not? Most speakers are not deliberately taught on all the variations allowed in a schematic idiom. Yet they tend to have some intuitions upon the acceptability of given instances. Constructional understanding of schematic idioms provides a theoretical foundation in accounting for this constrained productivity and how it can be learned in accordance.

2.2 Chinese schematic idioms

Chinese idioms, traditionally termed as *Chengyu*, are composed mostly of four-syllabic expressions, e.g. *qing-mei-zhu-ma* (green-plum-bamboo-horse), *xiao-ti-da-zuo* (small-topic-big-article). They have been traditionally held as lexicalized chunks that should be memorized item by item. Few studies have systematically addressed the productivity of Chinese idioms (Ching 1964; Ji 2007; Shi 1979; Zhang 2002). Yet semi-fixed templates are abundant in idiom formation. More than 270 patterns are identified in the dictionary by Jiang (2000). Moreover, schematic idioms are widely used and function for various purposes in communication. The central reason for this wide use might rest on their productivity and plasticity. They are not fully specified in lexemes and are capable in forming novel phrases to fit new and emergent contexts. Schematic idioms therefore deserve particular attention not only because of their abundance in language, but also because of speakers' productive use.

How to represent these schematic idioms in grammar? Constructional approach is equally applicable in Chinese context. The present study argues that Chinese schematic idioms should also be represented as form-meaning pairings with its own characteristics. In this way, the productive behavior and the learning of these schematic idioms can be accounted for.

Chen (2001) examined three semi-fixed idiomatic expressions in Chinese: *A-lai-A-qu* (A-come-A-go), *bu-A-bu-B* (not-A-not-B) and *bu-A-er-B* (not-A-but-B), and argued for a constructional claim to account for their unique features and productivity. The *A-lai-A-qu* (A-come-A-go) expressions semantically imply *a repetition of action A*, which is not predictable from the combination of the

constituents alone. Lexical items filling in the open slots of *A-lai-A-qu* are mostly verbs denoting repeatable and durable actions, such as *fei* (fly), *zou* (walk), which are determined by the aspectual feature of the whole pattern. The whole expressions can serve as transitive verb phrases but the object has to be placed in front of the whole phrase, which is different from the normal post-verbal position of objects in Chinese. The above evidence supports that *A-lai-A-qu* is a form-meaning pairing with its own characteristics. In addition, Chen (2001) pointed out the constructional meaning of *bu-A-bu-B* (not-A-not-B) actually varies in accordance with the semantic relationship between A and B, which can be synonyms, near synonyms or antonyms. In *bu-A-er-B* (not-A-but-B), the selection of the lexical items in the open slots is conditioned by the constructional meaning of the whole expression "*to carry out the action Y without the condition X*". Based on the evidence in syntactic and semantic features, Chen (2001) concluded that these expressions should be viewed as constructional idioms, instead of independent lexical items with arbitrary meanings. Ji (2007) and Gan (2008) proposed similar views in understanding the partially productive behavior of schematic idioms.

Yet whether this constructional claim for schematic idioms is also psychologically real during learning process is an issue that has not been properly addressed in current literature. The present study will take three related Chinese schematic constructions: XAXB, XAYB and AXBY (in which X and Y stand for the fixed morphemes; while A and B are the morphemes in the open slots) as targets to illustrate their productive use and corresponding constraints and will then empirically testify whether young and adult native speakers do form any generalizations on their productive use.

3. Distributional analysis of three related schematic constructions

XAXB, XAYB and AXBY are three schemas used very productively in Chinese idiom formation. In XAXB, the fixed morpheme X is identical; A, B are morpheme selections that fill the position of open slots. In XAYB, the fixed morphemes are varied as X and Y, while the position of fixed morphemes and open slots are reversed in AXBY. Different morphemes of X and Y make up

various schematic types in each schema, which generates various numbers of instantiations with different selection of A, B morphemes in the open slots. See the listed examples in Table1. The selection of A, B morphemes in each schema reflects both the structural and semantic requirements imposed on the lexemes in the open slots. In the following sections, we will look at the productivity in each schema and its corresponding constraints.

Table 1: Examples of schematic types and instances

schema	Schematic types	Instances in each type
XAXB	*bu-A-bu-B* (not-A-not-B)	*bu-zhi-bu-jue* (not-aware) *bu-nan-bu-nü* (not-male-not-female) *bu-tu-bu-kuai* (no-confess-no-comfort)
	da-A-da-B (big-A-big-B)	*da-qi-da-luo* (big-up-big-down) *da-bo-da-lan* (big-waves) *da-lu-da-f*a (big-road-big-money)
	yu-A-yu-B (more-A-more-B)	*yu-yan-yu-lie* (more-growing-the-worse) *yu-xia-yu-da* (the-longer-the-more) *yu-sheng-yu-gao* (more-rising-the-higher)
XAYB	*qian-A-wan-B* (hundred-A-thousand-B)	*qian-xin-wan-ku* (hundred-hardship-thousand-difficulties) *qian-tou-wan-xu* (hundred-cues-thousand-signs) *qian-jia-wan-hu* (hundred-houses-thousand-families)
	tian-A-di-B (sky-A-earth-B)	*tian-luo-di-wang* (sky-net-earth-trap) *tian-chang-di-jiu* (sky-lasting-earth-enduring)
	bai-A-bu-B (hundred-A-not-B)	*bai-si-bu-jie* (hundred-think-not-understand) *bai-wen-bu-yan* (hundred-questioning-not-bored)

AXBY	*A-yan-B-yu* (A-words-B-utterances)	*hua-yan-qiao-yu* (flowery-words-cunning-utterances) *tian-yan-mi-yu* (sweet-words-honey-utterances)
	A-chan-B-duan (A-long-B-short)	*yang-chang-bi-duan* (enhance-long-avoid-short) *shuo-chang-dao-duan* (say-long-talk-short)
	A-er-B-zhi (A-in-order-to-B)	*fen-er-zhi-zhi* (to-rule-by-dividing) *zong-er-yan-zhi* (to-speak-by-summarizing)

3.1 Productivity and constraints in XAXB schema

In terms of structure, XAXB schema is less complex than XAYB and AXBY due to its identical fixed morpheme. There are more than 43 schematic types identified in XAXB schema based on dictionaries and corpus data (see some common types in Table 1 above). This paratactic structure in XAXB expressions entails two types of relation in meaning: coordinate and subordinate. The majority of schematic types find XA and XB as coordinate in meaning; whereas a total of 8 schematic types express subordinate relations: conditional (as in *bu-da-bu-zhao* no-beating-no-confess) or resultative (*bai-fa-bai-zhong* hundred -shooting-hundred-targeted).

Structurally, in coordinate types, A and B morphemes must be of the same grammatical category, as in *bu-pang-bu-shou* (not-fat-not-thin), *you-mai-you-mai* (some-buy-some-sell). In subordinate types, A and B form a Verb-Result sequence, as in *bu-tu-bu-kuai* (no-confess-no-comfort), *lü-zhan-lü-bai* (always-fight-always-defeated). It would be unacceptable to say **lü-zhan-lü-chi* (always-fight-always-eat) as B morpheme (*eating*) does not denote a necessary result of the action (*fighting*) expressed in A morpheme.

Semantically, A and B morphemes belong to the same conceptual field. In *yi-yan-yi-xing* (each-utterance-each-act), *bu-leng-bu-re* (not-cold-not-warm), the open lexemes *yan-xing* (utterance-act), *leng-re* (cold-warm) all relate to

each other in the same semantic filed. Thus it is not acceptable to say: **yi-yan-yi-hua* (each-utterance-each-flower) or **bu-leng-bu-gao* (not-cold-not-tall), as the variations in the open slots do not refer to the same concept.

In addition, the meaning of A, B must be compatible with the meaning of the schema they reside in. For example, *chui-lei* (boasting-exaggerating) naturally co-occurs with *big-A-big-B* but not *small-A-small-B*. Notice that coercion happens in which the semantic compatibility overrides general structural principles: canonically inseparable compounds in Chinese, such as *yan-yu* (words-utterances), *si-hao* (a-shred-of), *bo-lan* (ripples-waves) can be separated and form schematic expressions of *bu-yan-bu-yu* (no-words-no-utterances), *yi-si-yi-hao* (a-tiny-bit) and *da-bo-da-lan* (big-ups-big-downs). The reason why these lexical chunks can be noncanonically separated in these expressions is primarily for discourse purpose: to emphasize or to reinforce the meaning of the overall phrases.

Finally, the use of numbers in schematic expression is conventional and not replaceable, as in *yi-wu-yi-shi* (one-five-one-ten), *bu-san-bu-si* (not-three-not-four). The order of A, B morphemes is also conventional. For example, it is more frequent to say *wu-fu-wu-mu* (no-father-no-mother) than *wu-mu-wu-fu* (no-mother-no-father).

Thus, it can be seen that the lexical selection of A and B in XAXB schema is an intricate interaction of requirements simultaneously operating at structural, semantic and conventional levels.

On checking the statistical distribution of the top 100 most frequent expressions in CCL corpus (Center for Chinese Linguistics PKU), we find that among these 100 expressions, 97 are coordinate instances while only 3 are subordinate expressions (*yu-yan-yu-lie* get-worse-and-worse, *yue-kuai-yue-hao* the-fast-the-better, *duo-lao-duo-de* more-work-more-pay). The paratactic structure in XAXB expressions makes it more compatible in expressing coordinate relations. When a subordinate meaning is entailed, there seems to be a mismatch between syntax-semantics: paratactic structure with conditional or resultative semantics, as in *bu-shuo-bu-ming* (no-explanation-no-understanding) (Culicover and Jackendoff 1999). This might partly explain why the number of

subordinate expressions is disproportionally lower in XAXB schema. At the same time, it can be predicted that the learning of subordinate expressions in XAXB schema would be much more difficult due to this syntax-semantic mismatch.

Among the 97 coordinate ones, 47% are phrases in which A and B combine as compounds, as in *quan-xin-quan-yi* (whole-heart-whole-mind), *ge-hang-ge-ye* (various-trades-various-industries). These compounds generally resist separation in use but noncanonically separated here to fill the open slots of the expressions. This echoes Bybee and Hopper's observation (2001) that high token frequency is usually associated with irregularities in a language. Empirical study also reported that compound and binding words are recognized more quickly and easier to access (Liu and Peng 1997; Myers 2006). Based on these findings, it is predicted that the chunk effect of A-B combination may facilitate the decoding of XAXB expressions.

3.2 Productivity and constraints of XAYB schema

XAYB is the most productive schema among the three introduced here as more than 46 schematic types of XAYB expressions are identified in this study, see some common types in Table 1 above. The varied X, Y increase the pattern's capability in combining with possible lexical selections, especially its potential in expressing subordinate relations. For instance, the pattern of *bu-A-er-B* (not-A-but-B) expresses an adversative relation: *B happens despite the non-occurrence of A*. Hence *bu-qing-er-lai* (not-being-invited-but-come) means *coming even without being invited*; schematic type of *wu-A-ke-B* (no-A-to-B) expresses the meaning that *B cannot be achieved because of the unavailability of A*. So *wu-mi-ke-cui* (no-rice-to-cook) means *there is no rice for cooking*.

Broadly speaking, the productivity of XAYB depends on similar factors affecting that of XAXB. For the coordinative type, the fixed X and Y must be of identical grammatical categories and they are either similar or opposite in meaning, so are the variable morphemes, A and B. The combinations of X-Y and A-B may form lexical chunks, such as *tian-di* (sky-earth), *luo-wang* (net-trap) in

tian-luo-di-wang (sky-net-earth-trap) or stand as independent free morphemes, as *jia* (house) and *hu* (family) in *qian-jia-wan-hu* (hundred-house-thousand-family).

For the subordinate type, X and Y are mainly functional words and can be extended as clauses that express various relations, such as conditional, concessive, adversative, etc. A and B in the open slots are morphemes of different parts of speech depending on the specific relations denoted in X and Y. For example, *bai-A-bu-B* (hundred-A-not-B) can be extended as a concession clause: *in spite of A happening a hundred times, B still does not occur*. A-B combinations in this pattern are required to form a *Verb-Result* sequence to conform to the meaning of the schema. Thus *bai-kan-bu-yan* (hundred-see-not-bored) means *in spite of seeing it a hundred times, the person does not show any boredom*. For another example, *wu-A-bu-B* (no-A-not-B) expresses a non-exclusion relation: *There is no A that is not B-ed; to B every A*. Accordingly *wu-e-bu-zuo* (no-crime-not-do) means *to commit all kinds of crimes*. A-B combinations in this pattern make up a reversed *Verb-Object* sequence.

In sum, the selection of A, B constitutes in the open slots is not arbitrary at all, but is determined by the specific structural and semantic requirement of the schematic type that they reside in.

The statistical distribution of the top 100 most frequent XAYB expressions is also checked in CCL corpus. There are 50 coordinate expressions, and also 50 subordinate ones. Nevertheless, the total token of coordinate expressions is slightly higher than subordinate ones (22356: 21807). Moreover, among the 50 coordinate instances, 30 are found with chunk selection in the open slots, as in *yi-qing-er-chu* (as-clear-as-one-and-two) and *wu-yan-liu-se* (of-various-colors). It is conjured that XAYB coordinate expressions may also be learned better than subordinate ones since they are semantically less complex and a bit more frequent in usage.

3.3 Productivity and constraints of AXBY schema

Almost every instance of AXBY belongs to the coordinate type except one: *A-er-B-zhi* (A-in-order-to-B). In *A-er-B-zhi*, the fixed *er* is a functional word with no content meaning; *zhi* (it/him/her) is a third person pronoun in classical

Chinese. The whole pattern delivers a meaning: *to B by means of A*. The inserted A, B usually form a *modifier-verb* sequence. For example, the expression *fen-er-zhi-zhi* (divide-to-rule-it) means *to rule by the means of getting it divided.* All the rest are coordinate expressions. See Table 1 above for some common types of AXBY expressions.

For the coordinate types, the fixed lexemes X and Y belong to the same grammatical categories: often nouns, adjectives, locatives, but seldom verbs. The selection of A, B in the open slots must be grammatically agreeable to X and Y, which usually function as heads in the subparts of AX and BY. The combinations of XY and AB usually stand as lexical chunks, such as *qian-hou* (front-rear) and *si-xiang* (think-over) in *qian-si-hou-xiang* (think-before-and-after). Other X, Y and A, B may still be independent free morphemes, such as *zui* (mouth) and *she* (tongue) in *qi-zui-ba-she* (seven-mouths-eight-tongues). Semantically X and Y have similar or opposite meanings. The meaning of A, B morphemes should be compatible with the template meaning they are in.

Therefore, in both subordinate and coordinate AXBY expressions, the selections of A and B in the open slots are also far from arbitrary, but an interactive process of the structural and semantic requirements of a specific template.

When the statistical distribution of the 100 frequent AXBY instances is checked in CCL corpus, coordinate expressions predominate in both overall token and type frequency. The rate of coordinate and subordinate expressions is 90:10, which is within expectation considering that there is only one subordinate type in AXBY schema. Furthermore, among the 90 coordinate phrases, 46 have chunk selections of open morphemes, as in *jing-tian-dong-di* (shake-heaven-startle-earth) and *da-jie-xiao-xiang* (high-streets-back-lanes). It is conjured that the lower type and token frequency may lead to difficulty in the learning of A-er-B-zhi expressions compared with other coordinate phrases.

To summarize, the distributional and statistical analysis of XAXB, XAYB and AXBY schemas reveals that schematic idioms are not simply lexical items with arbitrary meanings. Their productivity and the selection of A, B lexemes in the open slots are subject to a process of intricate structural and semantic

interactions and exhibit sub-regularities on their own. It is better to view them as constructional idioms; that is, form-meaning pairings with their own characteristics to fully account for their partially productive behaviors. The major findings in the above analysis can be summarized as follows:

1. XAXB schema is structurally simpler due to its identical fixed morphemes; non-XAXB schemas; that is, XAYB and AXBY, are structurally more complex.
2. Coordinate instances predominate in each schematic construction, especially in XAXB and AXBY.
3. There is a strong consistency between high token frequency and the noncanonical separation of AB chunks in the open slots in each schematic construction, especially in XAXB.

4. Usage-based language learning theory

In Constructional approach, the mental grammar of a native speaker is hypothesized to consist of "a network of schematic and substantive constructions" (Hoffmann and Trousdale 2013). The learning of these constructions at various levels is inherently usage-based (Barlow and Kemmer 2000; Goldberg 2006, 2013; Langacker 1988). Acquiring a language is thus not the discovery of the syntactic-semantic mappings, but the mastery of various linguistic symbols and constructional schemas. Children gradually build up constructions from the actual use of specific utterances. In this process, children resort to human general cognitive mechanisms such as culture learning, analogy making and structure combining to acquire language (Tomasello 2000, 2003). Generalizations of language are simply formed on the basis of regularities among concrete exemplars, as in the way categorizations are formed in non-linguistic areas. Factors such as frequency in their language experience are regarded as playing a key role in forming representations of linguistic units (Bybee 2003; Ellis 2002, 2009; Goldberg et al. 2004). Other factors that are reported to contribute to construction learning include complexity, consistency (or contingency), chunking, age, etc.

Studies based on spontaneous speech analysis and elicited experiments on

novel words have shown that learners are somehow highly sensitive to the statistical distributions from very early stage of language development (Brooks and Tomasello 1999; Casenhiser and Goldberg 2005; Goldberg et al. 2007; Lieven and Tomasello 2008). Frequency (both token and type) can be manipulated in a way to produce different learning outcomes. For example, Casenhiser and Goldberg (2005) concluded that one verb with high token frequency better facilitated construction learning than each verb with the balanced token frequency; Childers and Tomasello's study (2001) found that providing both pronouns and concrete nouns in the subject and object slots facilitated the generalizations of the transitive constructions more than providing only concrete nouns did; Goldberg et al. (2007) proved that initial skewed input, rather than random skewed input were efficient in promoting construction learning and generalizations; Ambridge et al. (2006) also confirmed a significant distributed learning effect for the learning of past-tense object-cleft construction.

Frequency alone cannot explain all construction learning phenomenon. Construction complexity is another factor that interacts during the course of learning. Diessel and Tomasello's study (2005) reported that though the frequency of S/A-relatives (relatives with a subject gap) is equal to or even higher than P-relatives (relatives with a direct object gap) in children-directed speech, S/A relatives showed few errors than P-relatives in children's repeating performance. This is because S/A relatives are syntactically less complex due to their similarity to simple non-embedded sentences.

Usage-based approach allows an ontogenetically developmental account for language learning. It proposes that children start learning item-based, concrete constructions, and then gradually abstract grammar and create inheritance hierarchies of constructions with general cognitive mechanisms (Tomasello 2000). Literature on idiom learning reported that children began to show comprehension of idiomatic expressions as young as 5 (Cain et al. 2009; Caillies and Sourn-Bissaoui 2006; Laval 2003) and their figurative competence does not reach adult stage even until adolescence (Nippold 1998). This shows learning of idiomatic expressions is also a gradual and continuous developmental process.

5. The experimental studies

This pilot study is to focus on the learning of the partial productivity of three related Chinese schematic idioms: XAXB, XABY and AXBY construction for L1 speakers from a cognitive perspective. The productivity of semi-fixed schematic idioms is highly selective and constrained. Apparently learners cannot memorize each instance by rote. How speakers build up the exact generalization that produces only grammatical expressions, no more and no less, poses a particularly interesting issue here. How and when could children come to learn the productivity of these expressions and the inherent structural and semantic constraints? What factors may facilitate the generalization process? These general questions arise as the central issues in the present study.

5.1 Hypotheses

Based on our findings in the distributional and statistical analysis and the usage-based learning principles, four factors are expected to affect the learning of partial productivity in schematic idioms: input frequency, structural complexity, inherent semantic relations and chunking effect of lexical selections. Expressions with higher frequency will be easier to learn; so do the XAXB expressions due to their simpler structure. Coordinate expressions are predicted to be easier than subordinate ones. And the chunk effect of A, B open morpheme may facilitate learning. To empirically test the effect of these factors for both young and adult speakers, the following hypotheses are raised:

- **Hypothesis 1**: Productivity of XAXB schema is learned earlier and better than the other two schemas for both young and adult speakers.
- **Hypothesis 2**: Productivity of coordinate instances is better learned than subordinate ones within each schema for both young and adult speakers.
- **Hypothesis 3**: For coordinate types in each schema, productivity of instances with AB chunk selections is learned earlier and better for both young and adult speakers.

Two age levels are targeted in the present study: 4th graders aged 9-10 years old, and the adult group. Current studies on learning of idiomatic expressions

mainly focus on 9-10 year-olds, such as Cain et al. (2009), Gibbs (1987), Hsieh and Hsu (2010), Laval (2003), Levorato and Cacciari (1992), Nippold and Duthie (2003), etc. Therefore, targeting on 4th graders makes the results of the present study easily comparable with current research.

The Acceptability Judgment Task is used to elicit speakers' knowledge on partial productivity in schematic idioms. For Acceptability judgments, Liker-scale judgment task is a common and valid instrument in testing linguistic knowledge with both adult and young speakers (Ambridge et al. 2008; Theakston 2004; Weskott and Fanselow 2008). A 4-point Liker-scale judgment task, ranging from *strongly acceptable* to *strongly unacceptable*, is used in this study. Since idiomaticity and schematicity are both gradient features, such scale can catch the naturalness and nativelikeness of schematic idioms more than the simple binary distinction of acceptable or unacceptable. Participants are required to indicate how much they find the given expression acceptable by ticking the responding item on the scale: 1. strongly unacceptable; 2. unacceptable; 3. acceptable; 4. strongly acceptable. For the younger participants, a more kids-friendly version of instructions is given. For example, two smiling faces signals that the expression is strongly acceptable; one smiling face signals acceptable; one crying face means unacceptable; and a cross signals completely unacceptable (See Appendix 1 and 2).

To test the three hypotheses, two experimental studies are conducted. Study One addresses Hypothesis1 and 2; Hypothesis 3 is tested in Study Two. The following part gives detailed description on each study.

5.1 Study One

5.1.1 Participants

20 participants in both young and adult age groups took part in the Grammaticality Judgment Task. Young participants are randomly chosen from a class of 4th graders in a local primary school in Guangzhou and their average age is 9.8. Adults are voluntary students in a university in Guangzhou, with an average age of 21.

5.1.2 Variables

There are 3 independent variables in Study One: structural complexity, inherent semantic relations and age. So this is a 2 (complexity) × 2 (semantic relations) × 2 (age) factorial design. Input frequency is also included as a control variable. We will see the effect of the 3 variables in the learning of both frequent and less-frequent schematic expressions. Frequency of an expression is determined by its record in CCL corpus. High frequent items are defined as the top 10% in their corresponding type. Low frequency refers to the items which rank within the bottom 10%. Table 2 lists the levels for each variable.

Table 2: Levels of Independent Variables in Study One

Variables		**Levels**	
Independent Variables	Complexity	2 levels	Structurally simple: *XAXB schema*
			Structurally more complex: *XAYB and AXBY schemas*
	Semantic relations	2 levels	Coordinate relations, e.g. *tong-xin-tong-de* (same-heart-same-moral)
			Subordinate relations, e.g. *bu-tu-bu-kua*i (no-confess-no-comfort)
	Age	2 levels	4th graders
			adults
Control Variable	Frequency	2 levels	Frequent expressions
			Less-frequent expressions

5.1.3 Materials

There are altogether 80 items in the Grammaticality Judgment task. 40 are acceptable ones; the other 40 are novel expressions. Novel expressions refer to these expressions in which the lexical selections are not legitimate candidates in the open slots thus they never occur in natural settings. For example, w*u-ju-wu-shu* (no-restriction-no-hindrance) is an acceptable expression, while *wu-ju-wu-jin* is a novel expression since *ju* (restriction) and *jin* (caution) are not legitimate morphemes in the schema of *wu-A-wu-B* and hence unacceptable. For the 40

acceptable ones, half are frequent and half are less-frequent. Within each frequency range, there are 10 XAXB expressions, 5 of coordinate and 5 of subordinate relation; and 10 non-XAXB expressions, again 5 of coordinate and 5 of subordinate relation. See Appendix 1 for all testing items.

The actual performance consists of two sessions, with each containing 40 judgments. There is a break of 15 minutes between the sessions. 4th graders spend about 30 minutes on the task in total and adults spend about 15-20 minutes to complete it.

5.2 Study Two

5.2.1 Participants

20 participants were recruited in each age level. The 4th graders were students randomly chosen from another local school in Guangzhou, whose average age was 9.7. Adults were another group of voluntary college students from the same university as in Study One. Their average age was 21.

5.2.2 Variables

There are 3 independent variables in Study Two: structural complexity, chunk effect of open morphemes and age. So this is also a 2 (complexity) × 2 (chunk effect) × 2 (age) factorial design. Input frequency is also included as a control variable, as in Study One. Table 3 lists the levels for each variable.

Table 3: Levels of Independent Variables in Study Two

Variables		**Levels**	
Independent Variables	Complexity	2 levels	Structurally simple: *XAXB schema*
			Structurally more complex: *XAYB and AXBY schemas*
	Chunk effect	2 levels	Chunk A, B morphemes, e.g. *tian-yan-mi-yu* (sweet-words-honey-utterances)
			Non-chunk A, B morphemes, e.g. *qian-jia-wan-hu* (hundred-houses-thousand-families)

	Age	2 levels	4th graders
			adults
Control Variable	Frequency	2 levels	Frequent expressions
			Less-frequent expressions

5.2.3 Materials

There are altogether 80 items in the Grammaticality Judgment task, half acceptable ones; half novel expressions. For the 40 acceptable ones, half are frequent and half are less-frequent. Within each frequency range, there are 10 XAXB expressions, 5 with chunk A, B morphemes and 5 with non-chunk A, B morphemes; and 10 non-XAXB expressions, again 5 with chunk and 5 with non-chunk A, B morphemes. See Appendix 2 for detailed items.

The actual performance again consists of two sessions, with each containing 40 judgments. There is a 15-minute break between the sessions. It took the young participants a total of 30 minutes to complete the task while adults spend 15-20 minutes on it.

6. Results and discussion

6.1 Results of Study One

The Average ratings for High Frequent (HF), Low Frequent (LF) and Novel Expressions (NE) for the 4th graders and adults are listed in Table 4.

Table 4: Average ratings for High Frequent, Low Frequent and Novel Expressions

Participants	**HF**	**LF**	**NE**
4th graders	3.41	2.78	2.30
adults	3.75	3.15	1.95

A repeated 3 (frequency) × 2 (age) ANOVA was conducted. Results revealed that there was a significant main effect of frequency ($F(1, 38) = 437.75$, $p < .001$). In addition, the interaction between frequency and age was also significant ($F(1, 38) = 32.83$, $p < .001$). From the average rating in Table 4, we can see that for adults, HF expressions scored higher than LF expressions, which in turn rated much higher than Novel expressions. On the other hand, 4th graders' judgment

ratings on LF and Novel expressions were quite similar though the difference between HF and Novel expressions was quite obvious. This suggests that even 9-year olds are able to tell which morphemes are allowed in the open slots of schematic idioms and which are not. But they still cannot tell exactly the acceptability difference between unfamiliar and unacceptable expressions.

Participants' ratings also differed significantly according to age ($F(1, 38) = 10.67$, $p < .001$). In addition, 4th graders gave much lower ratings than adults on HF, LF expressions. But their ratings on Novel expressions were still a lot higher. This means 9 year-olds are still in the developmental process of learning schematic idioms.

Table 5 listed the average ratings in terms of complexity, semantic relation for different age groups. 2 (complexity) × 2 (semantic relation) × 2 (age) analyses of variance were subsequently carried out under both the HF and LF conditions.

Table 5: Average ratings in accordance to complexity, semantic relation for 4th graders and adults

Frequency	**Participants**	**XAXBCo**	**XAXBSub**	**Non-XAXBCo**	**Non-XAXBSub**
HF	4th graders	19.25	16.10	15.90	16.95
	adults	19.40	18.50	19.90	19.70
LF	4th graders	13.25	13.10	15.80	13.50
	adults	14.15	16.75	16.70	16.30

Under the HF condition, the effect of complexity was not found as significant. But the interaction of complexity by age was significant ($F(1, 38) = 24.18$, $p < .001$): being of the XAXB pattern facilitated judgments of 4th graders, but not the adults'. It might be that the HF expressions are more entrenched in adults' representation and tend to be more autonomous as chunks in processing regardless of the structural complexity involved. Moreover, the main effect of semantic relation was also significant ($F(1, 38) = 9.36$, $p < .005$): being coordinate in relation facilitated the judgments of both groups. The two-way interaction of complexity by semantic relation was also significant ($F(1, 38) = 8.44$, $p < .01$):

4th graders scored highest with coordinate XAXB expressions, while adults rated coordinate non-XAXB expressions the highest. It seems adults can benefit more from more complex structures in making their judgments. More complex structures contain more information clues and it seems adults are better at analyzing the internal structure while making the judgments. Both 4th graders and adults scored the lowest with the subordinate XAXB expressions. This implies that the mismatch between paratactic structure and subordinate semantic relations in XAXB expressions did cause much difficulty for their judgments. The three way interaction was significant ($F(1, 38) = 4.31$, $p < .005$) and the between-subject effect of age was also significant ($F(1, 38) = 53.87$, $p < .001$).

Under the LF condition, the results of the analysis showed that main effect of complexity was significant for both young and adult participants ($F(1,38) = 21.13$, $p < .001$). Both age groups scored higher with non-XAXB expressions. The effect of semantic relation was not significant. But the interaction of semantic relation and age was significant ($F(1, 38) = 14.99$, $p < .001$): 4th graders scored higher with coordinate expressions while adults gave higher ratings to subordinate ones. The number of subordinate types is much lower than coordinate types, which might partly explain why adults performed better with subordinate expressions. The interaction of complexity and semantic relation was also significant ($F(1, 38) = 9.29$, $p < .005$). It can be seen that both young and adult participants rated coordinate non-XAXB expressions much higher.

In sum, Hypothesis 1 was only partly confirmed. The effect of complexity was significant for 4th graders when judging frequent expressions: being structurally simpler facilitated the judgments of 4th graders'. One the other hand, when dealing with less frequent schematic expression, it was the ratings on non-XAXB pattern that scored higher for both adults and 9 year-olds. The reason for this unexpected result might be due to the fact that the varied morphemes of X and Y in non-XAXB phrases provides more information and linguistic clues for decoding. It also implies that speakers may rely more on the internal structure analysis of schematic idioms when dealing with low frequent expressions (Gibbs 1987).

Hypothesis 2 was also partly confirmed. The effect of semantic relations

was significant with high frequent expressions: being coordinate in semantics facilitated the judgments of both 4th graders and adults, which confirms to Hypothesis 2. But there was no general main effect of semantic relations in judging the acceptability of less frequent expressions. As reported above in the results, 4th graders rated higher with coordinate expressions and adults scored higher with subordinate ones, which might balance the result and explain why no effect of semantic relation was found.

6.2 Results of Study Two

Table 6 lists the average ratings for HF, LF and Novel expressions by both age groups. Results of a repeated 3 (frequency) × 2 (age) ANOVA showed that there were significant main effect of frequency ($F(1,38) = 395.71$, $p < .001$). Ratings on HF expressions were higher than ratings on LF expressions, which were in turn higher than those on Novel expressions. The two-way interaction between age and frequency was also significant ($F(1,38) = 24.04$, $p < .001$). This was caused again, by 4th graders' similar judgments on LF and Novel expressions, as we can see from the average scores in Table 6. This echoes results in Study One and suggests 4th graders' were likely to reject an expression if it appeared unfamiliar.

Ratings on each type of expressions also differed significantly according to age ($F(1,38) = 41.48$, $p < .001$). Again this shows that 4th graders' competence on the partial productivity of schematic idioms is still in the developing stage.

Table 6: Average ratings for HF, LF and Novel Expressions

Participants	**HF**	**LF**	**NE**
4th graders	3.43	2.39	2.33
adults	3.79	3.15	2.49

To confirm the chunk effect in making judgments, two 2 (complexity) × 2 (chunk effect) × 2 (age) analyses of variance were further conducted based on Table 7.

For HF expressions, results showed that chunk effect was statistically significant ($F(1, 38) = 26.40$, $p < .001$). Expressions with chunk lexical selections

scored much higher for both adults and 9 year-olds. There was no significance of complexity. But the interaction of complexity and age was marginally significant ($F(1, 38) = 4.81$, $p < 0.5$): 4th graders gave higher rating to XAXB expressions, as can be seen from Table 7. Complexity by chunk was not significant. The three way interaction was not significant either. The between-subject effect of age was significant ($F(1, 38) = 30.09$, $p < .001$)

Table 7: Average ratings in terms of complexity and chunk effect for 4th graders and adults

Frequency	**Participants**	**XAXB Chunk**	**Non-XAXB Chunk**	**XAXB nonChunk**	**Non-XAXB nonChunk**
HF	4th graders	18.55	18.05	16.50	15.45
	adults	19.30	19.80	18.15	18.50
LF	4th graders	13.10	14.80	10.65	9.30
	adults	16.00	18.75	17.15	11.10

For LF expressions, chunk effect was also proved to be statistically significant for both age groups ($F(1, 38) = 79.03$, $p < .001$). This indicates the effect of chunk was strong and consistent as it was observed with both high and low expressions, for both young and adult speakers. There was also a main effect of complexity ($F(1, 38) = 6.82$, $p < .01$): adults rated XAXB expressions significantly higher than non-XAXB ones while ratings of 4th graders are not significantly different. Two-way interaction of complexity by age ($F(1, 38) = 10.44$, $p < .001$), complexity by chunk ($F(1, 38) = 148.99$, $p < .001$) were both significant. For both age groups, non-XAXB expressions with chunk selections rated the highest; whereas non-XAXB expressions with non-chunk selections had the lowest ratings. The three-way interaction was significant ($F(1, 38) = 35.08$, $p < .001$). The effect of age was also significant ($F(1, 38) = 66.45$, $p < .001$).

To sum up, Study Two largely supports Hypothesis 3: the consistent facilitating effect of chunk selection was witnessed for both 9 year-olds and adults in judging the acceptability of both high frequent and low frequent expressions. This chunk effect was more obvious with non-XAXB expressions, which might due to their distributional feature that both X, Y and A, B often

combine as chunks.

Study Two also confirmed the main effect of complexity in judging low frequent phrases: adults rated XAXB expressions much higher than non-XAXB ones. This sets contrast to the findings in Study One, which showed that non-XAXB expressions scored higher. In Study One, XAXB expressions includes both coordinate and subordinate expressions; while there was no subordinate testing items in Study Two. This contrast in results implies that the chunk effect may be particularly benefiting in judging the XAXB expressions with coordinate relations.

7. Conclusion

Chinese idiomatic expressions deserve our particular attention in the research of figurative languages due to its salient linguistic feature that the sophisticated figurative meaning is located in a highly institutionalized four-word pattern. So far few empirical studies were conducted regarding the learning of Chinese idioms (Hsieh and Hsu 2010; Liu and Cheung 2014). Even fewer studies have addressed the semi-fixed schematic idioms in Chinese, especially on its productive use and the learning of its productivity in acquisition.

The present study attempted to fill this research gap by exploring the learning on the productivity of Chinese schematic idioms via a graded Grammaticality Judgment task with young and adult native speakers. Schematic idioms do not specify all the lexical elements in them and have open slots for variations. We have demonstrated in the distributional analysis that schematic idioms are not arbitrary combinations, but exhibit specifications on the selection of open morphemes. They need to be treated as constructional idioms with their own characteristics (Croft and Cruise 2004; Fillmore et al. 1988; Goldberg 1995, 2006; Taylor 2002). Speakers, especially young kids, need to identify the requirements imposed on the open morphemes and generalize the sub-regularities across schemas.

The results of the present study have demonstrated that native speakers, even as young as 9 years old, can tell what kind of morphemes are acceptable in the open slots for a certain schema. That means, school-age children, as early as 9

years old have already started to form generalizations on which are legitimate instances for a certain schematic pattern and which are not, even for those unfamiliar variations. During this process, frequency of exposure plays an important role in forming generalizations, as it was observed in both studies that HF expressions scored much higher than LF expressions for both young and adult speakers. In addition, the linguistic features of the schematic constructions also have impact during the generalizing process, including the structural complexity of the expression involved, the internal semantic relation entailed in the expression, and chunk selection of its open morphemes. Those factors further interact with age in shaping the acceptability judgment of schematic idioms in the following manner:

---High frequent expressions are judged better than low frequent ones for both young and adult speakers.

---Structural simplicity does not necessarily indicate better accuracy in judgments, especially for adult speakers.

---Chunk selections in the open slots consistently facilitate the judgment of schematic idioms for both young and adult speakers.

These results revealed both general learning principles such as input frequency and specific linguistic features are responsible for the learning of productivity of schematic constructions. The results clearly demonstrate that schematic idioms are not learned as individual lexical items; but as form-meaning parings with their own sub-regular rules. Young speakers do form generalizations regarding the legitimacy of an expression. Thus the study provides empirical evidence for the psychological reality of constructional claim of schematic idioms during language learning.

Our findings also lend indirect evidence for the compositional view during idiom processing: figurative expressions are processed in a compositional manner rather than represented as big lexical chunks, even for Chinese idioms which happen to be highly institutionalized four-word-long patterns. Speakers as young as 9 years old began to make internal analysis on the constituents to make inference on the requirements imposed on each one. It suggests idioms are not

learned by rote memory as long words, but rather in a compositional way, that is, processing of individual components automatically occurs during idiom interpreting (Gibbs et al. 1989; Levorato and Cacciari 1992; Nippold 1998; Titone and Connine 1999).

References

Ambridge, B., J. M. Pine, C. F. Rowland, and C. R. Young. 2008. The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgments of argument-structure overgeneralization errors. *Cognition* 106.1:87-129.

Ambridge, B., A. L. Theakston, E. V. Lieven, and M. Tomasello. 2006. The distributed learning effect for children's acquisition of an abstract syntactic construction. *Cognitive Development* 21.2:174-193.

Barlow, M., and S. Kemmer. 2000. *Usage-based Models of Language*. California: CSLI publications.

Brooks, P., and M. Tomasello. 1999. How young children constrain their argument structure constructions. *Language* 75:720-738.

Bybee, J. 2003. Mechanisms of change in grammaticization: the role of frequency. *The Handbook of Historical Linguistics*, eds. by B.D. Joseph & R.D. Janda, 602-623. Oxford: Blackwell.

Bybee, J., and P. Hopper. (eds.) 2001. *Frequency and the emergence of linguistic structure*, Vol. 45. Amsterdam: John Benjamins Publishing.

Caillies, S., and S. Le Sourn-Bissaoui. 2006. Idiom comprehension in French children: a cock-and-bull story. *European Journal of Developmental Psychology* 3.2:189-206.

Cain, K., A. S. Towse, and R. S. Knight. 2009. The development of idiom comprehension: an investigation of semantic and contextual processing skills. Journal of Experimental Child Psychology 102.3:280-298.

Casenhiser, D. M., and A. E. Goldberg. 2005. Fast mapping between a phrasal form and meaning. *Developmental Science* 8.6:500-508.

Chen, K. Y. 2001. *A Construction Grammar Approach to Chinese Four-Character*

Idiomatic Expressions. Taipei: National Taiwan University MA thesis.

Childers, J., and M. Tomasello. 2001. The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* 37:739-748.

Ching, E. 1964. *Four-Syllable Expressions in Chinese*. Columbus: Ohio State University, Research Foundation.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Croft, W., and D. A. Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Culicover, P. W., and R. Jackendoff. 1999. The view from the periphery: the English comparative correlative. *Linguistic Inquiry* 30.4:543-571.

Diessel, H., and M. Tomasello. 2005. A new look at the acquisition of relative clauses. *Language* 81.4:882-906.

Ellis, N. C. 2002. Frequency effects in language acquisition: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24:143-188.

Ellis, N. C. 2009. Optimizing the input: frequency and sampling in usage-based and form-focused learning. *The Handbook of Language Teaching*, eds. by Michael H. Long, and Catherine J. Doughty, 139-158. Oxford: Blackwell.

Fillmore, C. G., P. Kay, and M. C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* 64.3:501-538.

Gan, L. H. 2008. Constructional meaning and pragmatic tendency of bu4-A-bu4-B pattern. *Rhetoric Learning* 2:56-60.

Gibbs, R. W. 1987. Linguistic factors in children's understanding of idioms. *Journal of Child Language* 14:569-586.

Gibbs, R. W., and N. P. Nayak. 1989. Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology* 21:100-138.

Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago and London: The University of Chicago Press.

Goldberg, A. E. 2006. *Construction at Work: The Nature of Generalization in Language.* New York: Oxford University Press.

Goldberg, A. E. 2013. Constructionist approaches. *The Oxford Handbook of*

Construction Grammar, eds. by T. Hoffmann, and G. Trousdale, 15-31. Oxford: Oxford University Press.

Goldberg, A. E., D. M. Casenhiser, and N. Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15.3:289-316.

Goldberb, A. E., D. M. Casenhiser, and T. R. White. 2007. Constructions as categories of language. *New Ideas in Psychology* 25:70-86.

Hilper, M. 2014. *Construction Grammar and Its Application to English*. Edinburgh: Edinburgh University Press

Hoffmann, T., and G. Trousdale. (eds.) 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.

Hsieh, S.C-Y, and C-C. N. Hsu. 2010. Idiom comprehension in Mandarin-speaking children. *Journal of Psycholinguist Research* 39:505-522.

Ji, M. 2007. From etymology to modern phraseology: a corpus-based study of structural variants of Chinese idioms in naturally-occurring contexts. *Journal of Chinese Language and Computing* 17.2:67-82.

Jiang, D. W. 2000. *Dictionary of Chinese Disyllable Expressions.* Beijing: Beijing Language and Culture University Press.

Katz, J. J., and P. Postal. 1963. The semantic interpretation of idioms and sentences containing them. *MIT Research Laboratory of Electronics Quarterly Progress Report* 70:275-282.

Kay, P., and C. J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the *What's X doing Y* construction. *Language* 75:1-33.

Langacker, R. W. 1988. A usage-based model. *Topics in Cognitive Linguistics* 50:127-163.

Langacker, R. W. 1991. *Foundations of Cognitive Grammar*, Vol. 2: *Descriptive Application*. Stanford, CA: Stanford University Press.

Laval, V. 2003. Idiom comprehension and metapragmatic knowledge in French children. *Journal of Pragmatics* 35.5:723-739.

Levorato, M. C., and C. Cacciari. 1992. Children's comprehension and production of idioms: the role of context and familiarity. *Journal of Child Language* 19:415-433.

Lieven, E, and M. Tomasello. 2008. Children's first language acquisition from a

usage-based perspective. *Handbook of Cognitive Linguistics and Second Language Acquisition*, eds. by Peter Robinson, and Nick C. Ellis, 168-196. New York and Landon: Routledge.

Liu, L., and H. T. Cheung. 2014. Acquisition of Chinese quadra-syllabic idiomatic expressions: effects of semantic opacity and structural symmetry. *First Language* 34.4:336-353

Liu, Y., and D. L. Peng. 1997. Meaning access of Chinese compounds and its time course. *Cognitive Processing of Chinese and Related Asian Languages*, ed. by Chen, H. C., 219-232. Hong Kong: The Chinese University Press.

Moon, R. 1998. *Fixed Expressions and Idioms in English.* Oxford: Clarendon Press.

Myers, J. 2006. Processing Chinese compounds: a survey of the literature. *The Representation and Processing of Compound Words*, eds. by G. Libben, and G. Jarema, 71-95. Oxford: Oxford University Press.

Nippold, M. A. 1998. *Later Language Development: The School-age and Adolescent Years*. Austin, TX: Pro Ed.

Nippold, M. A., and J. K. Duthie. 2003. Mental imagery and idiom comprehension: a comparison of school-age children and adults. *Journal of Speech, Language, and Hearing Research* 46:788-799.

Shi, S. 1979. *Hanyu Chengyu Yanjiu* [*Chinese Idioms*]. Chengdu: Sichuan People Publishing House.

Taylor, J. R. 2002. *Cognitive Grammar.* Oxford: Oxford University Press.

Theakston, A. L. 2004. The role of entrenchment in children's and adults' performance on grammaticality judgment tasks. *Cognitive Development* 19.1:15-34.

Titone, D. A., and C. M. Connine. 1999. Descriptive norms for 171 idiomatic expressions: familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity* 9:247-270.

Tomasello, M. 2000. Do young children have adult syntactic competence? *Cognition* 74:209-253.

Tomasello, M. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition.* Cambridge, Massachusetts and London: Harvard University Press.

Weskott, T., and G. Fanselow. 2008. Variance and informativity in different

measures of linguistic acceptability. *Proceedings of the 27th West Coast Conference on Formal Linguistics* 431-439. Somerville, MA: Cascadilla Press.

Wulff, S. 2008. *Rethinking Idiomaticity: A Usage-based Approach*. London: Continuum.

Zhang, B. 2002. *New Modern Chinese.* Shanghai: Fudan University Press.

[審查：2014.11.4 修改：2014.12.24 接受：2015.2.5]

劉麗
Li LIU
香港新界大埔中心露屏路 10 號 香港教育學院語言學與現代語系 B2-35
Department of Linguistics and Modern Language Studies
Hong Kong Institute of Education
10 Lu Ping, Taipo, HK
kristenliul@hotmail.com

張顯達
Hin Tat CHEUNG
香港新界大埔中心露屏路 10 號 香港教育學院語言學與現代語系
Department of Linguistics and Modern Language Studies
Hong Kong Institute of Education
10 Lu Ping, Taipo, HK
hintat.htc@gmail.com

中文半固定式成語的習得研究

劉麗
廣東外語外貿大學
英語教育學院

張顯達
香港教育學院
語言學與現代語系

摘要

中文成語裡有很多半固定的格式，如不-A-不-B，千-A-萬-B，A-言-B-語，等等，在使用中會呈現出一定的生成性（Ching 1964；Ji 2007；張 2002）。但這些格式的生成性是有限制的：空格上的 A，B 詞語選擇必須符合句法和語義等方面的要求。學習者通常並不對所有半固定的成語逐一記憶，一個有趣的問題是兒童又是如何習得這些限制規則的呢？本研究通過一項分級的可接受度判斷，調查了兒童和成人對三種不同成語構式的多產性及其限制的習得狀況，和影響他們判斷的因素：頻率，結構複雜度，內部語義關係和 A，B 詞素是否構成複合詞。結果顯示 9 歲左右的孩子已對何種的 A，B 詞素進入構式有了抽象的規則概括。此研究對中文成語的構式理解提供了心理真實性的實證依據。

關鍵詞：半固定成語　構式　生成性