Article

Implicit Learning of Chinese Numeral Classifiers Using the Stimulus Equivalence Paradigm

Yee Pin Tio*
Southern Illinois University Carbondale, USA

Usha Lakshmanan
Southern Illinois University Carbondale, USA

Abstract

Chinese Numeral Classifiers (NCs), which are obligatory in the quantification of nouns, have been observed to pose serious challenges to learners of Chinese as a second language. Learners typically have difficulties generalizing the use of classifiers to new nouns and produce collocation errors due to: (1) the inherent complexity and inconsistencies of the Chinese NC system and (2) insufficient exposure to input containing classifier-noun collocations. The current study adopted a pre-test-post-test randomized group design to investigate whether implicit instruction using the stimulus equivalence (SE) paradigm (Sidman, 1971) would have a positive impact on the learning of Chinese NCs and whether the variability in classifier category membership (determined via typicality ratings of native Chinese speakers) would influence the learning process. Eighty-Seven native-English speakers, assigned to one of two learning conditions (an implicit learning [i.e., SE] group or an explicit learning group), were systematically exposed to four Chinese NCs. Only the group that learnt Chinese NCs using the SE paradigm showed a significant improvement in overall classifier retention one week after training. More fine-grained analyses revealed that the implicit-learning (SE) group was significantly more accurate in generalizing the use of the classifiers to new objects than the explicit-learning group when matching classifier-object pairs with lower typicality ratings. In the case of classifier-object pairings with high prototypicality ratings, the overall effectiveness of implicit learning of Chinese NCs (based on the SE paradigm) was comparable to the explicit instruction condition. The results highlight the SE paradigm's success vis-à-vis explicit instruction in facilitating the inferential learning of fuzzy rules based on less prototypical classifier-object pairings. The findings of the study contribute to our understanding of the Chinese NC system as well as semantic categorization from a cognitive linguistic perspective and have significant implications for the development of effective instructional methods to optimize language learning.

Keywords

Chinese numeral classifiers, implicit-explicit instruction methods, stimulus equivalence paradigm, typicality effect, object categorization

*Corresponding Author

Address: School of Psychological and Behavioral Sciences, Southern Illinois University Carbondale, Life Science II, Room 281, 1125 Lincoln Drive, Carbondale, Illinois, 62901, USA
Email: yeepintio@siu.edu

1 Introduction

According to a recent report published by the Modern Language Association (Looney & Lusin, 2019) on enrollments in languages other than English, in the fall semester of 2016, approximately 53,069 students were enrolled in Chinese language programs in universities in the United States. Notwithstanding the recent spurt in interest in learning Mandarin Chinese as a second/foreign language in the U.S. and other countries in the world, the acquisition of Chinese as a second/ foreign language has continued to remain a relatively under-researched area, particularly in comparison to other languages such as French and Spanish (Cai & Zhu, 2012). Learning Chinese as a second/foreign language can be challenging for learners, especially given the tonal and orthographic systems that set it apart from languages such as English (Liu, 2013; Yu & Watkins, 2008). To facilitate the teaching and learning of Mandarin Chinese (hereafter referred to as Chinese), it is crucial to improve our understanding of the language-specific features that pose a challenge for learners.

The current study focuses on Numeral Classifiers (NCs)—a considerable challenge in learning Mandarin Chinese (Gao, 2010; Zhang & Lu, 2013). NCs are grammatical elements that are attached to a count noun when quantity is specified. For example, in Chinese, the translation equivalents of the English phrases 'one pen' and 'one rope' would be 一支笔 *yì-zhī-bǐ* 'one-**classifier**-pen' and 一条绳子 *yì-tiáo-shéngzi* 'one-**classifier**-rope' respectively, where the classifier occurs between the numeral and the noun. Semantically, NCs agree with the physical (and other) characteristics of the associated noun. As illustrated by the two examples provided (i.e., 'one pen' and 'one rope'), the classifier 支 *zhī* denotes a long and rigid object (e.g., pen, cigar, or flute) and the classifier 条 *tiáo* denotes a long and flexible object (e.g., rope, fish, or tie).

Some crosslinguistic studies on semantic (object) categorization have found that on similarity judgment tasks, Chinese (i.e., classifier language) speakers are more likely to evaluate an object's composition (e.g., solids, non-solids) than English (i.e., non-classifier language) speakers (Li et al., 2009). The bias towards an entity's composition, as observed in the Chinese speakers, supported the position that when evaluating category membership of objects, the use of NCs can potentially influence speakers' object categorization by drawing their attention to language-specific features or properties. A fundamental property of categorization involves comprehending and grouping different concepts based on similarities. The ability to selectively attend and match features of an object to a pre-existing categorical system is crucial in how we learn to name and recognize entities around us (Berlin, 2017; Imai & Gentner, 1997; Lucy, 1992). For second language (L2) learners of a classifier language like Chinese, the ability to successfully handle the purely grammatical and semantic aspect of language-specific features such as NCs is crucial. There is a need for research on the effectiveness of instructional methods to facilitate successful learning of this demanding aspect of the Chinese language.

Native Chinese speakers and L2 Chinese learners alike struggle with the acquisition and generalization of the Chinese NC system (Uchicda & Imai, 1999). This may be due to the system's complexity, particularly concerning the choice of the target classifier based on the semantic features of the noun. For example, some Chinese classifiers have an incoherent and loose semantic organization where the classifier rules are complicated and unintuitive. They are termed as "arbitrary" classifier categories by Gao and Malt (2009), which do not bear defining features nor any prototypical features (e.g., 顿 *dùn* 'pause' is commonly used to classify the noun 'meal'). These arbitrary classifiers tend to be acquired through rote memorization, and native speakers often do not have a clear intuition of the properties or features of the objects associated with these classifiers. A commonly observed strategy among L2 Chinese learners is avoidance—by either omitting the target NC where it is required or by using the general NC 个 *ge* as a default to replace a specific classifier (Zhang & Lu, 2013). The latter strategy has been observed among native speakers of Chinese as well.

The current study focused on classifiers that share semantic features with the associated nouns. It experimentally investigated factors that mediate the learning of NCs. Specifically, focusing on the

semantic features associated with object categorization, the study compared the effectiveness of implicit vis-à-vis explicit instruction and assessed the typicality effect in relation to semantic categorization on the learning of Chinese NCs. The typicality effect captures the graded structure of a classifier category, reflecting the variation of typicality among different classifier-noun pairings.

The remainder of this paper is organized as follows: Sections 1.1 to 1.5 provide the relevant background information. The first three of these sections include an overview of theories of semantic categorization with a focus on prototype theory, followed by a description of the Chinese NC system and previous research on the critical aspects involved in the learning of NCs. The subsequent two sections include a review of the relevant literature on explicit and implicit learning methods and a detailed description and justification of the stimulus equivalence (SE) paradigm used for the implicit learning of NCs in the study. Section 2 describes the methodology, and Section 3 presents the results. The last two sections discuss the theoretical and pedagogical implications of the findings and highlight recommendations for future research.

1.1 Cognitive perspectives on semantic categorization

A crucial theoretical development in cognitive linguistics relates to the study of categorization. Categorization enables us to summarize, order, and adapt the vast amount of information to our limited cognitive capacities (Alvarez & Franconeri, 2007; Lakoff, 1990) and is often studied to assess cognitive processes underlying conceptual acquisition. Additionally, category representations can be used to classify new items, rate an item's typicality, and make predictions about items based on their category membership.

According to prototype theory, as Eleanor Rosch and colleagues (1976) proposed, many categories are characterized by fuzzy boundaries and the typicality effect. For example, with the category BIRD, the SPARROW would be considered by most people as a typical member of the category, whereas PEACOCK would be considered as an atypical member. These asymmetries between category members are called typicality effects. In previous studies by Rosch and her colleagues, typicality was operationally defined as a semantic measure of family resemblance within a category (Rosch & Mervis, 1975; Rosch et al., 1976). In other words, typical members have more semantic features in common while atypical members show more significant variation in their semantic features. Additionally, category verification occurs more quickly for typical members than atypical members (Rosch et al., 1976).

From a linguistics perspective, every language indicates noun classes and nominal concepts differently. The nominal classification system, presented as a lexical-grammatical continuum, demonstrates the various methods devised to classify nouns in different languages (Craig, 1986). One end of the system focuses on the lexicon, which consists of independent words and morphemes that are semantically composed. For example, the class term *berry* in *strawberry*, *blueberry*, and *raspberry* indicates small and fleshy fruit entities. In contrast, the other end of the continuum consists of a classification of nouns embedded in the language as a morpho-syntactic structure in the form of affixes, clitics, or base modification. For example, Spanish, a language with grammatical gender marking, classifies nouns based on grammatical gender, as in *el libro* 'book' (masculine) vs. *la revista* 'magazine' (feminine). Unlike classifiers, grammatical gender markers are not associated with real-world semantics (e.g., biological sex of the entity) but serve a purely grammatical function.

The classifier system rests midway on this lexical-grammatical continuum, as classifiers incorporate both lexical (semantic) and grammatical features (Grinevald, 2000). Classifier systems also differ from purely grammatical systems (e.g., grammatical gender markers) in that the use of a classifier is determined predominantly by the semantic context, which is closely linked to the associated lexical noun and its pragmatic function (Seifart, 2010). Various theories have been put forth to explain classifier systems regarding their semantic structure (Allan, 1977; Denny, 1976; Lee, 1987). Allan (1977) concluded that classifier categories are derived from human perceptual capabilities: Inherent

and perceivable properties of an object, such as material composition (animacy, abstractness, & entity construal), shape (1D, 2D, & 3D), consistency (flexible, rigid, non-discrete), and size (small, medium, & large), are the common semantic features captured across more than 50 classifier languages.

1.2 Chinese NCs

NCs are used in many languages globally (e.g., East Asian, Austronesian, Mesoamerican, West African) but not in the Indo-European language family (e.g., English). An exciting aspect of the NC construction relates to the classification of countable nouns. As stated in Section 1, in Chinese, a classifier is necessitated in nominal phrases containing a numeral quantifier. They commonly occur adjacent to a numeral (e.g., definite quantifiers such as *one* or *two*) within a noun phrase (Seifart, 2010). Chinese classifiers are also obligatory in noun phrases containing an indefinite quantifier (e.g., *some* or *many*) or demonstrative (e.g., *this* or *that*). Importantly, NCs encode semantic features in agreement with the noun heading the noun phrase, as shown in examples (1) and (2).

(1) 一   本   书

yì   běn   shū

one   Classifier   book

'one/a book'

(2) 一   册   书

yí   cè   shū

one   Classifier   book

'one book series'

The classifier 本 *běn* describes an item bound together, which highlights a semantic feature of the noun 'book'. Similarly, the NC 册 *cè* indicates a book series consisting of books with a common theme and are formally identified together as a group. Without the NC 本 *běn* or 册 *cè*, the noun phrase will be grammatically incorrect in Chinese. In certain conditions, the noun can be omitted, as shown in example (3) below. The deictic purpose of NCs also contributes to their semantic importance in the language.

(3) 几   位?

jǐ   wèi

how many   Classifier for people

'How many (people)?'

In addition to their function in language, NCs are cognitive constructions of our sensation and perception (Allan, 1977). Lakoff (1987) proposed that classifier categories reflect speakers' conceptual representation and form conceptual categories through linguistic categorization. Since then, several studies have explored the motivating factors for the categorization of Chinese classifiers. For example, 条 *tiáo* 'branch' represents a specific type of human categorization based on the perceptual property of "extension in length," which stems from the original meaning of the Chinese character (Tai & Wang, 1990) and 张 *zhāng* represents the property of 'extension in length and width,' a human categorization for flat objects (Tai & Chao, 1994).

Given that each object or noun has multiple semantic facets, NCs provide information along several semantic dimensions, including shape, consistency, dimensionality, and size. Chinese speakers will need to analyze an object's composition, shape, dimensionality, and size to accurately categorize the object with the appropriate NC, especially for less commonly used classifier-noun pairs. With this obligatory classification embedded in the language, classifier language speakers must selectively attend to specific features of an item to form a grammatically accurate noun phrase. For example, when evaluating an

object, Y, that is long and flexible, the Chinese NC 条 *tiáo* must be expressed in front of the object, thus "one *tiáo* Y."

Inversely, NCs also signal the class of a referred noun (Huang & Chen, 2014). With the knowledge of the NC, we can infer the characteristics of the noun that follows. For example, when the classifier 条 *tiáo* is mentioned without the noun, it will be apparent that the speaker refers to a slender and flexible item. Knowing the NC is sufficient to infer information and characteristics of an entity. This inductive nature of the classifier systems corresponds to the general cognitive process of categorization.

1.3 Classifier learning

Given the relevance of NCs for conceptual representation, research indicates vital differences in children's developmental path between nouns and classifiers. Children acquire classifiers between the ages of four to five, as shown in Uchida and Imai (1999), which is about two years after they can fast-map the meaning onto novel words (Carey & Bartlett, 1978; Haryu & Imai, 1999; Markman & Wachtel, 1988). Four- and five-year-old children attain a significant leap in their ability to extract rules from exemplars and to generalize the rules to new items (Uchida & Imai, 1999). In addition to children's developmental abilities, one factor contributing to this discrepancy is the semantic organization for nouns and classifiers. Nouns have a more transparent semantic structure where children can have more explicit expectations of how the semantic organization works and how the noun extends.

In contrast, classifiers have an incoherent and loose semantic organization where the classifier rules are often complicated and unintuitive (DeKeyser, 2005). The extension of these classifier rules is often inconsistent across different classifiers within a language or different dialects of the same language. For example, in Chinese, the classifier 条 *tiáo* can indicate long and flexible objects (e.g., tie or rope), but this classifier can also describe animate entities (e.g., shark) and abstract entities (e.g., news or text messages). In addition, there are also dialectal and regional differences in the use of Chinese NCs among native Chinese speakers (Yi, 2011). For instance, the noun 学校 *xuéxiào* 'school' can be paired with the classifier 间 *jiān* (classifier for rooms) or 所 *suǒ* (classifier for institutions or places) depending on the speaker's dialectal preference.

The complexity and inconsistencies within the classifier system pose considerable challenges to young children and college-level Chinese as a foreign language (CFL) learners. Despite being older, CFL learners, similar to younger children, often overused the general classifier 个 *ge* in their writing; furthermore, beginner-level CFL learners often omitted classifiers (Zhang & Lu, 2013). Compared to native Chinese speakers, CFL learners seem to lack knowledge of classifier-noun collocations, which may be due to their lack of adequate exposure to co-occurrences of a classifier with a noun in their experience with the language. Consequently, learners are unable to generalize the classifiers to new nouns, and they are more likely to produce ill-formed collocations as they experiment with the new classifiers or resort to overusing the general classifier 个 *ge* (Myers & Tsay, 2002).

Uchida and Imai (1999) and Zhang and Lu (2013) have stated that first and second language learners seem to acquire the meanings of each classifier via bottom-up processing, which relies mainly on the input and constant exposure to classifier-noun collocations. According to Uchida and Imai (1999), learners do not tackle classifiers with strong top-down expectations, which is a process where learners could induce prior knowledge about a grammatical rule to predict the meaning of a new encounter. When learners become aware of the grammatical function of classifiers, they do not utilize these very complex and opaque semantic systems right away. Instead, they continue to use the general classifier 个 *ge*, because it is the classifier they hear most frequently. Compared to the acquisition of nouns, learners are more conservative regarding classifier systems and take a longer time to acquire them fully. In other words, although learners are aware of the ambiguity within the classifier systems, they seem to be prepared to encounter exceptions to the rules and often avoid the use of proper NCs.

Given the challenges that learners faced when acquiring NCs, Zhang and Lu (2013) raised a key concern: the lack of or insufficient exposure to input relevant for the learning of classifiers in textbooks and other pedagogical materials. Factoring in the word frequency effect on lexical learning (Bybee & Hopper, 2001; Ellis, 2002; see discussion in Section 1.4), it stands to reason that the learning of classifiers can be facilitated if learners encounter sufficient exemplars of classifier-noun collocations in the input. It is crucial to gain a deeper understanding of the NC system from the language learner's perspective and develop instructional methods and materials to ensure optimal language learning.

1.4 Implicit and explicit language learning

Within the fields of psycholinguistics and second language acquisition, a substantial amount of research has evaluated the effectiveness of explicit and implicit instruction (see Ellis, 2002 and Norris & Ortega, 2000 for review). This interest in evaluating instructional methodology stems from the fact that language learning can take place implicitly (i.e., an unconscious and automatic abstraction of the structural nature of the language through the experience of instances) and explicitly (i.e., selective learning of metalinguistic rules). While many studies have documented the positive effects of explicit language learning (Doughty & Long, 2008; Ellis, 2002; Hulstijn, 2005; Spada, 2011), research into the nature and mechanism of implicit learning is still developing.

Empirical research on implicit learning falls mainly into two categories: artificial grammar learning and statistical learning. The artificial grammar learning paradigm (Reber, 1967) involves exposing participants to a string of stimuli generated by an artificial system and testing them on what they have learned. In contrast to artificial grammar learning, statistical learning (also known as usage-based learning) is more current and focuses on the frequency effects in language processing (Diessel, 2007; Rebuschat & Williams, 2012; Trousdale & Hoofman, 2013). Theories of statistical learning proposed that learners are sensitive to the distributional regularities in the input, including sounds, syllables, and syntactic categories, and the information learners accumulate based on their regular or daily usage of the language guides the acquisition process (Erickson & Thiessen, 2015).

An important factor influencing statistical learning is frequency of use. In relation to lexical processing, the findings based on many experimental studies have found a word-frequency effect—it is the words with a high frequency of occurrence are learned more efficiently and processed more fluently (e.g., Bybee & Hopper, 2001; Ellis, 2002). The importance of the frequency effect in language learning supports the idea that language processing involves computations based on statistical probabilities and that language learning is exemplar-based (Ellis, 2015). In other words, language knowledge involves a collection of memories of previously experienced utterances rather than grammar and abstract rules or structures.

The ambiguous nature of the implicit process continues to be debated among scholars. Some researchers have questioned the implicitness of learning (Jiménenez & Mendez, 1999) and the role of consciousness in implicit learning and retrieval (DeKeyser, 1995; Ellis, 2015). Thus, in the current research, explicit instruction is operationalized as a rule-based explanation as a part of the instruction. In contrast, implicit instruction is operationalized as the *absence* of rule presentation and instruction that draws learners' attention to a particular form. This definition, adapted from DeKeyser (1995), has been used as a standardization method in several reviews and meta-analyses of explicit and implicit language instruction research (Ellis, 2002; Norris & Ortega, 2000).

Furthermore, other studies have sought to assess the differential effects of implicit and explicit learning. For example, based upon his investigation of morphological rules in an artificial language, DeKeyser (1995) further proposed that clear-cut categorical rules are learned more successfully in the explicit condition. In contrast, prototypical and fuzzy rules [e.g., English past tense forms for irregular verbs as documented in Bybee and Slobin (1982)] were learned slightly better in the implicit condition without any grammar explanation. Moreover, Reber (1976, 1993) and Krashen (1982, 1994) argued that

implicit learning is particularly advantageous for complex structures. Williams (1999) indicated that learning of semantically redundant agreement rules correlated strongly with what he calls passive and implicit 'data-driven processes.' Despite the differences in findings, most scholars concur that explicit instruction and implicit instruction are differentially effective for language learning, which motivates the current research on the learning of Chinese NCs.

1.5 Stimulus equivalence (SE) paradigm and implicit learning

The goal of the current research was to compare the effects of explicit and implicit approaches to the learning of NCs on object categorization. For the implicit-learning condition, the study developed and adopted an SE task. The SE paradigm was first used with individuals with developmental disabilities to help them learn the 'equivalent relations' (i.e., the association) between pictures, spoken words, and written words (Sidman, 1971). The principle underlying this paradigm is that associations will form among the stimuli through repeated exposure and differential feedback (i.e., reinforcement vs. no reinforcement).

The SE paradigm emphasizes learning relationships among stimuli and demonstrates that new relations (i.e., derived relations) emerge after the experience with a series of match-to-sample tasks. The stimulus equivalence task consists of three main components (a) reflexivity, in which items are matched to themselves (A=A); (b) symmetry, in which each pair of related items are learned (A=B, B=A; A=C, C=A); and (c) transitivity, in which the unconditioned relations between the previously learned items (B=C; C=B) are formed. Although the relations between B and C were never explicitly taught, these relations emerge due to the implicit training resulting from performing the match-to-sample tasks (Sidman & Tailby, 1982). In the context of Chinese NCs, an example of an "[A]" item for learning the Chinese NC for long and flexible objects would be the Chinese character 条. Equivalent classes indicate that each member of the category 条 *tiáo* (e.g., rope [B] & shoelace [C])) are interchangeable or can substitute for one another (Lynch & Cuvo, 1995). The most valuable property of SE is the emergence of a categorical relationship between items "rope [B]" and "shoelace [C]," which has not been taught explicitly.

Since its development, the use of the SE paradigm has expanded, and it has been adapted to teach a variety of skills to individuals with and without disabilities, in a range of content areas, including mathematics (Hall et al., 2006; Lynch & Cuvo, 1995; Ninness et al., 2006), geography (LeBlanc et al., 2003), undergraduate rehabilitation courses (Walker et al., 2010), and neuroanatomy (Fienup et al., 2016). The SE paradigm has also been applied to improve various language skills, including pre-reading skills (Connell & Witt, 2004; Lane & Critchfield, 1998), reading, and spelling (de Rose et al., 1996). This paradigm has also been adopted in language revitalization efforts of Native American languages such as Navajo and Ojibwe (Haegele et al., 2011). The successful application of the stimulus equivalence paradigm in language learning indicates a potentially valuable application for teaching complex Chinese NCs.

2 The Current Study

The present study investigated the role of explicit-implicit instructional methods and the typicality effect in relation to object categorization on the learning of Chinese NCs. We first assessed the effectiveness of implicit and explicit instruction on the learning of NCs. Explicit instruction is more in line with the standard top-down approach to language teaching, where the tasks focus on classifier rule explanation. In contrast, implicit instruction corresponds to a bottom-up approach, where participants implicitly learn classifiers through exemplars and classifier-noun pairings.

The SE task was adopted for the implicit-learning condition for two reasons. First, the paradigm's design allows for implicit learning to occur without rule explanation. The paradigm is simplistic and

can be easily adapted to include a large variety of stimuli, in this case, visuals of different objects and Chinese NCs. Second, the repetitive sample-match tasks used in this paradigm utilizes a more exemplar-based approach that might be more reflective of the statistical learning of the Chinese NC system, which consists of fuzzy morphological rules. As Uchida and Imai (1999) and Zhang and Lu (2013) have observed, there is greater reliance on bottom-up processing when learning NCs because of their ambiguous characteristics. The SE paradigm is potentially helpful in learning fuzzy morphological rules, as it places less emphasis on strict and rigid classifier rules and facilitates a more implicit and accommodating application. This methodology allows for flexible expansion of an existing category by conditioning additional stimuli with pre-existing class members, creating a presentation of various classifier-noun collocations required during the classifier learning process.

We, therefore, predict that there would be differences in classifier retention as a function of the two learning conditions (Hypothesis 1). Specifically, participants in the implicit condition should benefit from their experience in processing a large number of examples in the input provided through the SE task during the training phase and should retain and generalize that information better than the participants in explicit instruction conditions.

The second part of the study examined the effects of item-category typicality on the learning of NCs. The current study adopted the rating procedure from Rosch (1973) to determine the typicality (also known as 'goodness-of-exemplar') for various objects in relation to the Chinese NCs. In the influential Rosch (1973, 1975) study, items with high typicality scores are good examples of a category, and these items are often seen as superior compared to those with low typicality scores, as they are responded to and verified more quickly. Given their advantage in relation to the speed of processing and ease of verification, high typicality items should lead to better predictions from the learner participants and facilitate their generalization to new items. Therefore, we expect that there would be differential effects in generalizing to new classifier-based objects depending on the typicality ratings (Hypothesis 2).

3 Method

The current study consisted of a training phase and a testing phase. Participants were randomly assigned to one of the two learning groups (i.e., explicit-learning or implicit-learning). During the training phase, all participants learned four Chinese NCs with either explicit or implicit instruction. The testing phase was conducted twice, with the first testing session occurring immediately after the training phase and the second testing session occurring a week after the training phase. The purpose of the testing phase was to (1) assess the retention of classifier knowledge of the two groups and (2) assess the generalization of the classifiers to new objects. Participants were naïve to the purpose of the study and were only told that the study aimed to assess their responsiveness to different stimuli. The training and testing phases were administered individually via E-Prime 2.0 software (Psychology Software Tools, Pittsburg, PA) in a lab setting, and behavioral responses (key pressing) were collected. The entire experiment took approximately 45 minutes to complete.

3.1 Participants

Eighty-seven native English speakers aged 18-25 years participated in this study (see Table 1). Participants were undergraduate students from the introduction to psychology or introduction to linguistics courses at a Midwestern US university. All participants received course credit or extra credit (upon the course instructor's discretion) for their participation. Every effort was made to include only monolingual native speakers of English. However, many participants had learned another language besides English. Only those who indicated a very low level of proficiency (self-rating of 1 or 2) (1=

very little knowledge of or very low ability in the relevant language; 7= excellent, native, or native-like knowledge and ability) in the other language was included. Furthermore, individuals who had prior exposure to any Chinese languages (e.g., Mandarin, Cantonese) and any other classifier language (e.g., Malay, Thai, American Sign Language, Korean) were excluded. A participant background questionnaire was given at the end of the study to collect participants' general information (e.g., language background, reading difficulties) to ensure that the participants met the criteria to be included in the data analysis.

Table 1

Participant Demographic Information

Group	n	Age (Years) Mean (*SD*)	Gender	
			Male	Female
Implicit-Learning	43	20.23 (*1.823*)	20	23
Explicit-Learning	44	19.87 (*1.679*)	17	27

3.2 Materials

3.2.1 Main experiment: Chinese NCs and object stimuli

In the main experiment, four Chinese NCs (支 *zhī*, 条 *tiáo*, 张 *zhāng*, 颗 *kē*) were selected based on a review of previous literature on Chinese NCs and were introduced to the participants in the two experimental conditions. Table 2 shows the prototypical features and examples of associated nouns for the four NCs. These four NCs are the most commonly used NCs among Chinese children and adults (Erbaugh, 1986), apart from the general classifier 个 *ge*. Furthermore, 支 *zhī*, 张 *zhāng*, and 颗 *kē* correspond to the three primary shapes, long-rigid, flat-flexible and rounded, and are common in classifier systems across Southeast Asia and East Asia (Adams & Conklin, 1973). Thus, the classifiers selected for this study are not arbitrary linguistic elements but are frequently used in Chinese.

Table 2

Chinese NCs and Associated Noun Objects

Classifiers	Prototypical Features	Associated Noun Objects
支 *zhī*	Long and Rigid	Pen, Candle
条 *tiáo*	Long and Flexible	Fish, Rope
张 *zhāng*	Flat Surface	Paper, Table
颗 *kē*	Small and Rounded	Pearl, Bean

The object stimuli appeared as images on the computer screen. The use of word labels for the object stimuli was thus avoided to minimize the use of English input that could potentially influence the results. All images of the object stimuli used in this study were black and white to control for any influence from color perception.

The object stimuli appeared in three forms: a target object, a classifier-based object, and a distracter object. The target object was the reference object that shared the same classifier with the classifier-based object. In order to ensure that there was no predisposing bias towards classifier-based objects or distracter objects, the selection of the object stimuli was based on similarity ratings obtained from a separate group of native-English speakers in a preliminary study (see Section 3.2.2). Crucially, only those objects (whether belonging to the distracter or classifier-based object category) that were rated as not similar

to the target object in the preliminary study were included. Additionally, the findings based on a second preliminary study were also used to determine the typicality ratings of objects within each of the four classifier categories (see Section 3.2.3).

3.2.2 Pre-experiment: Pairwise similarity ratings

A preliminary study (Tio, 2016) obtained similarity ratings for a total of 140 pairs of object stimuli (i.e.,70 target object and classifier object pairs; 70 target object and distracter object pairs) from an independent group of 38 native speakers of English (mean age= 18.68 years; 23 women and 15 men). The 38 English native speakers rated the level of similarity between pairs of pictures using a 1-7 scale (1= *least similar*, 7= *most similar*). Similar to the participants recruited for the main study, none of the participants in the preliminary study had prior knowledge of any classifier language, as reflected in their self-reported information in the questionnaire. An average similarity rating score was computed for each object pair, and the score was used to determine the object stimuli selected for the main study. The extension task used in the experiment only included object pairs that received a low similarity rating (i.e., 4 and below) to ensure that participants would have similar baseline preferences (i.e., rated as not similar) for all object stimuli prior to any classifier exposure. Therefore, the expectation was that both experimental groups (i.e., implicit and explicit conditions) would have equal preference for classifier object and distracter object pairings before the training phase.

3.2.3 Pre-experiment: Chinese NC typicality ratings

The typicality of the items within each of the four classifier categories (支 *zhī*, 条 *tiáo*, 张 *zhāng*, 颗 *kē*) was determined based on a typicality rating (see Appendix I) by six native Chinese speakers (mean age= 19.18; 3 women and 3 men). Participants indicated how well each item (15 items for each classifier) reflects the corresponding classifier category using a 0-7 scale (0 = *the object shown would not be used with that classifier*; 1= *very poorly*; 7= *perfectly*). All participants in this preliminary study had prior knowledge of the Chinese classifiers, as reflected in their self-reported information in the background questionnaire and classifier assessment. Each participant saw a list of objects and was required to determine how well each noun reflected the idea of the corresponding classifier. An average typicality rating score was computed for each object item, and the score was used to determine the typicality of the object stimuli in the main experiment. The cut-off point for the high and low typicality rating was determined based on the mean ratings on all items (i.e., 5.5). In other words, the cut-off point for high typicality items was an average rating of 5.5 and above, while low typicality items have an average typicality rating of 5.4 and below.

3.3 Procedures for the main experiment

3.3.1 Training phase

During the training phase, the explicit and implicit learning groups were systematically exposed to four target NCs using computer-based tasks. The explicit-learning group received explicit instruction on the characteristics of the NCs. In contrast, the implicit learning group completed a series of SE tasks (Sidman, 1971) that presented the NCs without explicitly explaining the categorical rules involved. None of the object stimuli that appeared during the training phase appeared in the testing phase. In other words, participants were not tested on the same stimuli used in the training phase.

3.3.1.1 Explicit-learning group

For participants in the explicit-learning group, the training tasks utilized multisensory input (auditory, visual, and haptic) to maximize participants' exposure to Chinese NCs (Tio & Lakshmanan, 2017). The training phase introduced the explicit-learning group to Chinese NCs via four short video clips and 12 comprehension questions. The explicit teaching of NCs used a story context involving a native English-speaking college student being taught the four target Chinese classifiers (支 *zhī*, 条 *tiáo*, 张 *zhāng*, 颗 *kē*) by another student, a native Chinese-speaking friend. Chinese NCs were displayed on the screen, and the native Chinese speaker in the video presented them orally and explained the use of each NC with different objects. The videos explained the grammatical function of classifiers, including the use of classifiers with countable nouns, their adjacency to a numeral (e.g., quantifiers such as 'one') within a noun phrase, and appropriate classifier-noun collocation (e.g., the classifier 支 *zhī* is appropriate for the object *pen*, *toothbrush*, and *candle*, because it denotes an object that is long and rigid). The videos presented the four target NCs and the numeral one (一 *yī*) in Chinese (visually using Chinese characters and orally in Chinese.) The object labels (e.g., *toothbrush*) were presented visually using English labels and showing the physical object, and the remainder of the interaction (including the explanation part) was in English. Each NC was presented with three object pairings. The comprehension questions were presented on the screen after each video (e.g., Which of the following classifiers was mentioned in the video you just viewed?). Participants responded to the comprehension questions by selecting one of the two options. Feedback was displayed on the screen as "correct" or "incorrect' after each question.

3.3.1.2 Implicit-learning group

In contrast to the explicit classifier description for the explicit-learning group, the implicit-learning group learned about the classifiers through match-to-sample procedures by applying the SE paradigm. These procedures sought to assist participants to understand and learn the relationships among NCs and their corresponding classifier-based objects. The target stimulus appeared on top, above the choice options (i.e., three in the reflexivity task, two in the symmetry and transitivity tasks, see Section 1.5). Participants were instructed to choose which of the stimuli "matched" the target stimulus by clicking on the corresponding key on the keyboard. The placement of the choice stimuli was counterbalanced to ensure all stimuli appeared equally on the left versus the right side of the screen. The implicit-learning group completed the four SE tasks in the following order: transitivity pre-test, reflexivity, symmetry, and transitivity post-test. Participants' selection was recorded, and feedback was provided only in the reflexivity and symmetry task. Feedback was given as auditory input: a pleasant sound indicated a correct response, while a buzzer sound indicated a wrong response.

A pilot study was conducted to assess the object stimuli and the protocols for the implicit and explicit instructions. The results from the pilot study indicated that participants were able to learn the Chinese NCs within the 20 minutes of training and showed heightened sensitivity towards classifier-based categorization (Tio & Lakshmanan, 2018).

3.3.2 Testing phase

All participants participated in the testing phase a week after their initial exposure to classifier knowledge in the training phase. The testing phase consisted of two tasks: the extension task and the classifier assessment.

3.3.2.1 The extension task

The extension task was a two-alternative forced-choice task that assessed participants' ability to generalize the use of the classifiers to new objects (i.e., those not mentioned during the training phase). It consisted of

96 trials, with four practice trials, eight filler trials, and 72 experimental trials. Three objects appeared on the screen in each trial, with the target object placed above two objects, one of which was a classifier-based object (sharing the same classifier as the target object), and the other was a distracter object (see Figure 1). The target stimuli consisted of an object that would require the use of one of the four target NCs.

Additionally, to address the typicality effect with classifier group membership, we further grouped the classifier-based objects based on the Chinese NC typicality ratings (see Section 3.2.3). Half of the classifier-based objects displayed in the extension task consisted of high typicality rating items (i.e., an average rating of 5.5 and above on a scale of 1 to 7), while the remaining half consisted of low typicality rating items (average rating of 5.4 and below).

Figure 1

An Example of the Extension Task
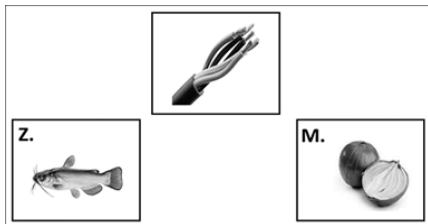
*Note.* In this example, Z is the classifier object, and M is a distracter object.

Participants were instructed to select (as quickly as possible) one of the two objects (classifier-based or distracter) that matched (i.e., that they perceived as being more similar to) the target object displayed on top. The placement of the classifier-based objects and the distracter objects was counterbalanced to ensure that both object types appeared equally on the left or the right side of the screen. Each participant's percentage of classifier object selection was recorded and reflected the participant's preference in categorizing the target object. A higher rate of selecting classifier-based objects would reflect participants' heightened sensitivity towards classifier-based categorization after exposure to NCs during the training phase.

3.3.2.2 The classifier assessment

The classifier assessment was administered twice, once prior to the training phase and once at the end of the testing phase (i.e., a week after training). Both groups completed the classifier assessment. The task consisted of 24 forced-choice questions to establish a baseline before the training and measure the groups' classifier knowledge retention a week after the training. Participants were asked to select one of the three object pairings appropriate for the corresponding NC presented as the Chinese character on the top of the screen (see Figure 2). Participants did not receive any feedback on this task. None of the object stimuli that appeared during the pre-training assessment appeared in the post-training assessment. In other words, participants were not tested on the same stimuli twice during the pre-and post-training classifier assessment.

Figure 2

Example of Classifier Assessment
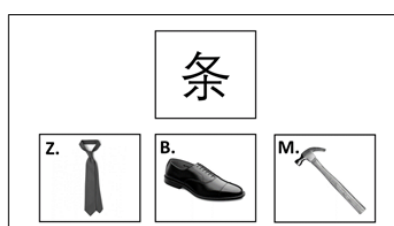
4 Results

A MANOVA analysis was conducted to compare the groups' performance on the classifier assessment and extension task. The group (implicit vs. explicit) served as the independent measure. Typicality effect (i.e., selection rate of high vs. low typicality rating objects in the extension task) and classifier retention (i.e., during pre- vs. post-training classifier assessment) served as the dependent measures. Using Pillai's trace, there was a significant effect of instructional method on participant's typicality ratings and classifier retention, $V = 0.175$, $F(4, 85) = 4.353$, $p = .003$, $\eta_p^2 = .175$.

With respect to the first hypothesis on classifier retention, a separate mixed between-within ANOVA on the outcome variable revealed a significant interaction effect between group (implicit vs. explicit) and classifier retention (pre- vs. post-training), $F(1,85) = 4.532$, $p = .036$, $\eta_p^2 = .051$. In other words, both groups demonstrated a similar level of classifier knowledge pre-training but showed varying levels of classifier retention in the post-training classifier assessment (see Figure 3). The implicit-learning group ($M = 0.506$, $SD = 0.173$) but not the explicit-learning group ($M = 0.429$, $SD = 0.165$) displayed a significantly higher classifier retention rate after the training.

Figure 3

Instructional Method and Classifier Retention
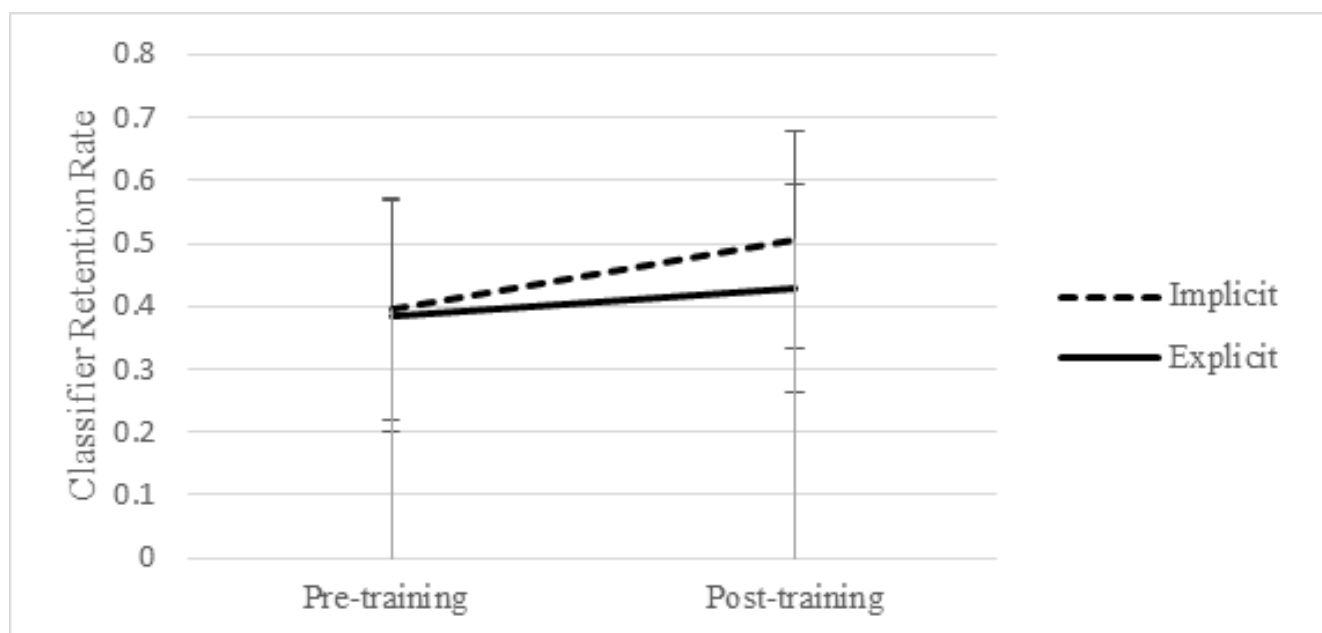
In line with the second hypothesis on typicality effect, a separate mixed between-within ANOVA on the outcome variable also revealed a significant interaction effect between group (implicit vs. explicit) and typicality rating (high vs. low), $F(1,85) = 12.79$, $p < .01$, $\eta_p^2 = .131$. Both groups performed similarly in relation to the high typicality items but responded differently towards the low typicality items (see Figure 4). Specifically, the implicit-learning group ($M = 0.579$, $SD = 0.204$) was significantly more accurate than the explicit-learning group ($M = 0.425$, $SD = 0.198$) in matching the classifier-object pairs with lower typicality ratings but not with objects that have high typicality rating.

Figure 4

Instructional Method and Typicality Rating
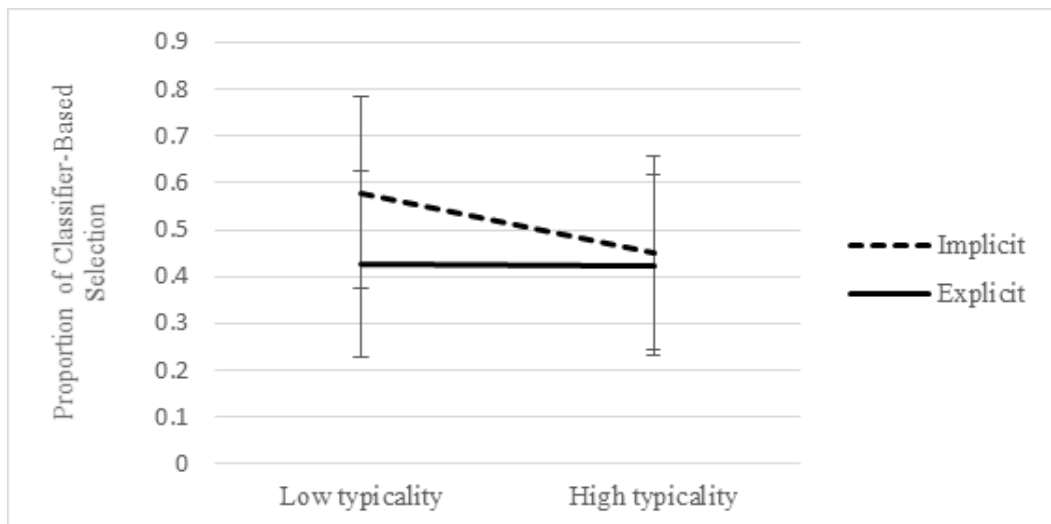
5 Discussion

The current study investigated the effects of instructional method and item-category typicality on native English speakers' learning of Chinese NCs. We will first address the interpretation of the results on the effects of implicit and explicit instruction (Hypothesis 1) and discuss how item-category typicality plays a role during the learning process (Hypothesis 2). The second part of the discussion addresses the theoretical and pedagogical implications of the findings.

5.1 Explicit and implicit instructional methods

The first hypothesis explored the effectiveness of implicit and explicit instruction on the learning of NCs. Despite the small effect size, the findings indicated that participants successfully acquired the untrained relations, specifically the association between two objects that share the same classifier, with the SE paradigm. The results supported our prediction on the differential effects of implicit and explicit instructions for Chinese NCs on the delayed post-training classifier assessment. Additionally, the implicit learning group showed better retention of classifier knowledge than their counterparts in the explicit learning condition. The results support the observations made by Uchida and Imai (1999) and Zhang and Lu (2013) regarding the learning of Chinese numeral classifiers and have significant implications for the relationship between explicit-implicit instructional methods and Chinese language pedagogy.

In addition to the higher retention rate, implicit instruction via the SE paradigm contributed to a significantly higher rate of generalization (based on classifier-based categorization) to object exemplars that were less typical members (i.e., which had lower typicality ratings). The retention and generalization of classifier-based categorization among those who received implicit instruction with delayed testing provide essential insights into the potential benefits of a more bottom-up approach when learning fuzzy and complex classifier rules. The SE paradigm contributed to the implicit-learning group's understanding of classification via the formation of arbitrary equivalence relations through repeated exposure and differential feedback, suggesting that the task design facilitates the occurrence of implicit learning of classifiers. It helps participants learn classifiers through exemplars and classifier-noun pairings without relying on the often complex and ambiguous classifier rules. Learners will have more frequent exposure to a classifier's co-occurrence with a noun and be able to experiment with new combinations. However, future studies are needed to examine the retention and generalization of classifier information on a longitudinal scale, evaluating the change in classifier retention over a more extended period of time (i.e., several weeks, months, or years).

Although the current research adopted the SE paradigm as an implicit instructional task, there are many alternative implicit tasks based on a different operational definition of implicit processing. These tasks include artificial language learning (Kerz et al., 2017; Williams, 2005), probabilistic classification task (Reber et al., 1996), and sequence learning (Granena & Yilmaz, 2019; Moody et al., 2004). More studies are needed to investigate if other implicit learning tasks would yield similar results in relation to the interaction between the instructional method and the typicality effect.

Furthermore, the SE paradigm promotes inclusive learning, where the task can cater to learners with developmental or learning disabilities. As previously mentioned, this paradigm has been adapted for a wide range of learning capabilities, including individuals on the autism spectrum (Hassler, 2018; Yamamoto, 1986) and those with severe intellectual disabilities (Clayton et al., 1999; McKeel & Matas, 2017). From an L2 learning perspective, the SE task can be modified to cater to language learners' needs based on their age and learning style and can be adopted not only in mainstream classroom settings but also for special needs and clinical settings.

Notwithstanding the success of implicit instruction via the SE paradigm, it is crucial to point out that our results also found that explicit instruction also had similar success. Both approaches were equally effective with "good members or exemplars" (i.e., object exemplars that had high typicality ratings). Only in the case of the less typical cases or exemplars, the implicit method was more effective than the explicit instruction. Our data concur with the findings from previous studies that explicit instruction in language learning is effective (Norris & Ortega, 2000), mainly when the items unambiguously reflect the classifier rules. With a more transparent semantic structure, learners can have more explicit expectations of how the semantic categorization works and how the classifier rule extends. That is, when there is consistency underlying the classifier rule extension, learners can more easily rely on top-down processing to induce prior knowledge about the classifier-based categorization and generalize the rules to new items (Uchida & Imai, 1999).

At the same time, an essential question that arises based on the results is why the implicit method was not more successful in the case of high typicality items. In addition to the lack of explicit rules, we suspect that the lack of multisensory input with the SE paradigm is another factor that influences the generalization of classifier rules. One limitation of the SE paradigm is that during the training phase for the implicit-learning group, the exposure to the language input (i.e., Chinese NCs) was restricted to the visual mode (i.e., Chinese orthography). As the NCs were only presented visually and not orally, there was no auditory input in the training phase, aside from the feedback (pleasant or buzzer sound) provided to the participants based on their responses. In contrast, the presentation of the language input in the training phase for the explicit-learning condition was multimodal, involving visual, auditory, and tactile information. Crucially, the explicit-learning group heard the NCs being pronounced as well, which made the input more naturalistic. One way of strengthening the SE method would be to adopt a dual input modality by having not only visual but also an oral presentation of the NCs in the SE task trials.

Overall, our findings support the use of a combination of instructional methods. Explicit instruction could be effective for those with good exemplars of a category, as would implicit training. However, in the case of exemplars that are the less good or atypical members of a category, learners would likely benefit more from implicit learning.

5.2 Item-category typicality rating

The results partially supported our hypothesis that participants' ability to generalize to new classifier-based objects is mediated by both the instructional method and typicality effects. Compared to the explicit-learning group, participants who received implicit instruction on the four target Chinese NCs responded significantly better towards classifier-object pairings with low typicality ratings. The differential effects of the implicit (i.e., heighten responses towards low typicality items but not high

typicality items) and explicit instructional methods (i.e., similar responses toward low and high typicality items) on participants' generalization of classifier-based categorization concur with the literature on how categories are learned in the clinical context, specifically the Complexity Account of Treatment Efficacy (CATE) (Thompson et al., 2003). As mentioned earlier, items with a lower typicality rating are more complex than those with high typicality ratings due to more significant semantic variation and slower category verification. CATE suggested that therapy will produce greater generalization when more complex items are trained. Kiran and Thompson (2003) reported that research involving patients with aphasia found that training with more complex and less typical items resulted in generalization to untrained typical items within the category. However, no generalization was found to less typical items after training with only typical items.

CATE also explains the lack of differential preference for high typicality items between the two learning conditions in our study. The two groups' generalization of classifier use is similar for objects with high typicality ratings (i.e., items deemed as better exemplars of a specific NC), which indicates that instructional methods do not affect the generalizations of highly typical items. This disconnect between category typicality and generalization in relation to highly typical classifier objects could be crucial when addressing the challenges many learners face in making appropriate generalizations when learning NCs. As illustrated previously, the NC system often has ambiguous and exceptional rules that can be difficult to master. Learners can benefit from this more bottom-up approach of the SE paradigm and learn the related relations between NCs to classifier objects with varying item-category typicality through repeated exposure. Further research is needed to investigate the pedagogical effectiveness of category typicality on NC learning.

5.3 Implications for future research

The findings from the current study provide crucial theoretical implications for future research. There is a considerable variation with NCs in the Chinese language. Based on Gao and Malt's (2009) taxonomy of 126 Chinese classifiers, the four target NCs introduced in the current study are labeled as salient and well-defined classifiers, which have relatively more definite extension rules and are predominantly shape-based (i.e., long, flat, rounded). Future studies should evaluate the learning of Chinese NCs' various extension rules to capture a more comprehensive view of classifiers' fuzzy and prototypical morphological nature. The analysis should include not only the well-defined classifier extensions (e.g., 一条绳子 *yì-tiáo-shéngzi* 'a rope') but also those that are abstract (e.g., 一条新闻 *yì-tiáo-xīnwén* 'a piece of news') and arbitrary 一条好汉 *yì-tiáo-hǎohàn* 'a true man').

An approach to further explore Chinese NCs is to examine the role of the various semantic dimensions embedded in NC constructions, such as dimensionality, composition, and size. A very insightful finding emerged when analyzing participants' categorization preferences based on the four NCs: Regardless of group membership, the frequency with which a classifier-based object was selected was highest in the classifier 张 *zhāng* condition (i.e., objects with a flat surface) and lowest in the 条 *tiáo* condition (i.e., long and flexible objects). Notably, among the four target NCs, only classifier 张 *zhāng* denotes a 2-dimensional aspect of object features: the length (long) and width (flat surface) (Tai & Chao,1994). In contrast, 颗 *kē* emphasizes the 3-dimensional feature of roundedness while 条 *tiáo* and 支 *zhī* emphasize the 1-dimensional feature of length Tai and Wang (1990).

This pattern of responses suggests a possible predisposition toward categorization based on the object's dimensionality, where 2-dimensional features (length and width) were more salient, followed by 1-dimensional (length) and then 3-dimensional objects (length, width, height). (Gelman et al., 1998; Schlosser, 2003). While the relationship remains unclear, some research suggests that language-modulated categorization reflects neural representation that corresponds to the organizational dimensions for object representation in the visual ventral pathway (Kemmerer, 2017). Therefore, the interaction

between representation dimensionality, object categorization, and neurobiology remains a question for further investigation.

6 Conclusion

This study systematically assessed Chinese NC learning by drawing upon theories and methodologies from linguistics, cognitive science, and behavior analysis. It was found that the implicit instructional method (i.e., SE paradigm) is comparable to traditional explicit-instructional methods for Chinese NC learning, especially concerning classifier-object pairings with a high typicality rating. At the same time, our findings highlight the SE paradigm's success vis-à-vis explicit instruction in facilitating the generalization of fuzzy rules, especially towards less typical classifier-object pairings. Based on our findings, we propose that a combination of implicit and explicit instructional methods based on the item-category typicality ratings would be desirable for the teaching of NCs. The current study has developed and established a methodological foundation that can be extended to and replicated with other grammatical rules (e.g., grammatical gender, articles, and verbal aspect) or other L2s.

Appendix I Typicality Rating

Teaching American students studying Chinese about classifiers (量词) is often difficult. It is helpful to show students good examples of the use of classifiers in words.

There will be a list of nouns (objects) on the right side of the table.

Using a scale of 0 to 7, indicate how well each noun reflects the idea of the classifier (量词):
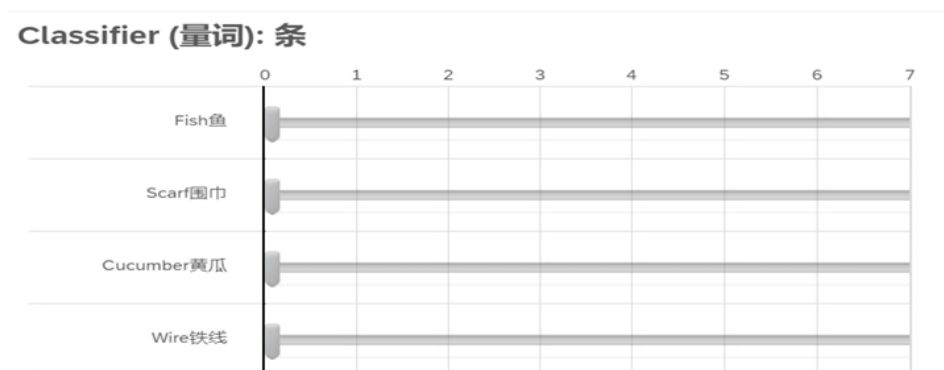
0: 'the object shown would not be used with that classifier'

1: 'very poorly'

7: 'perfectly'

Figure 5

Sample Typicality Rating Scale

Acknowledgments

References

Adams, K. L. & Conklin, N. F. (1973). Toward a Theory of Natural Classification. In C. Corum, T.C. Smith-Stark and A. Wiser (Eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, (pp.1-10). Chicago: Chicago Linguistic Society.

Allan, K. (1977). *Classifiers. Language*, 285-311.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), 14.1-14.10. https://doi.org/10.1167/7.13.14

Berlin, B. (2017). *Tzeltal numeral classifiers*. De Gruyter Mouton.

Bybee, J. L., & Hopper, P. J. (Eds.). (2001). *Frequency and the emergence of linguistic structure* (Vol. 45). John Benjamins Publishing.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, *58*(2), 265-289. https://doi.org/10.1353/lan.1982.0021

Cai, S., & Zhu, W. (2012). The impact of an online learning community project on university Chinese as a foreign language students' motivation. *Foreign Language Annals*, *45*(3), 307-329. https://doi.org/10.1111/j.1944-9720.2012.01204.x

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17-29.

Clayton, M. C., & Hayes, L. J. (1999). Conceptual differences in the analysis of stimulus equivalence. *The Psychological Record*, *49*(1), 145-161. https://doi.org/10.1007/BF03395312

Connell, J. E., & Witt, J. C. (2004). Applications of computer-based instruction: using specialized software to aid letter-name and letter-sound recognition. *Journal of Applied Behavior Analysis*, *37*(1), 67-71. https://doi.org/10.1007/BF03395312

Craig, C. G. (Ed.). (1986). *Noun classes and categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983* (Vol. 7). John Benjamins Publishing Company.

DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, *17*(3), 379-410. https://doi.org/10.1017/S027226310001425X

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *55*(S1), 1-25. https://doi.org/10.1111/j.0023-8333.2005.00294.x

de Rose, J. C., de Souza, D. G., & Hanna, E. S. (1996). Teaching reading and spelling: Exclusion and stimulus equivalence. *Journal of Applied Behavior Analysis*, *29*(4), 451-469. https://doi.org/10.1901/jaba.1996.29-451

Denny, J. P. (1976). What are noun classifiers good for? In *Papers from the Regional Meeting. Chicago Linguistic Society, Chicago, Ill*, *12* (pp. 122-132).

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology, 25*(2), 108-127. https://doi.org/10.1016/j.newideapsych.2007.02.002

Doughty, C. J., & Long, M. H. (Eds.). (2008). *The handbook of second language acquisition* (Vol. 27). John Wiley & Sons.

Ellis, N. 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, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P.

Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–24). John Benjamins.

Erbaugh, M. S. (1986). Taking stock: The development of Chinese noun classifiers historically and in young children. *Noun Classes and Categorization*, 399-436.

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66-108. https://doi.org/10.1016/j.dr.2015.05.002

Fienup, D. M., Mylan, S. E., Brodsky, J., & Pytte, C. (2016). From the laboratory to the classroom: The effects of equivalence-based instruction on neuroanatomy competencies. *Journal of Behavioral Education*, *25*(2), 143-165. https://doi.org/10.1007/s10864-015-9241-0

Gao, H. H. (2010). A study of Swedish speakers' learning of Chinese noun classifiers. *Nordic Journal of Linguistics*, *33*(2), 197-229. https://doi.org/10.1017/S0332586510000156

Gao, M. Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes*, *24*(7-8), 1124-1179. https://doi.org/10.1080/01690960802018323

Gelman, S. A., & Ebeling, K. S. (1998). Shape and representational status in children's early naming. *Cognition*, *66*(2), B35-B47. https://doi.org/10.1016/S0010-0277(98)00022-5

Granena, G., & Yilmaz, Y. (2019). Corrective feedback and the role of implicit sequence-learning ability in L2 online performance. *Language Learning*, *69*, 127-156. https://doi.org/10.1111/lang.12319

Grinevald, C. (2000). A morphosyntactic typology of classifiers. *Systems of Nominal Classification*, *4*, 50-92. https://doi.org/10.1515/9783110172782.2.13.1016

Haegele, K. M., McComas, J. J., Dixon, M., & Burns, M. K. (2011). Using a stimulus equivalence paradigm to teach numerals, English words, and Native American words to preschool-age children. *Journal of Behavioral Education*, *20*(4), 283-296. https://doi.org/10.1007/s10864-011-9134-9

Hall, S. S., DeBernardis, G. M., & Reiss, A. L. (2006). The acquisition of stimulus equivalence in individuals with fragile X syndrome. *Journal of Intellectual Disability Research*, *50*(9), 643-651. https://doi.org/10.1111/j.1365-2788.2006.00814.x

Haryu, E., & Imai, M. (1999). Controlling the application of the mutual exclusivity assumption in the acquisition of lexical hierarchies. *Japanese Psychological Research*, *41*(1), 21-34. https://doi.org/10.1111/1468-5884.00102

Hassler, E. A. (2018). The effects of tact training for students diagnosed with Autism: Will stimulus equivalence be demonstrated through the acquisition of intraverbal responses? Master's Thesis. https://commons.lib.jmu.edu/master201019/541

Huang, S. & Chen, J.Y (2014). The effects of numeral classifiers and taxonomic categories on Chinese and English speaker's recall of nouns. *Journal East Asian Linguist*, *23*, 27-42. https://doi.org/10.1007/s10831-013-9108-0

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, *27*(2), 129-140. https://doi.org/10.1017/S0272263105050084

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*(2), 169-200. https://doi.org/10.1016/S0010-0277(96)00784-6

Jiménez, L., & Mendez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 236-259. https://doi.org/10.1037/0278-7393.25.1.236

Kemmerer, D. (2017). Categories of object concepts across languages and brains: The relevance of nominal classification systems to cognitive neuroscience. *Language, Cognition and Neuroscience*, *32*(4), 401-424. https://doi.org/10.1080/23273798.2016.1198819

Kerz, E., Wiechmann, D., & Riedel, F. B. (2017). Implicit learning in the crowd: Investigating the role of awareness in the acquisition of L2 knowledge. *Studies in Second Language Acquisition*, *39*(4), 711-734. https://doi.org/10.1017/S027226311700002X

Kiran, S., & Thompson, C. K. (2003). The role of semantic complexity in treatment of naming deficits: training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research: JSLHR*, *46*(4), 773–787. https://doi.org/10.1044/1092-4388(2003/061)

Krashen, S. (1982). *Principles and practice in second language acquisition*. Prentice Hall.

Krashen, S. (1994). The pleasure hypothesis. *Georgetown University round table on languages and linguistics*, 299-302.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Lakoff, G. (1990). Some empirical results about the nature of concepts. *Mind and Language*, *4*(1/2), 103-129.

Lane, S. D., & Critchfield, T. S. (1998). Classification of vowels and consonants by individuals with moderate mental retardation: Development of arbitrary relations via match-to-sample training with compound stimuli. *Journal of Applied Behavior Analysis*, *31*(1), 21-41. https://doi.org/10.1901/jaba.1998.31-21

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with autism. *Behavioral Interventions*, *18*(4), 279-289. https://doi.org/10.1002/bin.144

Lee, M. (1987). The cognitive basis of classifier systems. In *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society* (pp. 395–407). Berkeley Linguistics Society.

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology*, *58*(4), 487-524. https://doi.org/10.1016/j.cogpsych.2008.12.001

Liu, N. (2013). Implicit priming effects in Chinese word recall: The role of orthography and tones in the mental lexicon. *International Journal of Computational Linguistics & Chinese Language Processing*, *18*(3), 1-20.

Looney, D. & Lusin, N. (2019). Enrollments in languages other than English in United States institutions of higher education, Summer 2016 and Fall 2016: Final report. https://www.mla.org/Resources/Research/Surveys-Reports-and-Other-Documents/Teaching-Enrollments-and-Programs/Enrollments-in-Languages-Other-Than-English-in-United-States-Institutions-of-Higher-Education

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge University Press.

Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis*, *28*(2), 115-126. https://doi.org/10.1901/jaba.1995.28-115

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, *20*, 121-157. https://doi.org/10.1016/0010-0285(88)90017-5

McKeel, A., & Matas, J. (2017). Utilizing PEAK relational training system to teach visual, gustatory, and auditory relations to adults with developmental disabilities. *Behavior Analysis in Practice*, *10*(3), 252-260. https://doi.org/10.1007/s40617-017-0194-8

Moody, T. D., Bookheimer, S. Y., Vanek, Z., & Knowlton, B. J. (2004). An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. *Behavioral Neuroscience*, *118*(2), 438-442. https://psycnet.apa.org/doi/10.1037/0735-7044.118.2.438

Myers, J., & Tsay, J. (2002). Grammar and cognition in Sinitic noun classifier systems. In *Proceedings of the First Cognitive Linguistic Conference* (pp. 199-216)

Ninness, C., Barnes-Holmes, D., Rumph, R., McCuller, G., Ford, A. M., Payne, R., ... & Elliott, M. P. (2006). Transformations of mathematical and stimulus functions. *Journal of Applied Behavior Analysis*, *39*(3), 299-321. https://doi.org/10.1901/jaba.2006.139-05

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528. https://doi.org/10.1111/0023-8333.00136

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). https://support.pstnet.com/

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855-863. https://doi.org/10.1016/S0022-5371(67)80149-X

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(1), 88-94. https://doi.org/10.1037/0278-7393.2.1.88

Reber, A.S. (1993). *Implicit learning and tacit knowledge*. Oxford University Press.

Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: differences in the flexibility of declarative and nondeclarative knowledge. *Behavioral Neuroscience*, *110*(5), 861-871. https://doi.org/10.1037/0735-7044.110.5.861

Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 111-144). Academic Press. https://doi.org/10.1016/B978-0-12-505850-6.50010-4

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192. https://doi.org/10.1037/0096-3445.104.3.192

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573-605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 491-502. https://doi.org/10.1037/0096-1523.2.4.491

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, *33*(4), 829-856. https://doi.org/10.1017/S0142716411000580

Schlosser, A. E. (2003). Experiencing products in the virtual world: The role of goal and imagery in influencing attitudes versus purchase intentions. *Journal of Consumer Research*, *30*(2), 184-198. https://doi.org/10.1086/376807

Seifart, F. (2010). Nominal classification. *Language and Linguistics Compass*, *4*(8), 719-736. https://doi.org/10.1111/j.1749-818X.2010.00194.x

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, and Hearing Research*, *14*(1), 5-13. https://doi.org/10.1044/jshr.1401.05

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*(1), 5-22. https://doi.org/10.1901/jeab.1982.37-5

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching*, *44*(2), 225-236. https://doi.org/10.1017/S0261444810000224

Tai, J., & Wang, L. (1990). A Semantic Study of the Classifier Tiao. *Journal of the Chinese Language Teachers Association*, *25*(1), 35-56.

Tai, J. H., & Chao, F. Y. (1994). A semantic study of the classifier Zhang. *Journal of the Chinese Language Teachers Association*, *29*(3), 67-78.

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia. *Aphasialogy*, *19*, 1021-1036. https://doi.

org/10.1044/1092-4388(2003/047)

Tio, Y. P. (2016). Cognitive effects of learning Mandarin Chinese numeral classifiers (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (SIUC No. 10163448) Southern Illinois University at Carbondale.

Tio, Y.P. & Lakshmanan, U. (2017). *The Effects of Exposure to Chinese Numeral Classifiers on the Categorization of Objects by Native-English Speakers*. Proceedings of the 2016 UK Cognitive Linguistics Conference, Bangor, UK.

Tio, Y.P. & U. Lakshmanan (2018). The Effects of Explicit and Implicit Learning of Mandarin Chinese Numeral Classifiers on Native English Speaker's Object Categorization. Paper presented at the 10th Annual Meeting of The Illinois Language and Linguistics Society, Urbana-Champaign, IL.

Trousdale, G, & Hoffmann, T., (Eds.). (2013). *The Oxford handbook of construction grammar*. Oxford University Press.

Uchida, N., & Imai, M. (1999). Heuristics in learning classifiers: The acquisition of the classifier system and its implications for the nature of lexical acquisition. *Japanese Psychological Research*, *41*(1), 50-69. https://doi.org/10.1111/1468-5884.00104

Walker, B. D., Rehfeldt, R. A., & Ninness, C. (2010). Using the stimulus equivalence paradigm to teach course material in an undergraduate rehabilitation course. *Journal of Applied Behavior Analysis*, *43*(4), 615-633. https://doi.org/10.1901/jaba.2010.43-615

Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, *21*(1), 1-48. https://doi.org/10.1017/S0272263199001011

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*(2), 269-304. https://doi.org/10.1017/S0272263105050138

Yamamoto, J. (1986). Establishing stimulus equivalences in autistic children. *Japanese Journal of Behavior Analysis*, *1*, 1-21.

Yi, B. U. (2011). What is a numeral classifier? *Philosophical Analysis*, *23*, 195-258.

Yu, B., & Watkins, D. A. (2008). Motivational and cultural correlates of second language acquisition: An investigation of international students in the universities of the People's Republic of China. *Australian Review of Applied Linguistics*, *31*(2), 17-1. https://doi.org/10.2104/aral0817

Zhang, J., & Lu, X. (2013). Variability in Chinese as a foreign language learners' development of the Chinese numeral classifier system. *The Modern Language Journal*, *97*(S1), 46-60. https://doi.org/10.1111/j.1540-4781.2012.01423.x

Yee Pin Tio, Clinical Assistant Professor, School of Psychological and Behavioral Sciences, Southern Illinois University Carbondale, USA. Research interests: Language and Cognition, the effects of language learning across the lifespan.

Usha Lakshmanan, Professor, School of Psychological and Behavioral Sciences (Program in Brain and Cognitive Sciences) and Professor of Linguistics, Southern Illinois University Carbondale, USA. Research interests: Psycholinguistics, First and Second Language Acquisition, Bilingualism and Multilingualism, Language and Cognition.

基于刺激对等范式的汉语量词内隐学习研究

赵裕萍
南伊利诺大学，美国

乌沙·拉沙马兰
南伊利诺大学，美国

摘要

汉语量词作为学习名词量化的必要知识，给汉语作为第二语言的学习者带来了严峻的挑战。学习者通常错误地将量词应用到新学的名词，因而产生错误的搭配，究其原因主要在于：（1）汉语量词系统本身存在复杂性和不一致性；（2）学生缺乏对量词和适用名词搭配的输入性训练。本研究采用前测－后测随机分组设计来考察采用刺激对等范式的内隐教学法是否对学习者的汉语量词学习产生正面效益，同时研究量词范畴内的变化性（根据母语为汉语者的典型性评估）是否会影响学习者的量词习得。87 名母语为英语的受试者被分配到两个学习组别（内隐学习组／外显学习组）来系统学习四个汉语量词。研究发现，只有使用刺激对等范式的内隐学习组在经过一周的训练后对量词的记忆有显著改善。进一步的分析表明，使用刺激对等范式的内隐学习组在面对非典型的量词名词配对时，比外显学习组更能准确地将量词推广到新学的名词。对于典型的个体量词名词搭配，内隐学习组的整体效果与外显学习组相当。研究结果表明，刺激对等范式的内隐教学法相较于外显学习法更能促进非典型的量词名词配对的推理学习。透过认知语言学的视角，本研究有助于进一步理解汉语量词的功能和语义范畴，同时为量词的教学提出一套更为有效的学习方式。

关键词

汉语量词，内隐教学，外显教学，刺激对等范式，典型效应，物体分类

作者简介

赵裕萍博士现任美国南伊利诺伊大学心理和行为科学学院临床助理教授，研究兴趣包括语言及认知功能，以及语言学习在不同年龄段产生的效益。

乌沙·拉沙马兰博士现任美国南伊利诺伊大学心理和行为科学学院（大脑与认知项目）和语言学系教授，研究兴趣包括心理语言学、第一及第二语言习得、双语和多语能力，以及语言和认知功能。