



Published in final edited form as:

*Lang Learn Dev.* 2008 October 1; 4(4): 249. doi:10.1080/15475440802333858.

## Classifiers as Count Syntax: Individuation and Measurement in the Acquisition of Mandarin Chinese

**Peggy Li,**

Department of Psychology, Harvard University

**David Barner,** and

Department of Psychology, University of Toronto

**Becky H. Huang**

Psychological Studies in Education, University of California–Los Angeles

### Abstract

The distinction between mass nouns (e.g., *butter*) and count nouns (e.g., *table*) offers a test case for asking how the syntax and semantics of natural language are related, and how children exploit syntax-semantic mappings when acquiring language. Virtually no studies have examined this distinction in classifier languages (e.g., Mandarin Chinese) due to the widespread assumption that such languages lack mass-count syntax. However, Cheng and Sybesma (1998) argue that Mandarin encodes the mass-count at the classifier level: classifiers can be categorized as “mass-classifiers” or “count-classifiers.” Mass and count classifiers differ in semantic interpretation and occur in different syntactic constructions. The current study is first an empirical test of Cheng and Sybesma’s hypothesis, and second, a test of the acquisition of putative mass and count classifiers by children learning Mandarin. Experiments 1 and 2 asked whether count-classifiers select individuals and mass classifiers select nonindividuals and sets of individuals. Adult Mandarin-speakers indeed showed this pattern of interpretation, while 4- to 6-year-olds had not fully mastered the distinction. Experiment 3 tested participants’ syntactic sensitivity by asking them to match two syntactic constructions (one that supported the mass or portion reading and one that did not) to two contrasting choices (a portion of an object and a whole object). A developmental trend was observed for the syntactic knowledge from 4-year-old children into adulthood: adults were near perfect and the older children were more likely than the younger children to correctly match the contrasting phrases to the objects. Thus, in three experiments we find support for Cheng and Sybesma’s analysis, but also find that children master the syntax and semantics of Mandarin classifiers much later than English-speaking children acquire knowledge of the English mass-count distinction.

### INTRODUCTION

Natural language encodes representations of sets and individuals. When speakers utter a word such as *cat*, we know that they intend to refer to a discrete, bounded individual (e.g., an individual feline), and not arbitrary portions thereof (e.g., pieces of cat). In English and other Indo-European languages, one cue to individuation is count syntax: words used directly with numerals (e.g., *one cat*), in singular or plural forms (e.g., *a cat*, *some cats*), or with quasicardinal determiners (e.g., *these/those cats*) all signal reference to sets of countable things. However, many languages, such as Chinese and Japanese, lack count syntax and instead use classifiers to explicitly signal reference to individuals. Recently, several investigations in linguistics have

suggested that count syntax and classifiers may be two variations on a single syntactic phenomenon (Borer, 2005; Cheng & Sybesma, 1998), and may express identical underlying structures and semantic distinctions. Here, we explore this hypothesis by testing the interpretation and acquisition of classifiers in Mandarin Chinese.

In all languages that have count syntax, there are also mass nouns (Allan, 1980). Words used in mass syntax, unlike count nouns, cannot be used directly with numerals, singular-plural morphology, or quasicardinal determiners. Also, mass and count constructions selectively specify different quantifiers (e.g., \*much/many tables vs. much/\*many wood). It is frequently claimed that this syntactic difference between mass and count corresponds to a clear-cut semantic distinction, whereby count nouns denote individuals but mass nouns do not (Quine, 1960; Bloom, 1994, 1999; Gordon, 1985, 1988; Link, 1983; Landman, 1991; Wisniewski, Imai & Casey, 1996). That is, count nouns (e.g. *cat*, *table*, *idea*) refer to things that have “atomic structure,” with “atoms” or “individuals” that can be counted. Mass nouns (e.g., *water*, *wood*, *fun*) refer to homogenous things that do not have atomic structure. As Wisniewski et al. (1996, p. 271) puts it, speakers “conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as nondistinct, uncountable, unindividuated things”. Thus, according to this hypothesis, count nouns individuate but mass nouns do not.

Many previous studies have argued that classifier languages like Mandarin lack syntactic structures equivalent to those found in mass-count languages, and that all nouns in these languages are syntactically mass (Allan, 1980; Chierchia, 1994, 1998; Krifka, 1995). Three main facts have been cited in support of this conclusion. First, nouns in Mandarin, like English mass nouns, cannot simultaneously occur directly with number words, but require classifiers for counting (i.e., words akin to English measure words like *piece* and *bit*—e.g., *two pieces of toast*). Classifiers reflect information such as the shape, animacy, functionality, and the unit of measure of the noun’s referent (e.g., *san zhi bi* = three *stick* pen, or “three pens”). Second, in Mandarin, pluralization is almost nonexistent in speech when referring to plural sets. The plural marker “men” is restricted to animate beings and is only obligatory on personal pronouns (e.g., *wo* = I/me, *women* = we/us). Third, Mandarin quantifiers can be used both with nouns that refer to individuals and those that do not (e.g., *hen duo bi* = very many pens or *hen duo shui* = very much water). Thus, Mandarin nouns exhibit all of the hallmarks of English mass nouns.

The hypothesis that all nouns in classifier languages are mass has important consequences for understanding both the syntax and semantics of nouns cross-linguistically. If all Mandarin nouns are mass, and mass nouns cannot denote individuals, then it follows that Mandarin nouns should not be able to individuate unless they are accompanied by a classifier. This view is made explicit by Borer (2005):

[T]he need for a classifier projection to license counting vs. the absence of classifiers in the context of mass interpretation confirms the claim that in the absence of classifiers, [noun] predicates in Chinese are interpreted as mass. Thus, at least in Chinese there is direct evidence that count interpretation must be structurally licensed, but mass interpretation need not be. (Borer, 2005, p. 108)

Some psychologists have expressed a similar view. For example, it has been suggested (1992; Lucy & Gaskins, 2001; 2003) that nouns in classifier languages refer to an unindividuated essence, and that speakers of classifier languages may consequently think differently about objects and other individuals relative to speakers of mass-count languages such as English. A series of studies show that speakers of classifier languages (Yucatec-Mayan, Japanese, and Mandarin) are less likely to construe stimuli as individuals than speakers of English, resulting in differences in how speakers extend novel words, their ability to recall the number of items in a set, and so on (see Lucy, 1992; Imai & Gentner, 1997; Imai & Mazuka, 2003; cf. Li,

Dunham, & Carey, in press). By this view, since nouns in classifier languages are by default mass (and lack the individuating structure of English count nouns), speakers of these languages make different assumptions about the meanings of new words.

To summarize, it is commonly argued that (1) nouns that lack count syntax do not encode individuation, (2) nouns in classifier languages are by default mass, and (3) nouns in classifier languages do not individuate. This view, however, has not gone unchallenged. First, it has been questioned whether mass syntax specifies an unindividuated interpretation. Second, studies have challenged the idea that bare nouns in classifier languages cannot individuate. A third challenge (and the main focus of the present paper) concerns the question of whether classifier languages truly lack mass-count syntax. Below, we review these challenges with a focus on elucidating the effect of noun phrase syntax on the interpretation of nouns cross-linguistically.

Is count syntax required for individuation or can nouns also individuate when used in mass syntax? Several researchers have argued that English mass nouns are not limited to denoting nonindividuals, and that some (if not all) mass nouns denote sets of individuals (Barner & Snedeker, 2005; Chierchia, 1998; Gillon, 1999). Take, for example, the English mass noun “furniture.” “A piece of furniture” cannot refer to just a leg of a chair, but must denote a whole individual (e.g., *a chair, a table, a bookshelves*, etc.). Only “a piece of a piece of furniture” can refer to the leg of a chair. This suggests that mass nouns like furniture do provide natural atomic units for counting; namely, any thing that counts as a “piece” (Doetjes, 2007). This intuition has been supported by experimental studies that probe how mass-count syntax affects judgments of amount. When asked to decide which of two sets contains “more furniture” or “more mail,” subjects base quantity judgments on number (e.g., judging that six tiny pieces of mail are *more mail* than two large pieces), despite basing judgments on volume for other mass nouns that denote solid pieces of stuff (e.g., *stone*; Barner & Snedeker, 2005, 2006). Similarly, mass nouns that denote instantaneous or “punctual” events, such as *jumping* and *kicking*, lead to quantity judgments based on number; while mass nouns that denote nonpunctual or “durative” events, such as *swimming* and *running*, lead to judgments based on dimensions such as time and distance (Barner, Snedeker, & Wagner, in press). Studies of word extension find similar results. When an unfamiliar object is named with a novel count noun, subjects almost always extend the word to other objects that share the same shape (but not to those with the same substance and a different shape). However, mass syntax leads to mixed judgments, and subjects do extend novel mass nouns on the basis of shape (especially for complex-shaped objects; Soja, 1992; Subrahmanyam, Landau, & Gelman, 1999; Samuelson & Smith, 1999).

Given that mass nouns in languages like English can denote individuals, one might question the related proposal that all nouns in classifier languages denote unindividuated entities. Crucially, using Barner and Snedeker’s (2005) quantity judgment task, recent studies have found evidence that many nouns in classifier languages also supply criteria for individuation. In the absence of classifiers, Japanese subjects base quantity judgments on number to the same extent as English speakers for count nouns such as *cup, plate, and candle*, and for mass nouns such as *furniture, jewelry, and mail*, but respond at chance for words that can be used as either mass or count nouns in English, such as *string, paper, and chocolate* (Barner, Inagaki & Li, in preparation; Inagaki & Barner, 2006). Further, only small differences are found between Japanese, Chinese, and English speakers in word extension tasks (Barner et al., in preparation; Imai & Mazuka, 2003; Li, Dunham, & Carey, in press), and even these differences may not be attributable to lexical differences as suggested by Lucy (1992). Contrary to Lucy, others have argued that it is the relatively high presence of count syntax that leads English speakers to guess that novel nouns referring to unknown objects tend to be count nouns (see Imai & Mazuka, 2003; Gleitman & Papafragou, 2005; Li, Dunham, & Carey, in press).

These observations regarding the interpretation of mass nouns in English and of bare nouns in classifier languages lead us to the final challenge regarding the purported privileged relation between count syntax and individuation. A common observation among researchers studying Chinese and other classifier languages is that not all classifiers are created equal (e.g., Lyons, 1977; Senft, 2000; Grinevald, 2004). As noted by Lyons (1977), some classifiers appear to perform a sortal function, occurring in constructions that pick out individuals or sets of individuals, whereas other classifiers perform a measuring function and denote specific measures of individuals or unindividuated stuff (these latter words are sometime referred to as “mensural classifiers” or “measure words”). Cheng and Sybesma (1998) have argued that this difference between classifier types may be equivalent to the distinction between mass and count constructions in Indo-European languages such as English. Though not obligatory (and often restricted to counting contexts), classifiers may provide Mandarin with a syntactic distinction that is homologous to the mass-count distinction in English.

Cheng and Sybesma (1998) distinguish between what they call “count” classifiers and “mass” classifiers on the basis of three criteria. First, most count classifiers are related to nouns by rote memorization, whereas mass classifiers can be used productively with a range of nouns. Count classifiers truly *classify* nouns. Second, mass classifiers, unlike count classifiers, are typically derived from nouns. For example, the word “wan” functions as a classifier in “yi wan tang” (a bowl of soup) and as a noun in “yi ge wan” (a bowl). As a result, mass classifiers form an open class; practically any noun that denotes measures of stuff, such as bowls, cups, or buckets, can be used as a mass classifier. Count classifiers, on the other hand, are a closed class. Third, Cheng and Sybesma note that the two types of classifiers occur in different syntactic environments. For one, the insertion of the modifier “de” is grammatical in the classifier phrase number-CL-N for mass classifiers, but not for count classifiers (see Example 1). The insertion of “de” modifies the meaning of the phrase subtly by placing emphasis on the total amount of stuff and away from reference to the individual counting units. With *de*, the example in 1a means “three bottles worth of wine” and can refer to that amount of wine regardless of the form of its container(s). Thus, hearing an expression such as 2a, one imagines three distinct bottles of wine, whereas hearing 2b, the presence of actual bottles is not necessarily (e.g., the wine may be spilt on the table).

#### Example 1

- a. san ping (de) jiu  
three CL-bottle DE liquor  
“three bottles of wine”
- b. san ge (\*de) ren  
three CL-individual DE person  
“three people”

#### Example 2

- a. zhuo shan you san ping jiu  
table top has three CL-bottle liquor  
“there are three bottles of wine on the table”
- b. zhuo shan you san ping de jiu  
table top has three CL-bottle DE liquor  
“there are three bottles worth of wine on the table”

Finally, only mass classifiers permit the insertion of adjectives between the numeral and the classifier,<sup>1</sup> which relates to the fact that mass classifiers are derived from nouns, and nouns are also subject to adjectival modification (see Example 3).

Example 3

- a. yi da ping jiu  
one big CL-bottle liquor  
“one big bottle of wine”
- b. \*yi da ge ren  
one big CL-individual person  
“one big person”

Based on this analysis, Cheng and Sybesma conclude that Mandarin makes a clear distinction between two kinds of classifiers and that this distinction is nearly isomorphic to the mass-count distinction in English. In their words:

[Count classifiers] form a closed class and they solely classify nouns that are cognitively singularizable ... nouns such as pens, dogs, etc. [Mass classifiers] do not form a closed class, and they classify (create counting units for) nouns that are cognitively masses, such as water and sand, (plural) pens and dogs. (Cheng & Sybesma, 1998, p. 403)

Therefore, analogous to English, in which count syntax picks out individuals, Mandarin count classifiers signal reference to countable individuals. Mass classifiers, on the other hand, can appear with nouns that refer to unindividuated stuff, as well as those that denote pluralities of individuals.

We suspect that an even simpler analysis may be in order, wherein only so-called count classifiers are relevant to making the mass-count distinction in Mandarin. Many previous studies have noted that what Cheng and Sybesma call mass classifiers function as measure words (Tai, 1990; Kuo, 1998; Tai & Chao, 1994; Huang & Ahrens, 2003). In Mandarin, and in mass-count languages such as English, measure words such as *inch* or *cup* designate units of measurement for nouns that refer to unindividuated stuff (e.g., *an inch of rain*, *a cup of milk*) and also to portions of individuals (e.g., *an inch of beads*, *a cup of berries*). In English, these measure words can be applied to either mass or to count nouns, and thus are not predictive of the syntactic category of the noun as mass or count. Similarly, in Mandarin they can occur with most nouns regardless of whether those words can also be used with count classifiers. Thus, measure words in Mandarin more closely resemble measure words in English than they resemble mass syntax.

Cheng and Sybesma’s claim that Mandarin makes a mass-count distinction is rooted most strongly in the observation that count classifiers provide a specialized structure for signaling reference to things *qua* individuals. When paired with a noun, count classifiers—such as count syntax—specify individuation. In contrast, nouns that lack count classifiers (i.e., bare nouns) do not obligatorily specify anything about (non)individuation; some bare noun phrases refer to individuals and others to nonindividuated stuff. Finally, what Cheng and Sybesma call “mass classifiers,” and what we will call “measure words” (to emphasize their similarity to English measure words), pick out measurements of portions and pluralities. These words can occur

<sup>1</sup>To be precise, only a restricted set of adjectives may modify mass classifiers. Specifically, following Hoekstra (1988), only adjectives (e.g., *da* – big, *xiao* – small, *man* – full, *zheng* – whole, *chang* – long, etc.) that modify the abstracted function of a container may modify the mass classifiers. Additionally, the mass classifiers must be container-based and not exact measures (e.g., *bang* – pound).

with any type of noun to refer to measurements of unindividuated stuff or of multiple individuals. See Table 1 for a summary.

This framework proposes an important homology between mass-count syntax and the classifier system in Mandarin, a homology which may exist in other numeral classifier languages (e.g., see Lyons, 1977; Croft, 1994; Tai 1990; who argue for distinguishing between “sortal” and “mensural” classifiers across a host of languages, including Japanese, Thai, Vietnamese, etc.). If correct, this observation of cross-linguistic similarity is relevant to researchers studying relations between language and thought, language acquisition, and conceptual development, and to those interested in characterizing points of variation and universality across the world’s languages. The present paper, therefore, tests Cheng and Sybesma’s linguistic analysis by evaluating the semantics of Mandarin classifiers and whether or not their interpretation resembles that of mass-count syntax in English.

To assess whether classifiers and mass-count syntax support similar interpretations cross-linguistically, we investigated Mandarin speaking adults’ interpretation of count classifiers and measure words. We also investigated the acquisition of the classifier system in children learning Mandarin. If languages such as Mandarin and English draw on common mechanisms for acquiring count syntax and classifiers, we might expect to find similar signatures in the acquisition of each. For example, children in the two cultures might acquire count syntax and count classifiers at the same rate, pass through similar stages, use similar cues, and respect the same biases and constraints. While many experiments have investigated the interpretation and acquisition of mass-count syntax in English (Gordon, 1985, 1988; Bloom, 1994, 1999; Barner & Snedeker, 2005, 2006; Barner, Wagner, & Snedeker, 2008; Gathercole, 1985; Samuelson & Smith, 1999; Soja, 1992; Soja, Carey, & Spelke, 1991; Lucy, 1992; Imai & Gentner, 1997; Subrahmanyam, Landau, & Gelman, 1999; Wisniewski, Imai & Casey, 1996), virtually no studies have tested for similar effects in the development of classifiers in Mandarin. For the most part, studies of classifier acquisition have focused on issues such as the content of early acquired classifiers (e.g., a preference for unmarked, general classifiers early in acquisition, when other classifiers are still being mastered; Muraishi, 1983, reported in Yamamoto & Keil, 2000; Matsumoto, 1987; Uchida & Imai, 1996; Carpenter, 1991; Erbaugh, 1986), and the resilience of animate classifiers relative to classifiers for inanimates over generations of speakers (Sanches, 1977). Other studies have examined children’s ability to choose the correct classifier for specific known words (Muraishi, 1983; Yamamoto & Keil, 2000; Erbaugh, 1986). However, there is almost no experimental evidence to support the claim that count classifiers and count syntax have analogous effects on interpretation. Furthermore, there is little evidence regarding the emergence of count classifiers and measure words in language acquisition. By exploring the developmental course of classifier acquisition, we stand to gain insight into the conceptual foundations of the distinction and to determine whether learning one language or another alters the fundamental properties of conceptual resources (e.g., object construal), which are made available to children across different cultures.

Only one empirical study (Chien, Lust & Chiang, 2003) has explored the acquisition of count classifiers and measure words in children learning Mandarin. With the help of a puppet who requested things (“*I want one CL something*”), Chien et al. asked children to select one of three objects, where only the intended object was consistent with the particular classifier used. They found that children performed better than chance for a number of classifiers at three years of age, the youngest age they tested. They also found that children were more likely to pick solid objects as referents for count classifiers and nonsolid substances as referents for measure words. This is consistent with the common assumption that solid objects (e.g., balls, plates) are good individuals and that nonsolid substances (e.g., water, sand) are not (e.g., Soja, Carey, & Spelke, 1991; Imai & Gentner, 1997). Based on these findings, Chien et al. concluded that Chinese children distinguish between the syntax and semantics of count classifiers and measure



words by as early as 3-years-old. However, their conclusion may be premature. Since Chien et al. presented children with familiar, nameable items (which they named for the children), children may have solved the task by using knowledge of how particular familiar nouns are typically paired with classifiers, without actually knowing the meanings of the classifiers. As a result, previous studies leave open what, if anything, children knew about either the syntax or the semantics of count classifiers and measure words.

Our first two experiments investigated Mandarin-speaking adults' and children's extension of count classifiers and measure words using novel referents. We asked whether count classifiers specify individuation and whether measure words pick out both portions of stuff and sets of individuals. Like many other studies (Huntley-Fenner, Carey, & Solimondo, 2002; Soja, Carey, & Spelke, 1991; Imai & Gentner, 1997; Samuelson & Smith, 1999; Yoshida & Smith, 2005; Chien et al., 2003), we used solidity as a proxy for individuation: piles or portions of nonsolids are not typically referred to as individuals by nouns, whereas solid objects are. In our experiments, like Chien et al., a puppet requested for "one CL/MW (count classifier/measure word) something," and the participants had to choose among novel referents. Since the noun "something" was used instead of a specific noun that contributes meaning, only the classifiers could be used to restrict interpretation and select between the visible alternatives. If Cheng and Sybesma's hypothesis is correct, then we should expect that the count classifiers will pick out referents that are good individuals (i.e., solids), and measure words will pick out referents that are either portions of nonsolid stuff or pluralities of individuals.

Finally, our third experiment tested adults' and children's sensitivity to classifier syntax. Specifically, we asked whether our Mandarin speakers are sensitive to insertion of *de* and the position of the adjective relative to the classifier.

## EXPERIMENT 1: COUNT CLASSIFIERS AND REFERENCE TO INDIVIDUALS

The first experiment tested Mandarin-speaking adults' interpretation of what Cheng & Sybesma call "count classifiers." Many count classifiers can only be used with nouns that denote things of certain shapes. For example, *gen* is used with nouns that denote rod-shaped things (pencils, carrots), *zhi* with names for stick-shaped things (twig, matchstick), and *kuai* with names for chunk-shaped things (dinner roll, stone). Thus, adult speakers who understand these words should associate them with objects of certain shapes. Further, if this shape information is carried by the classifiers themselves and not exclusively by the nouns with which they are paired, then participants should extend classifiers by shape even for novel stimuli. They should also be sensitive to substance and solidity information, since a match in shape alone may not be sufficient for a stimulus to qualify as an individual in the context of these classifiers. Thus, we probed whether count classifiers individuate by testing whether adults would extend them to shape-matched physical objects.

In addition to testing adults, we also tested young children learning Mandarin. This allowed us to test the age at which children acquire adult-like interpretations of the words, and what hypotheses they make about the words early in acquisition.

### Methods

We recruited twenty 4-year-olds (mean = 4.9, range = 4.1–5.1) and sixteen 6-year-olds (mean = 6.0; range = 5.8–6.4) from preschools in Chiayi, a city in Taiwan. Previous studies (see Chien et al., 2003 for a review) indicate considerable development in Taiwanese children's classifier knowledge between three and seven years of age. By seven, children are comparable to adults in knowing which classifiers to choose for *familiar* objects. We thus selected two groups of children within the three to seven age range who were likely to have acquired the shape specifications of some count classifiers, and asked whether they would use these classifiers to

pick out only solid objects (and not nonsolid substances). As multilingualism is common in Taiwan, we checked with the teachers to ensure that Mandarin was the primary language for each child tested. The children's performance was compared to that of 12 adults ( $M = 20$ -years-old, 19–24) recruited from the student population at Toku University in Chiayi County, Taiwan. The adults participated voluntarily without compensation, while all children were given a small gift for their participation.

Adapting Chien, Lust, and Chiang's (2003) methodology, we presented three boxes as choice options for each test trial. Each box contained an item. A puppet then asked for one of the items ("I want one CL something"). The participant was asked to select a box based on the classifier (CL) used. The three options always included (1) an open box containing an unfamiliar solid object (canonical individual), (2) an open box containing an unfamiliar nonsolid substance (canonical nonindividual), and (3) following the method of Huang and Snedeker (under review) a closed box. The purpose of the closed box was to circumvent an interpretive problem with two-item forced choice methods: when presented with only two visible options, participants may believe that neither of the two options presented fulfils the request made of them, yet make a choice nonetheless. The closed box allows subjects to opt out of the two visible options based on the assumption that the correct choice is concealed in the closed box. The participants were told that if the requested thing was not visible (i.e., inside one of the open boxes), it would be inside the closed box. The closed boxes were never opened at any time during the study. In sum, for all trials, one open box always contained a solid object substance and the other open box always contained a nonsolid substance.

There were three test trial types: *no-shape match* trials, *solid shape match* trials, and *nonsolid-shape match* trials (see Figure 1). On *no shape match* trials, the shape of neither the solid object nor the nonsolid substance matched the shape specified by the count classifier in the puppet's request. Thus, we expected participants who knew the shape specification of the classifier to choose the closed box. On remaining trials, the classifier matched the shape of either the solid object (*solid shape match* trials) or the nonsolid substance (*nonsolid shape match* trials), but never both. For the *solid shape match* trials, we expected participants to pick the shape-matched solid only if they knew the shape specification of the classifier and believed that it specifies solidity as well. Finally, for the *nonsolid shape match* trials, we expected that if participants believe that count classifiers specify solidity, they should choose the closed box despite the availability of a shape-matched nonsolid.

We tested six count classifiers: *gen* (rod), *zhi* (stick), *tiao* (line), *zhang* (sheet), *pian* (slice), and *kuai* (chunk). Each classifier was tested once for each type of trial (*no match*, *solid shape match*, and *nonsolid shape match*), for a total of 18 test trials. Trials were presented in two random orders (one order was the reverse of the other). Distractor stimuli (i.e., stimuli that did not match the classifiers in shape) were randomly chosen from the shape-matches of the other classifiers with the restriction that the items for *gen*, *zhi*, *tiao* did not cooccur within one trial, since these classifiers specify similar shapes. Likewise, items for *zhang* and *pian* did not cooccur within one trial. All items occurred with equal frequency in the 18 trials (see Appendix A for photographs of the same-shape matches of each of the six classifiers).

Prior to the actual test trials, four additional practice trials were conducted to familiarize participants with the instructions. Two of the trials involved picking the correct color from three different colored Legos and the other two involved picking the correct shape from a cube, a ball, and a triangular block. The closed box was correct for two of these trials. Unlike the test trials, feedback was provided in practice trials for incorrect responses, and the closed box was opened to reveal the contained item when the participants chose wrongly.



## Results

**Manipulation check: No shape match trials**—We first examined the *no shape match* trials, which serve as control trials. For these trials, neither the solid choice nor the nonsolid choice matched the shape specified by the classifier. Thus, if the experimental paradigm is valid, we should expect adults, if not children, to choose the closed box most of the time. Confirming this prediction, all three age groups chose the closed box more than the two alternative choices combined and significantly above the chance-level of 33% (see Figure 2a). A univariate ANOVA with Age (4-year-olds, 6-year-olds, and adults) as a between-subjects variable and percentage choosing closed box as dependent variable indicated an overall effect of age ( $F(1, 45) = 379.5, p < .001$ ), which reflected the fact that the older participants were more likely to choose the closed box relative to the younger participants. The 4-year-olds picked the closed box slightly less often (61% of the time) than the 6-year-olds (79% of the time;  $t(34) = 1.71, p = .09, n.s.$ ), and less often than the adults (100% of the time;  $t(30) = 4.46, p < .001$ ). Also, the 6-year-olds picked the closed box significantly less than the adults ( $t(26) = 2.10, p = .045$ ). The fact that 100% of adults picked the closed box when no shape match was visible validates the three box method. Also, the results reveal a developmental difference in the choice of the closed box, perhaps due to children's developing knowledge of how classifiers specify shape.

**Shape Match Trials: Solid vs. Nonsolid Shape Match Trials**—We next examined the two-shape match trials. On *solid shape match* trials, the solid matched the shape specified by the classifier. On *nonsolid-shape match* trials, the nonsolid matched the specified shape (see Figure 2b for the *solid shape match* trials and Figure 2c for the *non-solid shape match* trials). We asked whether the participants would be more likely to pick the closed box when the shape match was a nonsolid, relative to when it was solid. Since the two trial types differed only in whether the solid or nonsolid matched the classifier in shape, a preference for shape-matched solid objects should result in fewer choices of the closed box on these trials, relative to the nonsolid match trials.

The percentage of closed box choices was entered as the dependent variable into an Age (4-years-old, 6-years-old, and adults)  $\times$  Trial Type (solid shape match trials, nonsolid shape match trials), ANOVA with Trial Type as a within-subjects variable. There was a main effect of Age ( $F(2, 45) = 3.67, p = .03$ ), indicating that the older participants were more likely to select the closed box (24% of the time for the 4-year-olds, 36% for the 6-year-olds, and 44% for the adults).

Importantly, a main effect of Trial Type ( $F(1, 45) = 139.0, p < .001$ ) confirmed that the adults and children were overall more likely to select the closed box for the *nonsolid shape match* trials (60% of the time) than the *solid shape match* trials (10% of the time). Finally, there was an interaction between Age and Trial Type as well ( $F(2, 45) = 13.67, p = .001$ ). This was due to a developmental difference on *nonsolid shape match* trials, but no equivalent difference on *solid shape match* trials. At every age participants preferred the closed box more on *nonsolid shape match* trials compared to *solid shape match* trials. For the *solid shape match* trials, the 4-year-olds, 6-year-olds, and adults chose the closed box only 8%, 18%, and 3% of the time, respectively. They chose the shape matched solid 83%, 81%, and 97% of the time, and the nonsolid distractor only 9%, 1%, and 0% of the time. For each group the difference between the closed box and shape-matched solid choices was significant (t-tests,  $p < .001$ ), as was the difference between the visible shape matched solid and nonsolid distractor (t-tests,  $p < .001$ ).

For the *nonsolid shape match* trials, the 4-year-olds, 6-year-olds, and adults chose the closed box 39%, 54%, and 86% of the time, respectively. They chose the shape-matched nonsolid 48%, 38%, and 14% of the time, and the solid distractor only 13%, 8%, and 0% of the time. The difference between closed box and shape-matched nonsolid choices was significant only

for the adults ( $t(11) = 8.99, p < .001$ ), while the difference between the visible shape matched nonsolid and solid distractor was significant for all groups ( $t$ -tests,  $p < .001$ ). This last result indicates that participants considered shape to be more important than solidity when interpreting the classifiers.

Thus, adults and children as young as 4-years of age are aware of the shape specifications for certain count classifiers. For adults this specification is strongly associated with solid objects, though shape is more important to them than solidity. Younger children's preference for the closed box on *nonsolid shape match* trials vs. *solid shape match* trials also indicates a preference to match classifiers to solid shaped matched objects. However, children chose the nonsolid shape match as often as the closed box, and preferred it over the solid distractor, suggesting that shape is more strongly associated with classifiers than solidity.

## Discussion

As expected, adults picked the closed box when no shape match was visible, and chose solid shape matches whenever possible. In general, they rejected shape matches that were nonsolid, suggesting that for adults, count classifiers specify reference to individuated entities—in this case, solid objects with nonarbitrary physical structures.

Children exhibited the same pattern as adults, except that they frequently chose nonsolid shape matches. This tendency to choose nonsolid shape matches declined with age, indicating a developmental change in children's sensitivity to solidity information. Still, at all ages children preferred shape-matched solids over the shape-matched nonsolids, suggesting that they were not completely oblivious to solidity. These data indicate that children begin acquiring classifiers by attending to shape, and become sensitive to solidity over a course of several years.

## EXPERIMENT 2: MASS CLASSIFIERS AND REFERENCE TO NONINDIVIDUALS AND PLURALITIES

Experiment 1 demonstrated that certain count classifiers specify reference to shape-matched solid objects, but not to shape-matched nonsolid substances, and that the association of count classifiers to solidity grows as children's language develops. The second experiment probed whether these shape and solidity specifications are specific to count classifiers, by contrasting them directly with measure words (i.e., what Cheng and Sybesma call “mass classifiers”).

Count classifiers and measure words appear to differ in three ways. First, as in English, and in contrast to count classifiers, measure words specify reference to portions of nonsolid stuff (e.g., one pile of sugar). Second, when used with the numeral “one,” only measure words could refer to sets with multiple objects (e.g., one pile of coins). Count classifiers used with “one” should always refer to singleton sets (e.g., one coin). Finally, measure words, unlike count classifiers, should not be used to refer to single objects (e.g., “one pile of a table” does not amount to “one table”). The question we ask here is whether measure words exhibit these constraints and whether they can be shown experimentally to contrast with count classifiers. Further, we ask how knowledge of the distinction between count classifiers and measure words emerges developmentally.

## Methods

We tested thirteen 4-year-olds ( $M = 4.10$ ; range = 4.5–5.1) and sixteen adults ( $M = 20$ ; range 19–22) who had not participated in Experiment 1, using the same experimental paradigm.

Two shape-based count classifiers, *gen* (rod) and *pian* (slice), were tested and compared to two shape-based measure words, *dui* (pile) and *tuan* (wad/ball). As in Experiment 1, we expected

adult participants to pick the closed box for the *no shape match* trials for both count classifiers and measure words. However, the difference between measure words and count classifiers should emerge for the shape-match trials (compare Figure 3 for measure words with Figure 1 for count classifiers). The predictions for the measure words are as follows: (1) For the *nonsolid shape match* trials, we expected participants to select the nonsolid substance (e.g., a pile of unfamiliar nonsolid substance for “*dui*”). (2) For the *solid shape match* trials, we expected the participants to pick the closed box and refrain from picking the single shape-matched solid (e.g., a solid object in the shape of a pile for “*dui*”). The predictions differ for count classifiers, where we expected participants to pick the closed box for the *nonsolid shape match* trials and the solid for the *solid shape match* trials.

In addition to the three trial types of Experiment 1 (*no shape match*, *solid shape match*, and *nonsolid shape match* trials), we tested participants with trials that we called *multiple solids* trials. This fourth trial type tested whether measure words in the context of “one MW something” should pick out a set of individuals (e.g., one pile of coins), whereas a count classifier in the context of “one CL something” should only be able to pick out one individual object (e.g., one coin).

The three options for the *multiple solids* trials included an open box with multiple solids, an open box with a nonsolid substance, and a closed box (Figure 4). Of the open boxes, the box with the multiple solids matched the shape specified by the classifiers/measure words. More specifically, the box contained a set of identical solid objects whose aggregate shape matched the classifier. Each individual solid of the group, however, did not match the specified shape. For example, for the measure word *dui* (pile), the individual solid objects were rectangular blocks and the blocks were stacked into a pile. For the count classifier *gen* (rod), several individual one centimeter tubes with one centimeter diameters were aligned horizontally to create a rod shape. Performance for measure words was compared to that for count classifiers. In the case of the measure words we expected participants to select the multiple solids box, since the aggregate shape of the multiple objects matched the description of “one MW something.” In the case of count classifiers, we expected participants to choose the closed box, since, although the aggregate shape of the objects in the multiple solids box matched the classifier specification, no individual object in any of the open boxes did.

Both of the count classifiers (*gen/rod*, and *pian/slice*) and both of the measure words (*dui/pile*, and *tuan/wad-ball*) were tested once for each of the four trial types, totaling 16 trials (see Appendix B for photographs of the shape match choices). The order of trials, selection of the nonshape match items, and the positioning of the boxes for the no match, solid shape match, and solid were randomized in the same way as Experiment 1.

## Results

**Manipulation check: No shape match trials**—For the *no shape match* trials, we expected that adults, if not children, would choose the closed box over the other two choices, since neither the solid nor the nonsolid matched the shape specified by the count classifiers or the measure words. Confirming this prediction, all age groups chose the closed box more than either of the two alternative choices (see Figures 5a and 5d). Preference for the closed box was above the 33% criterion of chance on both the count classifiers and measure words for all groups ( $p$ s < .05).

Data were submitted to an ANOVA with Syntax (count classifiers vs. measure words) as a within subjects variable, Age (children vs. adults) as a between subjects variable, and percentage of closed box choices as the dependent variable. There were no significant main effects (Age  $F(1, 27) = 3.23$ ,  $p = .08$ , n.s.; Syntax,  $F(1, 27) = 3.60$ ,  $p = .07$ , n.s.) and no interaction effects (Age x Classifier Type,  $F(1, 27) = .21$ ,  $p = .67$ , n.s.). Thus, both adults and children

showed the same preference for the closed box, regardless of syntax. The marginal effect of Age, was due to adults preferring the closed box slightly more than the children, such as in Experiment 1 (92% vs. 77%). The marginal effect of Syntax resulted from subjects' slightly greater tendency to choose the closed box for the count classifiers (91%) than for the measure words (78%).<sup>2</sup>

Overall, the manipulation check confirmed that adults and children prefer the closed box when no shape match is visible for both count classifier and measure word trials.

**Shape match trials: Solid vs. nonsolid shape match trials**—We next compared the pattern of responses between count classifiers and measure words for the shape-matched trials (i.e., contrasted the pattern for Figure 5b and 5c with Figures 5e and 5f). Data were submitted to an ANOVA with Age (children vs. adults) as a between subjects factor, Syntax (count classifiers vs. measure words), and Trial Type (solid shape match vs. nonsolid shape match) as within subjects factors, and with percentage of closed box responses as the dependent variable.

Adults picked the closed box more often than the children (41% vs. 26% of the time), resulting in a main effect of Age ( $F(1, 27) = 5.65, p = .03$ ). There were no other main effects, though there were significant interactions between Trial Type and Syntax ( $F(1, 27) = 87.89, p < .001$ ) and between Trial Type, Syntax and Age ( $F(1, 27) = 13.0, p < .001$ ).

The Trial Type  $\times$  Syntax interaction reflects the fact that responses for *solid shape match* and *nonsolid shape match* trials were in the opposite direction for count classifiers and measure words. For count classifiers, subjects preferred the closed box on nonsolid shape match trials (62% of the time), but not for the *solid shape match* trials (7% of the time). Conversely, when given measure words, subjects preferred the closed box for the *solid shape match* trials (58% of the time) but not for the *nonsolid shape match* trials (7% of the time). These results confirm the hypothesis that measure words used with the numeral “one” select portions of stuff, but not single solid objects, whereas count classifiers show the opposite pattern.

The three-way interaction between Trial Type, Syntax, and Age reflects differences between children and adults for only the critical trials, i.e., count classifiers for the *nonsolid shape match* trials and measure words for the *solid shape match* trials. For count classifiers, the children and adults do not differ in picking the closed box for the *solid shape match* trials (children: 12% vs. adults: 3%,  $t(27) = 1.3, p = .21, n.s.$ ), but they do differ for the *nonsolid shape match* trials (children: 46% vs. adults: 78%,  $t(27) = 2.48, p < .05$ ). This result replicates the findings of Experiment 1. First, for the *solid shape match* trials, both adults and children were equally unlikely to pick the closed box and more often preferred the shape-matched solid choice instead (respectively, 12% vs. 85% for children and 3% vs. 94% for adults; paired  $t$ -tests,  $p < .001$ ). Second, children were less likely to choose the closed box for *nonsolid shape match* trials, relative to adults. Adults chose the closed box more often than the shape-matched nonsolid (respectively 78% vs. 9%,  $t(15) = 6.01, p < .001$ ), while the children did not (46% vs. 42%,  $t(12) = .19, p = .86, n.s.$ ).

Conversely, for measure words, there was a difference in the extent to which the children and adults picked the closed box for the *solid shape match* trials (children: 38% vs. adults: 78%,  $t(27) = 2.93, p < .01$ ), but no such difference for the *nonsolid shape match* trials (children: 8%

<sup>2</sup>The almost significant Classifier Type difference could be attributed to the particular classifiers chosen for the study. Participants might have sometimes opted to select the non-shape matching non-solids when given mass classifiers, *dui* (pile) and *tuan* (wad/ball), for two reasons. First, perhaps the two classifiers' shape requirements overlap. What is a pile is sometimes acceptable to some as a wad, and vice versa. Second, what is acceptable in shape as a *dui* (pile) or a *tuan* (wad) is less stringent in comparison to what is acceptable as a *gen* (rod) or a *pian* (slice), the two count classifiers.

vs. adults: 6%,  $t(27) = .22$ ,  $p = .83$ , n.s.). More specifically, for the *solid shape match* trials, the children did not pick the closed box more often than the solid choice (38% vs. 27%,  $t(12) = .67$ ,  $p = .51$ , n.s.) or the nonsolid choice (38% vs. 35%,  $t(12) = .18$ ,  $p = .86$ , n.s.), while adults overwhelmingly did (78% closed box vs. 9% solid choice and 13% nonsolid choice, paired  $t$ -tests,  $p < .001$ ). For the *nonsolid shape match* trials, both adults and children equally dispreferred the closed box and instead more often selected the shape-matched nonsolid choice (respectively, 8% vs. 85% for children; 6% vs. 94% for adults,  $t$ -tests,  $p < .001$ ).

**Classifier type and choice of multiple individuals**—Next, we examined the *multiple solids* trials in which the solid option consisted of multiple solids whose aggregate shape matched the classifier’s specified shape. The findings from these trials (Figures 6b and 6d) are contrasted with the findings from the *solid shape match* trials (Figures 6a and 6c, replotted from Figures 5b and 5d) in which the solid option consisted of only one individual solid. For present purposes, we will refer to the *solid shape match* trials as *single solid* trials. From the analyses already presented above, we know that for these *single solid* (i.e., *solid shape match*) trials that measure words used in the context of “*one MW something*,” do not pick out single solid objects, whereas count classifiers do (see Figures 8a and 8c). Thus, the main question here was whether measure words, but not count classifiers, are consistent with sets of individuals (i.e., pluralities) when used in the context of “*one CL/MW something*.”

Data were submitted to an ANOVA with Syntax (count classifiers vs. measure words) as a within subjects factor, Number (singular vs. plural set) and Age (adults vs. children) as between subjects factors, and percentage of shape-matched solid choice as the dependent variable (see Figure 7). There was a main effect of Number ( $F(1, 27) = 6.79$ ,  $p = .015$ ) and no other significant main effects. The main effect of Number indicated that the single solid choice was selected more often than the multiple solid choice (54% vs. 42%). More importantly, the analysis confirmed the expected interaction between Syntax and Number ( $F(1, 27) = 164.93$ ,  $p < .001$ ). Plural choices were more frequent for measure words (68% overall children & adults) than for count classifiers (15% overall children and adults;  $t(28) = 5.68$ ,  $p < .001$ ), and singular choices were more frequent for count classifiers (89% overall children & adults) than for measure words (15% overall children and adults;  $t(28) = 13.63$ ,  $p < .001$ ). Further, a significant interaction between Syntax, Number, and Age indicated a developmental trend in this difference between measure words and count classifiers ( $F(1, 27) = 19.561$ ,  $p < .001$ ). This three-way interaction is driven by a lack of difference between children and adults for the single solid trials, but differences for the multiple solids trials. On single solid trials there was no significant difference between children and adults for the count classifiers (85% vs. 94% for children and adults respectively;  $t(27) = 1.20$ ,  $p = .24$ , n.s.) or for the measure words (27% vs. 9% for children and adults respectively;  $t(27) = 1.79$ ,  $p = .08$ , n.s.). For multiple solid trials, children and adults differed in their judgments for both count classifiers and measure words. On multiple solid trials there was a significant difference between children and adults for both the count classifiers (31% vs. 0% for children and adults respectively;  $t(27) = 3.80$ ,  $p = .001$ ) and measure words (54% vs. 81% for children and adults respectively;  $t(27) = 2.14$ ,  $p < .05$ ). While both children and adults were more likely to pick multiple objects for measure words than for count classifiers, and single objects for count classifiers than for measure words, adults were more categorical in their judgments overall.

## Discussion

Experiment 2 replicated the pattern of results found in Experiment 1 for count classifiers, and established how they differ from Mandarin measure words. Children interpreted count classifiers in an adult-like manner. They chose solid objects for the *solid shape match* trials and the closed box for *no match* trials, indicating their sensitivity to the shape information encoded by classifiers. As in Experiment 1, children often chose the shape matched nonsolid



on *nonsolid shape match* trials, again suggesting that they acquire the shape specifications of count classifiers before they associate them with solidity. However, children are not entirely ignorant of the relation between count classifiers and solidity. This is evidenced by their far lower preferences for the shape match for *nonsolid shape match* trials than for the *solid shape match* trials.

The pattern for measure words differed from that of count classifiers in the way predicted by Cheng and Sybesma, particularly for adults. Adults overwhelmingly preferred the closed box for the *no shape match* trials, selecting it 88% of the time. They also preferred the nonsolids for the *nonsolid shape match* trials, selecting it 94% of the time. Critically, they showed a strong preference for the closed box and avoided the solid (78% vs. 9%) for the *solid shape match* trials. Thus, there was a stark contrast between adults' interpretation of measure words and count classifiers. Children's behavior resembled that of adults for all trials except the *solid shape match* trials, where they showed no preference between the three available options. In this respect, children's performance for the measure words resembled their performance for count classifiers: when presented shape matched nonsolids for count classifiers and shape matched solids for mass classifiers, children were more willing than adults to ignore the solidity specifications of the count classifiers and measure words and choose the shape-matched choice.

Experiment 2 also examined whether measure words can be used to refer to pluralities. According to Cheng and Sybesma's analysis, measure words (or mass classifiers), but not count classifiers, can pick out a set consisting of multiple individuals in the context of "*one MW/CL something*." This prediction is confirmed by the adults' performance. For count classifiers, adults overwhelmingly preferred the shape-matched solid choice on *single solid* trials (94%), but entirely refrained from choosing the shape-matched solid on *multiple solids* trials (0%). The choice pattern revealed their belief that count classifiers only pick out single and not multiple individuals. Adult responses for measure words displayed the opposite pattern; when given measure words, adults preferred the shape-matched multiple solids choice on the *multiple solids* trials (81%) and refrained from choosing the shape-matched solid on the *single solid* trials. Taken together, these results indicate that count classifiers, when used with the numeral "one," pick out single individuals, but not pluralities. Conversely, mass classifiers can be used to refer to pluralities but not single individuals.

Taken together, the findings from Experiments 1 and 2 show that children learn the shape specifications of count classifiers and measure words relatively early. Also, by four years of age, children show some evidence of understanding the solidity specifications of count classifiers and measure words, although their understanding is far from complete even by age six. These data suggest that the syntactic encoding of individuation in Mandarin is a relatively prolonged process compared to the case of English; even at six years of age, when English speaking peers appear to have grasped the semantics of mass-count syntax, Mandarin speaking children are not fully like adults.

### EXPERIMENT 3: CLASSIFIER SYNTAX

Cheng and Sybesma's argument for distinguishing between count and mass classifiers rests in part on the argument that the two types of *classifiers* appear in different syntactic environments. They point out that the insertion of *de* after a mass classifier subtly changes the meaning of the noun phrase to emphasize the total volume of stuff (see examples 2a and 2b in the Introduction). They also note that only mass classifiers can be modified by adjectives (see examples 3a and 3b). However, no study has ever tested speakers' sensitivity to these syntactic environments, much less when children become sensitive to the syntax. Experiment 3 therefore tested Mandarin-speakers' knowledge of classifier syntax.

Count classifiers can be followed by an adjective that modifies the noun—e.g., one-CL-ADJ-N (*yi ge da ren* = one CL<sub>indiv.</sub> big person). Typically, a count classifier construction of this kind picks out the whole object (i.e., individual) denoted by the noun. The modification of a count classifier with an adjective—e.g., one-ADJ-CL—is highly atypical by comparison and is likely to be judged ungrammatical without a supporting context (see example 3b in the Introduction).<sup>3</sup> However, if Cheng and Sybesma’s syntactic analysis is valid, one might expect that a count classifier used in a measure word frame could force quantification over stuff, just like English count nouns used in mass syntax (“There is dog all over the road”). If children and adults are sensitive to the syntactic difference, they should exhibit a difference in their interpretation of the count classifier construction when the adjective precedes the classifier, compared to when it follows the classifier.

We tested children and adults by asking them to match constructions to one of two stimuli. The constructions took the form “one-ADJ-CL-de-N” (normally consistent with only measure words) and “one-CL-ADJ-N” (consistent with both measure words and count classifiers). However, participants were presented only count classifiers used in both constructions.

## Methods

We recruited eighteen 4-year-olds (mean = 4.11, range = 4.4–5.3), sixteen 5-year-olds (mean = 5.11, range = 5.6–6.5), and 16 adults ( $M = 20$ ; range 19–22) who had not participated in the previous experiments. The children were from preschools in Chiayi and Taipei. The adults were again students from Toku University.

For each test trial, the participants were presented with two choices (e.g., a whole CD and a broken portion of a CD, as in Figure 8a) and asked to match two sentences, one in neutral syntax (“one CL small N”) and one in measure word syntax (“one small CL de N”), to the two choices. They were told “One side has one CL small N. The other side has one small CL de N.” Then they were asked to point to the side that matched the queried sentence, “Which side has one CL small N?” To ensure that the children understood that the nonqueried sentence had to match the other side, on the first three trials the experimenter confirmed whether the nonmatching sentence then should go with the other choice (“Then this side has one small CL de N?”). Most children answered in the affirmative. If the children did not, which was rare, the trial was repeated.

Participants were tested with three count classifiers, *gen* (rod), *zhi* (stick), and *pian* (slice). For these classifiers, the two choices always involved a whole individual object and a portion or part of an object (see Appendix C for examples). The two objects were of the same kind denoted by the noun. Consequently, we predicted that if the children and adults were sensitive to the syntactic manipulation, the count classifier in the neutral frame should be matched to the whole object and the count classifier in the MW frame should be matched to a portion of the whole object.

We also included *ba*, a classifier that can be used as either a count- or measure word (meaning “handle” and “handful,” respectively). The two choices for *ba* differed from the whole versus part contrast for the other classifiers. For *ba*, the two contrasting choices always involved an individual object and a handful of that object (e.g., one fan vs. one handful of fans, as in Figure 8b). The rationale is that when used in a MW context, the *ba* classifier should always take on its mass meaning (i.e., a set of individuals = a “handful”). Participants should thus select the handful option. The contrasting neutral sentence, on the other hand, should be matched to the

<sup>3</sup>There are some exceptions. In a footnote, Cheng and Sybesma (1998) pointed out that some speakers of Mandarin are willing to accept limited instances of adjectives modifying count classifiers as being grammatical (e.g., *yi da tiao yu*, or “one big CL fish.”). The grammaticality judgment relies on the availability of the substance reading provided by the noun (i.e. “fish meat”).

count meaning (i.e., one individual = “handle” shape individual). See Appendix C for the whole list of nouns (N) and classifiers (CL) along with photographs of the choices used in the experiment.

Each of the four classifiers was tested with three different noun/choice sets, making 12 trials in total. Four additional irrelevant and easy trials checked whether the participants were attending to the task, by contrasting only the words “big” and “small.” Specifically, the experimenter posed the following question: “One side has one CL small N. The other side has one CL big N. Which side has one CL small/big N?” The two object choices consisted one big and one small whole object of the same kind (i.e., apples, cups, candles, crayons). These trials were interspersed among the actual test trials so that every fourth trial was an attention check trials, and their order and the side of the big and small choices was counterbalanced.

The order of test items was randomly assigned with the caveat that half of the trials had the whole single object on the right side, and half on the left. As for the order of the two sentences, half of the participants from each age group heard a neutral frame first and half heard a MW frame first. In other words, the neutral frame group heard the experimenter say “One side has one CL small N. The other side has one small CL de N. Which side has one CL small N?” The MW frame group heard, “One side has one small CL de N. The other side has one CL small N. Which side has one small CL *de* N?” The participants were then asked to point to the choice that matched the queried sentence. The presentation order of the two types of sentences and the query question were kept identical throughout for each participant, because the difference between the two frames was minimal and we wanted to ensure that the children heard and remembered the difference.

## Results

With the exception of one child, all participants responded correctly on the attention check trials, indicating that they understood the task and were paying attention. The one child was incorrect on only one of the four attention check trials.

We calculated the percent correct for the twelve test trial questions (i.e., matches of neutral frames to whole objects and MW frames to portions or pluralities). The data were submitted to an ANOVA with Age (4-year-olds, 5-year-olds, and adults) and Query group (neutral frame first, measure word frame first) as between-subjects factors. The analysis revealed a significant effect of Age ( $F(2, 44) = 4.86, p = .01$ ). Pairwise *t*-tests indicate that the adults made a higher percentage of correct judgments compared to both the 5-year-olds (88% vs. 76%,  $t(30) = 2.15; p = .04$ ) and the 4-year-olds (67%,  $t(32) = 2.68, p = .01$ ). The two groups of children did not differ from each other ( $t(32) = 1.18, p = .25, n.s.$ ).

The only other significant effect was the Age  $\times$  Query group interaction ( $F(2, 44) = 4.29, p = .02$ ). Post hoc Fisher’s LSD tests comparing the difference between two query groups at each age showed that the 4-year-old children made a greater percentage of correct judgments for the neutral frame than for the MW frame (81% vs. 54% correct,  $p < .01$ ). This difference was not significant for either the older children (neutral: 78% vs. MW: 74%,  $p = .64$ ) or for the adults (neutral: 81% vs. MW: 95% for MW,  $p = .15$ ). The effect in 4-year-olds was due to a tendency to match the first queried phrase to the whole object, resulting in higher correct responses for the neutral frame group and higher incorrect responses for the mass frame group. As a result, the performance for the neutral frame group was similar across all ages (adults: 81%, 5-year-olds: 78%, and 4-year-olds: 81%, pairwise LSD  $p$ ’s  $> .50$ ). In contrast, the performance for the MW frame differed across all age groups (adults: 95%, 5-year-olds: 74%, 4-year-olds: 54%, pairwise LSD  $p$ ’s  $< .05$ ). Thus, there appears to be a developmental progression from 4-year-olds to adulthood.

We can also see this developmental trend by classifying the behaviors of individuals. Using the criterion of nine or more trials correct (out of twelve) as “success” (the probability of getting nine out of twelve is  $p = .05$  by binomial distribution), we find that the 4-year-olds were split between nine passers ( $M = 89\%$ ) and nine nonpassers ( $M = 46\%$ ). For the 5-year-olds, there were twelve passers ( $M = 83\%$ ) and four nonpassers ( $M = 56\%$ ). For the adults, there were fourteen passers ( $M = 94\%$ ) and two nonpassers ( $M = 46\%$ ).

## Discussion

The results show that adults are sensitive to the syntactic manipulation and understand that the use of count classifiers in a MW construction specifies reference to portions (i.e., of stuff or plural sets), rather than single individual objects. Six-year-olds, but not 4-year-olds, are also sensitive to this syntactic manipulation. For younger children, the meaning of the noun and its conventionally paired count classifier appeared to outweigh the meaning contributed by the incompatible MW syntax. It remains to be seen how same age English speaking children treat equivalent constructions in which nouns that are typically paired with count or mass syntax are paired with the opposite syntax (e.g., “There is dog on the road”). Perhaps in these contexts, English speaking children would also fail to arrive at novel interpretations.

The developmental trend observed here matches that in Experiments 1 and 2. At roughly the same time that children show evidence of associating classifiers with solidity and quantity information, they also begin to exhibit sensitivity to syntactic shifts such as the one in Experiment 3. Future studies should examine more closely the relationship between the knowledge of the semantics and syntax within individual children in order to further spell out the acquisition process—how exactly do individual children figure out the mapping between syntax and semantics? For example, does distributional information regarding adjective and classifier cooccurrence, or the use of *de* with measure words cause children to seek a semantic distinction between the two types of classifiers? Does the realization that measure words can also be used as nouns lead children to accept their use with adjectives? Finally, does the realization that measure words pick out portions or collections lead to the willingness to apply the *de* modifier, which functions similarly to the partitive construction in English?

## GENERAL DISCUSSION

These experiments explored the possibility of a homology between count classifiers in Mandarin and count syntax in English by asking how Mandarin-speaking adults interpret count classifiers and measure words. Children’s interpretation of these words was also compared to that of adults in order to examine the emergence of adult-like interpretation in language acquisition. In particular, we were interested in how children learn the shape, solidity, and number specifications of classifiers and measure words, and when each type of knowledge emerges. The course of acquisition may provide clues not only to how Mandarin-speaking children initially acquire classifier and measure word meanings, but also whether the acquisition of these words resembles the acquisition of mass and count syntax in English.

Findings from Experiments 1 and 2 indicate that for Mandarin-speaking adults and children, count classifiers and measure words can be readily applied to novel things in the world. However, there were differences between the groups. Adults selected shape-matched items for count classifiers and measure words when possible, and otherwise rejected items that did not match in shape, preferring to select the closed box (whose contents were unknown). However, this preference for shape-matched items was mediated by solidity. Adults preferred the shape match for count classifiers *only* when the match was a solid object and for measure words *only* when the match was a nonsolid substance.

Four- and 6-year-olds also chose on the basis of shape; the children picked the shape match when possible and rejected items that did not match in shape, preferring the closed box in such cases. However, children's understanding of count classifiers and measure words was not fully adult-like by four, nor even by six years of age. Although children were highly sensitive to the shape specifications of words tested in this study, their application of the words was not as strongly influenced by solidity as adults. For example, they often accepted nosolid shape-matches for count classifiers. This suggests that children acquire shape information for classifiers and measure words prior to solidity information and that associations between classifiers and solidity continue to develop even beyond the sixth birthday. Further, a number of 6-year-olds in our study (corroborated by Experiment 3) had yet to fully comprehend the syntactic differences between count classifiers and measure words, suggesting that both the syntax and meaning of these words have a protracted acquisition course.

These results extend previous findings that focus on how children interpret count classifiers in the context of familiar nouns. In our study, children demonstrated knowledge of classifier-shape relations by correctly picking out novel items that matched the shape specifications of the classifiers. It also extends previous studies by testing children's extension of shape-based measure words, as previous studies have focused primarily on count classifiers (see Chien, Lust & Chiang, 2003, for review).

Finally, the results appear to suggest, contrary to Chien et al., that children reach an adult-like understanding of count classifiers and measure words relatively late in language acquisition and have only begun to interpret them *productively* by three years of age. One question is why Mandarin-speaking children, unlike adults, did not show strong associations between solidity and classifiers/measure words. Previous studies show that when learning nouns, even at 2-years-old, Japanese and Mandarin speaking children were quite sensitive to the relationship between shape and solidity (Colunga & Smith, 2005; Imai & Gentner, 1997; Imai & Mazuka, 2003; Li, Dunham, & Carey, in press). Solidity affected how children chose to extend *nouns*, either as words naming object kinds or words naming substance kinds. Yet, in the experiments presented here children did not always consider solidity when interpreting classifiers and measure words. One possibility is that initially count classifiers and measure words are treated alike as they often appear in the same surface structure (one CL/MW N), and shape is a more systematic property that defines both types of word. For example, this study shows that measure words can be used to refer to collections or pluralities of objects, as long as those pluralities occur in appropriately shaped portions, like piles.

How do children begin the process of acquiring classifiers and measure words? We see several alternatives and believe that processes of both syntactic and semantic bootstrapping may be at work at different moments in acquisition. For each classifier, children may use syntactic bootstrapping (Fisher, Gleitman, & Gleitman, 1991; Landau & Gleitman, 1985) to infer the semantics of classifiers by analyzing the range of nouns with which they cooccur, and their cooccurrence with numerals and demonstratives. In doing so, children may begin to notice that, for classifiers, shape is often a relevant dimension. Subtler aspects of the distinction between count classifiers and measure words may rely initially on semantic bootstrapping (Grimshaw, 1981; Macnamara, 1982; Pinker, 1984). For example, count classifiers and measure words differ in their capacity to refer to single discrete objects (or singleton sets more generally). Children might assume that unitizing words used in NPs that refer to singularities are of one kind (e.g., count classifiers), while those that cannot refer to singularities are of another kind (e.g., measure words). Based on this initial semantic distinction they could then commence a distributional analysis for fully distinguishing the two types of words. Interestingly, this process could act universally and might also support the acquisition of the mass-count distinction in English. Since count nouns, but not mass nouns, can explicitly pick out single individuals, children acquiring English could distinguish count syntax from mass



syntax by noting which NPs can and cannot refer to singularities. Indeed, English-speaking 4-year-olds infer that nouns used in mass-count ambiguous syntax are count nouns when they refer to a single thing, but infer that they are mass nouns when they refer to more than one object (see Barner & Snedeker, 2006; Barner & McKeown, 2005; see also Gordon, 1985). Further studies examining the sensitivity of Mandarin-speaking children to the distinction between singular and plural sets are currently underway.

In sum, the results from this study demonstrate the way in which count classifiers and measure words can be used to apply to novel entities and how this emerges in acquisition. Beyond this, the study supports the conclusion that speakers of Mandarin have at their disposal a syntactic system for specifying individuation that closely parallels that found in English. In the Introduction we considered the hypothesis of Cheng and Sybesma (1998, 1999) that the distinction between classifiers and measure words in Mandarin in some way parallels the mass-count distinction in English. We pointed out that count syntax in English specifies reference to individuals, but sets of individuals can still be named in its absence by mass nouns like *furniture*, *luggage*, and *silverware* (Chierchia, 1998; Gillon, 1999; Barner & Snedeker, 2005, 2006). Analogously, count classifiers signal reference to individuals, though sets of individuals can still be referred to using measure words. Based on the observation that English also has measure words, and that in both languages these words can be used with almost any noun, we suggested that the analogy between Mandarin and English is carried entirely by the parallel between count classifiers and count syntax.

Future studies should investigate additional classifier languages to determine whether the semantic distinction between classifier types reported here extends beyond Mandarin. Classifier languages are subject to important variation in how classifiers are used. For example, in some Chinese dialects, classifiers can appear in the absence of numerals to express definite and indefinite reference (Cheng & Sybesma, 2005), which influences the frequency of classifiers in the languages (Erbaugh, 2002). Such differences could have important consequences for the course of acquisition, since children would not only encounter classifiers more frequently, but also could begin producing classifiers regardless of whether they have acquired numeral meanings (a process which typically extends beyond children's fourth birthday; e.g., Wynn, 1992).

Future work should also expand the analysis of classifiers beyond concrete entities to actions and abstract individuals. Classifiers support an "individual" reading of a noun, and can specify detailed information about the referent even for abstract things (Tai & Wang, 1990; Huang & Ahrens, 2003). For example, *tiao* specifies long, cylindrical, flexible objects, but also applies to abstract individuals like laws, units of news, and clues (see 4a-4c). The classifier *zhi* also specifies shape for long thin objects, but can do so equally for abstract individuals such as dances, songs, home runs, and hits in baseball (see 4d-4f):

- a. yi **tiao** fa lu (一法) = a law
- b. yi **tiao** xin wen (一文) = a piece of news
- c. yi **tiao** xiansuo (一线索) = a clue
- d. yi **zhi** wu (一舞) = a dance
- e. yi **zhi** quzi (一曲) = a tune
- f. yi **zhi** quanleida (一垒打) = a home run

Additional research is also required to fully test the parallels between Mandarin and English systems of individuation, since no previous studies have directly compared the developmental course of both systems using a single method that is sensitive to individuation. With this study, there is now evidence regarding how children in each language encode shape and solidity with

syntax (see also Dickinson, 1988; Imai & Gentner, 1997; Imai, Gentner, & Uchida, 1994; Landau, Smith, & Jones, 1988; Samuelson & Smith, 1999; Soja, 1992; Soja, Carey & Spelke, 1991), but no previous study has systematically compared the development of the *quantificational* properties of each syntactic system, which, as many have argued, are at the core of mass-count syntax and other linguistic structures that encode individuation (Gordon, 1985; Bloom, 1999; Barner & Snedeker, 2005). Studies currently under way are comparing Mandarin and English to examine this question.

## Acknowledgments









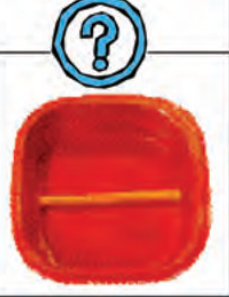
We thank Rosie Williams for constructing the stimuli and piloting Experiment 1, and Eva Xia for constructing Experiment 3 stimuli. We thank Yi-ting Huang for suggestions on how to run Experiments 1 and 2. Huang, Snedeker, & Spelke (2004) inspired our adaptation of Chien, Lust, & Chiang (2003). Special thanks also goes to C. -T. James Huang, Susan Carey, and members of her laboratory for discussions on classifiers and mass-count syntax. Thanks to Toku Univeristy, LinShen Elementary School Kindergarten, JiaDian Kindergarten, TseKuan Kindergarten, ChenHsin Kindergarten, and KuangFu Preschool for allowing us to test at their schools. This research was funded in part by an NRSA NIH postdoctoral fellowship to Peggy Li (#F32HD043532).

## References

- Allan K. Nouns and countability. *Language* 1980;56(3):541–67.
- Barner, D.; McKeown, R. The syntactic encoding of individuation in language and language acquisition. BUCLD 28: Proceedings of the 28th annual Boston University Conference on Language Development; Somerville, MA: Cascadilla Press; 2004.
- Barner D, Snedeker J. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 2005;97:41–66. [PubMed: 16139586]
- Barner D, Snedeker J. Children’s early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development* 2006;2:163–194.
- Barner D, Wagner L, Snedeker J. Events and the ontology of individuals: Verbs as a source of individuating nouns. under revision.
- Borer, H. Structuring Sense. Vol. 1. Oxford University Press; 2005. In Name Only
- Bloom, P. Semantic competence as an explanation for some transitions in language development. In: Levy, Y., editor. *Other children, other languages: Theoretical issues in language development*. Hillsdale, NJ: Erlbaum; 1994.
- Bloom, P. Theories of word learning: Rationalist alternatives to associationism. In: Ritchie, WC.; Bhatia, TK., editors. *Handbook of child language acquisition*. San Diego, CA: Academic Press; 1999. p. 249-278.
- Carpenter, Kathie. Later than sooner: Extralinguistic categories in the acquisition of Thai classifiers. *Child Language* 1991;18:93–113.
- Cheng LLS, Sybesma R. Yi-wan tang, yi-ge tang: Classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies, New Series* 1998;28(3):385–412.
- Cheng LLS, Sybesma R. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 1999;30(4):509–542.
- Cheng, LL-S.; Sybesma, RPE. Classifiers in four varieties of Chinese. In: Cinque, G.; Kayne, RS., editors. *The Oxford handbook of comparative syntax*. New York: Oxford University Press; 2005. p. 259-292.
- Chien YC, Lust B, Chiang CP. Chinese children’s comprehension of count-classifiers and mass-classifiers. *Journal of East Asian Linguistics* 2003;12:91–120.
- Chierchia, G. Syntactic bootstrapping and the acquisition of noun meanings: The mass-count issue. In: Lust, B.; Suñer, M.; Whitman, J., editors. *Syntactic theory and first language acquisition: Cross-linguistic perspectives: Heads, projections, and learnability*. Hillsdale, NJ: Lawrence Erlbaum; 1994. p. 301-318.
- Chierchia G. Plurality of mass nouns and the notion of “semantic parameter. *Events and Grammar* 1998;70:53–103.
- Croft, William. Semantic universals in classifier systems. *Word* 1994;45(2):145–171.

- Dickinson DK. Learning names for material: Factors constraining and limiting hypotheses about word meaning. *Cognitive Development* 1988;3:15–35.
- Doetjes, J. PhD dissertation. Leiden University; 1997. Quantifiers and selection; On the distribution of quantifying expressions in French, Dutch, and English.
- Erbaugh, MS. Taking stock: The development of Chinese noun classifiers historically and in young children. In: Craig, C., editor. *Noun classes and categorization*. Amsterdam: John Benjamins; 1986. p. 399-436.
- Erbaugh MS. Classifiers are for specification: Complementary functions for sortal and general classifiers in Cantonese and Mandarin. *Cahiers de Linguistique—Asie Orientale* 2002;31(1):33–69.
- Fisher C, Gleitman H, Gleitman LR. On the semantic content of subcategorization frames. *Cognitive Psychology* 1991;23:331–392. [PubMed: 1884596]
- Gathercole VC. “He has too much hard questions”: The acquisition of the linguistic mass-count distinction in much and many. *Journal of Child Language* 1985;12:395–415. [PubMed: 4019610]
- Gillon, BS. English indefinite noun phrases and plurality. In: Turner, K., editor. *The semantics/pragmatics interface from different points of view*. New York: Elsevier; 1999. p. 127-147.
- Gordon P. Evaluating the semantic categories hypothesis: The case of the mass/count distinction. *Cognition* 1985;20:209–242. [PubMed: 4064607]
- Gordon P. Count/mass category acquisition: Distributional distinctions in children’s speech. *Journal of Child Language* 1988;15(1):109–128. [PubMed: 3350873]
- Grimshaw, J. Form, function, and the language acquisition device. In: Baker, CL.; McCarthy, JJ., editors. *The logical problem of language acquisition*. Cambridge, MA: MIT Press; 1981.
- Grinevald, C. Classifiers. In: Lehmann, C.; Booij, G.; Mugdan, J., editors. *Morphology: A handbook on inflection and word formation*. Vol. 2. Berlin: Walter de Gruyter; 2004. Article 97
- Huang CR, Ahrens K. Individuals, kinds and events: Classifier coercion of nouns. *Language Science* 2003;25:353–373.
- Huang, YT.; Snedeker, J.; Spelke, ES. What exactly do number words mean?. Paper Presented at 26th annual meeting of the Cognitive Science Society; Chicago, IL. 2004.
- Imai M, Gentner D. A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 1997;62:169–200. [PubMed: 9141906]
- Imai M, Gentner D, Uchida N. Children’s theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development* 1994;9(1):45–75.
- Kuo, Y-J. Unpublished MA thesis. National Chung Cheng University; Jiayi, Taiwan: 1998. A semantic and contrastive analysis of Mandarin and English measure words.
- Krifka, M. Common nouns: A contrastive analysis of Chinese and English. In: Carlson, G.; Pelletier, FJ., editors. *The generic book*. Chicago: University of Chicago Press; 1995. p. 398-411.
- Landau, B.; Gleitman, LR. *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press; 1985.
- Landau B, Smith L, Jones S. The importance of shape in early lexical learning. *Cognitive Development* 1988;3:299–321.
- Landman, F. *Structures for semantics*. Dordrecht: Kluwer Academic Publishers; 1991.
- Link, G. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In: Bauerle, R.; Schwarze, C.; Stechow, A., editors. *Meaning, use, and interpretation of language*. Berlin: de Gruyter; 1983.
- Li P, Dunham Y, Carey S. Of substance: The nature of language effects on entity construal. To appear in *Cognitive Psychology*. accepted.
- Lucy, JA. *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, MA: Cambridge University Press; 1992.
- Lyons, J. *Semantics*. Vol. 1 & 2. Cambridge, MA: Cambridge University Press; 1977.
- Macnamara, J. *Names for things: A study of human learning*. Cambridge, MA: MIT Press; 1982.
- Matsumoto, Y. Order of Acquisition in the lexicon: Implication from Japanese numeral classifiers. In: Nelson, KE.; Kleeck, A., editors. *Children’s language*. Vol. 7. Hillsdale, NJ: Lawrence Erlbaum Associates; 1987. p. 229-260.

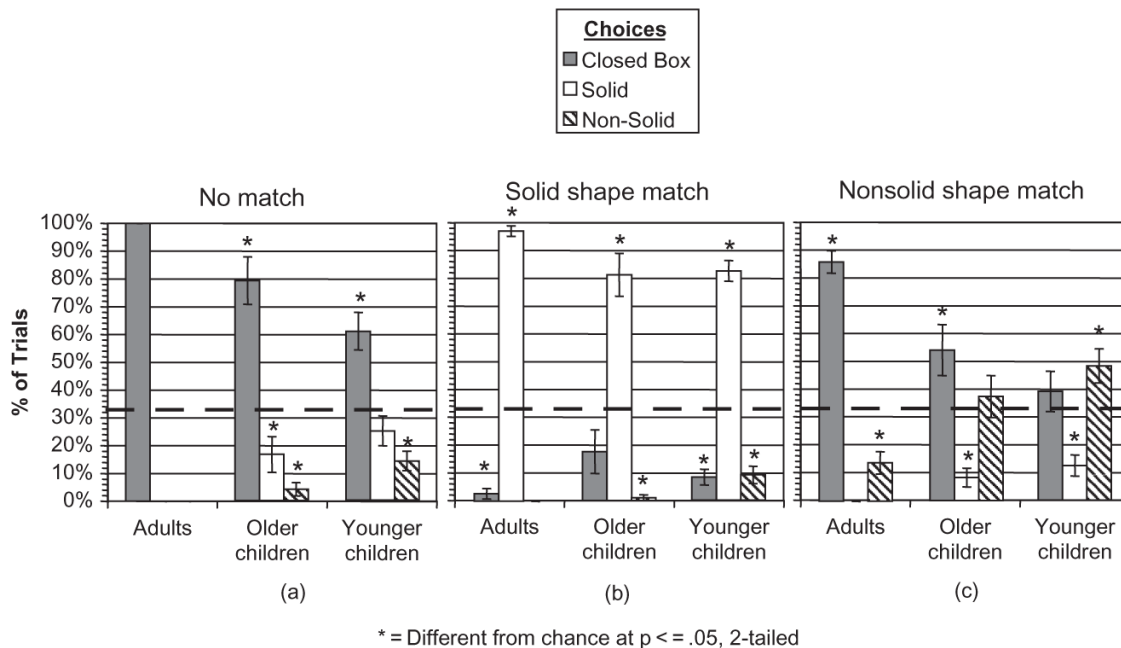
- Muraishi, S. Josuushi tesuto [Classifier tests]. In: Kenkyuusho, Kokuritsu Kokugo, editor. Yooji/Jidoo no Gengokeisei to Gengo [Conceptual development and Language in children]. Tokyo, Japan: Tokyo Shoseki; 1983. p. 273-316.[The National Language Research Institute],
- Mazuka R, Friedman R. Linguistic relativity in Japanese and English: Is language the primary determinant in object classification? *Journal of East Asian Linguistics* 2000;9:353–377.
- Pinker, S. *Language learnability and language development*. Cambridge, MA: Harvard University Press; 1984.
- Samuelson LK, Smith LB. Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition* 1999;73:1–33. [PubMed: 10536222]
- Sanches, M. Language acquisition and language change: Japanese numeral classifiers. In: Snaches, M.; Blount, B., editors. *Sociocultural dimensions of language change*. New York: Academic Press; 1977. p. 51-62.
- Senft, Gunter. *Systems of nominal classification*. Cambridge, MA: Cambridge University Press; 2000.
- Soja NN. Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development* 1992;7:29–45.
- Soja N, Carey S, Spelke E. Ontological categories guide young children’s inductions of word meaning: Object terms and substance terms. *Cognition* 1991;38(2):179–211. [PubMed: 2049905]
- Subrahmanyam K, Landau B, Gelman R. Shape, material, and syntax: Interacting forces in children’s learning in novel words for objects and substances. *Language and Cognitive Processes* 1999;14:249–281.
- Tai, JH-Y. Proceedings of the First International Symposium on Chinese Languages and Linguistics. 1990. Variation in classifier systems across Chinese dialects: Towards a cognition-based semantic approach; p. 308-322.
- Tai J, Wang LQ. A semantic study of the classifier Tiao. *Journal of the Chinese Language Teachers Association* 1990;25:35–56.
- Tai JHY, Chao FY. A semantic study of the classifier Zhang. *Journal of the Chinese Language Teachers Association* 1994;29:67–78.
- Uchida N, Imai M. A study on the acquisition of numeral classifiers among young children: The development of human-animal categories and generation of the rule of classifiers applying. *Japanese Journal of Educational Psychology* 1996;44:126–135.
- Wisniewski EJ, Imai M, Casey L. On the equivalence of superordinate concepts. *Cognition* 1996;60:269–298. [PubMed: 8870515]
- Wynn K. Children’s acquisition of the number words and the counting system. *Cognitive Psychology* 1992;24:220–251.
- Yamamoto K, Keil F. The acquisition of Japanese numeral classifiers: Linkage between grammatical forms and conceptual categories. *Journal of East Asian Linguistics* 2000;9:379–409.

	Closed Box	Solid	Nonsolid
No shape match – Closed box correct			
(e.g. CL-rod)			
Solid shape match – Solid correct			
(e.g. CL-rod)			
Nonsolid shape match – Closed box or nonsolid?			
(e.g. CL-rod)			





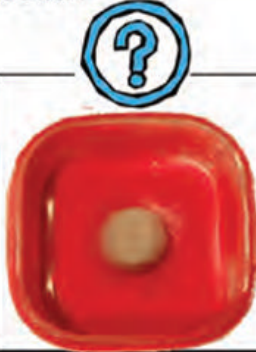




**FIGURE 1.**

Examples of the three trial types (Experiment 1 & 2) when given a count classifier such as “gen” (rod). For the *no shape match* trial, neither the solid object nor the nonsolid substance matches the shape specified by the count classifier. For the *solid shape match* trial, the solid object matches the shape specified by the classifier. For the *nonsolid shape match* trial, the nonsolid substance matches the shape specified by the nonsolid substance.



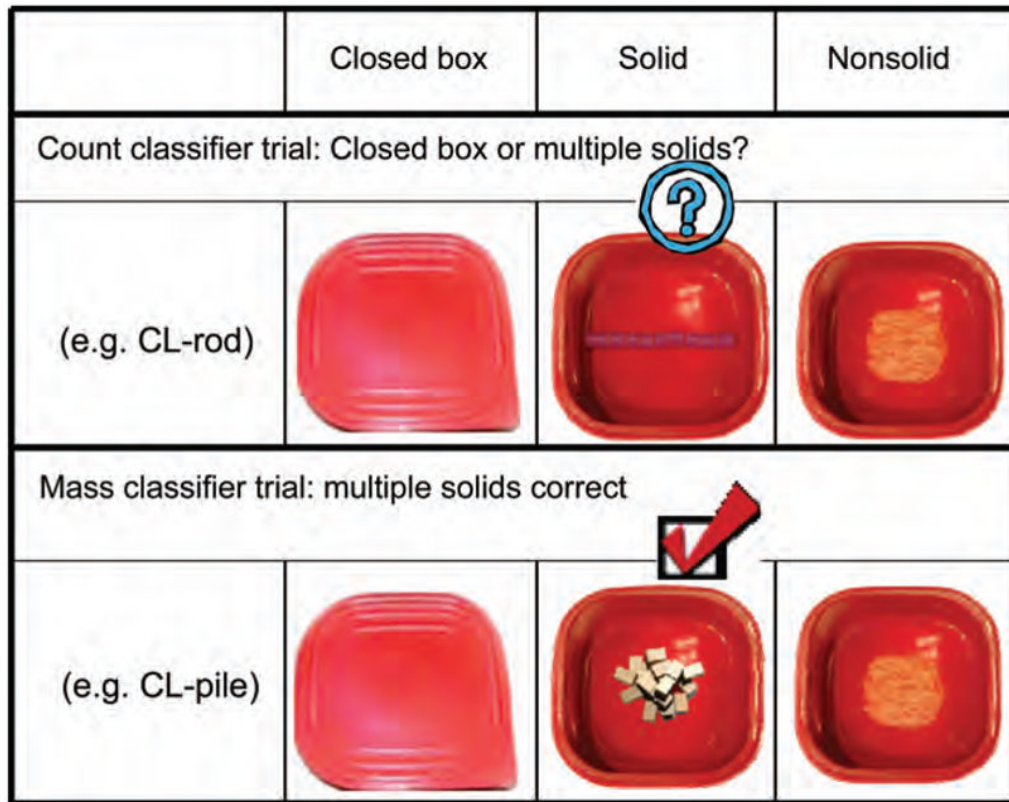


**FIGURE 2.** Results for Experiment 1 comparing adults’ and children’s box choices across trial types. Figures (a)–(c) are respectively for the *no shape match*, *solid shape match*, and *nonsolid shape match* trial types. Each figure shows the percentages of trials participants selected each of the three boxes.

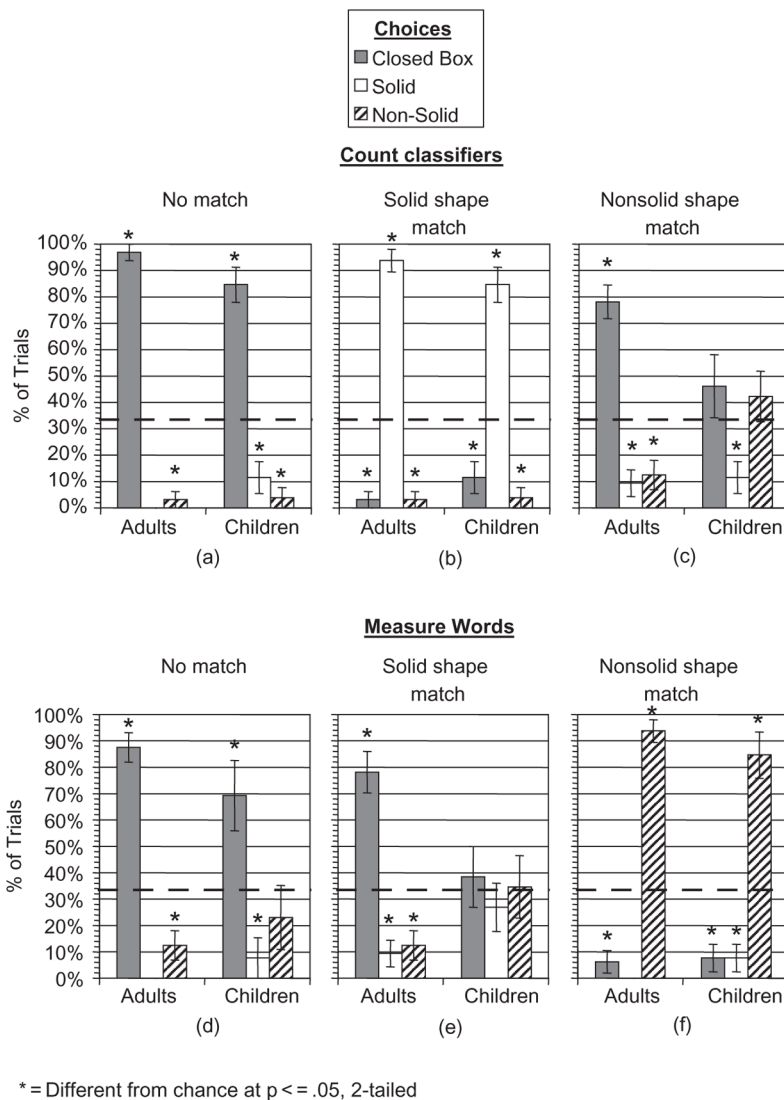
	Closed Box	Solid	Nonsolid
No shape match – Closed box correct			
(e.g. CL-pile)			
Solid shape match – Closed box or solid?			
(e.g. CL-pile)			
Nonsolid shape match – Nonsolid correct			
(e.g. CL-pile)			

**FIGURE 3.**

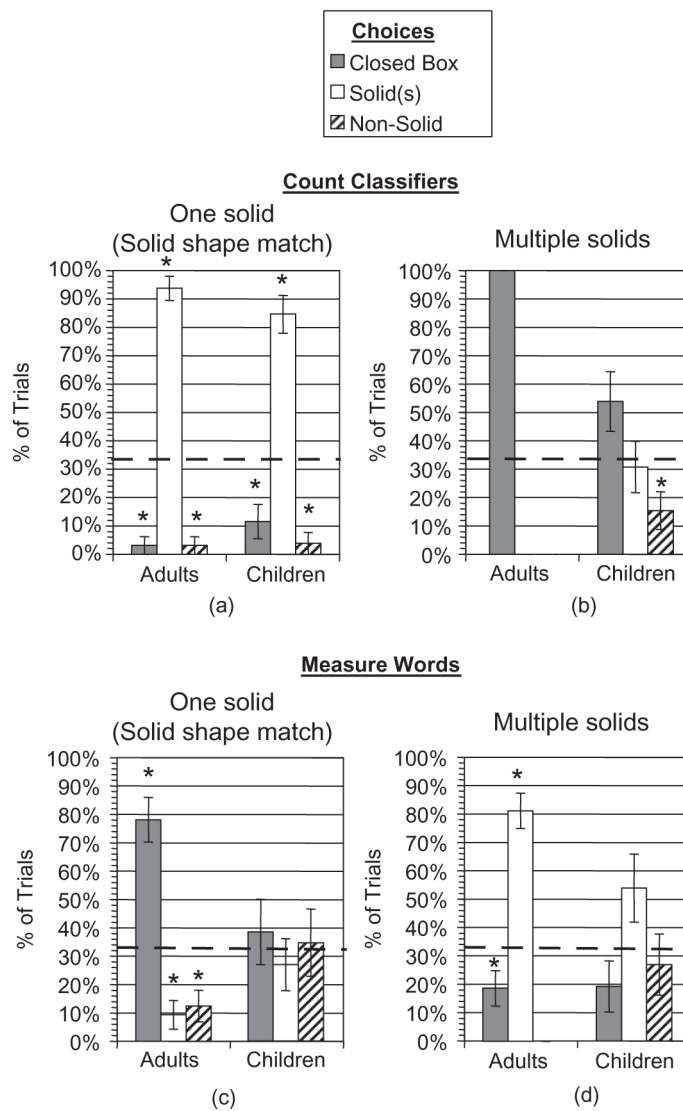
Examples of the three trial types (Experiment 2) when given a measure word such as “dui” (pile). For the *no shape match* trial, neither the solid object nor the nonsolid substance matches the shape specified by the count classifier. For the *solid shape match* trial, the solid object matches the shape specified by the classifier. For the *nonsolid shape match* trial, the nonsolid substance matches the shape specified by the nonsolid substance.

**FIGURE 4.**

Examples of the multiple solids trials (Experiment 2) when given a count classifier (e.g., “gen”) or a measure word (e.g., “dui”). Only the multiple solids choice matches the shape specified by the classifier.



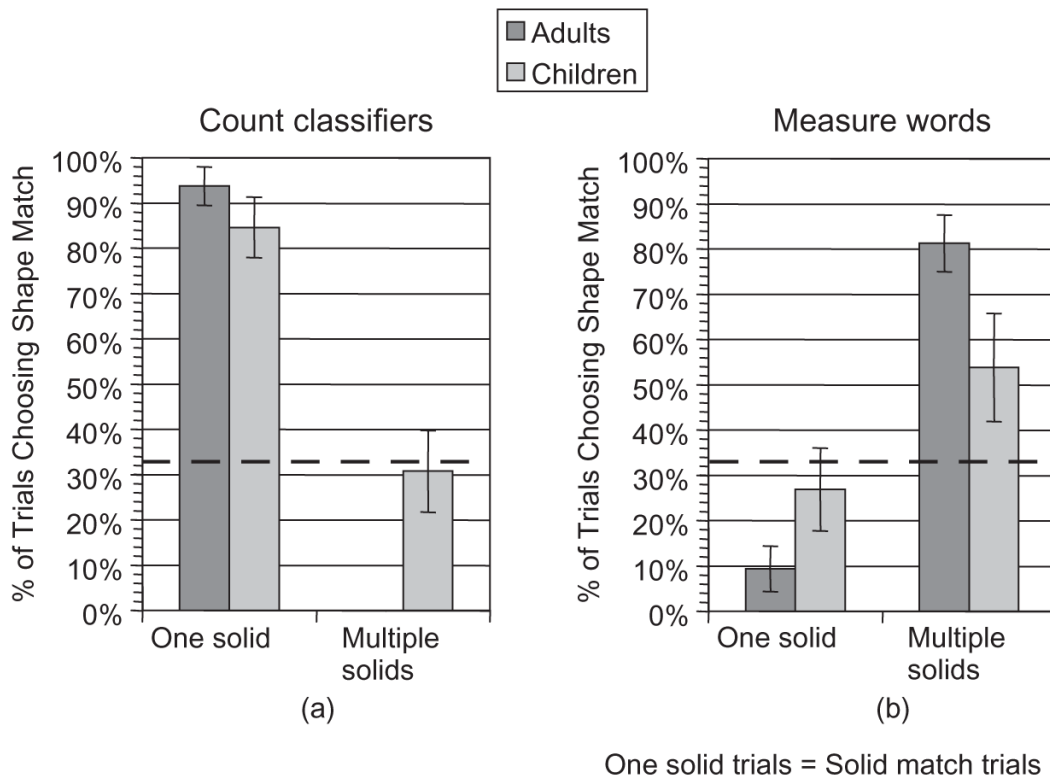
**FIGURE 5.** Results for Experiment 2 comparing adults' and children's box choices across three trial types: *no shape match*, *solid shape match*, and *nonsolid shape match* trial types. Figures (a)–c) are respectively for count classifiers and figures (d)–e) for measure words.



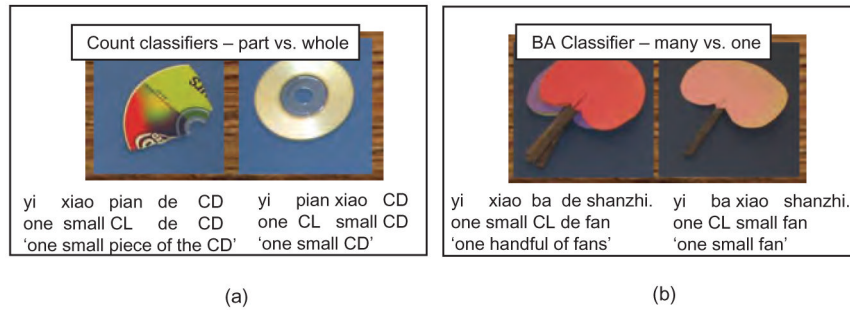
\* = Different from chance at  $p \leq .05$ , 2-tailed

**FIGURE 6.** Results for Experiment 2 comparing adults' and children's box choices for the trials when the solid(s) match the shape specified by the count classifier (figures a and b) or measure word (figures c and d).





**FIGURE 7.** Results for Experiment 2 plotting just the percentage of trials choosing the solid box when the shape of the solid(s) match the shape specified by the count classifier (figure a) and measure words (figure b).



**FIGURE 8.** Examples of stimuli that the participants had to match to phrases they heard. Panel (a) shows an example for the classifier “pian”, and panel (b) shows an example for the classifier “ba.”


**TABLE 1**


## Parallels between English and Mandarin

<b>English</b>	<b>Mandarin</b>
Noun + count syntax → individuals	Noun + count classifier → individuals
Noun + no count syntax (i.e., mass noun phrase) → unindividuated stuff or individuals	Noun + no count classifier (i.e., bare noun phrase) → unindividuated stuff or individuals
Noun + measure word → measure of stuff or set of individuals	Noun + measure word → measure of stuff or set of individuals


**Appendix A**



Experiment 1 Stimuli. Count classifiers and their shape matched choices.


Classifier	Solid
<i>gen (rod)</i>	 <p data-bbox="964 1535 1062 1568"><i>rigid plastic</i></p>

Classifier	Solid
<i>zhi (stick)</i>	 <p data-bbox="928 445 1101 470"><i>baked Super Sculpey</i></p>




Classifier	Solid
<i>tiao (strand)</i>	 <p data-bbox="954 441 1075 466"><i>flexible plastic</i></p>


Classifier	Solid
<i>zhang (sheet)</i>	 <p data-bbox="990 1480 1039 1507"><i>cork</i></p>
<i>pian (slice)</i>	 <p data-bbox="925 1701 1104 1734"><i>dired Crayola Magic</i></p>

Classifier	Solid
<i>kuai (chunk)</i>	 <i>baked Super Sculpey</i>


**Appendix B**


Experiment 2 Stimuli. Count classifiers and measure words and their shape-matched choices.

	Solid	
<i>gen (rod)</i>	 <i>rigid plastic</i>	

	Solid	
<i>pian (slice)</i>		
	<i>dired Crayola Magic</i>	









	Solid	
<i>dui (pile)</i>		
	<i>plaster</i>	

	Solid	
<i>tuan (wad)</i>	 <i>soap</i>	

**Appendix C**

Photographs of Experiment 3 stimuli. The stimuli are categorized by the classifier that modified them and the left most column list the corresponding noun.

Noun	Classifier: pain	
<p><i>CD</i></p>	 <p data-bbox="389 554 441 579"><i>A CD</i></p>	 <p data-bbox="1110 1520 1305 1545"><i>A piece of a broken CD</i></p>
<p><i>Yezi</i></p>	 <p data-bbox="389 1709 441 1734"><i>A leaf</i></p>	 <p data-bbox="1143 1709 1273 1734"><i>A part of a leaf</i></p>


Noun	Classifier: pain	
<i>Mianbao</i>	 <p data-bbox="349 409 483 436"><i>A slice of bread</i></p>	 <p data-bbox="1107 1075 1315 1102"><i>A part of a slice of bread</i></p>

Noun	
<i>Mianbao</i>	

*A roll of bread*

Noun	C
<i>Feizao</i>	 <p data-bbox="922 1390 1052 1417"><i>one whole soap</i></p>

<b>Noun</b>	
<i>Zhuan</i>	 <p data-bbox="922 1270 1055 1297"><i>one piece of tile</i></p>

<b>Noun</b>		<b>Classifier: gen</b>
<i>luobo</i>	 <p data-bbox="831 1549 909 1575"><i>A carrot</i></p>	





Noun	Classifier: gen	
yumao		



*A feather*

Noun	Classifier: gen	
<i>bi</i>		

*A pencil*

Noun	Classifier: ba	
<i>Shanzi</i>	 <p data-bbox="370 394 425 424"><i>A fan</i></p>	 <p data-bbox="1117 1333 1263 1360"><i>A handful of fans</i></p>

Noun	Classifier: ba	
Daozi	 <p data-bbox="365 394 430 422"><i>A knife</i></p>	 <p data-bbox="1112 1129 1274 1157"><i>A handful of knives</i></p>

Noun	Classifier: ba	
<i>San</i>	 <p data-bbox="349 394 446 424"><i>A umbrella</i></p>	 <p data-bbox="1096 1365 1291 1394"><i>A handful of umbrellas</i></p>