Published in final edited form as:
J East Asian Ling. 2010 November ; 19(3): 207–230.

Learning that classifiers count: Mandarin-speaking children's acquisition of sortal and mensural classifiers

Peggy Li,
Laboratory for Developmental Studies, Harvard University, Shannon Hall, 25 Francis Ave, Cambridge, MA 02138, USA pegs@wjh.harvard.edu

Becky Huang, and
Harvard University, Cambridge, MA, USA

Yaling Hsiao
University of Wisconsin-Madison, Madison, WI, USA

Abstract

Two experiments explored two-to five-year-old Mandarin-speaking children's acquisition of classifiers, mandatory morphemes for expressing quantities in many Asian languages. Classifiers are similar to measure words in English (e.g., a *piece* of apple; a *cup* of apples), with the main difference being that classifiers are also required when counting sortals (e.g., yi *ge* pinguo or "one *unit* apple" in Mandarin means "one apple"). The current study extended prior studies (e.g., Chien et al., J East Asian Linguist 12:91–120, 2003) to examine Mandarin-speaking children's understanding of classifiers as indicating units of quantification. Children were also tested on their knowledge of numerals to assess the relationship between children's acquisition of numerals and classifiers. The findings suggest that children first notice that sortal classifiers specify properties such as shape. Only after learning some numerals do they begin to work out how classifiers indicate units of quantification. By age four, children scored above chance on most classifiers tested.

Keywords

Language acquisition; Mandarin; Classifiers; Number; Quantification

1 Introduction

In order to talk about exact quantities, one needs to know the unit of quantification. Some kinds of things come in natural units; there is a clear sense of what constitutes a whole unit or an individual and what constitutes a portion or part. *Cats* and *tables* are instances of such whole units. Our knowledge and understanding of cats and tables allow us to determine that the tail of a cat or the leg of a table cannot be a whole cat or a whole table, respectively, but are instead parts of whole individuals of these kinds. On the other hand, substances, such as *water* or *wood*, do not have easily determinable natural units. Noun phrases provide the means to talk about sets and quantities of all kinds of things, regardless of whether the nouns refer to things with natural units. In learning language, a child must figure out how nouns combine with elements such as numerals, quantifiers, measure words, and morpho-syntax to

Correspondence to: Peggy Li.

talk about sets and quantities. This paper explores the development of Mandarin-speaking children's abilities to talk about exact quantities.

1.1 Mandarin as a classifier language

In many languages with count-mass syntax, such as English, a numeral can directly co-occur with nouns for which the criterion for the whole unit is clear (e.g., three cats, five tables). Here, the implicit and unspecified unit of quantification is whole individuals of that kind, and the numeral in combination with the noun specifies the number of individuals. This direct co-occurrence is in contrast to talking about quantities of substances, in which the individual unit is difficult to discern. A measure word is required to specify the unit of quantification (e.g., three *bowls* of water, *three waters).

For many other languages (e.g., Asian languages like Mandarin Chinese or Japanese, and Mayan languages like Yucatec or Tzotzil), the unit of quantification has to be explicitly specified regardless of what the noun denotes. For example, the noun in Mandarin for table (*zhuozi*) provides the criterion for individuation—Mandarin-speaking adults know that the leg of a table is not a whole table, or that two tiny tables are more tables than three gigantic pieces of one broken table (Li, Cheung, and Barner, in prep.). Yet, the unitizer *zhang* 'sheet' is needed when counting the number of whole tables. Thus, the expression for "three tables" is *san zhang zhuozi* 'three sheet table'. Mandarin and other languages that always require a unitizer are called "classifier languages" because the required unitizer (i.e., classifier) classifies nouns according to properties such as shape, animacy, functionality, flexibility, and size (Allan 1977). In the case of tables, *zhang* is a shape-based classifier that is generally used to count entities labeled by nouns referring to flat objects.

Many researchers suggest that there are two general types of classifiers: sortal and mensural classifiers (e.g., Allan 1977; Tai 1994; Cheng and Sybemsa 1998, 1999; Zhang 2007).[1] To borrow the term "sortal" from Locke (1975) and other philosophers, sortal classifiers indicate quantificational units for sortal nouns (i.e., nouns that name things with natural quantificational units). Sortal classifiers belong to a closed class, and sortal nouns usually have a relatively fixed and rigid association with a particular sortal classifier; using the correct classifier often requires rote memorization. That is, simply knowing the meaning of a noun and the meanings of a range of sortal classifiers would not necessarily allow one to select the proper classifier. For example, although both *yezi* 'leaf' or *tiezhi* 'sticker' denote flat objects, the two nouns take different classifiers that classify flat objects (*pian* and *zhang*, respectively).

Mensural classifiers, on the other hand, form a more contingent or temporary relationship with the nouns (Tai 1994). These classifiers are often open-class words and are typically nouns that are co-opted as units of quantification. For instance, *wan* 'bowl' is a noun, and like any noun, it can be counted with its associated sortal classifier (CL_{sortal}), e.g., *san ge wan* 'three CL_{sortal} bowl'. However, *wan* can also be used as a classifier to indicate quantity, as in *san wan shui* 'three bowls of water' or *san wan bianshi* 'three bowls of pennies.'

In summary, classifiers indicate the unit of quantification and function like measure words in English. The main difference between English and Mandarin in this regard is that sortal classifiers are required to indicate the quantificational unit of sortal nouns. It should be noted, however, that sortal classifiers can also provide units of quantification for (1) substances that do not have readily individuated units (e.g., *san zhang/kuai/pian sujiao* 'three sheets/chunks/slices of plastic' or (2) sortal nouns that can flexibly refer to the substance or

[1]Instead of sortal and mensural classifiers, some linguists adopt the terms count- and mass-classifiers to describe these two types of classifiers.

part of wholes (e.g., *san pian pinguo* 'three slices of apple'. In these cases, the sortal classifier is not the one that is memorized and associated with the sortal noun to specify the whole. These uses of sortal classifiers are analogous to uses of English measure words such as "piece" or "chunk," to refer to parts.

As noted earlier, a child must figure out how nouns combine with elements such as numerals, quantifiers, and classifiers in order to talk about sets and quantities. In this paper, we explore how Mandarin-learning children come to that understanding. Specifically, we chart the ages and phases when children come to understand the function of classifiers. The long-term goal, however, is to explain what leads to the development of children's abilities to talk about sets and quantities cross-linguistically. In the conclusion of this paper, we relate our findings to English-speaking children's language acquisition and speculate on how children's growing conceptual capacities might support the acquisition of words to express quantities.

1.2 Past to present studies

Extensive research on children's acquisition of classifiers in Chinese, Japanese, Thai, and other languages has been conducted over the past few decades using a range of methodologies including diary or corpus analysis, elicited production tasks, and comprehension tasks (e.g., Sanches 1977; Erbaugh 1986; Matsumoto 1987; Yamamoto and Keil 2000; Carpenter 1991; Uchida and Imai 1996; Sumiya and Colunga 2006; see Chien et al. 2003 for a review). However, many of these studies focused only on the acquisition of sortal classifiers. These researchers use the acquisition of sortal classifiers as a means to explore categorical learning; the relative amount of time and effort needed to learn certain classifiers indicates which kinds of conceptual categories are difficult for children. For example, researchers asked whether ontological-based classifiers (classifiers for animals) are easier to learn than shape-based classifiers, whether the number of dimensions matters for shape-based classifiers (1-D vs. 2-D vs. 3-D; see, e.g., Loke and Harrison 1986; Erbaugh 1986; Tse et al. 2007), and whether more specific classifiers (i.e., classifiers that specify more features) are harder to learn than general classifiers (e.g., Hu 1993; Tse et al. 2007).

Despite the extensive research on children's acquisition of classifiers, surprisingly very little research has been conducted on children's understanding of classifiers as units of quantification. Only recently have researchers contrasted the acquisition of sortal classifiers with that of mensural classifiers (Chien et al. 2003; Li et al. 2008). Inspired by linguistic analyses that drew parallels between the sortal-mensural distinction and the count-mass distinction (see Cheng and Sybesma 1999; Li et al. 2008), these studies focused on whether the count-mass distinction exists at the level of classifiers in Mandarin Chinese, and whether children and adults are sensitive to the distinction.

In Chien et al. (2003) study, native Mandarin-speaking children were introduced to a Mickey Mouse puppet who came from the United States and spoke limited Mandarin Chinese. For each trial, the experimenter brought out three items as choices and named them for the child. Mickey, with his limited Chinese and inability to remember the names for the items would ask for one of them without the noun ("I want one CL something"). The child's task was to interpret for the experimenter which choice Mickey indicated among the three. Crucially, the classifier (CL) used was the only means for distinguishing between the three choices as to what Mickey wanted.

Chien and colleagues found that children's performance on sortal and mensural classifiers did not differ. Even at the youngest age tested (i.e., three-year-olds), children often scored above chance in selecting what Mickey requested based on the classifiers. They also more often chose an individual object over a container of substances for sortal classifiers and vice

NIH-PA Author Manuscript

NIH-PA Author Manuscript

versa for mensural classifiers. They suggest this finding parallels studies showing that English-speaking children are more likely to pick objects over substances when novel words are used in count syntax than in mass syntax (e.g., Soja et al. 1991). The authors interpreted this finding as indicating that Mandarin-speaking children honor the count-mass distinction. However, there is no reason why mensural classifiers must serve as quantificational units only for substances and not for collections of objects. Chien et al.'s finding was driven by their choice of stimuli (see Li et al. 2008 for a critique and new experiments testing for and showing parallels between English and Mandarin).

Importantly for the current study, Chien et al.'s choice of stimuli was not ideal for one reason. As with many classifier studies, their test items consisted of familiar objects and nouns. As such, children's success at matching the classifier with the noun/test item could be due purely to associations formed from prior exposures to the classifier and noun/test item pairings in the speech input. In fact, there is reason to suspect that children's success was driven by pairings that are frequent in the input. For example, when testing mensural classifiers such as *wan* 'bowl' or *bei* 'cup', no bowls or cups were present. The substances, cooked rice (*fan*) and water (*shui*), were placed into square containers that were neither bowl- or cup-shaped. Yet, children selected the cooked rice and water for classifiers *wan* 'bowl' and *bei* 'cup', respectively, at above-chance level. Because children's success on the mensural classifiers can only be explained by prior learned associations between the classifier and the noun/test item, Chien et al.'s data leave open whether children truly know the meanings of the classifiers and can extend the use of these classifiers to novel test items or situations.

The current study modified Chien et al.'s paradigm to assess children's understanding of classifier meanings. In the new paradigm, children cannot simply use common classifier and noun/test item pairings to determine their answers. For example, Experiment 1 tested children's sortal classifier knowledge by using novel, handcrafted objects that they had never seen before and did not have names for. The purpose was to assess children's knowledge of how sortal classifiers classify nouns and their referents based on shape and function properties. Experiment 2 tested children's classifier knowledge as units of quantification by providing the same kind of object for all three choices. For example, in one trial, children might see a toy orange, toy oranges in a cup, and toy oranges in a pile. Only understanding the meaning of classifiers *ge* 'individual', *bei* 'cup', or *dui* 'pile' would allow the children to pick the correct choice.

2 Experiment 1

In this experiment, we asked whether children pick the novel object that matches the properties specified by the classifier. For example, do children know that for a classifier like *zhang*, which classifies 2-D flat objects, the novel object has to be flat?

2.1 Methods

2.1.1 Participants—Based on the work of Chien et al., who tested three-to seven-year-olds, we tested two-to five-year-old children. Since the three-year-olds in their study scored above chance on two-thirds of the classifiers tested, we included two-year-olds to explore younger children's knowledge of classifier meanings. We tested children up to five years of age because their six-and seven-year-olds were near or at adult-level performance, scoring in the ninety percent correct range for the classifiers.

Altogether, we tested 69 children plus 16 adults as a control group. The number of children and their mean ages for the four age groups were: 21 two-year-olds (*M* = 2;7; range = 2;1–2;10); 20 three-year-olds (*M* = 3;6; range = 3;0–3;10); 16 four-year-olds (*M* = 4;4; range =

4;0–4;10); and 12 five-year-olds (*M* = 5;4; range = 5;0–5;9). All children were acquiring Mandarin Chinese as their first language and were recruited from daycare centers and kindergartens in greater Taipei, Taiwan. The adult participants consisted of teachers working in the daycare centers and kindergartens. The adults participated voluntarily without compensation whereas the children received stickers for their participation.

2.1.2 Stimuli and procedure—The instruction followed Chien et al.'s script (see Appendix A for the Mandarin instruction). Children were introduced to Mickey Mouse, a puppet who did not speak Mandarin very well. The task was to figure out what Mickey wanted based on his requests. For each test trial, the experimenter presented three objects, and Mickey would then request one of them using a classifier (i.e., "I want one *CL* something").

Ten high-frequency sortal classifiers were selected from Chien et al. (2003) and Myers (2000) for the test trials. Among the 10 classifiers, one selected for animals (*zhi*), one for shape functionality (*ba*, for artifacts with handles), and eight for shapes varying in number of dimensions (e.g., *zhang* for two dimensions, *kuai* for three dimensions) or sizes (e.g., *ke* for small round objects). Table 1 lists the classifiers used in the experiment (column 1) and provides information about their meanings (column 2). The last two columns show, for each classifier, the test items designated as the preferred target response. These two columns are divided on the basis of two conditions, one in which children received known objects as choices (column 3) and one in which they received novel objects (column 4). For each trial, in addition to the designated target item, two other items from the same condition were randomly selected as foils with the rule that each item must appear as the foil in exactly two trials. This ensured that each item was presented with equal frequency, once as a target and twice as a foil.

Altogether, children received twenty test trials because each of the ten classifiers was tested twice, once with known objects and once with novel objects. The order of the two conditions (known objects vs. novel objects) was counter-balanced, such that half of the children were tested first on ten known object trials, and half were first tested on ten novel object trials. The order of the classifiers tested was randomized, such that half of the participants received the classifiers in one random order, and the other half received the classifiers in the reverse order. The locations of the target and the foils were also randomized beforehand.

Additionally, two practice trials were administered before each of the two conditions. These trials served to familiarize the children with the task and to check whether they were paying attention. The first practice trial always involved color, and the second always involved shape. For the color trials, children were presented with three toy bricks, each in a different color (e.g., red, yellow, blue), and Mickey would request one of them by color (e.g., "I want something red"). For the shape trials, children were presented with three objects, each in a different shape (e.g., round ball, rectangular block, and a triangular block), and Mickey would request one of them (e.g., "I want something round"). Children were near perfect on these trials, and so no further analyses were conducted on them.

2.2 Results

We first examined the adult speakers' performance to verify whether our selection of test items was appropriate for testing children. The majority of the adults agreed with our classifier-object pairings (see Table 2, last column). The adults were in 100% agreement for known objects, meaning that all adults selected what the authors designated as the "correct" answer. For the novel objects, four of the ten classifiers were not in 100% agreement, but only the classifier *zhi* (75%) differed statistically from 100%. The problem stems from the

fact that randomly paired novel foils could also be described by other target classifiers. Particularly, *zhi* is a homonym (
隻

and
枝

). Besides animals (
隻

), *zhi* (
枝

)can pick out a long, thin, and rigid object (e.g., a pen or a twig). One of the foils included a long, fat, cigar-shaped object meant for the classifier *gen*. Some of the participants (four of 16) chose the fat rod for *zhi* instead of the animal-looking novel object, lowering the overall average of correct choices for this classifier and novel object pair. We thus excluded the *zhi* trials from subsequent analyses. However, overall, the adult speakers' intuitions matched what we had anticipated.

Given that the adults' near-perfect performance verified that our stimuli were appropriate, we next asked how the children performed (see Table 2, which also lists the children's data by age). The question was whether children in Chien et al.'s study were succeeding merely because they had memorized the classifier–noun pairs. If so, our children should perform better for known objects than for novel objects. Alternatively, if they know the meanings of the classifiers and understand how these words pick out objects of similar shape or functionality, performance on the novel objects should be comparable to known objects. We therefore compared how children did with known objects versus unknown objects. Participants' response (either correct or incorrect) for each classifier was analyzed using a logistic regression model[2] to examine the effect of Familiarity (known objects and novel objects), the effect of Age (two, three, four, five, and adults), and the effect of Classifier (*ba, gen, tiao, kuai, duo, ke, pian, mian*, and *zhang*), as well as interaction effects.

The results showed no significant effect of Familiarity (known objects: 69% correct vs. novel objects: 69% correct; Wald χ^2 (1) = .081, p = .76, n.s.), suggesting that children were not merely succeeding through classifier-noun associations. They were capable of generalizing the classifier to pick out novel objects.

There was a significant effect of Age (Wald χ^2 (4) = 602.42, $p < .001$). The percentage correct for each age, combining known and unknown object trials, were 38% for the two-year-olds, 56% for the three-year-olds, 67% for the four-year-olds, 86% for five-year-olds, and 98% for the adults. Pairwise *t*-tests and Mann–Whitney *U* tests (MWUs) indicated a significant age-related improvement (all *p*'s < .05). That is, adults were better than the five-year-olds, five-year-olds were better than the four-year-olds, and so forth.

There was also an effect of Classifiers (Wald χ^2 (8) = 36.51, $p < .001$), indicating that some classifiers were more difficult to acquire than others. The average percentage correct ranged from 58% (for *mian*) to 78% (for *gen*). The variation in acquisition timetable among classifiers is likely attributable to factors such as frequency of the classifiers being heard and ease of figuring out the category based on common instances encountered in the input. There was also a Classifier × Age interaction (Wald χ^2 (30) = 101.42, $p < .001$), indicating that the

[2]We used the Generalized Estimating Equation Module of SPSS (Statistical Package for the Social Sciences).

ease of the classifiers differed for different age groups. Finally, there was a Familiarity × Classifier × Age interaction (Wald χ^2 (30) = 134.25, $p < .001$), indicating that Familiarity mattered for some classifiers and its effect also differed across each age group. However, this 3-way-interaction is likely to be noise since paired *t*-tests without correcting for multiple t-tests indicated that Familiarity mattered only for a few classifiers (notice that only three pairs out of 45 possible pairs[3] in Table 2 are significant as indicated by a '*', i.e., *ba* for two-year-olds, *gen* and *duo* for four-year-olds).

Finally, Table 3 lists the performance for each classifier, combining the results of known and novel object tasks by age, and each classifier was tested against chance of 33.3%. Most two-year-olds scored at chance on the majority of the classifiers, and with increasing age, children scored significantly above chance.

2.3 Discussion

Experiment 1 extended Chien et al.'s results to include data from younger children (i.e., two-year-olds). In general, the two-year-olds knew very little about classifiers; their performance was above chance on only one third of the classifiers. The three-year-olds were statistically better than the two-year-olds and scored above chance on two thirds of the classifiers. By four years of age the children scored above chance on all but one of the classifiers. The data corroborated Chien et al.'s data, in which three-year-old children also scored above chance on roughly two thirds of the classifiers tested, and four-year-olds scored above chance on all the classifiers tested.

Importantly, our experiment also tested children on novel objects and found no difference in whether the classifiers were paired with familiar objects or novel objects. It suggests that by 3 years of age, children can make generalizations regarding how a classifier categorizes a class of known objects/nouns and can spontaneously figure out whether a novel object would fit into that class of objects/nouns. Future studies could teach children novel classifiers to examine the ease at which children are able to make generalizations from limited instances (see Uchida and Imai 1999 for evidence that Japanese children can learn the meanings of animal- or shape-based novel classifiers with little input). In general, the present finding more conclusively supports Chien et al.'s (2003) claim that three-and four-year-old children have abstracted the fact that many classifiers often categorize things by shape similarities.

In the next experiment, we explored children's understanding of classifiers as units of quantification. Specifically, the experiment addressed at what age children begin to understand the meanings of mensural classifiers. While sortal classifiers are a closed class, mensural classifiers are often derived from open-class, basic-level nouns that are frequent in children's input (e.g., *wan* 'bowl', *bei* 'cup'). Since basic-level nouns are easily learnable (Gentner and Boroditsky 2001), perhaps mensural classifiers are acquired earlier than sortal classifiers by virtue of also being nouns.

3 Experiment 2

In Experiment 2, children were tested on four mensural classifiers. Two were container classifiers (*he* 'box' and *bei* 'cup') derived from common container nouns that toddlers know [see MacArthur-Bates CDI (Fenson et al. 1993; Hao et al. 2008)]. Two were configurational classifiers (*dui* 'pile' and *pai* 'row')that pick out collections of individuals in which the shape configuration matters. These two classifiers are also common verbs (i.e., 'to pile' and 'to line in a row', respectively). The mensural classifiers were contrasted with the

[3]There were nine classifiers after excluding *zhi*. 9 classifiers × 5 age groups = 45 pairs.

classifier *ge*. *Ge* is a sortal classifier memorized for many nouns (e.g., balls, people, dolls) whose referents do not necessarily share shape or size similarities or functionality. *Ge* is also used with nouns that denote abstract entities, such as hopes and wishes. As can be seen, *ge*, meaning 'individual', functions to indicate individuation. It is known as the default or general classifier because colloquially it can also substitute more specific sortal classifiers (Li and Thompson 1981) or be used to indicate the unit of substance kinds (e.g., *yi ge kafei* 'one coffee').[4] Not surprisingly, *ge* is the most frequent Mandarin classifier. It is also the earliest classifier to appear in children's speech production (Erbaugh 1986). However, production data alone cannot provide strong evidence for children's understanding of the classifier's meaning. Children could simply be inserting the classifier whenever they produce a numeral without appreciating why. The current experiment assesses whether children understand that the most commonly used classifier in their vocabulary (*ge*) picks out individuals.

Additionally, after being tested on classifier knowledge, children were tested on their number word knowledge to examine how number word acquisition is tied to classifier acquisition. The goal of counting is to provide an exact numerical quantity, and in order to know what to count, one needs to know the unit being counted. Given the dependent relationship of numerals to units being counted, the current study explores the relationship between children's number word knowledge and classifier acquisition.

Prior research established that children work out the meanings of number words between the ages of two and four (e.g., Wynn 1990; Le Corre and Carey 2007). The meanings of number words are learned one by one over the course of many months, starting with "one", then "two", and then "three." Some time after that, by the age of four, children realize how the count list they have learned to recite months earlier is used to determine the cardinality of sets. Given the overlapping acquisition timetable of classifiers and numerals, and the meaning dependence between these two kinds of words, working out how classifiers pick out units could very well impact how children acquire numerals and vice versa (see Bloom and Wynn 1997; Barner et al. 2009; Sarnecka et al. 2007 for various proposals on how classifier acquisition might slow numeral acquisition).

3.1 Methods

3.1.1 Participants—The same children and adults who participated in Experiment 1 also participated in Experiment 2. The order of the two experiments was counter-balanced, and they took place either within the same day with a few hours in between or within one day of each other.

3.1.2 Stimuli and procedure—The procedure again followed Chien et al.'s (2003) procedure. The children were introduced to Mickey, who cannot speak Mandarin very well. For each trial, Mickey requested something ("I want one *CL* N") and the participant had to select for the requested item among three choices. The three choices consisted of items that could be named by a single noun (N). Of the five classifiers in this experiment, each was tested twice, making a total of 10 test trials. For each trial, the two foils were randomly chosen with the requirement that one of the choices always had to be one whole object. The other choice could come from any of the remaining categories. (See Table 4 for the targets, foils, and nouns used.) As in Experiment 1, the stimuli consisted of physical three-dimensional objects. With the exception of the candies, rings, and pens, in which real objects were used, all other stimuli consisted of toy miniatures of real objects. There were

[4]Such uses are typically permissible in the context of ordering at a restaurant, much as in English: "I want a coffee."

also two practice trials in the same format as Experiment 1 to familiarize the children with the task.

3.1.3 Number word knowledge assessment: Give-a-Number Task—After participating in Experiments 1 and 2, the children were tested on their number word knowledge. To obtain children's number word knowledge, children were administered the Give-a-Number task (Wynn 1990; Le Corre and Carey 2007). The experimenter provided a tub of ten plastic strawberries; each time she asked the child to "Please give me N–*CL* strawberry", with N (number) starting at "one" and possibly ending at "six". A titration method was used to determine the next number requested. If the child succeeded at giving N strawberries, then the experimenter next requested N + 1. If the child failed at N + 1, N was requested on the next trial. The requested number continued to decrease until the child could not correctly give the smallest number at least two out of three times, or continued to increase until "six" was reached. Each time the child was thanked for his or her effort. In our Give-a-Number variant, children were not asked to check or correct their answers.

3.2 Results

The mean percentage correct for each of the age groups by classifier is shown in Table 5. The adults (Table 5, last column) were in 100% agreement with the authors' intuitions of how the classifiers indicate units. Having established that adults responded in line with our intuitions, we turned to two questions: (1) At what age did the children become aware of how classifiers indicate units of quantification? (2) Were some types of classifiers easier to learn than others? To address these questions, the percentage correct was entered into a logistic regression model to examine the effects of Age (two, three, four, five, and adult), Classifier Type (individual, configuration, and container), and the interaction between Age and Classifier Type. The analysis revealed a significant effect of Age (Wald χ^2 (4) = 518.33, $p < .001$). Pairwise comparisons using either *t*-tests or MWUs indicated that two-year-olds did not differ from three-year-olds (two: 33% correct vs. three: 38% correct, *t*-test: $p = .58$). However, all other comparisons were significant (*p*'s < .001), indicating improvement with increasing age (four: 78%, five: 91%, adults: 100%). There was also a significant effect of Classifier Type (Wald χ^2 (2) = 67.6, $p < .001$). Pairwise comparisons revealed that the general sortal classifier (*ge*: 86% correct) was better than the other classifiers (*t*-tests, *p*'s < .01). Container classifiers (*bei* 'cup': 67% and *he* 'box': 70%; average: 69%) were the next best. Finally, configuration classifiers were the worst (both *dui* 'pile' and *pai* 'row' were 58% correct) and differed from the other classifier types (all *p*'s < .01). There was also a significant Classifier Type × Age effect (Wald χ^2 (8) = 75.94, $p < .001$), suggesting that the amount of improvement for each classifier type was not uniform across the ages. For example, the improvement for container classifiers was quite big between three-year-olds and four-year-olds and was the smallest for sortal *ge* classifier (see Table 5).

In general, children under 4 years of age were not very knowledgeable of most mensural classifiers. In contrast, children were better at the general individuating classifier, *ge*. While testing, it was also apparent some children employed the strategy of always picking the choice with exactly one individual object. If this preference was prevalent, it would artificially increase children's performance on the *ge* 'individual' trials and worsen their performance on the configuration and container classifiers. To explore this possibility, we computed how often children selected the one object for the configuration and container classifiers trials. That is, we summed up how many times out of eight total trials the child selected the one object choice (e.g., one pen) over the multiple object choices (e.g., pile of pens or a box of pens).

We classified children as having a one-object bias if they picked the individual object in seven or more trials. The probability of picking the individual object in seven or more trials out of eight is .0025 by binomial distribution. Ten children (4 two-year-olds and 6 three-year-olds) fell into this group. We split the children into two groups based on their bias to pick the individual object and then examined how each group performed. Table 6 lists the mean percentage for the children classified by their bias (see also Table 5 for comparison). Percentage correct on the *ge* 'individual' classifier trials declined with the removal of the children with a bias to pick individuals. Two-year-olds no longer scored above chance on the *ge* trials, but three-year-olds still scored above chance (Table 6, first two columns). There was also a small improvement for the configuration and container classifiers. However, overall, children still did not score above chance level on these trials.[5] Thus, this new analysis again revealed that children under four knew very little about how mensural classifiers pick out units for counting. In comparison, these same three-year-old children, when tested on sortal classifier in Experiment 1, were beginning to understand how classifiers classify nouns and objects on the basis of shape.

3.2.1 Comparing Experiments 1 and 2—Next, we asked whether quantificational unit information (Experiment 2) or shape-based property information (Experiment 1) was acquired earlier (see Fig. 1 for data summary). The percentage correct was entered into a 5 Age (two, three, four, five, and adults) × 2 Experiment (Exp 1, Exp 2) ANOVA with Classifier type as within-subjects factor. The analysis revealed a significant main effect of Age ($F(4, 80) = 52.83$, $p < .001$), indicating that there was improvement across the ages (two: 36%, three: 47%, four 72%; five: 88%; adult: 99%). There was no effect of Experiment (Exp. 1: 69% correct vs. Exp. 2: 68% correct, $p = .47$). There was, however, an Experiment × Age interaction ($F(4, 80) = 8.75$, $p < .001$), indicating that the magnitude of difference between the two Experiments differed according to age. Pairwise *t*-tests at each age determined that the interaction effect was due entirely to three-and four-year-olds. Three-year-olds understood how classifiers select for nouns and referents based on shape information better than how they select for individuals or collection of individuals (Exp. 1: 56% correct vs. Exp. 2: 38% correct, $t(19) = 4.79$, $p < .001$). For four-year-olds, the performance was reversed (Exp. 1: 67% vs. Exp. 2: 78%, $t(15) = 2.273$, $p < .05$), possibly because the children who understood classifiers as quantificational units drew upon their lexical knowledge of the mensural classifiers as nouns or verbs (*bei* 'cup' or *dui* 'pile'). Eventually, by five years of age, children were equally proficient on either the sortal or the mensural classifiers.

3.2.2 Number word and classifier knowledge—Lastly, we explored the relationship between number word knowledge and classifier knowledge. Children were classified as N-knowers by their number word knowledge using Wynn's (1990) criterion. To be considered an N-knower, the child had to give N objects at least two thirds of the time correct when the experimenter requested N objects and not give N objects more than half as often when the experimenter asked for a different number. The criterion ensured that children reliably gave N objects only when N was requested. Figure 2 shows the distribution of number knowers by age.

Zero-knowers are children who cannot even give "one" reliably when the experimenter asked for "one." CP-knowers are Cardinality Principle Knowers; they can give all numbers requested. As Fig. 2 indicates, age was highly correlated with knower level (Pearson's $r(67)$

---

[5]The same generalized linear model excluding the biased children was conducted as before. Again, there was a significant effect of Classifier Type (Wald χ^2 (2) = 34.18, $p < .001$), Age (Wald χ^2 (4) = 343.52, $p < .001$) and Classifier Type × Age interaction (Wald $\times^2$ (8) = 43.91, $p < .001$). Pairwise comparisons indicated the same trend of improvement by Age, with only two-year-olds not differing from three-year-olds. The relative order of which classifier type was easier held also.

=.77, $p < .001$). Our data is consistent with prior experiments showing that by four years of age, the majority of children figure out how counting is used to determine the cardinality of sets (Wynn 1990; Sarnecka et al. 2007).

In our analysis, we focused on two-and three-year-old children since there was virtually no variation among the four-year-olds and five-year-olds in their knower-level or classifier knowledge. Because age was highly correlated with knower-level ($r(38) = .59$, $p < .0001$), we examined the relationship between knower-level and how children performed on the classifier tasks, controlling for age. The partial correlation analysis revealed that children's knower-level was not correlated with their scoreon Experiment 1 ($r(38) = .27$, n.s.) but with their score on Experiment 2 ($r(38) = .32$, $p < .05$). In other words, there was no correlation between children's number knower-level and sortal classifier knowledge. In contrast, knower-level was positively correlated with knowledge of how classifiers pick out units. We also examined the relationship between age and how children performed, controlling for knower-level. Here, the reverse pattern occurred. Children's age was positively correlated with Experiment 1 ($r(38) = .36$, $p < .05$) but not with Experiment 2 ($r(38) = -.06$, n.s.).

Since number knower-level and knowledge of classifiers as quantificational units correlated, we asked whether there was a clear causal direction of influence. Here, whichever one initially led in development could serve to benefit the one that lagged. It is more likely that the acquisition of some number words helped the acquisition of mensural classifiers as quantificational units rather than the reverse. The analyses above showed that there were children who did not seem to understand how classifiers indicate units of quantification but knew the meaning of at least one number word. Specifically, recall the ten children who always picked the one individual object over a collection of individuals (Experiment 2, Table 6). Of these biased children, all but one were at least a 1-knower. In contrast, there were no children who scored high on classifiers and were on the low end of their knower-level knowledge. Of the 0-and 1-knowers, no children remotely came close to scoring well on Experiment 2. Only two out of these 24 children got at least six questions out of 10 correct ($P(6 \text{ out of } 10) = .06$).

The whole pattern of data suggests that children can notice how sortal classifiers pick out items that share similarities in shape or shape functionality without knowing that they indicate units of quantification. It may be after learning some number words that children begin to realize that these classifiers indicate units of quantification.

3.3 Discussion

The experiment revealed that at three years of age, some children were aware that the sortal classifier *ge* picks out individuals (Tables 5, 6). On the other hand, as the percentages correct also revealed, these three-year-olds did not know many mensural classifiers. They often did not choose the correct item on the basis of the mensural classifier, despite the fact that the mensural classifiers are common nouns and verbs that the children ought to know. Their performance on the mensural classifiers in Experiment 2 was also worse than their performance on the sortal classifiers in Experiment 1; they did not score above chance on any of the mensural classifiers while they did so for some sortal classifiers. However, by four years of age, there was a sudden improvement on mensural classifiers (Table 5; Fig. 1).

The correlational analyses also showed that while both age and knower-level were highly correlated with each other and with classifier knowledge, they differed in which aspect of classifier meaning acquisition they best predicted. Age predicted how well children could classify the object on the basis of shape and shape-dependent functionality, possibly because age is a proxy for amount of exposure to classifiers. In contrast, knower-level best predicted understanding of classifiers as quantificational units. That is, children at higher knowerlevels were more aware that one could count units consisting of more than one individual. Below, we speculate on why this may be so.

4 General discussion

In order to express notions regarding sets and quantities, children are faced with the problem of figuring out the combinatorial nature of their language. Using Mandarin as a case study, this paper is one of the first few studies to explore children's understanding of how nouns combine with number words and classifiers to generate the meaning of noun phrases. Our starting place wasChien et al. (2003)'s seminal work contrasting sortal and mensural classifiers. Their study suggested that three-year-old children were at least aware of noun-classifier pairings for familiar nouns. We extended their paradigm and tested children on novel entities. We asked whether children truly appreciate the fact that sortal classifiers categorize nouns and their referents on the basis of properties such as shape and shape functionality (Experiment 1). Varying the number of individuals of a kind (one vs. a collection in a container vs. a collection in a configuration), we also asked whether children appreciate the fact that classifiers indicate quantificational units (Experiment 2).

Testing a wide age range, the emerging picture is that two-and three-year-olds are still working out the function of classifiers. It is not until 4 years of age that children begin to realize how classifiers specify units for quantification. Several new findings contribute to our knowledge of the developmental progression. First, learning how sortal classifiers categorize nouns and their referents based on properties such as shape and ontological categories may initially develop independent of any awareness that these classifiers indicate the natural unit of sortal nouns. At 3 years of age, children in general performed better when asked to pick out referents on the basis of shape properties than on unit information (Experiment 1 vs. 2). Acquisition of sortal classifiers also improves gradually with age, with no initial correlation to number word knowledge.

On the other hand, the acquisition of classifiers as quantificational units (measured by performance on Experiment 2) is correlated with number word knowledge. In fact, knowledge of classifiers as quantificational units may first require robustly knowing the meanings of some number words and possibly also knowing how the general classifier *ge* picks out individuals. The responses given by the two-and three-year-old children are consistent with this order of acquisition. First, children with little or no number word knowledge (0-knowers and 1-knowers) scored around one-third correct for Experiment 2, as expected by chance. This is despite the fact that some of these classifiers are derived from frequent nouns and verbs that children this age know (e.g., *bei* 'cup'). Among the children who knew some number words, there were even those who, in hearing "one *CL* N", always chose one individual object and ignored the classifier completely. Thus, it does appear that children are initially unaware of how number words combine with classifier and noun to indicate quantity. However, with increasing certainty of number word meanings and how they pick out discrete individuals determined by the noun (*yi* 'one' for one individual, *liang* 'two' for two individuals, etc.), children may begin to notice instances when the number and noun do not pick out the number of individuals they expected.[6] Here, knowing nouns and verbs for which they already know the meaning (*bei* 'cup' or *dui* 'pile') could help them come to the realization of how each morpheme jointly contributes to meaning. This realization may then lead children to reconsider the role of sortal classifiers as indicating whole objects are being counted rather than simply as words that classify referents according

[6]This account is similar to Bloom and Kelemen's (1995) account of how English-speaking children might use their knowledge of count-mass syntax to learn novel collection nouns when the noun phrase ("a blicket" or "one blicket") is used to reference a collection of individuals.

to properties such as animacy, functionality, or shape. This reanalysis likely does not occur until four years of age or after since works elsewhere indicate that children are quite willing to accept expressions with the numeral, sortal noun, and its associated sortal classifier as referencing broken pieces of objects (see, e.g., Brooks et al. 2010; Huang and Lee 2009; further discussion follows).

In summary, while Erbaugh's (1986) study on children's production of classifiers indicated that even young two-year-old children know to insert the default classifier alongside the numeral in uttering noun phrases, our study suggests that there is much more these young children have yet to learn about the functions of classifiers. Our findings align with more recent studies showing that the acquisition of the classifier system may take years (Li et al. 2008; Huang and Lee 2009) but differ from those ofChien et al. (2003), who suggested that "Chinese children as young as three years of age were fairly adept in comprehending both count-and mass-classifiers" (p. 112).

Note, however, our study and others do not really address why learning classifiers goes through these stages. Why is the period of acquisition protracted when it comes to realizing that classifiers specify units of quantification? Why does knowledge develop when it does? Here, it may be useful to draw upon prior developmental research. For example, research indicates that children do not necessarily possess the same abilities as adults in using measurement units in determining quantity. Huntley-Fenner (2001) showed that (English-speaking) three-year-olds and young four-year-olds do not make use of units such as cups and glasses to compare quantities of non-cohesive substances when the two amounts of sand are very close in volume. He demonstrated that although children were perfectly able to determine which side has "more glasses" when comparing three identical glasses full of sand with two other identical glasses full of sand, they scored at chance when asked to determine which side has "more sand" for the same visual display. It is not until children are four years old or older that they can successfully compare quantities in such a task. Thus, the use of containers in quantification may not be the kind of relationship that children typically process and notice. This is perhaps why our two-and three-year-old Mandarin-learning children sometimes ignored the container classifiers in their choice and preferred the one individual.

Researchers have also suggested that children privilege spatially discrete physical objects in word learning (Markman 1989; Bloom 2000; Gentner and Boroditsky 2001; Huntley-Fenner 1995, Chap. 4; Bloom and Kelemen 1995). If learning words for object kinds is conceptually easier than learning words for collections (e.g., "bunch", "pile", "army"), this could also explain why learning the configuration classifiers might pose some difficulties for Mandarin-speaking children. Future research using direct within-subject comparison of children's conceptual development and their linguistic skill could shed light on how language learning might rely upon conceptual development.

Comparison between Mandarin-speaking children and children learning a count-mass language, such as English, could also aid our understanding. For a count-mass language like English, children do not have to learn a complex system of sortal classifiers because numerals in English directly co-occur with most[7] sortal nouns. Unlike Mandarin, English also has obligatory plural markings that could facilitate acquisition of measure words (Sarnecka et al. 2007; Barner et al. 2009). For example, the plural on the noun "cow" in the phrase, "one herd of cows," might clue children into the fact that a collection of cows is being referenced and help children uncover the meaning of the novel word "herd." In sum,

[7]Exceptions are object-mass nouns such as "furniture" or "jewelry." "One piece of furniture" can reference one whole table but not a leg of a table.

cross-linguistic differences very likely affect the ease by which English-and Mandarin-speaking children come to understand how their native language encodes quantities. However, the two groups of children could still show relatively similar developmental trajectories despite cross-linguistic differences. This would suggest that conceptual underpinnings and developments that support language acquisition (e.g., ability to think about collections or package substances) are universally affecting both groups.

In two recent studies, Barner and colleagues (Melgoza et al. 2008; Brooks et al. 2010) made progress charting English-speaking children's evolving understanding of how quantifiers, numerals, and nouns contribute to meaning. They noted that children start using quantifiers, number markings, and number words very early, even prior to learning many nouns. For example, by 2 years of age, English-speaking children use singular-plural morphology to talk about individuals and sets of multiple individuals though nouns only make up a small percentage of their overall vocabulary of around 300 words (Kouider et al. 2006; Barner et al. 2007). Because children do not know many nouns, Barner et al. suggested that children might not initially rely on the criteria of individuation supplied by the noun to figure out quantifier or number meanings. Instead, children might figure out quantifier meanings based on their analysis of the number of spatio-temporally discrete individuals being talked about. Indeed, infant research has demonstrated that pre-linguistic infants represent and quantify spatio-temporally discrete individuals (e.g., Wynn 1992, 1996; Wood and Spelke 2005; Xu 2007), and such ability could initially support learning of quantifiers and numerals.

Barner and colleagues' hunch of how children learn quantifiers was influenced by Shipley and Shepperson (1990), who asked English-speaking preschoolers to count broken objects (e.g., forks broken into pieces among whole forks). Unlike adults, children would oddly count each individual piece. In other words, a fork broken into two pieces would be counted as "two forks" rather than "no forks" or "one (broken) fork." Children's tendency to count discrete individuals is also observed when counting events with temporally discrete subparts, attached parts of discrete individuals, and collections (Wagner and Carey 2003). For example, when asked to indicate number of families, some children are inclined to count the number of individuals in the family rather than the number of families. In general, the odd counting behavior does not subside until 5 or 6 years of age.

Importantly, Barner and colleagues explored whether the behavior is limited to counting.Melgoza et al. (2008) showed that even when making quantity judgments, the majority of preschoolers tend to select one fork broken into three pieces as having "more forks" than two whole forks. They are also more likely to pluralize and describe the three pieces as "forks" or call one broken piece "a fork."Brooks et al. (2010) have since replicated the finding, showing that English-speaking four-year-olds happily treat broken pieces as units of counting across a host of quantifiers tested (*more, every, both, a*, and singular-plural morphology). Thus, it does seem that English-speaking preschoolers take number of discrete individuals as sufficient units for counting and as the basis from which hypotheses about quantifier meanings are formed.

This is not to say that children are not paying attention to the noun. They realize that the noun factors in: when counting *dogs, cats* should not be considered in the count. For example, Wagner and Carey (2003) showed that despite strong tendencies to count the discrete individuals, children were sensitive to the nouns (see also Giralt and Bloom 2000). They were more likely to count families when asked to count families than when asked to count people. Furthermore,Brooks et al. (2010) showed that children understand that sortal nouns provide criteria for determining what constitutes a natural whole unit. When children were asked to count objects glued together, they counted each of the individual objects (i.e., two cups glued together was counted as "two cups"). Thus, it appears that children also realize that nouns provide criteria of individuation and can in some circumstances override their tendencies to count discrete physical individuals. They do not lack the conceptual knowledge about objects and can tell when an object is or is not a whole (see Brooks et al. 2010 for more evidence). What they may need to learn is divisity of reference (i.e., how parts are named). That is, if a fork is broken in half, a broken piece is no longer “a fork” but rather “a piece of fork.” Preliminary data from our laboratory show that this is a distinction that some four-year-old children do not make; they are perfectly happy in comprehension studies to accept the expressions “one fork” or “one piece of fork” interchangeably to describe either a broken piece or a whole fork.

In the case of Mandarin, the acquisition of quantifiers and numerals may also initially be through the analysis of discrete individuals. Some Mandarin-speaking children in the current study ignored the classifier but always picked one individual in the presence of the numeral *yi* ‘one’. Furthermore, new evidence from Huang and colleagues (Huang and Lee 2009; Huang 2009) and from our laboratory showed that Mandarin-speaking children also treat broken pieces as units of quantification. We robustly replicated Brooks et al. (2010) for Mandarin-speaking children. Preschoolers have a tendency to count each broken part and to make quantity judgments on the basis of broken parts (i.e., one fork broken into three pieces is “more” than two forks). This is in contrast to Mandarin-speaking adults who consider two whole forks as more than the three broken pieces. It is likely that once Mandarin-speaking children notice and realize that broken pieces can no longer be counted with the same classifier as whole objects they start to behave like adults.

More research remains to be conducted to directly compare English-and Mandarin-speaking children. Importantly, the current paper makes a first step in charting Mandarin-speaking children’s developmental progress in talking about quantities through the use of classifiers. It also makes contact with children’s conceptual development. We take the stance that children must have some preexisting conceptual abilities to support the acquisition of quantifiers and classifiers/measure words. As such, we try to link up children’s linguistic development with their conceptual development, drawing upon existing literature. However, it is possible that the process of understanding how numerals combine with measure words/ classifiers and nouns could benefit how children reason about quantities (Gordon 2004; Spelke 2003). Cross-linguistic differences in linguistic structures and discoveries of differences in acquisition timetable could allow us to eventually make inroads into the question of the relationship between language and conceptual development. We have no doubt that future work will further our understanding of this complex relationship.

Acknowledgments

We are indebted to Susan Carey, David Barner, and members of the Laboratory for Developmental Studies for their discussions with us on issues related to classifiers. Special thanks go to Aijun Huang for sharing her paper and to the teachers, children, and parents at the following schools in the Taipei area for participating in our study: Yongbin (
永濱

), Fanmei (
泛美

), Shengru (
聖如

), Jili (
吉利

), Hesana (
和撒那

NIH-PA Author Manuscript
NIH-PA Author Manuscript

), Minquan (
民權

), and Chunchi (
春池

). This research was funded by an NRSA NIH postdoctoral fellowship to Peggy Li (F32HD043532), an NSF grant (REESE 0633955) to Elizabeth Spelke and Susan Carey, and an NIH grant (R01HD038338) to Susan Carey.

Appendix: Instructions for Experiments 1 and 2 (taken from Chien et al. 2003)

Practice trials

Mickey: *Wo*
yao ADJ-*de* something.
I want ADJ-de something.
Experimenter: *Milaoshu*
yao ADJ-*de*
shemo.
Mickey Mouse want ADJ-de what
(Example ADJ: *yuan* = round)

Test trials

Mickey: *Wo*
yao
yi CL something.
I want one CL something.
Experimenter: *Milaoshu*
yao
yi CL *shemo*.
Mickey Mouse want one CL what
(Example CL: *gen* = rod shape sortal classifier; *dui* = pile, mensural classifier)

References

Allan K. Classifiers. Language. 1977; 53:285–311.

Barner, David; Inagaki, Shunji; Li, Peggy. Language, thought, and real nouns. Cognition. 2009; 111(3):329–344. [PubMed: 19345937]

Barner, David; Libenson, Amanda; Cheung, Pierina; Takasaki, Mayu. Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. Journal of Experimental Child Psychology. 2009; 103:421–440. [PubMed: 19162276]

Barner, David; Thalwitz, Dora; Wood, Justin; Carey, Susan. On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. Developmental Science. 2007; 10(3):365–373. [PubMed: 17444976]

Bloom, Paul. How children learn the meanings of words. Cambridge: MIT; 2000.

Bloom, Paul; Kelemen, Deborah. Syntactic cues in the acquisition of collective nouns. Cognition. 1995; 56:1–30. [PubMed: 7634763]

Bloom, Paul; Wynn, Karen. Linguistic cues in the acquisition of number words. Journal of Child Language. 1997; 24:511–533. [PubMed: 9519584]

Brooks, Neon; Pogue, Amanda; Barner, David. Piecing together numerical language: Children's use of default units in quantification. Developmental Science. 2010; 13(2):1–14. [PubMed: 20121858]

Carpenter, Kathie. Later rather than sooner: Extralinguistic categories in the acquisition of Thai classifiers. Journal of Child Language. 1991; 18:93–113. [PubMed: 2010507]

Chien, Yu-Chin; Lust, Barbara; Chiang, Chi-Pang. Chinese children's comprehension of count-classifiers and mass-classifiers. Journal of East Asian Linguistics. 2003; 12:91–120.

Cheng, Lisa Lai-Shen; Sybesma, Rint. Yi-wan tang, yi-ge tang: Classifiers and massifiers. Tsing Hua Journal of Chinese Studies. 1998; 28:385–412.

Cheng, Lisa Lai-Shen; Sybesma, Rint. Bare and not-so-bare nouns and the structure of NP. Linguistic Inquiry. 1999; 30(4):509–542.

Erbaugh, Mary S. 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.

Fenson, Larry; Dale, Phillip S.; Reznick, J Steven; Thal, Donna; Bates, Elizabeth; Hartung, Jeffrey P.; Pethic, Steven; Reilly, July. The MacArthur communicative development inventories: User's guide and technical manual. San Diego: Singular Publishing; 1993.

Gavin, Huntley-Fenner. The representation of objects, non-solid substances, and collections in infancy and early childhood. PhD diss. MIT; 1995.

Gavin, Huntley-Fenner. Why count stuff? Young preschoolers do not use number for measurement in continuous dimensions. Developmental Science. 2001; 4(4):456–462.

Gentner, Dedre; Boroditsky, Lera. Individuation, relational relativity and early word learning. In: Bowerman, M.; Levinson, S., editors. Language acquisition and conceptual development,. Cambridge: Cambridge University Press; 2001. p. 215-256.

Giralt, Nuria; Bloom, Paul. How special are objects? Children's reasoning about objects, parts, and holes. Psychological Science. 2000; 11:497–501. [PubMed: 11202496]

Gordon, Peter. Numerical cognition without words: Evidence from Amazonia. Science. 2004; 306:496–499. [PubMed: 15319490]

Hao, Meiling; Hua, Shu; Xing, Ailing; Li, Ping. Early vocabulary inventory for Mandarin Chinese. Behavior Research Methods. 2008; 40(3):728–733. [PubMed: 18697668]

Hu, Qian. The acquisition of Chinese classifiers by young Mandarin-speaking children. PhD diss. Boston University; 1993.

Huang, Aijun. Count-mass distinction and the acquisition of classifiers in Mandarin-speaking children. Masters thesis. Chinese University of Hong Kong; 2009.

Huang, Aijun; Thomas, Lee Hun-Tak. Quantification and individuation in the acquisition of Chinese classifiers. In: Otsu, Y., editor. Proceedings of the tenth Tokyo conference on psycholinguistics. Tokyo: Hituzi Syobo; 2009. p. 117-141.

Kouider, Sid; Halberda, Justin; Wood, Justin; Carey, Susan. Acquisition of English number marking: The singular plural distinction. Language Learning and Development. 2006; 2:1–25.

Le Corre, Mathieu; Carey, Susan. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. Cognition. 2007; 105:395–438. [PubMed: 17208214]

Li, Peggy; Barner, David; Huang, Becky. Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese. Language Learning and Development. 2008; 4(4):1–42. [PubMed: 21765807]

Li, Peggy; Cheung, Pierina; Barner, David. In prep. Empty words: An investigation of classifier meaning in Mandarin Chinese. Harvard University; Unpublished ms.

Li, Charles N.; Thompson, Sandra A. Mandarin Chinese: A functional reference grammar. Santa Barbara: University of California Press; 1989.

Locke, John. An essay concerning human understanding. Book III Chap. 3. Nidditch, PH., editor. Oxford: Clarendon Press; 1975.

Loke, Kit Ken; Harrison, Godfrey. Young children's use of Chinese (Cantonese and Mandarin) sortal classifiers. In: Kao, HSR.; Hoosain, R., editors. Linguistics, psychology, and the Chinese language. Hong Kong: Centre of Asian Studies, University of Hong Kong; 1986. p. 125-146.

Markman, Ellen M. Categorization and naming in children: Problems of induction. Cambridge: MIT; 1989.

Matsumoto, Yo. 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; 1987. p. 229-260.

Melgoza, Vicente; Pogue, Amanda; Barner, David. A broken fork in the hand is worth two in the grammar: A spatio-temporal bias in children's interpretation of quantifiers and plural nouns. In: McNamara, DS.; Trafton, JG., editors. Proceedings of the 30th annual conference of the cognitive science society. Austin: Cognitive Science Society; 2008. p. 1581-1585.

Myers, James. Rules vs. analogy in Mandarin classifier selection. Language and Linguistics. 2000; 1(2):187–209.

Sanches, Mary. Language acquisition and language change: Japanese numeral classifiers. In: Sanches, M.; Blount, B., editors. Sociocultural dimensions of language change. New York: Academic Press; 1977. p. 51-62.

Sarnecka, Barbara W.; Kamenskaya, Valentina G.; Yamana, Yuko; Ogura, Tamiko; Yudovina, Yulia B. From grammatical number to exact numbers: Early meanings of "one," "two," and "three" in English, Russian, and Japanese. Cognitive Psychology. 2007; 55:136–168. [PubMed: 17070794]

Shipley, Elizabeth F.; Shepperson, Barbara. Countable entities: Developmental changes. Cognition. 1990; 34:109–136. [PubMed: 2311354]

Soja, Nancy; Carey, Susan; Spelke, Elizabeth S. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. Cognition. 1991; 38:179–211. [PubMed: 2049905]

Spelke, Elizabeth S. What makes us smart? Core knowledge and natural language. In: Gentner, D.; Goldin-Meadow, S., editors. Language in mind: Advances in the investigation of language and thought. Cambridge: MIT; 2003. p. 277-311.

Sumiya, Hiromi; Colunga, Eliana. The effect of familiarity and semantics on early acquisition of Japanese numerical classifiers. In: Bamman, D.; Magnitskaia, T.; Zaller, C., editors. Proceedings of the 30th annual Boston University conference on language development. Somerville, MA: Cascadilla Press; 2006. p. 607-618.

Tai, James H-Y. Chinese classifier systems and human categorization. In: Chen, MY.; Tzeng, O., editors. In honor of William S.-Y. Wang: Interdisciplinary studies on language and language change. Taipei: Pyramid Press; 1994. p. 479-494.

Tse, Shek Kam; Hui, Li; Leung, Shing On. The acquisition of Cantonese classifiers by preschool children in Hong Kong. Journal of Child Language. 2007; 34:495–517. [PubMed: 17822137]

Uchida, Nobuko; Imai, Mutsumi. Heuristics in learning classifiers: The acquisition of the classifier system and its implications for the nature of lexical acquisition. Japanese Psychological Research. 1999; 41(1):50–69.

Wagner, Laura; Carey, Susan. Individuation of objects and events: A developmental study. Cognition. 2003; 90:163–191. [PubMed: 14599752]

Wood, Justin N.; Spelke, Elizabeth S. Infants' enumeration of actions: Numerical discrimination and its signature limits. Developmental Science. 2005; 8:173–181. [PubMed: 15720375]

Wynn, Karen. Children's understanding of counting. Cognition. 1990; 36:155–193. [PubMed: 2225756]

Wynn, Karen. Children's acquisition of number words and counting system. Cognitive Psychology. 1992; 24:220–251.

Wynn, Karen. Infants' individuation and enumeration of actions. Psychological Science. 1996; 7:164–169.

Xu, Fei. Sortal concepts, object individuation, and language. Trends in Cognitive Sciences. 2007; 11:400–406. [PubMed: 17698404]

Yamamoto, Kasumi; Keil, Frank. The acquisition of Japanese numerical classifiers: Linkages between grammatical forms and conceptual categories. Journal of East Asian Linguistics. 2000; 9:379–409.

Zhang, Hong. Chinese numeral classifiers. Journal of East Asian Linguistics. 2007; 16:43–59.

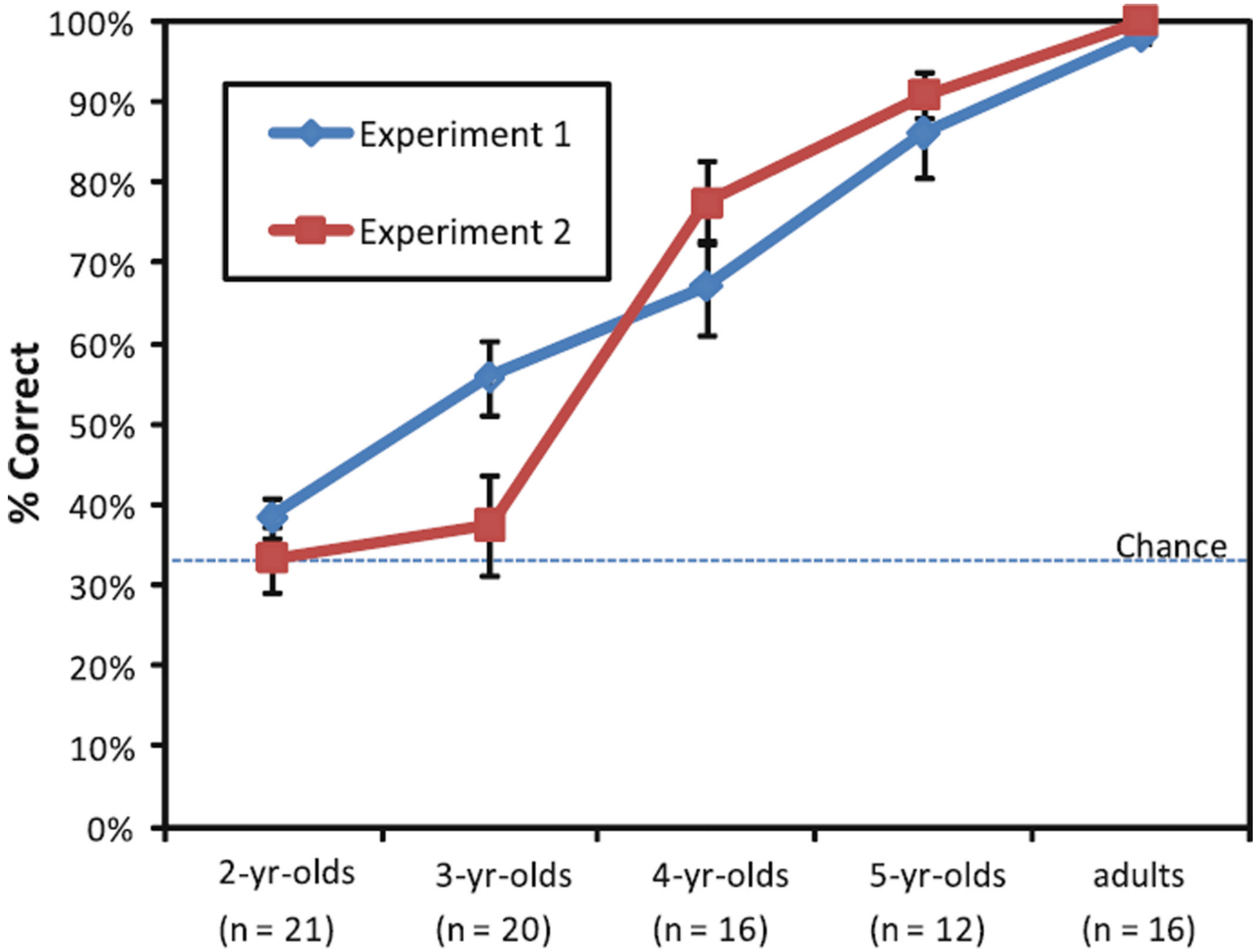


Fig. 1.

Comparison of Experiments 1 and 2 by Age

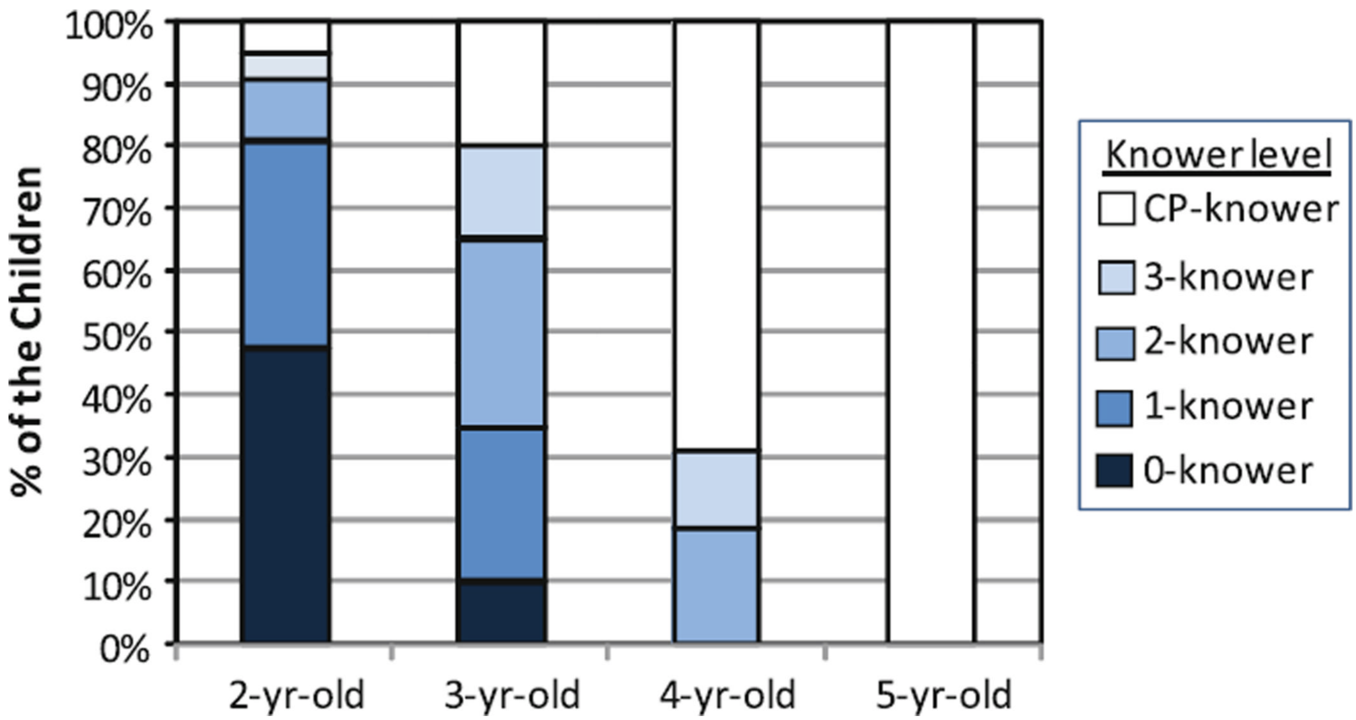


Fig. 2.
Children in each age group categorized by their knower-level

Table 1

Classifiers and stimuli for Experiment 1

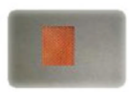
Sortal classifiers	Meanings	Known objects	Novel objects
把 (ba)	objects with handles (e.g., scissors, umbrella)		
根 (gen)	rod shape (e.g. cigar, banana)		
條 (tiao)	thin, flexible, long shape (e.g. string, river)		
塊 (kuai)	chunk, lump (e.g., toy block, rock)		
朵 (duo)	bloom (e.g., cloud, flower)		
顆 (ke)	small objects (e.g., ball, bead)		
片 (pian)	flat, slice (e.g. CD, leaf)		
面 (mian)	flat and smooth surface objects (e.g., mirror, wall)		
張 (zhang)	sheet, flat object (e.g., table, paper)		
隻 (zhi)	animals (e.g., dog, cat)		

Table 2

Percentage correct for known vs. novel object trials in Experiment 1

Classifier	2-year-olds	3-year-olds	4-year-olds	5-year-olds	Adults
把 ba	62 vs. 29 *	35 vs. 35	56 vs. 75	92 vs. 92	100 vs. 100
根 gen	62 vs. 52	70 vs. 70	63 vs. 88 *	75 vs. 100	100 vs. 100
條 tiao	33 vs. 33	70 vs. 65	81 vs. 69	100 vs. 92	100 vs. 100
塊 kuai	67 vs. 57	80 vs. 60	50 vs. 69	83 vs. 58	100 vs. 94
朵 duo	10 vs. 24	45 vs. 50	81 vs. 50 *	92 vs. 83	100 vs. 81
顆 ke	29 vs. 48	55 vs. 55	56 vs. 81	92 vs. 83	100 vs. 94
片 pian	33 vs. 52	35 vs. 60	69 vs. 63	100 vs. 92	100 vs. 100
面 mian	33 vs. 24	35 vs. 50	50 vs. 56	58 vs. 75	100 vs. 100
張 zhang	10 vs. 33	65 vs. 70	81 vs. 69	92 vs. 92	100 vs. 100
隻 zhi	48 vs. 38	60 vs. 50	81 vs. 50 *	92 vs. 75	100 vs. 75 *

Significance test of each pair:

* $p < .01$

Table 3

Percentage correct in Experiment 1 (known and novel trials combined)

Classifier	2-year-olds	3-year-olds	4-year-olds	5-year-olds	Adults
把 ba	45.24**	35.00	65.63**	91.67***	100.00***
根 gen	57.14***	70.00***	75.00***	87.50***	100.00***
條 tiao	33.33	67.50***	75.00***	95.83***	100.00***
塊 kuai	61.90***	70.00***	59.38*	70.83**	96.88***
朵 duo	16.67*	47.50	65.63**	87.50***	90.63***
顆 ke	38.10	55.00*	68.75**	87.50***	96.88***
片 pian	42.86	47.50	65.63**	95.83***	100.00***
面 mian	28.57	42.50	53.13	66.67*	100.00***
張 zhang	21.43	67.50***	75.00***	91.67***	100.00***
隻 zhi	42.86	55.00*	65.63**	83.33***	87.50***

Significance test against chance level of 33.3% correct:

* $p < .05$;

** $p < .01$;

*** $p < .001$

Table 4

Classifiers and stimuli for Experiment 2

Classifier		Meaning	Target	Foils
Individual	個 ge	Individual	one lemon	box of lemons, row of lemons
			one orange	cup of oranges, pile of oranges
Configuration	堆 dui	Pile	pile of bananas	one banana, row of bananas
			pile of stones	one stone, row of stones
	排 pai	Row	row of rings	one ring, pile of rings
			row of stars	one star, box of stars
Container	杯 bei	Cup	cup of shoes	one shoe, a pair of shoes
			cup of carrots	one carrot, bunch of carrots
	盒 he	Box	box of pens	one pen, pile of pens
			box of candies	one candy, pile of candies

Table 5

Percentage correct broken down by classifiers and age for Experiment 2

		2-year-olds	3-year-olds	4-year-olds	5-year-olds	Adults
Individual	個 ge	57.14*	77.50***	93.75***	100.00***	100.00***
Configuration	堆 dui	19.05	25.00	65.63***	79.17***	100.00***
	排 pai	26.19	20.00	62.50**	83.33***	100.00***
	Average	22.62*	22.50	64.06***	81.25***	100.00***
Container	杯 bei	33.33	27.50	84.38***	91.67***	100.00***
	盒 he	30.95	37.50	81.25***	100.00***	100.00***
	Average	32.14	32.50	82.81***	95.83***	100.00***

Significance test against chance level of 33.3% correct:

* $p < .05$;

** $p < .01$;

*** $p < .001$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 6

Comparison of percentage correct of non-biased and biased 2- and 3-year-olds

		No bias		Bias	
		2-year-olds	**3-year-olds**	**2-year-olds**	**3-year-olds**
		(*n* = 17)	**(*n* = 14)**	**(*n* = 4)**	**(*n* = 6)**
Individual	個 ge	50.00	71.43**	87.50*	91.67***
Configuration	堆 dui	23.53	32.14	0.00***	8.33*
	排 pai	32.35	28.57	0.00***	0.00***
	Average	27.94	30.36	0.00***	0.00***
Container	杯 bei	41.18	39.29	0.00***	0.00***
	盒 he	38.24	53.57	0.00***	0.00***
	Average	39.71	46.43	0.00***	0.00***

Significance test against chance level of 33.3% correct:

* $p < .05$;

** $p < .01$;

*** $p < .001$