

Running head: GRAMMATICAL CATEGORIZATION IN INFANTS

**Grammatical Categorization of Nouns and Verbs in Mandarin-Learning Infants**

Zhao Zhang\*

Nankai University, China

Institute of Linguistics, Chinese Academy of Social Sciences, China

Rushen Shi\*

Université du Québec à Montréal, Canada

&

Aijun Li

Institute of Linguistics, Chinese Academy of Social Sciences, China

\* Corresponding authors

## Abstract

This study inquired whether infants can categorize nouns and verbs at the early stage of language acquisition. Mandarin-Chinese-learning 11-14-month-old infants were tested in a preferential looking experiment. Targets were ambi-categorical words that can be nouns and verbs. These words do not belong to the infant-directed vocabulary in Mandarin, thus enabling us to test if infants have generalized representation of noun and verb categories. One group of infants was familiarized with the targets used as nouns in sentences, and another group with the targets as verbs. All infants then heard test trials that presented new sentences, one trial type presenting a familiarized target in a new noun context, and the other trial type presenting the target in a new verb context. Results revealed that infants in both familiarization groups distinguished between the two types of test trials. Thus, Mandarin-learning infants showed evidence of categorizing nouns and verbs at an early age, as did European-language-learning infants in previous perceptual studies. We suggest that it is a language-universal characteristic that infants use frequent function words to acquire the grammatical categorization of open-class words early in infancy.

Key words: language acquisition, grammatical categorization, syntax

### Grammatical Categorization of Nouns and Verbs in Mandarin-Learning Infants

Assigning words to grammatical categories (such as nouns, verbs, auxiliaries, demonstratives, etc.) is crucial for syntactic acquisition. These basic categories are the building blocks of larger, more sophisticated phrasal structures. Moreover, they are the primitive grammatical units accepted by all syntactic theories despite various differences among the theories.

Nouns and verbs are among the most basic grammatical categories. These two categories have attracted wide interest in the field of language acquisition. A few perceptual studies have been published on the initial categorization of nouns and verbs in infants aged 1-1.5 years (German: Höhle, et al., 2004; English: Mintz, 2006; French: Shi & Melançon, 2010). These studies used novel target words to assess whether infants have abstract, generalized knowledge of grammatical categories. Höhle and colleagues (2004) created non-word targets and used each as nouns and verbs in German. They familiarized one group of German-learning infants with the targets (e.g., *pronk*) as nouns (i.e., following a German determiner), and another group of infants with the non-words as verbs (i.e., following a pronoun). Both groups were then tested with two types of sentences: those containing the non-words being used as nouns and those containing the same non-words being used as verbs. Essentially, none of the contextual words in the test sentences had been used during the familiarization period. Results showed that the noun familiarization group discriminated between the two types of test sentences, indicating that infants categorized nouns.

Noun categorization was also shown in French-learning infants in a perceptual study. Shi & Melançon (2010) familiarized one group of infants with non-words following French determiners (e.g., *ton mige*, ‘your mige’; *des miges*, ‘some miges’; *ton crale*, ‘your crale’), and another group with the non-words following pronouns (e.g., *je mise*, ‘I mige’; *il mige*, ‘he miges’; *il crale*, ‘he crales’) during the familiarization phase. Test stimuli were the same non-words co-occurring with a different, non-familiarized determiner (e.g., *le mige*, ‘the mige’) in half of the trials and with a non-familiarized pronoun (e.g., *tu miges*, ‘you mige’) in the other half of the trials. The infants who were familiarized with the noun context discriminated between the two types of test trials. That is, they perceived the non-words as belonging to the noun category that could be preceded by different determiners, but not by pronouns.

Mintz (2006) tested English-learning infants’ categorization of nouns and verbs. Infants were familiarized with sentences containing non-words embedded in noun versus verb frequent frames (e.g., ... *his gorp on* ...; ... *the bist in* ...; ... *you deeg the* ...; ... *I lonk you* ...). Test sentences contained the same non-words and frames, but the pairing of specific non-words and specific frames were all different, some grammatical (e.g., ... *his bist on* ...; ... *you lonk the* ...) and some ungrammatical (e.g., ... *you bist the* ...; ... *I gorp you* ...; ... *his lonk on* ...) relative to the pairing in the familiarization phase. Infants discriminated some test trials. Mintz interpreted the results as suggesting that English-learning infants categorized verbs.

Existing perceptual studies on early grammatical categorization of nouns and verbs have only tested infants in European language environments. In the present study we inquired whether the same capability is present in the acquisition of non-European

languages such as Mandarin Chinese. Mandarin is typologically distinct from European languages. It has no inflectional morphology, although a small cohort of closed-class words do occur frequently as in other languages. Unlike many languages, pronouns can be dropped from various syntactic positions. We tested whether Mandarin-learning infants can categorize nouns and verbs. Our goal was to understand if the emerging ability of grammatical categorization ability in early infancy is largely language universal.

There are a few published studies on children's production of nouns versus verbs in Mandarin Chinese (Chen, Setoh, Meng, & Tardif, 2009; Peng, 2009; Tardif, 1996). They showed evidence that verbs in Mandarin begin to emerge no later than nouns in children's early speech, unlike the general verb delay in children of many other languages. It is, however, difficult to know how much this kind of data can reveal representations that are truly syntactic, since children's initial speech contains generally single-word utterances. Two recent perceptual studies also examined the acquisition of nouns versus verbs in Mandarin (Chan, Tardif, Chen, Pulverman, Zhu, & Meng, 2011; Imai, Li, Haryu, Okada, Hirsh-Pasek, Golinkoff, & Shigematsu, 2008), but focused on the question of whether children infer the noun meaning or the verb meaning when hearing a novel word, rather than on their syntactic categorization. The results of the two studies are mixed. In Chan, et al., Mandarin-learning infants as young as 14 and 18 months associated an isolated nonsense monosyllable to an action when both noun and verb meanings were logically possible. Imai, et al., studied older Mandarin-learning children. They showed that while children were successful in associating noun meaning,

the verb-meaning association was much poorer despite contextual syntactic support in the speech stimuli, even for 3- and 5-year-olds.

The above studies on Mandarin did not clearly address the question of children's grammatical categorization of nouns and verbs. In the present study the stimuli included variable syntactic contexts that surrounded nouns and verbs. This approach was therefore a direct test of the grammatical knowledge. Furthermore, noun and verb target words are not part of the infant-directed vocabulary, thus allowing us to examine whether infants' grammatical category knowledge can generalize to novel words in novel utterances.

We chose noun-verb ambi-categorical words as targets. An experimental advantage of ambi-categorical over uni-categorical words is that both noun and verb recordings of an ambi-categorical target are natural and fluent. During the familiarization period of the experiment, we exposed the same target in noun sentences to one group of infants, and in verb sentences to another group. Following familiarization, we tested both groups of infants on one new sentence using the target as a noun and on another new sentence using the same target as a verb. If infants attended to the syntactic contexts during familiarization and interpreted the target as belonging to a particular grammatical category, they should perceive the target appearing in that category during the test phase as cohesive and discriminate it from the use of this target in a different category.

## Method

### *Participants and stimuli*

Thirty-two monolingual Chinese-learning infants completed this experiment (22 boys, 10 girls, aged from 11 months 0 days to 14 months to 22 days, mean age = 13 months 5 days). We chose two noun-verb ambi-categorical target words *tun* (a type of

farming village; to stock up) and *sheng* (province; to save). These are not the kind of words that are used when addressing a child in Chinese. Indeed, according to a corpus of mother-child natural interactions (Zhang, 2008), these two words each had 0 occurrence (corpus size: 6204 total input word tokens) in mothers' speech to their 15- to 30-month-old infants. Thus, infants were not expected to know these words.

We used three noun-contextual expressions, each of which can have a noun embedded between the two elements, *zai...shang* (at...on), *ge...shi* (classifier...be) and *de...dou* (particle...all). Similarly, three verb-contextual expressions were used, *jiu...le* (just...aspect marker), *mei...guo* (not...aspect marker) and *hai...zhe* (still...aspect marker). Based on our analysis of the corpus mentioned above (Zhang, 2008; 2229 total word tokens produced by children), these contexts begin to emerge in children's natural speech between 23 to 30 months of age. However, these contexts all consist of closed-class words and are frequently used in Chinese. The present experiment thus tested if children younger than this age range can perceive these contexts and rely on them for syntactic categorization.

Four sentences were constructed for the noun-contextual expressions *zai...shang* and *ge...shi*. Each context surrounded a target word *tun* or *sheng*. Four other sentences were constructed with the same targets *tun* and *sheng* in the two verb-contextual expressions *jiu...le* and *mei...guo*. In addition, we created another two sentences, one with a new noun-contextual expression *de...dou*, another with a new verb-contextual expression *hai...zhe*, each using the target *tun*. Thus, there were a total of 5 sentences in noun contexts and 5 in verb contexts (see Appendix 1). To more reliably test syntactic categorization, we deliberately ensured that there was no semantic contextual relation

across the sentences. That is, the sentences containing the targets in the same category provided no plausible semantic cues that could support the grammatical categorization of the targets.

A native Chinese female speaker recorded multiple repetitions of the 10 sentences in an acoustic chamber, using a lively voice. The sampling frequency was 44.1 kHz, and the bit rate was 16 bits. We carefully selected two exemplars for each of the 10 sentences, ensuring that the prosodic properties of the target words were comparable in noun versus verb uses (see Appendix 2). The durations of the sentences ranged from 1.427 s to 1.707 s (noun sentences), and from 1.49 s to 1.64 s (verb sentences).

In addition to speech stimuli, we used water bubble sounds in the pre-test and post-test trials. A black-and-white checkerboard was presented during these trials. The pre-test served to introduce the infant to the procedure, and the post-test trial was presented after the test trials to mark the end of the experiment. The visual stimulus for the speech trials was a puppet. During familiarization the puppet moved her hands and legs while the sentences were played, and her mouth opened and closed in synchrony with the speech stimuli. The animation and audio-visual synchrony were created using Adobe® Flash® CS3 Professional software. During the test phase, the puppet was stationary. The attention-getter consisted of the sound of a cricket, accompanied by an animation of a bouncing ball. It served to attract infants' attention between trials.

### *Design*

As shown in Table 1, infants were divided into two groups, the noun familiarization group and the verb familiarization group. For the noun familiarization group, the target words *tun* and *sheng* were used as nouns. They were embedded in two



noun contexts (*zai...shang; ge...shi*), in a total of four sentences. For the verb familiarization group, the targets *tun* and *sheng* were used as verbs in two verb contexts (*jiu...le; mei...guo*), and there were also a total of four sentences. In the test phase, both groups of infants heard the same stimuli: a new sentence containing the target word *tun* in a noun context that was different from familiarization (*de...dou*) and another new sentence containing the target *tun* in a verb context that was different from familiarization (*hai...zhe*).

The test trials were of two types, Grammatical and Ungrammatical trials. For the noun familiarization group, the Grammatical trials presented the ...*de tun dou*... sentence (noun context), whereas the Ungrammatical trials presented the ...*hai tun zhe* ... sentence (verb context). The reverse was the case for the verb familiarization group. That is, the ...*hai tun zhe*... sentence was grammatical whereas the ...*de tun dou*... sentence was ungrammatical. The sentence within each trial was presented two times. The Grammatical and Ungrammatical trials were presented alternately for a total of six trials. Within each group, infants were divided into two subgroups, with one sub-group hearing a Grammatical trial first, and the other sub-group starting the test phase with an Ungrammatical trial.

### *Procedure*

Infants were individually tested in a visual preferential procedure. The infant sat on his or her parent's lap facing a monitor in an acoustic chamber. Loudspeakers were placed on each side of the monitor and simultaneously played the auditory stimuli during the experiment. The parent heard masking music through headphones and avoided interfering with the infant's responses. The researcher, blind to all stimuli presentation,

observed the infant in an adjacent room through a monitor, and administered the experiment using an experimental program. The software presented trials, while recording all looking times online.

The main part of the experiment included a familiarization phase and a test phase. When the infant looked at the monitor, the pre-test trial was presented (15 s), followed by the familiarization phase presenting the puppet and the familiarization sentences. The familiarization sentences were played continuously, and the total duration of this phase was 101 s. After familiarization, the experiment advanced automatically to the test phase. Each test trial lasted 9 seconds, and each was initiated upon the infant's look toward the monitor. Once a trial began, it continued for the entire trial duration. Throughout the experiment the silence interval between any two sentences was approximately 1 s. The blind observer pressed down a computer key to record the infant's looks toward the screen during each trial.

### Results

All infants received the same amount of exposure to the familiarization materials. To assess infants' active attention during the familiarization phase, we compared the looking time of the noun group and that of the verb group during the familiarization phase, using an unpaired T-test. The results revealed no difference in looking times for the two groups (Noun Group: mean 76.74 s, SE 4.51 s; Verb Group: mean 76.55 s, SE 4.04 s;  $t(30) = 0.032$ ;  $p = 0.975$ ), suggesting that the two groups of infants attended to the familiarization stimuli comparably.

For each infant, we calculated the average cumulative looking time per trial for Grammatical test trials versus Ungrammatical test trials. Following the standard practice

in this testing procedure (e.g., Vouloumanous & Werker, 2004), the first test trial of each type was removed from the analysis, as they are usually unstable. The analysis was a mixed ANOVA, with Familiarization Group (Noun Group vs. Verb Group) as the between-subject factor, and Grammaticality (Grammatical vs. Ungrammatical test trials) as the within-subject factor. The results showed no effect of Group ( $F(1,14) = 0.486, p = 0.491$ ) nor of Grammaticality ( $F(1,14) = 0.361, p = 0.552$ ). There was a significant effect of Familiarization Group X Grammaticality interaction ( $F(1,14) = 14.62, p = 0.001$ ). Thus, follow-up paired T-tests were conducted separately for each familiarization group. Infants in both groups showed a significant discrimination of the two types of test trials. Specifically, for the Noun Group, the looking times for Grammatical trials (Mean: 7.214 s; SE: 0.435 s) and Ungrammatical trials (Mean: 7.889 s; SE: 0.291 s) were significantly different,  $t(15) = -2.468; p = 0.026$ . For the Verb Group, the looking times for Grammatical trials (Mean: 7.662 s; SE: 0.361 s) and Ungrammatical trials (Mean: 6.734 s; SE: 0.447 s) were also significantly different,  $t(15) = 2.921; p = 0.011$ . These discrimination results suggest successful noun and verb categorization.

### Discussion

In this study we used a perceptual task to test whether Mandarin-Chinese-learning infants at the early stage of language acquisition represent abstract grammatical categories. Infants between 11 and 14 months of age showed evidence of categorization. After being familiarized with a novel word as a noun or as a verb in sentences, infants discriminated between new sentences that used the same target word as a noun versus as a verb. Our experiment presented variable syntactic contexts that surrounded the targets across familiarization and test, so that infants must possess knowledge of syntactic

categorization in order to succeed in the task. The use of target words unknown to infants eliminated any contribution of their semantic component, allowing us to strictly assess infants' syntactic knowledge. More importantly, infants showed evidence of generalizing their grammatical categorization to novel words and novel phrasal combinations, thus demonstrating abstract representations.

The Mandarin-learning infants in our study categorized nouns and verbs at a similar age to those in previous perceptual studies with infants in the environment of European languages. This suggests that the ability to perform grammatical categorization early in acquisition is likely a language universal characteristic. Certain crucial information supporting grammatical categorization must be universally available in the input to assist infants. It has been proposed that function words are an important cue to grammatical categorization (e.g., Christophe, Guasti, & Nespor, 1997; Morgan, Shi, & Allopenna, 1997; Shi, Morgan, & Allopenna, 1998). It has been shown in parental speech of different languages (including Mandarin Chinese) that open-class words (e.g., nouns) co-occur with a small set of function words, which have distinct acoustic and phonological characteristics and are extremely frequent relative to open-class words (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Monaghan, Christiansen, & Chater, 2007; Shi, Morgan, & Allopenna, 1998). Correspondingly, function words are recognized very early, during the first year of life (e.g., Hallé, Durand, & de Boysson-Bardies, 2008; Shi, Werker, & Cutler, 2006), and they serve as anchors for learning adjacent words (Shi & Lepage, 2008; Shi, Cutler, Werker, & Cruickshank, 2006; Valian & Coulson, 1988). In the present study target words were surrounded by function words, as is the case of previous studies of categorization in German, English and French (Höhle, et al., 2004;

Mintz, 2006; Shi & Melançon, 2010). Infants may use function words to discover the grammatical categories for open-class words.

The infants in our study showed different preferential looking directions for noun categorization versus verb categorization. The noun familiarization group showed a novelty preference during the test phase whereas the verb familiarization group showed a familiarity preference. Novelty and familiarity preferences are both commonly observed in infant studies. While it is usually not possible to clearly predict the looking direction in preferential looking procedures, the accepted view is that discrimination by a group of participants in either direction is sufficient as the key evidence of a capability/representation. In the present study both the noun and verb familiarization groups discriminated the two types of test trials, thus demonstrating that infants categorized both nouns and verbs. We might further interpret the different preference directions of the two groups as suggesting a stronger categorization ability for nouns than for verbs. This interpretation would be consistent with the observation that novelty preference may indicate better knowledge of the experimental stimuli, while familiarity preference may indicate emerging, less robust knowledge (Cyr & Shi, 2013; Hunter & Ames, 1988; Thiessen & Saffran, 2003). The key finding, however, is that both noun and verb categories in Mandarin emerges early in infancy.

It should be noted that the prosodic characteristics of the target words in noun versus verb contexts were deliberately balanced in our study. We carefully chose our recorded stimuli so that target productions in noun and verb sentences were equivalent. This manipulation was designed to test pure contextually driven, syntax-like categorization. Nevertheless, there is cross-linguistic evidence that nouns and verbs are

supported by distinct prosodic (e.g., Conwell & Morgan, 2012; Shi & Moisan, 2008) and phonological (e.g., Cassidy & Kelly, 1991; Kelly, 1992; Monaghan, Christiansen, & Chater, 2007; Sereno & Jongman, 1990) cues. With respect to Mandarin, nouns and verbs are produced with prosodic differences by parents (Li, Shi, & Zhang, 2011). It is also possible that certain phonological differences (e.g., syllabic structure) exist for nouns and verbs in Mandarin input. We expect that prosodic and phonological cues may assist early grammatical categorization, by enabling infants to link open-class words (such as nouns or verbs) that possess such speech cues with specific contextual function words (e.g., determiners, auxiliaries). Previous studies using an artificial language (Gomez & Lakusta, 2004) and a foreign language (Gerken, Wilson, & Lewis, 2005) showed that infants can indeed associate distributional and phonological cues for form class categorizations in artificial and unfamiliar languages. These findings are consistent with the view that acoustic and phonological cues may assist learners to break into the syntactic system (e.g., Christophe, Guasti, & Nespors, 1997; Morgan, Shi, & Allopenna, 1997; Shi, Morgan, & Allopenna, 1998). Future perceptual experiments can test the role of acoustical and/or phonological cues in the initial categorization of nouns and verbs.

The successful noun and verb categorization in the present study supports the theory of early syntactic knowledge in a few ways. First, our stimuli involved target words that are not used in speech addressed to children, and the word combinations with the targets were all novel to infants. Infants could not rely on any previously stored exemplars to perform the task. Instead, they were performing abstract grammatical categorization. Furthermore, the infants in our study were young in comparison to those in production-based grammatical categorization studies. Our infants were at the

beginning of the single-word production stage. Early syntactic representations are therefore far more sophisticated than what could be revealed by infants' speech production.

The ability to form grammatical categories at a very early age may play a crucial role in infants' learning of other linguistic tasks. For instance, syntactic representations can guide subsequent learning of meaning, as proposed in the Syntactic Bootstrapping theory (e.g., Gleitman, 1990). Basic categories are needed in order to acquire larger syntactic constituents. The present study on Mandarin-learning infants corroborates previous categorization studies on European-language-learning infants, suggesting that syntactic categories are universally present early in language acquisition.

## References

- Cassidy, K., & Kelly, M. 1991. Phonological information for grammatical category assignments. *Journal of Memory and Language*, 30, 348-369.
- Chan, C. C. Y., Tardif, T. Chen, J., Pulverman, R. B., Zhu, L., & Meng, X. (2011). English- and Chinese-learning infants map novel labels to objects and actions differently. *Developmental Psychology*, 47(5), 1459-1471.
- Chen, J., Setoh, P., Meng, X., & Tardif, T.(2009). The influence of adult input on Children's early word learning: A case study of a Mandarin-speaking child. *Acta Psychologica Sinica*, 41(8), 715–725.
- Christophe, A., Guasti, T., & Nespors, M. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12, 585–612.
- Conwell, E., & Morgan, J. L. (2012). Is it a noun or is it a verb? Resolving the ambicategoricity problem. *Language Learning and Development*, 8(2), 87-112.
- Cyr, M., & Shi, R. (2013). Development of abstract grammatical categorization in infants. *Child Development*, 84, 617–629.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249-268.
- Gervain, J., Nespors, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, 57, 56–74.
- Gleitman, L. R. (1990). The structural sources of verb meaning. *Language acquisition*, 1, 3–55.



- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7(5), 567–580.
- Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech*, 51, 23–44.
- Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5, 341-353.
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, pp. 69–95). Norwood, NJ: Ablex.
- Kelly, M. H. (1992). Using sound to solve syntactic problems. The role of phonology in grammatical category assignments. *Psychological Review*, 99, 349-364.
- Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, 79(4), 979 – 1000.
- Li, A., Shi, R., & Zhang, Z. (2011). Prosodic properties of verbs and nouns in infant-directed Mandarin Chinese. *Studies of the Chinese Language (ZHONGGUO YUWEN)*, 344, 387-397.
- Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action Meets Word: How Children Learn Verbs* (pp. 31-63). New York: Oxford University Press.

- Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, *55*, 259–305.
- Morgan, J., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 263–283), Hillsdale, NJ: Erlbaum.
- Peng, X. (2009). On noun bias in children's early lexicon. *Journal of Hunan University (Social Sciences)*, *23*(6), 81-84.
- Sereno, J. A., & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, *19*, 387–404.
- Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, *119*(6), EL61-EL67.
- Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, *11*(3), 407-413.
- Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, *15*, 517-533.
- Shi, R., & Moisan, A. (2008). Prosodic cues to noun and verb categories in infant-directed speech. In H. Chan, H. Jacob, & E. Kiparsky (Eds.). *BUCLD 32: Proceedings of the 32th annual Boston University conference on language development* (Vol. 2, pp. 450-461). Somerville, MA: Cascadilla Press.

- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language, 25*, 169-201.
- Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy, 10* (2), 187-198.
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*(3), 492 -504.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Statistical and stress cues in infant word segmentation. *Developmental Psychology, 39*, 706–716.
- Valian, V., & Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language, 27*, 71-86.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science, 7*(3), 270–276.
- Zhang, Z. (2008). *A study of Chinese children's acquisition of verb complement constructions*. M.A. thesis, Capital Normal University, Beijing, China.

## Appendix 1

*Stimuli and design of the familiarization and test phases*

Familiarization	
Noun Group	Verb Group
<p><i>zai ... shang</i> (at ... on) <i>mama zai tun shang shuijiao.</i> (mama at village on sleep) <i>Shushu zai sheng shang gongzuo.</i> (uncle at province on work)</p> <p><i>ge ... shi</i> (classifier ... be) <i>Zhege tun shi dongwuyuan.</i> (this classifier village is zoo) <i>Nage sheng shi fangzi.</i> (that classifier province is house)</p>	<p><i>jiu ... le</i> (just ... aspect-marker) <i>Baobao jiu tun le binggan</i> (baby just stock up <i>le</i>-asp.-marker cookie) <i>Gege jiu sheng le wanju.</i> (brother just save up aspect-marker toy)</p> <p><i>mei ... guo</i> (not ... aspect-marker) <i>Jiejie mei tun guo yifu.</i> (sister not save up <i>guo</i>-asp.-marker clothes) <i>baba mei sheng guo dongxi.</i> (dady not save up <i>guo</i>-asp.-marker thing)</p>
Test	
<p>A</p> <p><i>de ... dou</i> (particle ... all) <i>Hongse de tun dou piaoling</i> (red particle village all beautiful)</p>	
<p>B</p> <p><i>hai ... zhe</i> (still ... aspect-marker) <i>Nainai hai tun zhe pingguo.</i> (grandma still stock up <i>zhe</i>-aspect-marker apple)</p>	

## Appendix 2

Average acoustic values (and standard deviations) across multiple tokens of the target words in the experiment

Acoustic measure	Mean for noun use	Mean for verb use	Independent T-test
Target word duration (sec)	0.244 (0.046)	0.245 (0.039)	$t(18)=-0.1$ ; $p=0.922$
target word mean pitch (Hz)	213.974 (35.674)	223.557 (11.745)	$t(18)=-0.807$ ; $p=0.43$
Target word mean amplitude (dB)	68.405 (2.774)	68.327 (1.986)	$t(18)=0.072$ ; $p=0.943$