Structures of Function Words Guide Mandarin-Learning 19-Month-Olds in Backward Syntactic Categorization

Yuanfan Ying, Xiaolu Yang, and Rushen Shi

1. Introduction

Infants are sensitive to function words from a young age. Though they cannot use function words properly before three years of age (Brown, 1973), infants’ perception of function words starts much earlier according to the prosodic-functor bootstrapping hypothesis (Christophe et al., 2008; Shi, 2005, 2014). One-year-old infants already recognize frequent functional elements in their native language (Shi et al., 2006; Shi & Lepage, 2008) and track their immediate syntactic contexts (Dye, Kedar, & Lust, 2019; Mintz, 2003).

Shortly after their first birthday, infants can use function words alone to predict the syntactic category of a following word in a bigram. For instance, they tend to categorize novel words preceded by a determiner as nouns. In an experiment with German-learning infants, Höhle et al. (2004) used a variation of the Head-turn Preference Procedure (HPP) and demonstrated that after being familiarized with determiner-noun sequences like \( \text{ein}_\text{Det} \ X_N \ ‘\text{a } X’ \) (\( X \) represents a novel word), 14- to 16-month-olds, but not 12- to 13-month-olds, categorized \( X \) as a noun. To be specific, during the test phase, infants listened significantly longer to verb passages (i.e., \( X \) as a verb) than to noun passages (i.e., \( X \) as a noun). However, infants familiarized with pronoun contexts (e.g., \( \text{die}_\text{Pron} \ X_v \ ‘\text{she } X’ \)) did not show differentiation between the two types of passages. Similar results have also been found in another study with French-learning 14-month-olds (Shi & Melançon, 2010). Instead of listening to passages during the test phase, infants familiarized with novel words in either determiner-noun sequences (i.e., \( \text{ton} \ X \ ‘\text{your } X’ \), \( \text{des} \ X \ ‘\text{some } X’ \)) or pronoun-verb sequences (i.e., \( \text{je} \ X \ ‘\text{I } X’ \), \( \text{il} \ X \ ‘\text{he } X’ \)) listened to two types of bigram sequences containing a new function word that was either a determiner (\( \text{le} \ X \ ‘\text{the } X’ \)) or a pronoun (\( \text{tu} \ X \ ‘\text{you } X’ \)). Again, only infants familiarized with determiner-noun sequences discriminated the test trials, which provides robust evidence that infants use determiners to anticipate a following noun.

* Yuanfan Ying, Tsinghua University, jackying.yyf@yahoo.com; Xiaolu Yang, Tsinghua University, xlyang@mail.tsinghua.edu.cn; Rushen Shi, Université du Québec à Montréal, shi.rushen@uqam.ca. Many thanks to all members of the Language Acquisition Lab at Tsinghua University, the Chinese National Funding of Social Sciences #11BYY080, to all parents and children who participated in the study, and to all those who have contributed to the study.

Though function words marking the syntactic category of their preceding word are also common in human language, there is currently little evidence whether infants can categorize words backwardly, namely to categorize a novel word \( X \) preceding a function word. Previous studies looking at bigram distribution (Höhle et al., 2004; Shi & Melançon, 2010) lend support only to infants’ capability of categorizing \( X \) following a function word (or forward categorization), while those looking at frame distribution (e.g., Mintz, 2006) provide infants with both preceding and following contexts, from which we cannot tell whether backward categorization (\( aXb \)) alone plays a part.

This study aims to answer the question whether backward categorization is available to toddlers. To do this, we investigated syntactic categorization in Mandarin-learning 19-month-olds. Previous studies on syntactic categorization in Mandarin-learning infants looked either at forward categorization (Zhang, Shi, & Li, 2015) or at frame distribution (Zhang, Shi, & Li, 2014). Given the early emergence of the noun category in many languages, we first examined whether toddlers can use a function word marking a preceding noun for backward categorization.

2. Experiment 1: Backward categorization of nouns

To test whether toddlers are capable of backward categorization of nouns, we use the focus particle \( ye \) that typically co-occurs with a preceding noun as the focus (e.g., Constant & Gu, 2010; Hole, 2004). (1) is an example to illustrate the co-occurrence.

(1) [natali]_F ye xihuan mihoutao
Natalie  Foc  like  kiwi
‘Even Natalie likes kiwi.’ (Natalie is the least likely to like kiwi)

We also checked Tong’s corpus in CHILDES (Deng & Yip, 2018) for the bigram probability of \( XN-ye \).\(^1\) As the corpus analysis yielded 0.957 for the targeted bigram (see Appendix for the calculation method), the co-occurrence is indeed reliable.\(^2\)

2.1. Methods

2.1.1. Participants

Twelve Mandarin-learning toddlers (7 boys, 5 girls) with a mean age of 19 months 14 days (range = 18 months 18 days to 20 months 25 days) were tested. Data from 4 other toddlers were excluded in the analysis due to fussiness (\( n = 3 \)) and misoperation (\( n = 1 \)).

---

\(^1\) The corpus is based on the speech of caregivers to Tong, a Mandarin-learning boy between the age of 19 and 20 months.

\(^2\) The token frequency of \( ye \) is 22 (5465 total word tokens), which is equivalent to 4209 per million words.
2.1.2. Stimuli and Design

Sequences of X_N-ye-Y_V (‘even X Y’) were created for the familiarization phase. X represented the target word and Y the filler word. Two monosyllabic nonce words (shai1, man2) were used for X, both ambiguous between a noun (shai1 ‘sieve’; man2 ‘eel’) and a verb (shai1 ‘to sift’; man2 ‘to conceal’). The two disyllabic words (tong3ji4, jian3yan4) used for Y were also ambiguous between a noun (tong3ji4 ‘calculation’; jian3yan4 ‘examination’) and a verb (tong3ji4 ‘to calculate’; jian3yan4 ‘to examine’). Words used for X and Y were unfamiliar to young children. Sequences of shi-zhege-X_N (‘it’s this-CL X’) and dou-keyi-X_V (‘all may X’) were created for the test phase.

A Mandarin female speaker recorded all the sequences used for our stimuli (sampling frequency 44.1 kHz, bit rate 24 bits). There were a total of 12 tokens of X_N-ye-Y_V sequences (4 types, each with 3 exemplars of different intonation patterns) for the familiarization phase. Two strings were constructed for the test phase, one for grammatical trials (i.e., shi-zhege-X_N; 2 types × 4 repetitions = 8 tokens), and the other for ungrammatical trial (i.e., dou-keyi-X_V; 2 types × 4 repetitions = 8 tokens). Each string had a duration of 17.6 sec, with an interval of 1000 msec between any two tokens.

The visual stimuli for the familiarization and the test phases were a lip-sync puppet, with prerecorded materials being aligned with the auditory stimuli. The attention-getter was the scene of a rotating moon on which a rabbit and a fairy girl were chasing an animate carrot, accompanied by a piece of instrumental music. A static meadow with light music in the background was used as the stimuli for the post-test trial (for determining whether the toddler was still on task by the end of the experiment). Table 1 shows the stimuli and design.

**Table 1. The design of the experiment**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Audio stimuli</th>
<th>Visual stimuli</th>
</tr>
</thead>
</table>
| Familiarization (30 sec) | shaiN ye tongji  
manN ye tongji  
manN ye jianyan  
shaiN ye jianyan | ![image] |
| Test (10 trials) | Trial A  
shi zhege shaiN  
shi zhege manN (grammatical) | ![image] |
|              | Trial B  
dou keyi shaiV  
dou keyi manV (ungrammatical) | ![image] |
2.1.3. Procedure

The visual fixation procedure (Cooper & Aslin, 1990) was adopted for the experiment. Each toddler was led into a sound-proof test booth by an experimenter and instructed to sit between his or her parent’s laps on the sofa to watch a cartoon to be played on a TV screen in front of him or her. The parent was then asked to put on earphones to listen to music in a way that would not influence the toddler’s reaction to the stimuli. After the experiment began with an attention-getter, an experimenter outside the booth could monitor the toddler’s eye movement through an HD camera placed under the TV screen and press the spacebar to initiate the familiarization phase upon the toddler’s look to the screen. Presentation of the familiarization stimuli would be interrupted if the toddler looked away for more than 2 sec, with the attention-getter automatically appearing. As soon as the child looked back to the screen, stimulus presentation was resumed. The 12 tokens of familiarization sequences were pseudo-randomized to ensure maximal local variability in utterance types and intonation patterns. This variability of stimuli was also intended to engage the toddler. This familiarization phase only terminated when the accumulated looking time reached 30 sec (for the criterion, see Shi & Melançon, 2010). The test trials alternated between grammatical and ungrammatical trials. A trial would terminate if the toddler looked away for at least 2 sec or if the maximum length (17.6 sec) was reached. Whether the test phase began with grammatical or ungrammatical trials was counterbalanced between subjects. For the post-test trial presenting new stimuli, the toddler’s looking time should recover compared to his or her looking time for the last test trial if he or she was on task.

2.2. Results and Discussion

Figure 1 plots toddlers’ mean looking time to the two types of test trials.

![Figure 1. Mean looking time for grammatical and ungrammatical trials during the test phase in Experiment 1 (ye condition)]
Our paired t-test indicates that toddlers’ average looking time per trial during grammatical trials ($M = 12.547$ sec, $SE = .702$ sec) was significantly longer than that during ungrammatical trials ($M = 11.349$ sec, $SE = .754$ sec), $t(11) = 2.913$, $p = .014$, Cohen’s $d = .841$, 2-tailed.

This result demonstrates that Mandarin-learning 19-month-olds can use the focus particle $ye$ to categorize its preceding noun. This provides evidence that around one year and a half’s age, infants can use function words following a noun to categorize the noun backwardly.

One way to interpret toddlers’ success in Experiment 1 is that backward categorization is facilitated when the function word and its preceding target word co-occur within a phrase, or a syntactically natural cluster. In the case of the focus particle $ye$, the novel word $X$ is within a focus phrase or a FocP, as represented in $[\text{FocP } X\ y e \text{Foc} \ldots]$. Syntactic categorization based on within-phrase co-occurrences also seems to be privileged in forward categorization. That infants in previous studies (Höhle et al., 2004; Shi & Melançon, 2010) consistently use determiners in Det+Noun context for categorizing novel nouns but not pronouns in Pron+Verb context for categorizing novel verbs might be due to the fact that determiner-noun sequences form determiner phrases or DPs, while pronoun-verb sequences are beyond a syntactically local phrase. This possibility is borne out by another infant study on syntactic categorization (Hicks, 2006), where English-learning 14- to 18-month-olds succeeded in using auxiliary-verb sequences like $\text{can } X$ and $\text{will } X$ to categorize $X$ as a verb.

To find out whether backward categorization is also possible with verbs in within-phrase co-occurrences, we conducted Experiment 2.

3. Experiment 2: Backward categorization of verbs

This time, we tested with the aspect marker $le$ that follows a verb in indicating perfectivity (e.g., Li & Shirai, 2000; Lin, 2000; Pan & Lee, 2004). (2) is an example sentence.

(2) tomasi kan-le xiyouji
Thomas read-Asp Journey to the West
‘Thomas (has) read Journey to the West.’

It is worth mentioning that the bigram probability of $XV-le$ in the corpus is as low as 0.477, compared to that of $XN-ye$ (0.957) in distribution. Meanwhile, $XN-le$ sequences constitute a potentially formidable barrier to the reliability of $Xv-le$, as the bigram probability of $XN-le$ is notable (0.18). Given that there are multiple functions of $le$, the choice of it should also allow us to probe whether phrasal co-
occurrences suffice for backward categorization in spite of considerable noise in distribution.

3.1. Methods
3.1.1. Participants

12 Mandarin-learning toddlers (6 boys, 6 girls) with a mean age of 19 months 22 days (range = 19 months to 20 months 9 days) were tested. Data from 9 other toddlers were excluded in the analysis due to fussiness (n = 2), no recovery of looking time during the post-test (n = 6), and ceiling (n = 1).

3.1.2. Stimuli, Design and Procedure

The materials used in Experiment 2 were the same as those in Experiment 1, except that the auditory stimuli for the familiarization phase were replaced with X-le-Y sequences, as shown in Table 2.

Table 2. Auditory stimuli used in Experiment 2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Audio stimuli</th>
</tr>
</thead>
</table>
| Familiarization (30 sec) | shai\textsubscript{V} le tongji  
|               | man\textsubscript{V} le tongji  
|               | man\textsubscript{V} le jianyan  
|               | shai\textsubscript{V} le jianyan  |
| Test (10 trials) | Trial A  
|                | shi zhege shai\textsubscript{N}  
|                | shi zhege man\textsubscript{N}  
|                | (ungrammatical)  
|                | Trial B  
|                | dou keyi shai\textsubscript{V}  
|                | dou keyi man\textsubscript{V}  
|                | (grammatical)  

Note that the new familiarization stimuli in Experiment 2 reversed the grammaticality status of the two test strings. That is, given that the novel words (X) were presented in verb contexts (i.e., X\textsubscript{V}-le-Y\textsubscript{N} ‘have X-ed Y’) during familiarization, shi-zhege-X\textsubscript{N} (‘it’s this-CL X’) trials requiring X to be a noun were now ungrammatical, while dou-keyi-X\textsubscript{V} (‘all may X’) trials appropriate only for X as a verb were grammatical. The reversed grammaticality of the test stimuli should allow us to examine whether toddlers’ looking preferences for a particular type of test trials resulted from their syntactic categorization of the target novel
words (X) during familiarization or perchance, from their inclination to hear the novel words in a particular type of contexts.⁴

The target words (X) and the medial function words in Experiment 2 were recorded with similar acoustic features compared to those in Experiment 1, as shown in Table 3.

### Table 3. Average acoustic values of stimuli used in Experiment 1 and 2

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>Mean (Exp.1)</th>
<th>Mean (Exp.2)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total utterance duration (sec)</td>
<td>1.267 (.049)</td>
<td>1.256 (.051)</td>
<td>.662</td>
</tr>
<tr>
<td>TW duration (sec)</td>
<td>.372 (.070)</td>
<td>.396 (.071)</td>
<td>.514</td>
</tr>
<tr>
<td>TW vowel duration (sec)</td>
<td>.171 (.015)</td>
<td>.171 (.027)</td>
<td>.964</td>
</tr>
<tr>
<td>TW vowel mean pitch (Hz)</td>
<td>181.672 (32.973)</td>
<td>204.050 (12.513)</td>
<td>.094</td>
</tr>
<tr>
<td>TW vowel mean intensity (dB)</td>
<td>84.680 (.588)</td>
<td>84.400 (.460)</td>
<td>.306</td>
</tr>
<tr>
<td>FW duration (sec)</td>
<td>.143 (.009)</td>
<td>.137 (.009)</td>
<td>.168</td>
</tr>
<tr>
<td>FW mean pitch (Hz)</td>
<td>182.502 (26.273)</td>
<td>184.713 (17.353)</td>
<td>.845</td>
</tr>
<tr>
<td>FW mean intensity (dB)</td>
<td>83.387 (1.206)</td>
<td>83.993 (.784)</td>
<td>.254</td>
</tr>
</tbody>
</table>

Note: TW = target word; FW = function word.

The design and the procedure were the same as those of Experiment 1.

### 3.2. Results and Discussion

Figure 2 plots the looking data of Experiment 2.

![Bar chart showing mean looking time for grammatical and ungrammatical trials](image)

**Figure 2. Mean looking time for grammatical and ungrammatical trials during the test phase in Experiment 2 (le condition)**

⁴ It is very unlikely that one group of toddlers might simply prefer to listen to noun contexts and the other to listen to verb contexts.
Toddlers showed a looking pattern similar to the one in Experiment 1. As figure 2 shows, their average looking time per trial during grammatical trials ($M = 10.852$ sec, $SE = .917$ sec) was again significantly longer than that during ungrammatical trials ($M = 9.199$ sec, $SE = .966$ sec), $t(11) = 2.503, p = .029$, Cohen’s $d = .722$, 2-tailed.

Specifically, unlike toddlers in Experiment 1 who preferred listening to noun contexts during the test phase (i.e., $shi$-zhege-$X_N$ ‘it’s this-CL X’), toddlers in Experiment 2 preferred listening to verb contexts during the test (i.e., $dou$-keyi-$X_V$ ‘all may X’). This crossed preference for the corresponding grammatical trials rules out the possibility that toddlers may simply prefer listening to a certain type of contexts during test trials regardless of their auditory input during familiarization.

This finding indicates that Mandarin-learning 19-month-olds can use the aspect marker $le$ to categorize its preceding verb. Their ability to categorize novel words backwardly applies to both novel nouns and novel verbs, provided that the novel word co-occurs with the function word within the same phrase. In the current case with $le$, the structure can be represented as [$AspP X le Asp ...$], where $X$ and the aspect marker $le$ co-occur within the same phrase $AspP$. Moreover, the considerably low predictive reliability of $X_V-le$ does not prevent them from recognizing the corresponding structure.

Though within-phrase co-occurrences have so far reliably characterized the distributional regularities that infants are readily able to identify both in forward and backward categorization, one may wonder whether by 19 months of age children have become bold enough to attempt to identify distributional regularities in either directions and beyond phrasal co-occurrences. That is, 14- to 16-month-olds who did not use pronoun-verb sequences to categorize novel verbs in previous studies (Höhle et al., 2004; Shi & Melançon, 2010) might still be conservative learners, given their younger age.

In view of children’s burgeoning perception of finer-grained structures around 18 months of age (e.g., Santelmann & Jusczyk, 1998; Shi, Emond, & Badri, 2020), we performed Experiment 3 as a follow-up to further investigate whether 19-month-olds are able to categorize backwardly novel words that are beyond the local phrase marked by the function word.

4. Experiment 3: Backward categorization beyond local phrases

To test backward categorization beyond within-phrase co-occurrences, we selected the negation marker $bu$ that is typically preceded by a noun, as illustrated by the example in (3).

(3) keluoyi bu xiyang
Chloe Neg smoke
‘Chloe does not smoke.’
For the distributional characteristics of the negation marker *bu* in the same corpus we have previously examined, the bigram probability of Xₙ-*bu* is 0.429, thus close to that of Xᵥ-*le* for the aspect marker *le* regarding predictive reliability.⁵ Also, as suggested by one study on the use of negation markers by Mandarin-speaking children, sequences as Xₙ-*bu*-Yᵥ featuring a noun left to *bu* have already started to emerge in the spontaneous speech of 17- to 19-month-olds (Fan, 2007).

Altogether, the abovementioned facts ensure toddlers’ acquaintance with the kind of structures used in the experiment and allow us to probe whether they are more liberal than 14- to 16-month-olds in using function words for categorizing adjacent novel words beyond syntactic phrase boundaries.

4.1. Methods
4.1.1. Participants

12 Mandarin-learning toddlers (6 boys, 6 girls) with a mean age of 19 months 13 days (range = 18 months 5 days to 20 months 26 days) were tested. Data from 3 other toddlers were excluded in the analysis due to fussiness (n = 1) and no recovery of looking time (n = 2).

4.1.2. Stimuli, Design and Procedure

The auditory stimuli during the familiarization phase were replaced with Xₙ-*bu*-Yᵥ sequences, while all the other materials (including novel words and their contexts in the test phase) remained the same as in the previous two experiments, as is shown in Table 4.

<table>
<thead>
<tr>
<th>Table 4. Auditory stimuli used in Experiment three</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
</tbody>
</table>
| Familiarization (30 sec) | shaiₙ bu tongji  
manₙ bu tongji  
manₙ bu jianyan  
shaiₙ bu jianyan |
| Test (10 trials) | Trial A  
shi zhege shaiₙ  
shi zhege manₙ  
(grammatical)  
Trial B  
dou keyi shaiᵥ  
dou keyi manᵥ  
(ungrammatical) |

⁵ The token frequency of *bu* is 125 (5465 total word tokens), which is equivalent to 22873 per million words.
As in Experiment 1, the noun contexts during the test phase (i.e., *shi-zhege*-XN ‘it’s this X’) were grammatical, whereas the verb contexts (i.e., *dou-keyi*-XV ‘all may X’) were ungrammatical. As the negation marker *bu* syntactically selects or combines with its following word (Y), the X-*bu*-Y sequences used in this experiment should allow us to probe whether toddlers are able to infer whole structures of the sequences from their parts (i.e., *bu*-Y), given the same length of familiarization.

Besides, we also controlled for the acoustic values of the auditory stimuli, as illustrated in Table 5.

Table 5. Average acoustic values of stimuli used in the three experiments (one-way ANOVA)

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>Mean (Exp.1)</th>
<th>Mean (Exp.2)</th>
<th>Mean (Exp.3)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total utterance duration (sec)</td>
<td>1.267 (.049)</td>
<td>1.256 (.051)</td>
<td>1.255 (.096)</td>
<td>.924</td>
</tr>
<tr>
<td>TW duration (sec)</td>
<td>.372 (.070)</td>
<td>.396 (.071)</td>
<td>.383 (.046)</td>
<td>.761</td>
</tr>
<tr>
<td>TW vowel duration (sec)</td>
<td>.171 (.015)</td>
<td>.171 (.027)</td>
<td>.165 (.015)</td>
<td>.754</td>
</tr>
<tr>
<td>TW vowel mean pitch (Hz)</td>
<td>181.672 (32.973)</td>
<td>204.050 (12.513)</td>
<td>181.202 (31.778)</td>
<td>.187</td>
</tr>
<tr>
<td>TW vowel mean intensity (dB)</td>
<td>84.680 (.588)</td>
<td>84.400 (.460)</td>
<td>84.966 (.893)</td>
<td>.264</td>
</tr>
<tr>
<td>FW duration (sec)</td>
<td>.143 (.009)</td>
<td>.137 (.009)</td>
<td>.137 (.010)</td>
<td>.273</td>
</tr>
<tr>
<td>FW mean pitch (Hz)</td>
<td>182.502 (26.273)</td>
<td>184.713 (17.353)</td>
<td>180.47 (16.042)</td>
<td>.917</td>
</tr>
<tr>
<td>FW mean intensity (dB)</td>
<td>83.387 (1.206)</td>
<td>83.993 (.784)</td>
<td>84.615 (1.806)</td>
<td>.207</td>
</tr>
</tbody>
</table>

Note: TW = target word; FW = function word.

The design and the procedure were the same as those of Experiment 1.

4.2. Results and Discussion

Toddlers’ looking data in Experiment 3 were analyzed in a paired t-test. There was no significant difference in average looking time per trial for grammatical trials ($M = 10.461$ sec, $SE = 1.152$ sec) versus ungrammatical trials ($M = 11.415$ sec, $SE = 1.241$ sec), $t(11) = -.710$, $p = .493$, Cohen’s $d = -.205$, 2-tailed. The results are plotted in Figure 3.
According to our results, 19-month-olds are not able to use the negation marker *bu* to categorize the preceding novel word as a noun. This is predicted by categorization biased towards within-phrase co-occurrences, given that the negation phrase (NegP) headed by *bu* marks a phrase boundary that excludes the preceding novel word X from the following phrase, as in X [NegP *bu* Neg …].

This, together with findings from the previous two experiments, indicates that backward categorization is only possible when the novel word and the function word co-occur within a phrasal unit, even for 19-month-olds.

5. General Discussion

The three experiments reported above demonstrate that Mandarin-learning 19-month-olds are able to use function words alone to categorize novel words (X) backwardly in three-word sequences of X-functor-Y, where both the target words (X) and the filler words (Y) were unfamiliar to toddlers.

Importantly, our findings testify to young children’s capability of processing and recognizing backward-linking functional elements (e.g., suffixes and particles) in utterance-medial positions that denote the syntactic categories of words in the preceding contexts. Their success in using the focus particle *ye* for categorizing preceding novel nouns and the aspect marker *le* for categorizing preceding novel verbs provides robust evidence for their ability to recognize local structures featuring target-functor sequences.

One important caveat is that the feasibility of a syntactic categorization task using a function word is not necessarily associated with the syntactic category of the target word (i.e., noun vs. verb); it depends on the syntactic properties of the function word, namely whether the phrase delimited by the function word incorporates the target word. This is substantiated by 19-month-olds’ failure to use the negation marker *bu* for categorizing preceding novel nouns that are exterior to the negation phase. Presumably, this structural constraint on phrasal
co-occurrences also applies to findings on forward syntactic categorization, where 14- to 16-month-olds do not show signs of categorizing novel words following subject pronouns as verbs (Höhle et al., 2004; Shi & Melançon, 2010).

Our results further indicate that distributional reliability does not fully explain toddlers’ syntactic categorization. Though toddlers’ success in the ye condition might be attributed to the high bigram probability of XN-ye, it is not likely that this predictive reliability account applies to toddlers who categorize words preceding the aspect marker le as verbs since there is considerable noise in distribution with XV-le sequences in child-directed speech.

In brief, this study provides evidence for toddlers’ syntactic categorization of novel words preceding function words. We suggest that their tendency to categorize words forwardly or backwardly within the phrasal domain marked by the function word reflects a linguistic intuition available to them early in life that may assist them in assigning syntactic categories to words conservatively but tolerantly.

Appendix

The term “bigram probability” refers to the probability of a word X identified with the syntactic category k given its preceding word or its following word a. In the case of backward categorization, the bigram probability P(Xk|a) is the number of occurrence of the bigram Xk-a divided by the number of occurrence of a.

\[
P(X_k|a) = \frac{C(X_k a)}{C(a)}
\]

For instance, the bigram probability of XV-le, namely P(XV|le), is calculated using C(XV,le)/C(le).

References


