Hierarchical Structure Dependence in Infants at the Early Stage of Syntactic Acquisition

Rushen Shi, Emeryse Emond, and Sophia Badri

1. Introduction

Words in natural language are organized as abstract categories, which can be combined to form larger constituents such as phrases and sentences. Grammatical rules and principles determine how different levels of combinations are structured and related to one another. A commonly accepted view in most syntactic theories is that linguistic combinations are hierarchically structured, even though surface strings of speech appear as linear sequences. An issue which is of central interest to researchers concerns how the abstract knowledge of hierarchical structures become available to children. Within a few years of life after birth children acquire linguistic rules of their ambient language and demonstrate productivity and creativity in their comprehension and production. This is amazing since they only hear linear output of abstract structures and receive no explicit grammar lessons from their caregivers. Whether young children represent phrase structures hierarchically and how such knowledge is acquired are basic questions in the field of language acquisition.

One important aspect of syntactic knowledge is dependency. Grammatical categories and constituents are often dependent on one another, and the dependencies operate hierarchically. Subject-verb agreement is a well-known kind of example. The sentence *[*The girl sing*] in English is ungrammatical because there is a violation of subject-verb agreement in number. The rule cannot be a linear relation between the first noun and the first verb, as it will incorrectly render a grammatical sentence such as [*The girl who we like sings*] ungrammatical. The dependency operates between the higher subject phrase constituent (e.g., [*the girl who we like*]) and the verb (e.g., [*sings*]). The child must go beyond the linear input and acquire syntactic rules hierarchically.

The existing literature suggests that young children and infants can track adjacent and non-adjacent grammatical dependencies, although most studies were not designed to test linear versus hierarchical structures. It has been shown that infants shortly after their first birthday understand that a noun can co-occur with a determiner (Höhle, et al., 2004; Shi & Melançon, 2010), and that

* Rushen Shi, Emeryse Emond and Sophia Badri, Université du Québec à Montréal, shi.rushen@uqam.ca; emond.emeryse@courrier.uqam.ca; badrisophia@gmail.com. This research was supported by NSERC and CFI grants.

 Proceedings of the 44th Boston University Conference on Language Development, ed. Megan M. Brown and Alexandra Kohut, 572-585. Somerville, MA: Cascadilla Press.

subsequently, infants who learn French (a grammatical gender language) begin to encode gender feature agreement between the noun and the determiner in NPs (i.e., noun phrases) (Cyr & Shi, 2013; Van Heugten & Christophe, 2015). In studies that presented grammatical and ungrammatical dependences, English-learning infants perceived the ungrammaticality in sentences such as *[*Everyone can baking ...*] (Santelmann & Jusczyk, 1998), and German-learning infants can detect grammatical violations in *[*... kann ... geholt ...*] ('… can … fetched …') (Höhle, Schmitz, Santelmann, & Weissenborn, 2006). Infants are sensitive to number feature agreement, noticing agreement violations such as *[*Mommy book a reads*] in English (Soderstrom, White, Conwell, & Morgan, 2007), and *[*Les*$_{PL}$ Noun *il*$_{SG}$ ...] ('The$_{PL}$ Noun he$_{SG}$ …') and *[*Les*$_{PL}$ Noun *va*$_{SG}$ ...] ('The$_{PL}$ Noun will$_{SG}$ …') in French (Culbertson, Koulaguina, Gonzalez-Gomez, Legendre, & Nazzi, 2016; Van Heugten & Shi, 2010). These results, however, are consistent with both linear and hierarchical interpretations. A recent study tested subject-verb number agreement in French sentences containing conjoint subjects (i.e., NP+NP) versus those containing a simple NP subject (Koulaguina, Legendre, Barrière, & Nazzi, 2019), with the latter always following a separate, disjoint NP utterance. They found that 24-month-olds processed the two distinct kinds of agreement, showing that they used prosody to determine the beginning of a sentence and the relevant NP(s) for agreement tracking.

Grammatical gender feature is present in certain languages. The gender is specified as an inherent feature of a noun (e.g., masculine, feminine, or neutral). French has masculine and feminine features, and gender assignment to most nouns is semantically arbitrary, with no reliable phonological cues. Most nouns are specified with one gender, and must be learned item by item during word learning. Some other gender-agreeing categories (such as determiners, adjectives, pronouns, quantifiers) can have both genders for the same word, e.g., *le*$_M$/*la*$_F$ 'the'; *petit*$_M$/*petite*$_F$ 'small'. Gender agreement in phrases and sentences is determined by the noun gender and is structure-sensitive. The noun gender selects the determiner of the matching gender within the same NP (e.g., [*le*$_M$ *kiwi*$_M$] 'the kiwi'; [*la*$_F$ *pomme*$_F$] 'the apple'). French-learning toddlers begin to learn the gender features of specific nouns from mid second year of life (Cyr & Shi, 2013; Van Heugten & Christophe, 2015). Toddlers also track the gender agreement between a dislocated subject phrase and the repeated subject pronoun clitic (e.g. [*Le*$_M$ Noun$_M$ *il*$_M$ …] 'The Noun, it/he …') (Melançon & Shi, 2015).

These findings suggest that infants below age two possess impressive abilities to tracking dependencies and feature agreement. However, in nearly all existing studies (except Koulaguina, Legendre, Barrière, & Nazzi, 2019) the testing stimuli were not set up to assess linear versus hierarchical representations.

We recently set out to specifically test linear versus hierarchical representation in young children (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press). Using French subject-dislocation constructions containing distinct subject phrases (see Section 2 for details), we found that 30-month-old French-learning toddlers processed grammatical gender agreement by

activating hierarchical phrase structures, beyond linear sequences of gender-bearing words in our sentences. The question arising from that finding was whether the knowledge of hierarchical structures was present at the onset of agreement learning, or gradually established over time. The former would be consistent with the continuity hypothesis in acquisition research, which considers the child's internal system as crucial. The latter is compatible with the idea of input-based inductive learning. To address this question, we decided to test infants at the earliest stage of gender agreement acquisition. We chose the age of 17 to 18 months, the age when the earliest learning of the gender of some familiar nouns was observed (Van Heugten & Christophe, 2015).

The methods of the present experiment, as described in Section 2, are identical to those for the 30-month-olds in our prior study (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press). We hypothesized that if hierarchical knowledge is continuous, then 17- to 18-month-olds in the present study should succeed in the task, as did the 30-month-olds in our previous study.

2. The present study

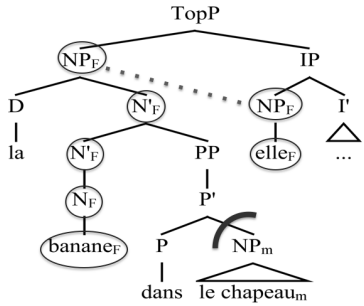
2.1. Two hierarchically distinct structures

In our study we consider feature agreement in two hierarchically distinct syntactic structures in French, (1) and (2). The feature agreement patterns of the two structures offer an ideal opportunity for examining the nature of children's initial representation of these structures. Both structures are subject-dislocation sentences, differing only in the internal organization of the dislocated subject phrase. The subject pronoun clitic of the main verb must agree with the dislocated subject phrase in features (e.g., number, grammatical gender). The gender marking of the subject phrase depends on its particular internal structure.

(1) a. **Structure 1:** The subject phrase of the sentence contains an NP (noun phrase) with a PP (prepositional phrase) modifier, e.g.,
(1) b. La banane$_F$ dans le chapeau$_M$, elle$_F$ demeure au fleuve.
(the banana in the hat, it stays in the river)

In (1b) the subject phrase of Structure 1 contains two nouns each with its inherent gender feature. In general, the noun gender determines the gender of the upper nodes that dominate the noun, through feature percolation. For (1b), the feminine feature of *banane*$_F$ is percolated all the way up the structural tree to the topmost NP, as shown in the tree representation (1c). The feature of *chapeau*$_M$ is also percolated up to the immediately dominating NP [*le chapeau*]$_M$, but feature percolation is blocked at the next level up. That is, the gender of a noun below a PP cannot get through to the higher phrasal nodes. The feminine feature of the subject phrase as a whole (i.e., the topmost NP$_F$: [*La banane dans le chapeau*]$_F$) is thus inherited from the noun *banane*$_F$, and it is this phrasal feature NP$_F$ with which the following subject pronoun clitic must agree (*elle*$_F$).

(1) c.



Feature percolation operates differently for Structure 2, with both nouns in the subject phrase relevant for determining the feature of the topmost ConjP node. In (2b)-(2d), the subject phrase each contains two nouns with their inherent genders. For (2b), the nouns *banane*$_F$ and *pomme*$_F$, both feminine, pass the feature all the way up the structural tree to the topmost ConjP$_F$ (i.e., [*La banane et la pomme*]$_F$), as shown in (2e). Likewise, in (2c) the nouns *camion*$_M$ and *chapeau*$_M$, both masculine, pass the feature all the way up to the topmost ConjP$_M$ (i.e., [*Le camion et le chapeau*]$_M$). The following subject pronoun clitic agrees with the feature of the topmost ConjP, i.e., *elles*$_F$ in (2b) and *ils*$_F$ in (2c).

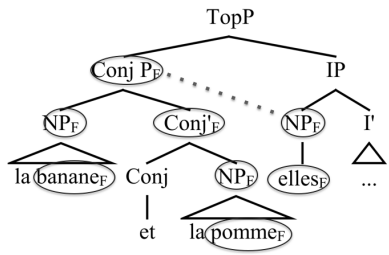
(2) a. **Structure 2:** The subject phrase of the sentence contains two conjoined NPs, e.g.

(2) b. La banane$_F$ et la pomme$_F$, elles$_F$ demeurent au fleuve.
(the banana and the apple, they stay in the river)

(2) c. Le camion$_M$ et le chapeau$_M$, ils$_M$ demeurent au fleuve.
(the truck and the hat, they stay in the river)

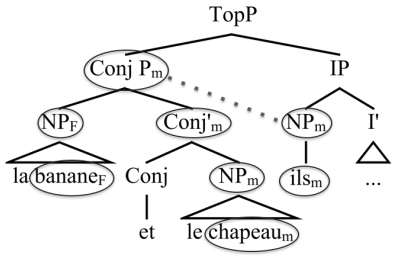
(2) d. La banane$_F$ et le chapeau$_M$, ils$_M$ demeurent au fleuve.
(the banana and the hat, they stay in the river)

(2) e.



In (2d) the two nouns *banane*$_F$ and *chapeau*$_M$ also pass their features up the nodes of the structural tree. However, since the two nouns have different genders, the topmost ConjP resorts to the default masculine gender (i.e., ConjP$_M$: [*La banane et le chapeau*]$_M$), as shown in (2f). The subject pronoun clitic must agree with this default gender (i.e., *ils*$_M$).

(2) f.



Linearly, the dislocated subjects of Structures 1 and Structure 2 contain two gender-bearing nouns, and their sequential positions are identical in the two structures, i.e., [Det Noun Prep/Conj Det Noun]. However, as explained above, the two nouns function differently for feature percolation and agreement due to their distinct hierarchical organizations.

We hence asked whether children during early language acquisition represent and process such strings and the features hierarchically or linearly. In our recent study (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press) we tested French-learning 30-month-olds on gender agreement in these two structures. The toddlers showed evidence in support of hierarchical knowledge. In this study we further asked if the hierarchical phrase structure knowledge shown in 30-month-olds in our prior study is present at the earliest stage of grammatical gender learning. An early presence of this knowledge would be consistent with the view of continuity, i.e., hierarchical syntactic representation possibly available in the child's internal system without needing to be built gradually and entirely from the input (Chomsky, 1965, 1988). As discussed above, grammatical gender agreement depends on the noun gender, but the gender of nouns only comes with the learning of individual nouns, and this would be the case especially in French, which has no clear phonological markings of gender on nouns. The child needs to first learn the gender of some nouns in order to process gender agreement in structures such as (1) and (2). Given that French-learning infants begin to learn the genders of nouns from around 18 months of age (Cyr & Shi, 2013; Van Heugten & Christophe, 2015), we decided to test French-learning infants of this age on the same experiment in which the 30-month-olds in our previous study had shown gender agreement.

2.2. Methods

2.2.1. Participants

Participants were monolingual French-learning 17- to 18-month-old infants (n=24; mean: 543 days; range: 528-559 days).

2.2.2. Stimuli

Using the above two structures of the subject-dislocation construction, we created sentences with correct gender agreement and sentences with agreement violations. The correct and incorrect sentences within each structure differed only in the subject pronoun clitic. For example, one of the six correct sentences in Structure 1 was (1b) *La banane*$_{F}$ *dans le chapeau*$_{M}$, *elle*$_{F}$ *demeure au fleuve*, and the corresponding incorrect sentence was **La banane*$_{F}$ *dans le chapeau*$_{M}$, *il*$_{F}$ *demeure au fleuve*. All six sentences shared the same initial dislocated subject phrase (i.e., *La banane*$_{F}$ *dans le chapeau*$_{M}$ 'the banana in the hat'). The words after the subject pronoun clitic (i.e., *elle*$_{F}$, *il*$_{F}$) varied across sentences.

In Structure-2 sentences, the words following the subject pronoun clitic were identical to those in Structure-1 sentences. As an example, one of the six correct sentences in this structure was (2d) *La banane*$_{F}$ *et le chapeau*$_{M}$, *ils*$_{F}$ *demeurent au fleuve*, and the corresponding incorrect sentence was **La banane*$_{F}$ *et le chapeau*$_{M}$, *elles*$_{F}$ *demeurent au fleuve*. Note that the singular and plural versions of the subject pronouns in our sentences are phonetically identical, with the *-s* ending being mute. The singular and plural forms of the main verbs in our sentences are also phonetically identical (for example, *demeure* and *demeurent* are pronounced the same) despite spelling differences. The dislocated subject phrase (i.e., *La banane*$_{F}$ *et le chapeau*$_{M}$ 'the banana and the hat'), same across all six Structure-2 sentences, contained a conjunction *et*. The full set of correct and incorrect sentences of both structures are listed in the Appendix.

For sound stimuli, we needed to avoid recording ungrammatical sentences since there could be disfluency artifacts that might affect infants' responses in our perceptual experiment. We thus recorded other grammatical sentences, and then used the cross-splicing method to assemble all our experimental stimuli, for both the correct and incorrect sentences. Using a Sound Device 702T, we recorded (at 44 kHz sampling frequency and 24 bits bitrate) the speech of a native speaker of French in an acoustic chamber. She produced multiple exemplars of sentences (3) and (4), each containing a left-dislocated phrase. The left-dislocated phrases of (3) and (4) consisted of identical words as in the left-dislocated phrases of Structures 1 (e.g., 1b) and Structure 2 (e.g., 2d), respectively. The dislocated phrase here in (3) and (4), however, corresponded to the object pronoun clitic of the main clause. The object clitic and the dislocated phrase must agree in gender (for (3)) and in number (for (3) and (4)); note that the plural form of the object pronoun *les* is not marked for gender and can agree with any gender. Our speaker produced a natural brief prosodic break at the end of the left-dislocated phrase. The following three words *on va toujours* were the same for (3) and (4). We were thus able to slice out the dislocated phrases from the same phonetic context of (3) and (4) and from all grammatical sentences. The cut was made from the sentence onset to the offset of the word *chapeau*. Multiple exemplars of the dislocated phrases in varying intonations were sliced out from (3) and (4). These sliced phrases were

subsequently put together with the second part of (5)-(8), to create the stimulus sentences for our test trials.

(3) [*La banane*$_{F}$ *dans le chapeau*$_{M}$]$_{F\text{-}SG}$, *on va toujours la*$_{F\text{-}SG}$ *voir*.
(the banana in the hat, we will always see it)
(4) [*La banane*$_{F}$ *et le chapeau*$_{M}$]$_{M\text{-}PL}$, *on va toujours les*$_{PL}$ *voir*.
(the banana and the hat, we will always see them)

To obtain the second part of our test sentences, we asked our speaker to produce six sentences of the set (5), which contained a feminine subject pronoun *elles*$_{F}$. These six sentences differed only in the words after the subject pronoun (i.e., the part 'VERB …'). The speaker also produced six sentences of the set (6) containing a masculine subject pronoun *ils*$_{M}$. Likewise, we recorded her production of six sentences of the set (7), and six sentences of the set (8). All of these sentences (6 x 2 sets for Structure 1, and 6 x 2 set for Structure 2) were produced with a natural brief prosodic break after the left-dislocated phrase.

(5) Structure 1: [*Les tisanes*$_{F}$ *dans les autos*$_{F}$]$_{F}$ [*elles*$_{F}$ VERB …]
(The tisanes in the cars, they VERB …)
(6) Structure 1: [*Les domaines*$_{M}$ *dans les autos*$_{F}$]$_{M}$ [*ils*$_{M}$ VERB …]
(The domains in the cars, they VERB …)
(7) Structure 2: [*Les tisanes*$_{F}$ *et les autos*$_{F}$]$_{F}$ [*elles*$_{F}$ VERB …]
(The tisanes and the cars, they VERB …)
(8) Structure 2: [*Les domaines*$_{M}$ *et les autos*$_{F}$]$_{M}$ [*ils*$_{M}$ VERB …]
(The domains and the cars, they VERB …)

Like (3) and (4), the sentences in (5)-(8) were all grammatical. The second part of each sentence in (5)-(8) was sliced out, starting from the beginning of the prosodic break (i.e., the end of *autos*) to the end of the sentence. The second part of each sentence in (5) and (6) was then connected with a sliced first part of (3), and the second part of each sentence in (7) and (8) was connected with a sliced first part of (4). This cross-splicing yielded the correct and incorrect stimulus sentences for our experiment. The final cross-spliced test sentences corresponded to those that we had conceptualized for our stimuli. Specifically, the sentences were organized into four separate audio files: Structure-1 correct (e.g., [*La banane dans le chapeau*]$_{F}$, *elle*$_{F}$ *demeure au fleuve*), Structure-1 incorrect (e.g., *[*La banane dans le chapeau*]$_{F}$, *il*$_{M}$ *demeure au fleuve*), Structure-2 correct (e.g., [*La banane et le chapeau*]$_{M}$, *ils*$_{M}$ *demeurent au fleuve*), and Structure-2 incorrect (e.g., *[*La banane et le chapeau*]$_{M}$, *elles*$_{F}$ *demeurent au fleuve*).

Our speaker also produced the two dislocated phrases as independent utterances (i.e., *La banane dans le chapeau* and *La banana et le chapeau*). Four exemplars of each utterance were chosen, and were used as the familiarization stimuli. This familiarization was intended to make the task easier for the infants.

We then created animation files of a cartoon character speaking the speech stimuli, using the Adobe Flash CS3 Professional software and exported in a QuickTime format with a 3 fps frame rate and a resolution of 1360 x 768 pixels. The speech files were synchronized with the mouth and body motions of the cartoon character. An animation consisting of colorful zooming bubbles and water bubble sound served as the stimuli for pre-trial and post-trial. The attention-getter stimuli were a moving star together with bird songs.

2.2.3. Design and predictions

Infants were randomly divided into two groups. Each group heard two types of test trials: Correct agreement in one structure, and incorrect agreement in the other structure. For the first group, Structure-1 sentences were correct in gender agreement, whereas Structure-2 sentences were incorrect:

(9) Correct trials (Structure 1), e.g.,
[*La banane*$_{F}$ *dans le chapeau*$_{M}$]$_{F}$ [*elle*$_{F}$ VERB …]
(The banana in the hat, it VERB …)
(10) Incorrect trials (Structure 2), e.g.,
*[*La banane*$_{F}$ *et le chapeau*$_{M}$]$_{M}$ [*elles*$_{F}$ VERB …]
(The banana and the hat, they VERB …)

In (9) and (10), the gender-bearing words in the two trial types were identical and had the same linear order. The only difference between the two types of sentences was the third word (*dans* versus *et*), which contained no gender feature but were crucial for the hierarchical structural differences.

The second group of infants heard the two types of test trials in the reverse arrangement: Correct agreement for Structure-2 sentences, and incorrect agreement for Structure-1 sentences. Again, the gender-bearing words in the two trial types, i.e., (11) versus (12), were identical and had the same linear order, but they were distinct in their hierarchical structures.

(11) Correct (Structure 2), e.g.,
[*La banane*$_{F}$ *et le chapeau*$_{M}$]$_{M}$ [*ils*$_{M}$ VERB …]
(The banana and the hat, they VERB …)
(12) Incorrect (Structure 1), e.g.,
*[*La banane*$_{F}$ *dans le chapeau*$_{M}$]$_{F}$ [*il*$_{M}$ VERB …]
(The banana in the hat, it VERB …)

This way of arranging the correct and incorrect agreement and the different structures across the two groups of infants enabled us to test competing predictions. If infants simply tracked the linear order of gender-bearing words, they should not discriminate the two types of trials; that is, their listening times to the two trial types should not differ (note that listening times were measured by looking times in our task, as described in Section 2.2.4). However, if infants

activated the distinct structures and processed the corresponding feature percolation/agreement patterns, then both groups should discriminate the two trial types (i.e., yielding different listening times to the two trial types) and show longer listening time towards the same trial type (i.e., both groups listening longer to correct trials, or both groups listening longer to incorrect trials).

2.2.4. Procedure

Infants were tested individually in a preferential looking procedure. In an acoustic chamber the child sat on the parent's lap about 1.5 meters in front of a 42-inch TV monitor. Loud speakers adjacent to both sides of the monitor played auditory stimuli. The parent heard masking music through headphones. A hidden camera filmed the baby and sent the video simultaneously to a monitor in the adjacent room, where a researcher observed the infant while running the experiment. The researcher could not hear or see any stimuli, and did not know the purpose of the study. The experiment was run by a computer program, which sent the auditory stimuli to the loud speakers and displayed the visual stimuli on the monitor, as well as recording simultaneously the baby's look to the visual display during each trial. Each trial was initiated by the child, i.e., a trial would start if the child looked at the display monitor. Speech and visual stimuli were presented as long as a trial was ongoing. A trial would end if the child looked away from the display monitor for at least 2 s or if the maximum trial length was reached. Between trials the attention-getter (moving stars with bird songs) were presented automatically to attract the baby back to the display monitor.

The task began with a pre-trial, presenting zooming bubbles and water bubble sound. This trial served the function of acquainting the baby with the audiovisual equipment. Then, two familiarization trials (each 11.1 s long) followed. In one trial the puppet said the exemplars of the utterance *La banane dans le chapeau* ('the banana in the hat'), and in the other trial he said the exemplars of the utterance *La banane et le chapeau* ('the banana and the hat'). We set up these familiarization trials to make the task easier, as our test sentences were quite complex, including two structures and many infrequent words in multiple sentences. The ISI between any two utterances within the familiarization trials was 879 ms. Half of the infants heard the *dans* trial first, and the other half heard the *et* trial first.

The test phase began automatically after the familiarization, with the puppet appearing again on the display monitor and speaking in each trial. For each group of infants, correct and incorrect test trials occurred alternatingly for a total of 12 trials (6 for each type). The first test trial presented *dans* sentences (i.e., Structure 1) for infants who heard the *dans* phrases as the first familiarization trial. The other half of the infants, who heard the *et* phrases as the first familiarization trial, also heard the *et* sentences (Structure 2) as the first test trial. This manipulation also covered the counterbalancing of the order of the correct versus incorrect trial as the first test trial. A child would hear all six sentences in a trial if he or she looked to the full trial length (29.3 s). The ISI between any

two sentences within a test trial was 1.2 s. Following the test phase, the post-trial was presented, indicating the end of the task.

2.3. Results

Looking times (i.e., listening times) during test trials were compiled. For every child, we calculated the average looking time of correct trials and that of incorrect trials, excluding the initial trial of each of the two types, as is generally done in this procedure. Given our design, a 2x2 mixed ANOVA was conducted, with Test Trial Type (correct-agreement vs. incorrect-agreement) as the within-subject factor, and Group (Group 1 vs. 2) as the between-subject factor.

Recall that for both groups of infants, the linear order of all gender-bearing words for correct and incorrect trials were identical. Thus, if infants only tracked these words linearly, they should show no main effect of Test Trial Type and no interaction between Test Trial Type and Group. However, if they recognized the grammaticality of the feature agreement by accessing the distinct hierarchical structures in the two trial types, they should show a main effect of Test Trial Type and no interaction of the two factors; that is, we should observe an overall looking/listening time difference and a uniform preference towards the same type of test trials for both groups of infants.

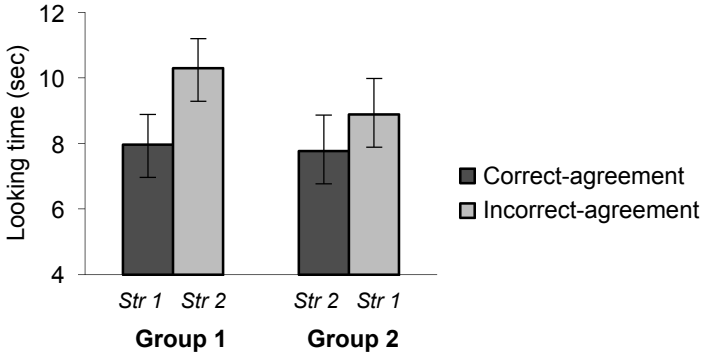


Figure 1. 17-month-olds' looking times (Means & SEMs) to correct vs. incorrect-agreement trials. The agreement status and the structures were yoked across two groups of infants. Both groups looked longer to incorrect-agreement trials.

The results confirmed the latter prediction. As shown in Fig. 1, both groups looked longer while listening to incorrect-agreement trials. The main effect of Test Trial Type was significant, $F(1,22)=5.824$, $p=.025$, with no interaction between Test Trial Type and Group, $F(1,22)=.719$, $p=.406$. The average difference in looking time between correct and incorrect trials across the two groups was 1.716 s; using the pooled within-group variance estimate, the 95% confidence interval went from 0.331 s to 3.101 s, and the *Cohen d* effect size

was 0.340. The main effect of Group was non-significant, $F(1,22)=.166$, $p=.688$. These results indicate that infants aged 17-18 months discriminated the two test trial types, and the two infant groups did not differ in their looking responses. These findings resemble those of our previous study (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press), in which 30-month-olds (n=24) yielded the same pattern of results in the same ANOVA (i.e., a significant main effect of Test Trial Type, but no interaction with Group) (see Fig. 2).

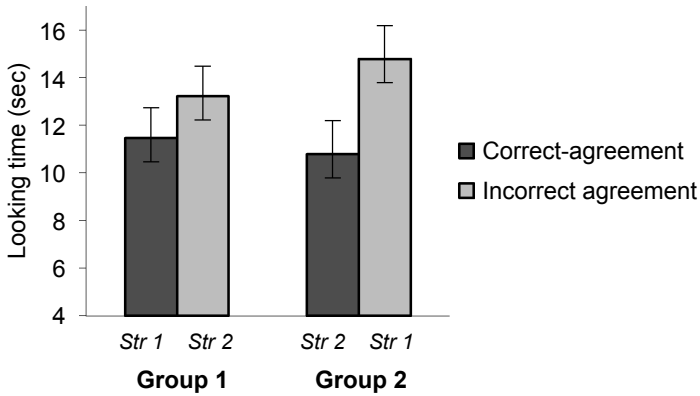


Figure 2. 30-month-olds' looking times (Means & SEMs) to correct vs. incorrect-agreement trials. The agreement status and the structures were yoked across two groups of infants. Both groups looked longer to incorrect-agreement trials.

We conducted further statistical analyses to address the question of continuity. Since the 17- to 18-month-olds in the present study and the 30-month-olds in our previous study (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press) were tested on the same experiment, we were able to directly compare the responses of the two ages in a 2x2x2 mixed ANOVA. The within-subject factor was Test Trial Type (correct-agreement vs. incorrect-agreement), and the between-subject factors were Group (Group 1 vs. 2) and Age (17-18 months vs. 30 months). We found a highly significant main effect of Test Trial Type ($F(1,44)=14.983$, $p=.000$), with longer looking to incorrect-agreement trials than correct-agreement trials. The average difference in looking time between correct and incorrect trials across the two groups was 2.301 s; using the pooled within-group variance estimate, the 95% confidence interval went from 1.133 s to 3.468 s, and the *Cohen d* effect size was 0.431. The main effect of Group was non-significant ($F(1,44)=.019$, $p=.891$). The main effect of Age was significant, $F(1,44)=8.124$, $p=.007$, due to an overall longer looking time in the older infants than in the younger ones. However, none of the interactions with Age was significant (Test Trial Type x Age: $F(1,44)=0.967$, $p=.331$; Group x Age: $F(1,44)=0.211$, $p=.648$; Test Trial Type x Group x Age: $F(1,44)=2.089$, $p=.155$); that is, the younger and older infants were comparable

in their responses. Importantly, Test Trial Type x Group interaction was non-significant, $F(1,44)=.186$, $p=.669$, indicating that Group 1 and Group 2 both discriminated the two test trial types, and both listened longer to incorrect trials.

3. Discussion

In the present study we tested 17- to 18-month-olds on the same experiment that 30-month-olds did in our prior study (Legrand & Shi, 2018; Shi, Legrand, & Brandenberger, in press). The results of the younger infants were the same as those of the older ones. Both ages discriminated the correct- and the incorrect-agreement sentences, and both looked longer while listening to the incorrect trials (that is, a novelty preference). In particular, across both ages, infants who heard correct Structure-1 sentences and incorrect Structure-2 sentences listened longer to Structure-2 sentences, while infants who heard correct Structure-2 sentences and incorrect Structure-1 sentences listened longer to Structure-1 sentences. Such systematic looking preference is extremely hard to be obtained by chance. Thus, the evidence is strong for structure-dependent agreement processing in infants. The highly significant novelty preference in both ages is a striking demonstration that there is continuity in learners' syntactic knowledge: Hierarchical representations are present from the onset of syntactic learning and guide children's feature agreement processing.

The 17-18-month-olds' performance is especially impressive. Grammatical gender is a language-specific featural property that requires time to acquire. Moreover, gender agreement depends on noun gender, but the genders of nouns must be learned one by one as the child builds a vocabulary. Our study shows that at an age when infants are just beginning to learn the gender of some nouns (Van Heugten & Christophe, 2015), they immediately track gender agreement and do so in a structure-dependent fashion. The amazing timing of the emergence of these two aspects of acquisition suggests that hierarchical phrase structures must already be available to toddlers, and that the child is likely constrained by the principle of structure dependence of the universal grammar, consistent with the view in Chomskyan theory (Chomsky, 1965,1988).

However, we probed further by asking whether the knowledge shown by our 17-18-month-olds might be the result of efficient inductive learning based on rich input. Rich input can be characterized by frequent occurrences of the two structures and the gender feature agreement used in our studies. To address this question, we searched the literature and also conducted our own analysis of input speech. Our analysis revealed that the two structures were rare in French children's input (see details in Shi, Legrand, & Brandenberger, in press). Furthermore, Koulaguina, et al. (2019) analyzed the frequency of occurrences of left-dislocated conjoined subjects (i.e., which is Structure 2 of our study) in French input. They also found the structure to be rare. These input analyses suggest that it is unlikely infants can rely on input-based inductive learning to create the representations of the two structures and track the gender agreement patterns. It is plausible that there exist built-in hierarchical phrase structure

representations that become available and guide infants' tracking of gender feature agreement, as soon as they start learning the gender of some nouns.

To conclude, our results suggest that infants represent phrase structures hierarchically from the onset of syntactic acquisition, and that they rely on phrase structures when tracking feature agreement. Our findings thus provide evidence in support of the principle of structure dependence in the universal grammar (Chomsky, 1965, 1988).

Appendix

Test Sentences and Design

Group 1	
Correct agreement (Structure 1) [*La banane*$_{F}$ *dans le chapeau*$_{M}$]$_{\underline{\mathbf{F}}}$ *elle*$_{\underline{\mathbf{F}}}$ ... (The banana in the hat, it ...)	Incorrect agreement (Structure 2) *[*La banane*$_{F}$ *et le chapeau*$_{M}$]$_{\underline{\mathbf{M}}}$ *elles*$_{\underline{\mathbf{F}}}$ ... (The banana and the hat, they ...)
La banane dans le chapeau, elle pèse les graminées.	*La banane et le chapeau, elles pèsent les graminées.*
La banane dans le chapeau, elle demeure au fleuve.	*La banane et le chapeau, elles demeurent au fleuve.*
La banane dans le chapeau, elle grasseye ardemment.	*La banane et le chapeau, elles grasseyent ardemment.*
La banane dans le chapeau, elle fuie les tentatives.	*La banane et le chapeau, elles fuient les tentatives.*
La banane dans le chapeau, elle tente leur pécule.	*La banane et le chapeau, elles tentent leur pécule.*
La banane dans le chapeau, elle résonne au couvent.	*La banane et le chapeau, elles résonnent au couvent.*
Group 2	
Correct agreement (Structure 2) [*La banane*$_{F}$ *et le chapeau*$_{M}$]$_{\underline{\mathbf{M}}}$ *ils*$_{\underline{\mathbf{M}}}$… (The banana and the hat, they ...)	Incorrect agreement (Structure 1) *[*La banane*$_{F}$ *dans le chapeau*$_{M}$]$_{\underline{\mathbf{F}}}$ *il*$_{\underline{\mathbf{M}}}$ ... (The banana in the hat, it ...)
La banane et le chapeau, ils pèsent les graminées.	*La banane dans le chapeau, il pèse les graminées.*
La banane et le chapeau, ils demeurent au fleuve.	*La banane dans le chapeau, il demeure au fleuve.*
La banane et le chapeau, ils grasseyent ardemment.	*La banane dans le chapeau, il grasseye ardemment.*
La banane et le chapeau, ils fuient les tentatives.	*La banane dans le chapeau, il fuie les tentatives.*
La banane et le chapeau, ils tentent leur pécule.	*La banane dans le chapeau, il tente leur pécule.*
La banane et le chapeau, ils résonnent au couvent.	*La banane dans le chapeau, il résonne au couvent.*

Note: The strings in contrasting structures for each group only differ in Word 3 (*dans* vs. *et*), with the remaining words being homophonous despite orthographic differences. For example, the subject clitics *il* and *ils* are homophonous, so are the verbs *pèse* and *pèsent.* The content words (verbs, nouns, adverbs) after the subject clitic are all infrequent words.

References

Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. (1988). *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

Culbertson, Jennifer, Koulaguina, Elena, Gonzalez-Gomez, Nayeli, Legendre, Géraldine, & Nazzi, Thierry. (2016). Developing knowledge of non-adjacent dependencies. *Developmental Psychology, 52*(12), 2174-2183.

Cyr, Marilyn & Shi, Rushen. (2013). Development of abstract grammatical categorization in infants. *Child Development, 84*(2), 617-629.

Höhle, Barbara, Schmitz, Michaela, Santelmann, Lynn M, & Weissenborn, Jürgen. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development, 2*(4), 277–300.

Höhle, Barbara, Weissenborn, Jürgen, Kiefer, Dorothea, Schulz, Antje, & Schmitz, Michaela. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, *5*(3), 341-353.

Koulaguina, Elena, Legendre, Géraldine. Barrière, Isabelle, & Nazzi, Thierry. (2019). Towards abstract syntax at 24 months: Evidence from subject-verb agreement with conjoined subjects. *Language Learning and Development*, *15*(2), 157-176.

Legrand, Camille & Shi, Rushen. (2018). Beyond linear non-adjacent or adjacent dependencies: Infants track hierarchical syntactic dependencies. Presentation at the 43rd Annual Boston University Conference on Language Development, Boston, USA.

Melançon, Andréane & Shi, Rushen. (2015). Abstract representation of feature agreement across grammatical categories in infants. In E. Grillo & K. Jepson (Eds.), *Proceedings of the 39th Annual Boston University Conference on Language Development* (Vol. 2, pp. 324-336). Boston, MA: Cascadilla Press.

Santelmann, Lynn M. & Jusczyk, Peter W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition, 69*(2)*,* 105-134.

Shi, Rushen, Legrand, Camille, & Brandenberger, Anna. (In press). Toddlers track hierarchical structure. *Language Acquisition*.

Shi, Rushen & Melançon, Andréane. (2010). Syntactic categorization in French-learning infants. *Infancy, 15*(5), 517-533.

Soderstrom, Melanie, White, Katherine S., Conwell, Erin, & Morgan, James L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy, 12*(1)*,* 1-29.

Van Heugten, Marieke & Christophe, Anne. (2015). Infants' acquisition of grammatical gender dependencies. *Infancy, 20*(6), 675-683.

Van Heugten, Marieke & Shi, Rushen. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *Journal of the Acoustical Society of America, 128*(5), EL223-EL228.

Proceedings of the 44th annual Boston University Conference on Language Development

edited by Megan M. Brown
and Alexandra Kohut

Cascadilla Press Somerville, MA 2020

ISSN 1080-692X
ISBN 978-1-57473-057-9 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com