

ARTICLE

Do infants have abstract grammatical knowledge of word order at 17 months? Evidence from Mandarin Chinese.

Jingtao ZHU^{1,2*}, Julie FRANCK³, Luigi RIZZI^{4,5}, and Anna GAVARRÓ¹

¹Departament de Filologia Catalana, Universitat Autònoma de Barcelona, Spain, ²ClicAsia, Centre d'Estudis Orientals, Barcelona, Spain, ³Laboratoire de Psycholinguistique, Université de Genève, Switzerland, ⁴Departament de Linguística, Université de Genève, Switzerland, and ⁵Centro Interdipartimentale di Studi Cognitivi sul Linguaggio, Università degli Studi di Siena, Italy

*Corresponding author: Jingtao Zhu, ClicAsia, Centre d'Estudis Orientals, Carrer de València, 359, 5^è2^a - 6^è2^a, 08009, Barcelona, Spain. E-mail: jtzhu@clicasia.com

(Received 17 May 2020; revised 7 October 2020; accepted 16 November 2020; first published online 8 February 2021)

Abstract

We test the comprehension of transitive sentences in very young learners of Mandarin Chinese using a combination of the weird word order paradigm with the use of pseudo-verbs and the preferential looking paradigm, replicating the experiment of Franck *et al.* (2013) on French. Seventeen typically-developing Mandarin infants (mean age: 17.4 months) participated and the same experiment was conducted with eighteen adults. The results show that hearing well-formed NP-V-NP sentences triggered infants to fixate more on a transitive scene than on a reflexive scene. In contrast, when they heard deviant NP-NP-V sequences, no such preference pattern was found, a performance pattern that is adult-like. This is at variance with some of the results from Candan *et al.* (2012), who only found evidence for canonical word order comprehension at almost age 3 when considering fixation time. Furthermore, within the age range tested, performance showed no effect of age or vocabulary size.

Keywords: eye-tracking; word order; Mandarin Chinese; sensitivity to ungrammaticality; abstract grammatical knowledge

1. Introduction

There is evidence that children show sensitivity to the properties of the language they are exposed to at the earliest observable stage of their syntactic productions. Results from corpus studies indicate that, already at the two-word stage, infants raised in Mandarin-speaking environments can produce the canonical Verb-Object order, as in (1a) (example taken from Zhou's corpus (2001) in CHILDES, MacWhinney, 2000), while their Japanese peers produce the Object-Verb order, as in (1b) (example taken from Yokoyama & Miyata's corpus (2017) in CHILDES).¹

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

- (1) a. Chi luo-bu. (Xu'er, 1;8)
eat radish.
'(They) eat radish.'
- b. Mikan tabe-yoo. (Kiichan, 1;8)
orange eat-ORT
'Let's eat orange.'

The fact that children's early multiword utterances deviate little from their target fulfils the predictions of two broad theoretical approaches: the nativist or grammatical approach (Chomsky, 1981, 1993 *et seq.*) on the one hand, which assumes that infants are born with innate linguistic knowledge, i.e., Universal Grammar; and the usage-based or lexical approach (e.g., Tomasello, 2000 *et seq.*) on the other, which rejects the existence of innate linguistic knowledge and claims that language learning is an item-based learning behavior building on general cognitive capacities, and that abstract syntax is not established before the second year of age (see Tomasello, 2003, Matthews *et al.*, 2005, although Ambridge & Lieven, 2011 find some sensitivity to word order by 21 months).

For the generative or grammatical tradition, since infants have innate knowledge of the building mechanisms of phrase structure, the work remains to fix the parameters of the language from the primary linguistic data. For an example like (1), the basic parameters associated with word order include a fundamental parameter determining the position of complements relative to heads formalized in various ways (Berwick & Chomsky, 2011; Chomsky, 1986; Kayne, 1994; Travis, 1984), leading to the contrasts between (1a) and (1b). Thus, at the time when they can produce and comprehend transitive sentences, they have correctly set a fundamental word-order parameter. Children's compliance with word order constraints led Wexler (1998) to formulate the Very Early Parameter Setting (VEPS) hypothesis, according to which basic parameters are correctly set already at the beginning of multiple word combinations.

On the other hand, the usage-based approach attributes the target production of (1) to imitation of the input, with no initial abstract syntactic knowledge. Unlike in the grammatical approach, the child's word order knowledge is triggered by usage, i.e., the frequent exposure to word order patterns for a particular verb that s/he encounters in the input. Thus, long-term exposure is required and only at later stages does the child generalize from memorized fragments to abstract syntactic notions, such as general word order properties.²

Thus, the two approaches make crucially different predictions on the child's capacity to generalize his/her knowledge to new items and structures around the two-word stage. According to the lexical approach, young children will not be able to comprehend new transitive sentences if they do not have a suitable lexically specific schema of the verb. In contrast, under the grammatical approach, since VEPS claims that fundamental word order parameters are already set in the two-word stage, infants are expected to understand new transitive sentences, provided that they contain a target transitive frame, even if the sentences include new verbs.

2. Early acquisition of word order

Starting with Naigles (1990), the preferential looking paradigm allows us to study the comprehension of sentences by infants. Naigles (1990) tested 2-year-old English infants' comprehension of both transitive and non-transitive actions with a novel

verb *gorp*. In the training phrase, half of the participants heard a transitive sentence (e.g., ‘*The duck is gorp*ing the bunny’) and half an intransitive sentence (e.g., ‘*The duck and the bunny are gorp*ing’). Both groups watched a scene in which a duck performed an action on a bunny on one screen and, on the other, the duck and bunny each performed a synchronous non-transitive, reflexive action. In the test phase, infants were asked to “find gorp”ing”. Naigles (1990; see also Naigles & Kako, 1993; Naigles, 1996 using slightly different methods) found that infants who had heard the transitive frame looked longer at the transitive scene than those who heard the intransitive sentences, while children who heard the conjoined-subject intransitive audio did not show any preference. Hirsh-Pasek and Golinkoff (1996) extended the results to 17-month-old infants and found that they can use word order to understand transitive sentences containing familiar words.

Moreover, a study conducted by Gertner *et al.* (2006) found that 21-month-old English-speaking children can use canonical word order to interpret transitive sentences containing pseudo-verbs. Test sentences were illustrated with two simultaneous videos with theta-role reversal: one representing the target SVO interpretation and the other the non-target OVS interpretation. Later, Franck *et al.* (2013) extended the results to French-speaking infants of 19 months using eye-tracking. In their experiment, they resorted to the weird word order paradigm (Akhtar, 1999) in which children heard well-formed (NP-V-NP) and deviant (NP-NP-V) sequences with pseudo-verbs. The distractor video critically differed from the study of Gertner *et al.* (2006): rather than illustrating reversed theta-roles, it illustrated the same action performed reflexively. The results indicated that infants only looked at the transitive scene when they heard the well-formed sequences, while they showed random behavior when they heard the deviant sequences. The preference for the SVO interpretation of NP-V-NP sentences provides strong evidence that infants know that the NP following V is its object, while the preference of English infants reported by Gertner *et al.* (2006) could be due to a preference for SVO over OVS, in line with the quasi-universal SO order found across languages. Moreover, the lack of a preference for NP-NP-V sentences shows that the NP preceding V is not interpreted as its object. Similar results were obtained by Gavarró *et al.* (2015) in the same framework. They tested 20 infants aged 19 months exposed to an OV language with case-marking, Hindi-Urdu, and the results show that infants can parse the well-formed SOV sequences as they looked significantly longer at the transitive video, but they failed to assign a consistent interpretation to the deviant VSO order. Taken together, these experiments indicate that the parameter responsible for the VO/OV alternation is set correctly by 19 months regardless of lexical knowledge of the verb.

These studies are in line with the grammatical approach. Nevertheless, a study of children’s early productions conducted in Cantonese claimed to provide counterevidence to it. Chan *et al.* (2009) used an act-out task and found that Cantonese children did not choose the first noun as AGENT in the canonical SVO sentences containing pseudo-verbs at above-chance levels until 3;6. Without going into the controversy between grammatical and usage-based approaches, Candan *et al.* (2012) was one of the few studies using the preferential looking paradigm to test the acquisition of early word order in Mandarin. Their study focused on how English-, Turkish- and Mandarin-speaking children differ in sentence comprehension when it depends on word order. Test stimuli consisted in two simultaneous videos with theta-role reversal (e.g., ‘*The horse is washing the bird*’ and ‘*The bird is washing the*

horse'). Since they wanted to look solely at the weight of word order, the Turkish nouns were produced without case-marking, even though Turkish is a language with morphological case marking. The results indicated that English children showed early sensitivity to canonical word order at age 1.5, earlier than Turkish (2-year-olds) or Chinese children (almost age 3). Importantly, data collection was incomplete for the Chinese 1-year-old group in Candan *et al.*'s work (2012). The authors attributed the delayed comprehension of canonical transitive sentences to the fact that, in Mandarin as well as in Turkish, both subject and object can be dropped, and the existence of varying word orders in these languages makes the canonical word order less prominent in the input. However, the measures of Candan *et al.* (2012) were not fully consistent, with gazing measures differing from number of switches of attention. Although Mandarin-speaking children were the less certain about matching sentences with scenes, they switched attention less frequently than Turkish-speaking children and did not differ from their English peers, which looked longer at the matching screen from very early on. A higher number of switches of attention is standardly interpreted as uncertainty in comprehension.

Interestingly, the experiment on another language with word order alternations and argument drop reached the same conclusion: Omaki *et al.* (2012) used the eye-tracking techniques and found that Japanese 19-month-olds fail to understand sentences with a canonical SOV order; since their corpus study revealed that 91% of child-directed speech was uninformative to identify canonical word order as case markers are often omitted (see also Matsuo *et al.*, 2012), they suggested that the sparseness of SOV in the input would delay language acquisition in Japanese.

Recent work by Hsu (2018) challenged Candan *et al.* (2012)'s study. Using the forced choice pointing paradigm, Hsu (2018) assessed Mandarin-speaking 2-year-olds' comprehension of canonical SVO and non-canonical SOV sentences with the object marker *ba* using pseudo-verbs. The results show that two-year-olds pointed to target trials 68% of the time for the canonical construction, and performed similarly with non-canonical constructions.

Whether Mandarin-speaking infants are delayed in parsing canonical transitive SVO sentences, as suggested by Candan *et al.* (2012), or they can process them just as French or Hindi-Urdu children before 2 years old, as claimed by Hsu (2018), is to this day an open question. The discrepancy between the results of Chan *et al.* (2009), Candan *et al.* (2012) and Hsu (2018) could be partly explained by methodological differences. Although act-out tasks like the one used in Chan *et al.* (2009) are easier than elicitation tasks, particularly for children with low MLUs, they are surprisingly difficult for very young children, since they require memory when planning an action (see Höhle *et al.*, 2009). Candan *et al.* (2012) used a less cognitively demanding methodology, but there was a high rate of missing data or non-responses in their youngest group (see Franck & Lassotta, 2012 on the problem of missing data in the context of experiments relying on the weird word order paradigm, Akhtar, 1999 and much related work). As already pointed out, the measures of Candan *et al.* (2012) were not consistent, with gazing measures differing from switches of attention. These seemingly contradictory results of the previous research reported motivate the present study. We address the question of comprehension of canonical SVO word order by Mandarin-speaking infants at an earlier period, using eye-tracking measures. In particular, we use the same experimental design as Franck *et al.* (2013) and Gavarró *et al.* (2015), combining the weird word order paradigm with the preferential looking paradigm (Hirsh-Pasek & Golinkoff, 1996), using

pseudo-verbs to ensure that infants can not rely on lexical information to process the sentences. Before we present the experimental design, we introduce some word order properties of Mandarin.

3. Word order in Mandarin Chinese

Word order and noun animacy have been considered the most reliable syntactic devices for sentence interpretation in Mandarin and Cantonese Chinese (Chang, 1992; Li *et al.*, 1993; Miao, 1981), given the lack of morphological markers such as agreement, number, gender or case in these languages. The basic word order in Mandarin is SVO³ (Li, 1990; Sun & Givón, 1985), illustrated in (2):

- (2) Xiao-tu-zi zhua le xiao-ya-zi.
 little-rabbit catch PERF little-duck
 ‘The little rabbit caught the little duck.’

Three other word orders, SOV, OSV, and VOS, are also possible in the spoken language with morpho-syntactic markers such as the object marker *ba* or passive *bei*. These three word orders are possible without any specific markers, but only under very special conditions. In particular, SOV without morphological markers is marked and mainly used when the object is contrastively focused, and also marked by special intonation (Tsai, 2008). Besides, when the object is animate, *ba* is obligatorily required in neutral contexts and with neutral intonation (Van Bergen, 2006). In a recent grammaticality judgment task (Yu & Tamaoka, 2018), the animate-animate-verb sentences without *ba* were judged of very low acceptability, and mostly regarded as uninterpretable among native speakers. We refer the reader to Huang *et al.* (2009) for analyses of SOV and the other non-canonical word orders in Mandarin. Quantitatively, canonical word orders are attested in about 90% of sentences in child-directed speech, according to the study of Yeh (2015).

Another feature of Mandarin is the presence of null arguments. Both subjects and objects can be omitted as long as reference can be recovered through the previous discourse context (Huang, 1984). (3) is an example of topic drop (with both subject and object drop).

- (3) Speaker A: Ni-men dou kan guo tai-tan-ni-ke-hao le ma?
 you all see EXP Titanic PERF SFP
 ‘Have you seen *Titanic*?’
 Speaker B: (Ø) Kan guo (Ø) le.
 see EXP SFP
 ‘(We) have already seen (it).’

In (3), both the subject ‘we’ and direct object ‘Titanic’ can be dropped, because they can be recovered through the discourse. Given this fact, Candan *et al.* (2012) argued the topic drop phenomenon can limit the reliability of canonical order. A recent corpus study quantified the omission of subject and object in Mandarin and revealed 49.83% subject omission and 34.42% object omission in child-directed speech (Zhu & Gavarró, 2019) while in other languages, such as Japanese, null subjects raise to 83%, according to Matsuo *et al.* (2012).

4. Present study

The present study adopted the experimental design and the procedure of the study by Franck *et al.* (2013) who tested infants acquiring French, used again in the study on Hindi-Urdu by Gavarró *et al.* (2015). This allows us to include our findings from Mandarin-speaking infants (Experiment 1) in a cross-linguistic comparison. Moreover, we conducted the same experiment with adults (Experiment 2), while no results for adults were reported by Franck *et al.* (2013). The experiments were approved by the university's ethical committee (CEEAH approval number 5071).

4.1. Experiment 1: Infant Mandarin

4.1.1 Method

Participants

Seventeen typically-developing, Mandarin-speaking infants (7 boys, 10 girls) with a mean age of 17 months and 4 days (age range = 1;1.3–1;9.0, $SD = 2.2$) participated in Experiment 1. Seven additional infants participated in the study but they were not included in the results because of large errors in calibration ($n = 4$) or because of the infants' lack of eye tracking samples ($n = 3$). They were recruited in Guiyang, China.

As a measure of the infants' linguistic development, their vocabulary was assessed using the Mandarin version of the Communicative Development Inventory (CDI, Hao *et al.*, 2008), which consists of two checklists: an infant checklist (used for infants between 12 and 16 months of age) and a toddler checklist (used for children between 17 and 30 months of age). Both checklists included the animals' names used in our study. Following Hao *et al.* (2008), for the toddler list parents were only asked to indicate whether their children had ever said the word, as is done in the English CDI (Fenson *et al.*, 1994), and so no comprehension scores are available.

For our study, infants from 13 to 16 months (the younger group, $n = 8$) achieved a mean score of production of 5 words ($SD = 5.5$, range from 0 to 11) and a mean score of comprehension of 25 words ($SD = 14.2$, range from 9 to 36). Infants from 17 to 21 months (the elder group, $n = 9$) achieved a mean score of production of 43 words ($SD = 31.4$, range from 0 to 102). The summary of their scores is shown in [table 1](#).

Materials

Following Franck *et al.* (2013) we created 3 well-formed (NP-V-NP) and 3 deviant (NP-NP-V) sequences (see [Table 2](#)). NP-NP-V is deviant, because, first, it is used in neutral contexts and with neutral intonation, whereas it is only possible in contexts that license contrastive focus and involves a special focal intonation; second, when both NP are animate, the non-canonical SOV and OSV sentences (i.e., the NP-NP-V strings) are not acceptable for native speakers (Yu & Tamaoka, 2018).

In Mandarin, aspectual information is systematically expressed; and the perfective marker *le* to mark the end of the action is used far more frequently than other markers in early speech (Erbaugh, 1982). For that reason, the perfective aspect *le* was selected to describe the scene.

The two monosyllabic pseudo-verbs *nú* 'to put a crown on someone's head' and *chēi* 'to put someone's head under a net' were devised in this study. Verbs in the phonological neighborhood (Luce & Pisoni, 1998) of these two pseudo-verbs showed a similar distribution of transitivity. Statistics computed on the number of verbs showed that 61.3% of the verbs in the phonological neighborhood of *nú* were

Table 1. Infants' vocabulary scores

Months	Comprehension (words)		Production (words)	
	Mean	Range	Mean	Range
13–16	25 (<i>SD</i> = 14.2)	9–36	5 (<i>SD</i> = 5.5)	0–11
17–21	-	-	43 (<i>SD</i> = 31.4)	0–102

Table 2. List of experimental sentences.

Well-formed	Xiao-gou chei le xiao-lv.
	The-little-dog PSEUDOV PERF the-little-donkey
	Shi-zi chei le xiao-ma.
	The-lion PSEUDOV PERF the-little-horse
	Xiao-niu chei le xiao-yang.
	The-little-cow PSEUDOV PERF the-little-sheep
Deviant	Xiao-niu shi-zi nui le.
	The-little-cow the-lion PSEUDOV PERF
	Xiao-lv xiao-gou nui le.
	The-little-donkey the-little-dog PSEUDOV PERF
	Xiao-yang xiao-ma nui le.
	The-little-sheep the-little-horse PSEUDOV PERF

transitive, while the distribution was 60% for *chéi*. *Chéi* was used in the NP-V-NP condition, whereas *nui* was used in the NP-NP-V condition.

To verify that our pseudo-verbs followed the phonological pattern and phonotactic constraints of Mandarin verbs, we asked 10 adult Chinese speakers to judge if each verb (which was presented embedded in a sentence) sounded familiar and whether they knew its meaning. The judgement was based on a binary scale (yes/no) and all 10 participants said the verbs sounded familiar but could not assign any meaning to them.

Mandarin being a tone language, the pseudo-verbs used in the test presented a high tone, and lexical tone interacts with sentential intonation. We compare the pitch movements and pitch range expansion in three of the test sentences, with lexical tone kept constant, using Praat (Boersma & Weenink, 2005). The intonational pattern of test SVO sentences is illustrated in Figure 1. In Figure 2, we show the intonational pattern of the deviant SOV sentences, which was the same as the intonational pattern in their well-formed counterparts in Figure 3, as both were rising-falling-rising, with pitch accent on the first NP. Thus, the ill-formedness of the SOV sequences in our experiment stemmed only from word order, rather than the intonational pattern imposed on the sequence.

Videos were the same as in Franck *et al.* (2013), and are illustrated in Figure 4, and the characters included a dog, a donkey, a lion, a horse, a cow and a sheep. We added the adjective *xiao* 'little' to most of the nouns, as is common in child-directed speech. All the children knew the name of the animals used in the experiment according to their vocabulary checklist. The sound track was pre-recorded by a Mandarin female

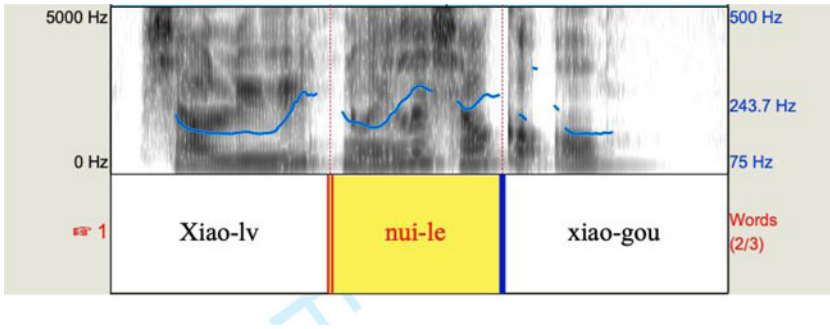


Figure 1. Intonational pattern of the well-formed sentence *xiao-lv nui-le xiao-gou*.

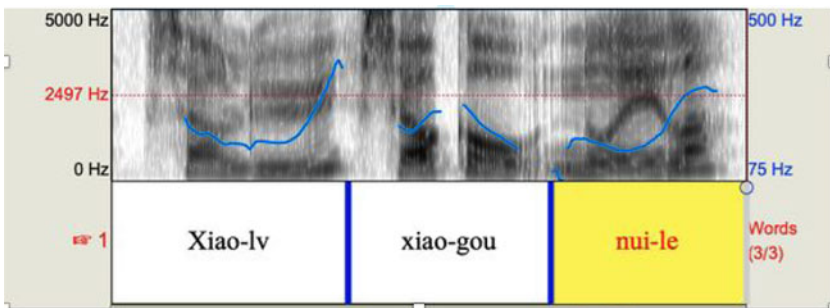


Figure 2. Intonational pattern of the deviant sentence *xiao-lv xiao-gou nui-le*.

native speaker. Utterances were chopped using Praat (Boersma & Weenink, 2005) to make sure all repetitions were the same and videos were re-edited with Adobe Premier Pro CC 2017 (v. 11.0.2).

In the experimental session, for each sentence the infants were presented with two simultaneous videos, one video showing the action carried out transitively with the first NP as AGENT and the second NP as PATIENT (e.g., the cow putting a crown on the lion's head), the other video illustrating the same action carried out reflexively with both NPs as AGENTS (e.g., the cow and the horse each putting a crown on their own head). The items were presented in random order with the presentation of the transitive and reflexive event counterbalanced across the left and right sides of the screen and across the well-formed and deviant conditions.

Procedure

The eye-tracker used was a Tobii Pro X3-120 (with a sampling rate of 120 Hz) and Tobii Studio™ (Version 3.4.8) was used as platform for the recording and analysis of the eye gaze data. The video stimuli were projected from a laptop and the stimuli ratio corresponds to the screen resolution (1920 x 1080). Each child sat on his or her caregiver's lap approximately 60 cm from the computer screen during the whole length of the experiment, such that the gaze angle did not exceed 40 degrees (the supported operating distance for the Tobii Pro X3-120 Eye Tracker is 50–90 cm)⁴.

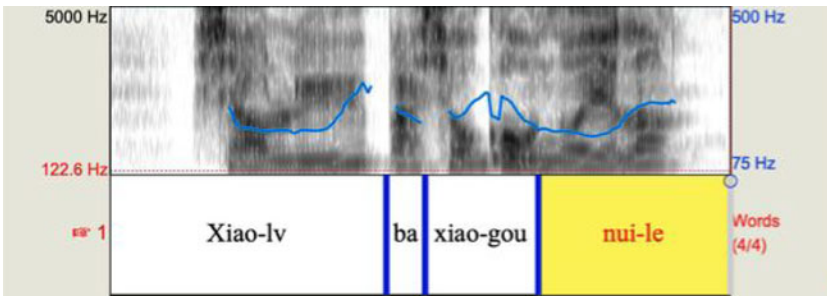


Figure 3. Intonational pattern of the well-formed sentence *xiao-lv ba xiao-gou nui-le*.

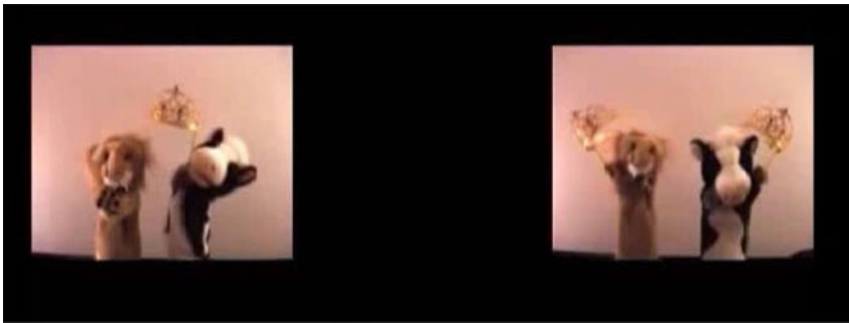


Figure 4. Visual stimuli used in the experiment (from Franck *et al.*, 2013).

The caregivers were asked to close their eyes and listen to music played through headphones during the test trials so as not to guide their children towards any of the videos. The test room remained isolated from sunlight and other uncontrolled light sources (300–350 Lux, Temperature 18–25°C).

The experimental session started with the procedure of eye calibration, then we proceeded to the training session. In the training session, first the participants went through a character-identification phase; all the puppets were presented once (e.g., *Bao-bao kuai kan, shei zai na-li? O, shi xiao-lv* ‘Look, who’s here? It’s the little donkey’), while half of the screen remained blank (6s). Next, the participants were introduced to the simultaneous presentation, which showed two different animals at the same time while the recorded voice asked them to find one of them (e.g., *Bao-bao kuai kan, kan-dao xiao-lv le ma? Xiao-lv zai na-li ya?* ‘Look, do you see the little donkey? Where is the little donkey?’). Finally, the participants saw the novel actions used; most importantly, novel actions were presented in neutral frames without the use of the novel verbs, paired with sentences like *Kan, fa-sheng le shen-me?* ‘Look, what happened?’ such that later understanding of the test sentences cannot be attributed to lexical learning during the training phase (see Ambridge & Lieven, 2011; Franck *et al.*, 2013 for discussion).

After the training session and a short transition cartoon, the experimental session started. A blank screen (2s) appeared between experimental items (six in total), and after items 3, 4 and 5 a clip of a Teletubbies landscape was shown to keep the child’s

Table 3. Mean looking time (in ms, standard deviations in parentheses) across the four critical ROIs in infants.

	Well-formed		Deviant	
	Transitive	Reflexive	Transitive	Reflexive
Baseline (0–4s)	1438(619)	1098(689)	1394(818)	1138(439)
Sentence 1 (5–9s)	1386(665)	928(724)	1198(779)	1072(569)
Sentence 2 (10–14s)	1266(903)	793(831)	974(680)	1041(1012)
Sentence 3 (15–19s)	931(772)	852(662)	1153(1021)	876(779)

attention. All videos started with a sentence to draw the child's attention (e.g., 'Look, what happened?') as baseline, and then the experimental sentences were played three times. Thus, the recording of gazing time took place in four windows: the baseline and three consecutive exposures to the target sentence starting at 5, 10, 15 seconds. The whole session lasted between 10–15 minutes. After the test session, the experimenter asked the infants' caregivers to fill out the Chinese version of CDI (Hao *et al.*, 2008).

Data analyses

Following Franck *et al.* (2013), only infants whose detected signal was more than 55% were taken into account. The number of participants analyzed was 17.

To provide an overview of the eye movement data, linear mixed-effects models were applied using the lme4 package (Bates *et al.*, 2014) from R (v3.5.2, R Development Core Team, 2015). We computed generalized linear mixed models with proportions of fixations to the transitive video (calculated over total looking time to the transitive and reflexive videos) as dependent variable and regions of interest (ROIs, Baseline, Sentence 1, Sentence 2, Sentence 3) and Condition (Well-formed, Deviant) as fixed effects with random intercept and slope for participants and items. We explored the effect of age and vocabulary on proportions of fixations to the transitive video using generalized linear models on two ROIs that showed a significant effect of Condition, with the proportion of looks to the transitive video as dependent variable and Condition, Vocabulary (as continuous variable) and Age (as continuous variable) as factors.

4.1.2 Results

Table 3 reports the mean looking times to each of the videos (transitive vs. reflexive) as a function of the well-formedness of the sentence in each of the four ROIs: the baseline window and the three consecutive windows corresponding to first, second and third exposure to the experimental sentence.

Visual inspection of the heat maps (which display the accumulated fixation duration on different locations in the video) for well-formed sentences across all infants and all ROIs suggests that they fixed their gaze longer on the transitive action as shown by the thicker red shade indicating intensity of gaze based on fixation durations (see Fig. 5), while this intensity effect was fluctuating in the deviant sentences as can be observed in Fig. 6.

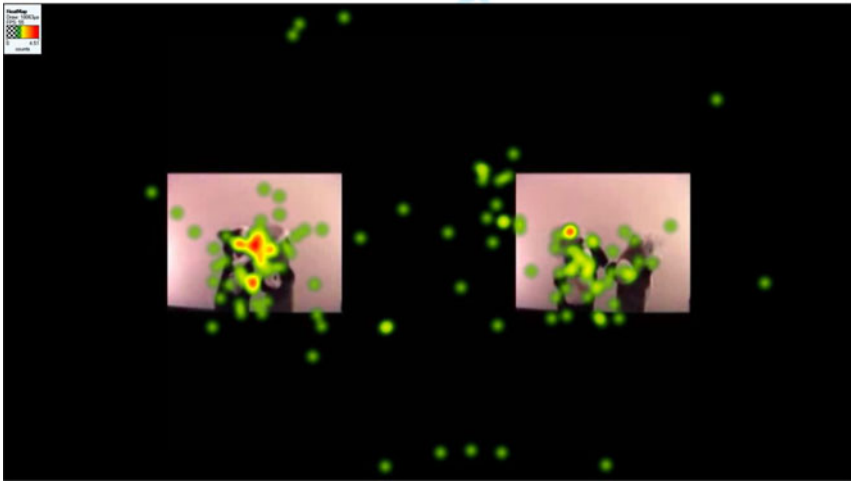


Figure 5. Heat map for the well-formed sentence *Xiao-gou chei le xiao-lv* ‘The little dog chei-ed the little donkey’. Red indicates the highest number of fixations or the longest time, and green the least. The left video represents the transitive event and the right video represents the reflexive event.

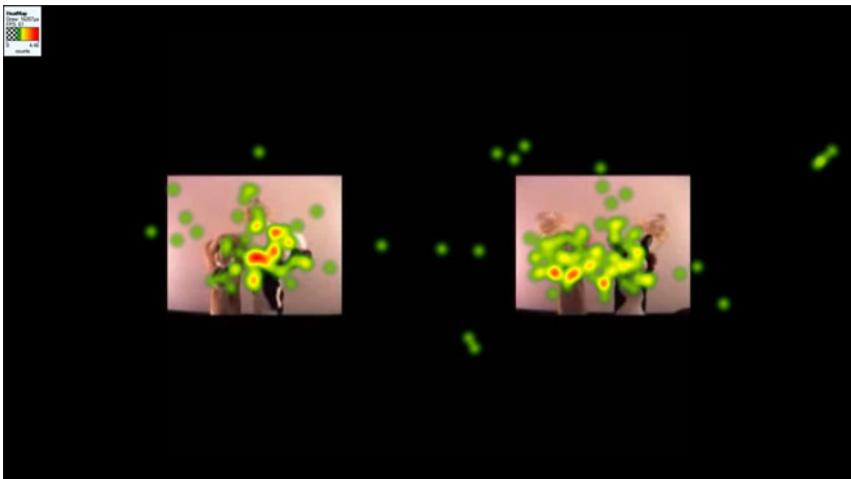


Figure 6. Heat map for the deviant sentence *Xiao-liu shi-zi nui le* ‘The little cow the lion nui-ed’.

Figure 7 illustrates the distribution of the proportion of looking time to the transitive scene as a function of well-formedness, across the four ROIs. Wilcoxon signed-rank analysis was conducted on proportions of looking time to the transitive video. The results showed a significant above chance effect (defined as 50%) in the well-formed condition during the first presentation of the test sentence in the 5–9s window only ($Z = -2.20, p = .028$) and marginally significant in the 10–14s window ($Z = -1.89, p = .058$). Looking time to other windows for the well-formed condition as well as for all the windows of the deviant condition were at chance level.

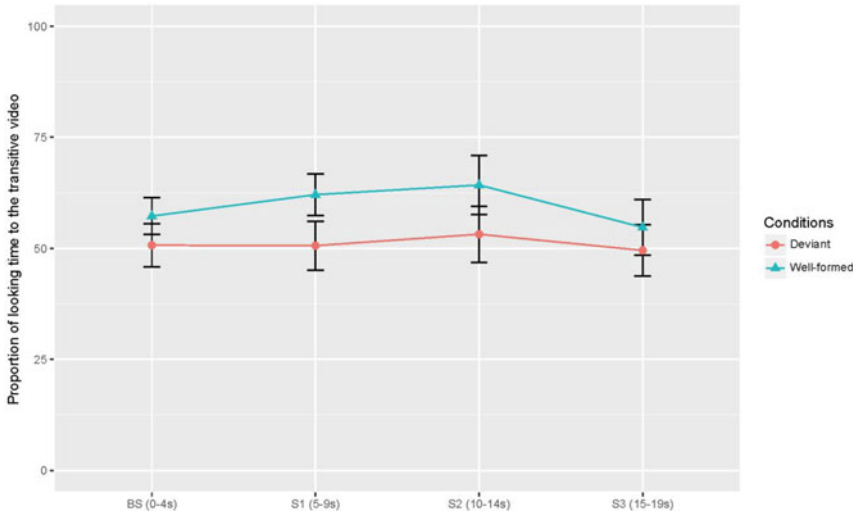


Figure 7. Proportion of looking time to the transitive video in the four critical ROIs in Experiment 1 (Infants).

The generalized linear model with the proportion of looking times to the transitive action as dependent variable and ROI (Baseline, Sentence 1, Sentence 2, Sentence 3) and Condition (Well-formed, Deviant) as factors showed a significant interaction between ROI and Condition ($z = .46$, $SE = .11$, $p = .045$), which allowed us to further explore the effect of Condition in each ROI. We found a significant effect of Condition after the first presentation of the sentence ($\beta = .29$, $t = 1.43$, $p = .016$) and the second presentation ($\beta = .11$, $t = 1.49$, $p = .027$), which means that infants showed an increased preference for the transitive video compared to the reflexive one when they heard a well-formed sentence compared to when they heard a deviant one. No effect of Condition was found in the baseline window ($\beta = .15$, $t = .84$, $p = .41$) nor after the third presentation of the sentence ($\beta = -.07$, $t = -.35$, $p = .72$).

Generalized linear models run on the two ROIs showed a significant effect of Condition (i.e., S1 and S2 together), with the proportion of looks to the transitive video as dependent variable and Condition, Vocabulary and Age as factors showed no effect of Age ($\beta = -.01$, $t = -.44$, $p = .66$), no main effect of Vocabulary ($\beta = .016$, $t = 1.74$, $p = .34$), and critically no interaction between Vocabulary and Condition ($\beta = -.016$, $t = -1.25$, $p = .22$), nor between Age and Condition ($\beta = -.021$, $t = -.69$, $p = .49$). This indicates that neither vocabulary nor age modulated the effect of well-formedness. The three-way interaction was not significant either ($\beta = .0009$, $t = 1.2$, $p = .24$). This confirms that the increased preference found for the transitive video over the reflexive video when a well-formed sentence is presented is independent from age and vocabulary.

4.2 Experiment 2: Adult Mandarin

4.2.1 Method

Participants – Eighteen native Mandarin-speaking adults (age range = 24–53, mean age = 29, $SD = 7.7$) participated in our study. They were recruited in Guiyang and Barcelona.

Table 4. Mean looking time (in ms, standard deviation in parentheses) across the four critical ROIs in adults.

	Well-formed		Deviant	
	Transitive	Reflexive	Transitive	Reflexive
Baseline (0–4s)	1765(511)	1437(503)	1833(533)	1567(556)
Sentence 1 (5–9s)	2585(617)	909(397)	1711(613)	1829(662)
Sentence 2 (10–14s)	3140(967)	325(504)	1626(1070)	1639(963)
Sentence 3 (15–19s)	3267(720)	379(451)	1730(740)	1790(711)

Materials, procedure and data analyses – The materials and procedure used were the same as those for infants. We adopted the same analysis for adults' data as had been adopted for the infant data. For all the adults tested, the detected signal was more than 75%.

4.2.2 Results

The mean looking time to each of the scenes in the four ROIs for adults can be found in Table 4.

The proportion of looking time to the transitive video in each ROI is shown in Figure 8.

The generalized linear model with the proportion of looking time to the transitive action as dependent variable and ROI and Condition as factors showed a significant interaction between the two factors ($z = 1.92$, $SE = .98$, $p = .04$). Thus, we explored the effect of Condition in each ROI separately. With the generalized model we found a significant main effect of Condition during the first presentation ($\beta = .86$, $t = 5.64$, $p < .001$), the second presentation ($\beta = 1.59$, $t = 5.54$, $p < .001$) and the third presentation of the sentence ($\beta = 1.56$, $t = 7.82$, $p < .001$), showing that the preference for the transitive video is increased when a well-formed sentence is presented. No effect of Condition was found in the baseline window ($\beta = .03$, $t = .23$, $p = .81$).

5. Discussion

The present study tested the comprehension of canonical transitive NP-V-NP sentences in very young learners of Mandarin combining the weird word order paradigm (with deviant NP-NP-V sequences) and the preferential looking paradigm using eye-tracking techniques (as in Franck *et al.*, 2013 and Gavarró *et al.*, 2015). Our work indicates that, just like Mandarin-speaking adults, 17-month-old infants acquiring Mandarin show a preference for the transitive scene when they encounter well-formed transitive NP-V-NP sequences with novel verbs, but that does not happen when they hear deviant NP-NP-V sequences. Besides, the results for adults are very similar to those for infants: with well-formed sequences, adults direct their gaze towards the transitive video, with deviant sequences they direct they gaze randomly across the two videos. The only difference between adults and infants is that adults maintain attention on the transitive video with a well-formed sequence until the last presentation of the sentence.

The preference observed for infants in the well-formed condition cannot be explained by usage-based approaches, since all sentences included pseudo-verbs.

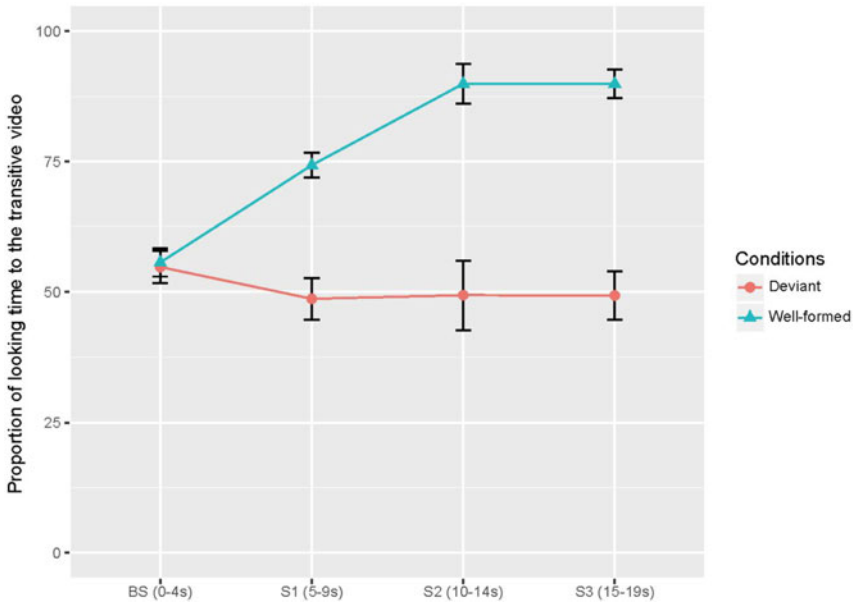


Figure 8. Proportion of looking time to the transitive video in the four critical ROIs in Experiment 2 (Adults).

Neither comprehension of the well-formed sequence, nor the difference in performance between well-formed and deviant sequences is predicted by the usage-based approach. The performance attested is not only contrary to the predictions of the usage-based account; it also runs against the predictions of a grammar-based approach which claims that infants follow an *AGENT*-first strategy in their parsing (Lidz *et al.*, 2001). In contrast to results from Gertner *et al.* (2006) which could indeed be interpreted as such, given that the two videos illustrated transitive actions with reversed theta-roles, if infants had proceeded in that way in our experiment, they would have performed identically with NP-NP-V and NP-V-NP sequences, since the two videos illustrated the first NP as the *AGENT*. Even if an *AGENT*-first strategy exists, our results show that it cannot override grammaticality: that is, it cannot be used to assign an interpretation to an ungrammatical sentence.

The well-formed NP-V-NP sentences include known nouns and an unknown verb, and the correct interpretation of such structures implies that infants can use the arguments in a sentence to infer the syntactic structure and take the unknown word to be the verb (i.e., by syntactic bootstrapping, Fisher *et al.*, 1994; Gleitman, 1990); in our study, Mandarin-speaking infants can infer the subcategorization frame of a novel verb based on the syntactic structure: namely, when they hear a verb describing a two-argument event in a target NP-V-NP manner, they infer that the verb has a transitive meaning.⁵ Infants exposed to Mandarin fail to parse the sequence in that way when the two arguments appear in a NP-NP-V frame: in particular, they do not identify the immediately preverbal NP as the object; neither infants nor adults consider animate NP-NP-V sequences as SOV. Besides, recall that the ill-formedness of the test items in our experiments arises from word order alone, since they have been produced just as their well-formed counterparts, which shows

again that the behavior observed in NP-NP-V sequences cannot be taken as a sign that infants lack syntactic competence, as with adults we attribute the same behavior to sensitivity to the ill-formed sequence.

As pointed out by an anonymous reviewer, if the looking preference for the transitive video were obtained when ruling out the reflexive video because the two characters are not carrying out a joint action, again we would expect the same preference to emerge upon hearing the deviant sentences, which was not the case.

Our study is in line with the original French experiment (French being an SVO language with little presence of null arguments), as well as the Hindi-Urdu experiment (Hindi-Urdu being an SOV language with generalized pro-drop). The only difference in the results is due to the grammatical difference between SVO and SOV languages: French and Mandarin-speaking infants increased significantly their fixations to the transitive video when they heard the well-formed NP-V-NP sequences compared to the NP-NP-V word order, ill-formed in these languages. By contrast, the NP-NP-V order gave rise to a looking preference for the target transitive event in infants acquiring Hindi-Urdu, an SOV language. This shows that infants are sensitive to the specific syntactic structures of the languages they are exposed to.

Due to the length of each experimental item, comparison between the exact timing of effects among the three languages is not really possible: in the French experiment, the 20-second video was split in 5 windows, while in both Hindi-Urdu and the present experiment there were 4 windows of analysis. However, we can still make some observations. Preference for the transitive over the reflexive action appears in the window at 8–12s in the case of French infants, and at 6–10s in the case of Hindi-Urdu, while in Chinese the effect emerges at 5–9s. All correspond to the first presentation of the sentences. Hindi-Urdu was the only language in which the effect persisted until the last 16–20s window, while in Chinese and French the effects disappeared in the last window, which should be due to tiredness; since, at least in Chinese, infants were younger.

Age is indeed one respect in which the three studies differ, since the Mandarin infants here were younger by almost two months (17.4 vs. 19). Thus the findings of Franck *et al.* (2013) are now replicated with infants younger than in previous studies. The vocabulary scores of the infants exposed to Mandarin were also lower than those of the French infants (for the infants exposed to French the mean was 87, and the range was 8–389). In Franck *et al.* (2013), children's lexical knowledge failed to predict individual preferences for the matching video. The same fact had also been observed in other studies (Candan *et al.*, 2012). Our results for infants corroborated the conclusion that vocabulary size did not relate in any systematic way to comprehension and to that we added a new result: within the age range of the infants tested for Mandarin, age was not a predictor of comprehension either.

The absence of an age effect suggests that the parameter has been fixed earlier than 17 months (if the parameter was fixed around this age, one would expect age to be relevant). This raises the further question of when children start to be sensitive to the word order of their target language. In earlier work Nespors *et al.* (1996) argued that headedness may be fixed in the basis of prosodic prominence patterns at the prelexical stage, as an instance of phonological bootstrapping, and Christophe *et al.* (2003) brought empirical evidence showing that, by the age of 3 months, babies are able to discriminate head-complement from complement-head languages on the sole basis of prosodic prominence differences. Gervain *et al.* (2008) further showed that

Italian and Japanese prelexical 8-month-old infants already show preferences for the order of lexical vs. functional elements of their language, a distributional property that correlates with head directionality across. In addition, recently neural evidence using near-infrared spectroscopy (NIRS) suggests that the ability to learn the sequential order of words is present even in newborn infants (Benavides-Varela & Gervain, 2017). If these studies are on the right track, it should come as no surprise that the infants in our study show sensitivity to canonical word order by 17 months and that no age effect is found within the age range tested. Testing non-canonical word orders (e.g., the *ba* construction) remains for future research, although evidence from French using the same experimental paradigm (Lassotta *et al.*, 2014) as well as English using different paradigms (Gagliardi *et al.*, 2016; Seidl *et al.*, 2003) suggests that young children are already able to parse some of those sentence types.

In the original French experiment replicated here, the pseudo-verbs in the test sentences did not involve any aspectual or functional information (4a), while in both Chinese and Hindi-Urdu, the verbs contained perfective aspect markers like *le* in Chinese and *-(y)aa* in Hindi-Urdu (with additional case markers in Hindi-Urdu, see (4b) and (4c)). This could help infants identify the verb. Previous studies have found that infants from 12–16 months are able to use function words to categorize novel words (Höhle *et al.*, 2004; Zhang *et al.*, 2015) and 18-month-olds can use function words to recognize verbs (Cauvet *et al.*, 2014). Still, the presence of overt functional elements is not essential, as witnessed by the original results from French, where infants at 19 months were able to parse the well-formed sentences with no overt functional element.

- (4) a. Le lion poune le cheval. (French)
 D lion PSEUDOV D horse
 b. Xiao-shi-zi chei le xiao-ma. (Chinese)
 the-little-lion PSEUDOV PERF the-little-horse.
 c. Sher -ne ghode -ko khalaayaa. (Hindi-Urdu)
 lion -ERG horse -ACC PSEUDOV-PERF

Finally, let us go back to previous work on Mandarin, in particular the results of Candan *et al.* (2012), and compare them to our results. Our results contrast with those from Candan *et al.* (2012) for Mandarin, since they only found evidence for word order acquisition around age 3. Although they also used the preferential looking paradigm, we hypothesize that the different results may be due to a combination of factors, the first of which relating to perfective marker *le*. A recent study by Yang *et al.* (2018) reveals that the perfective marker *le* did have an immediate effect on 30-month-old Mandarin-speaking children's looking behavior: as soon as they heard *le*, they looked at the scene in which the event began and terminated, while they showed latency in looking at scenes matching sentences with the imperfective marker *zhe*, which describes an on-going, progressive event. In Candan *et al.*'s items *le* was either absent or replaced by imperfective *zhe*, which does not ease comprehension when compared to *le* (Yang *et al.*, 2018). In the longitudinal study of Erbaugh (1982), *le* appeared earlier than *zhe* in child production and was used far more frequently than *zhe* in early speech. These studies converge in showing that *le* is acquired earlier than *zhe*, possibly due to its higher frequency in the input. A second difference between our study and Candan *et al.* (2012) is that, in ours, the target video depicted a transitive action, while the

distractor depicted a reflexive one, with no theta-role reversal, while distractors with theta-role reversal were used in Candan *et al.* (2012). As pointed out by Yang *et al.* (2018), reversibility of NPs may have complicated the processing task. Finally, children from Candan *et al.* (2012) were recruited in Taiwan, so that, apart from Mandarin, children might have been exposed to the Taiwanese Southern Min dialect, which is a strongly OV language (Huang & Roberts, 2017). This may have influenced their performance when confronted with target SVO and non-target OVS in Mandarin Chinese. These three factors are to be added to the lack of some measures for one-year-olds (see section 2).

It would seem, then, that Mandarin, French and Hindi-Urdu would pattern alike, and therefore there would be no grounds to establish a cross-linguistic difference in the emergence of early syntax, as far as basic word order properties are concerned, at least for languages like Mandarin, French and Hindi-Urdu.

6. Conclusion

Infants acquiring Mandarin preferentially look at transitive scenes when they hear well-formed NP-V-NP sequences, whereas no significant preference is observed when infants are confronted to ill-formed NP-NP-V sequences. We have observed this preference pattern with pseudo-verbs, of which the infants had no previous knowledge. We conclude from these results that infants acquiring Mandarin from age 1;5 at the latest have abstract knowledge that their target language is VO. Their response pattern thus appears to be grammar-based.

This finding is consistent with the evidence already gathered on Indo-European languages (French, Hindi-Urdu), albeit for a slightly older age (19 months, rather than the 17.4 months of the participants in the present research). Our result comes from a language which displays word order variation and the presence of null arguments. Thus, children acquiring Mandarin are sensitive to the canonical word order even before they have a sizeable lexicon, from around 1;5, in support of the VEPS hypothesis; the alternative view that infants do not have any early abstract knowledge of word order fails to predict the performance pattern encountered.

Acknowledgements. We are grateful to Roberta Golinkoff for pointing out the age issue in relation to Experiment 1, and to Aojun Chen and Maria del Mar Vanrell for discussion and help with the intonation of Mandarin. The work reported received the funding of projects FF2017-87699-P and 2017 SGR 634. Many thanks are due to the infants and adults who agreed to participate in the experiments reported. Many thanks are also due to two anonymous reviewers for their suggestions.

Notes

¹ The following abbreviations are used in the examples: ACC = accusative case, BA = ba construction, D = determiner, ERG = ergative case, EXP = experiential aspect, ORT = cohort modal, PERF = perfective aspect, PSEUDOV = pseudo-verb, SFP = sentence-final particle.

² Although how children go from lexically specific frames to abstract syntactic constructions still remains speculative, since, to our knowledge, there are very few studies that address this issue directly.

³ As pointed out by an anonymous reviewer, some word order patterns in Mandarin reveal head-final properties, like prenominal relative clauses.

⁴ See <https://www.tobiiipro.com/pop-ups/accuracy-and-precision-test-report-x3/?v=1.0> for more details about accuracy and precision performance of the Tobii system and the test methodology used.

⁵ The fact that infants, as adults, looked numerically longer at the transitive scene in the baseline has been observed previously by Naigles (1996), who found the same effect with 2-year-olds; however, looks at the transitive scene in the baseline were not significantly longer than looks at the reflexive scene.

References

- Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, 26, 339–356.
- Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: *Linear mixed-effects models using Eigen and S4*. R package version 3.5.2.
- Benavides-Varela, S., & Gervain, J. (2017). Learning word order at birth: A NIRS study. *Developmental Cognitive Neuroscience*, 25, 198–208.
- Berwick, R., & Chomsky, N. (2011). The biolinguistics program: The current state of its evolution and development. In A. di Sciullo & C. Boeckx (Eds.), *The biolinguistics enterprise: New perspectives on the evolution and nature of the human language faculty* (pp. 19–41). Oxford: Oxford University Press.
- Boersma, P., & Weenink, D. (2005). *Praat: Doing phonetics by computer*. Computer program. <http://www.praat.org/>.
- Candan, A., Küntay, A., Yeh, Y. C., Cheung, H., Wagner, L., & Naigles, L. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, 27, 205–221.
- Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, 10(1), 1–18.
- Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, 20(2), 267–300.
- Chang, H. W. (1992). The acquisition of Chinese syntax. In H-C Chen & O-J-L Tzeng (Eds.), *Language processing in Chinese* (pp. 175–206). Amsterdam: North Holland.
- Chomsky, N. (1981). *Lectures on government and binding: the Pisa lectures*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S.J. Keyser (Eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 1–52). Cambridge, MA: MIT Press.
- Christophe, A., Nespore, M., Guasti, M. T., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science*, 6(2), 211–220.
- Erbaugh, M. (1982). *Coming to order: Natural selection and the origin of syntax in the Mandarin-speaking child*. Doctoral dissertation, University of California at Berkeley.
- Fenson, L., Dale, P. S., Reznick, J.S., Bates, E., Thal, D. J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59, v-179.
- Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L.R. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92, 333–375.
- Franck, J., & Lassotta, R. (2012). Revisiting evidence for lexicalized word order in young children. *Lingua*, 122, 92–106.
- Franck, J., Millotte, S., Posada, A., & Rizzi, L. (2013). Abstract knowledge of word order by 19 months: An eye-tracking study. *Applied Psycholinguistics*, 34, 323–336.
- Gagliardi, A., Mease, T. M., & Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15- and 20-month-olds. *Language Acquisition*, 23(3), 234–260.
- Gavarró, A., Leela, M., Rizzi, L., & Franck, J. (2015). Knowledge of the OV parameter setting at 19 months: Evidence from Hindi-Urdu. *Lingua*, 154, 27–34.
- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstracts knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684–691. DOI: <https://doi.org/10.1111/j.1467-9280.2006.01767.x>

- Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, 57(1), 56–74. DOI : <https://doi.org/10.1016/j.cogpsych.2007.12.001>.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55.
- Hao, M. L., Shu, H., Xing, A. L., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods*, 40(3), 728–733.
- Hirsh-Pasek, K. R., & Golinkoff, R. M. (1996). *The Origins of Grammar*. Cambridge, MA: The MIT Press.
- Höhle, B., Weissenborn, J., Klefer, 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(3), 341–353.
- Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior & Development*, 32, 262–274.
- Hsu, D. B. (2018). Children's syntactic representation of the transitive constructions in Mandarin Chinese. *PLoS ONE*, 13(11), e0206788. DOI: <https://doi.org/10.1371/journal.pone.0206788>
- Huang, J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531–574.
- Huang, C. T., Li, Y. H. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge: Cambridge University Press.
- Huang, J., & Roberts, I. (2017). Principles and parameters of Universal Grammar. In I. Roberts (Ed.), *The Oxford handbook of Universal Grammar* (pp. 307–354). Oxford: Oxford University Press.
- Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Lassotta, R., Omaki, A., & Franck, J. (2014). Abstract knowledge of non-canonical word order by 21-month-olds. Paper presented at BUCLD 36. Boston.
- Li, Y. A. (1990). *Order and constituency in Mandarin Chinese*. Dordrecht: Kluwer Academic Publishers.
- Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of judgement interpretation in Chinese. *Journal of Memory and Language*, 32, 169–192.
- Lidz, J., Gleitman, H., & Gleitman, L. (2001). *Kidz in the Hood: Syntactic bootstrapping and the mental lexicon* (IRCS Technical Reports Series). Philadelphia, PA: University of Pennsylvania.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19, 1–36.
- MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Matsuo, A., Kita, S., Shinya, Y., Wood, G.C., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, 39, 637–663.
- Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, 20, 121–136.
- Miao, X.C. (1981). Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics*, 8, 109–122.
- Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: defining a role for syntax. *Child Development*, 64, 1665–1687. DOI: <https://doi.org/10.1111/j.1467-8624.1993.tb04206.x>
- Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374. DOI: <https://doi.org/10.1017/S0305000900013817>
- Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58, 221–251. DOI: [https://doi.org/10.1016/0010-0277\(95\)00681-8](https://doi.org/10.1016/0010-0277(95)00681-8)
- Nespor, M., Guasti, M. T., & Christophe, A. (1996). Selecting word order: The rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin: Akademie Verlag.
- Omaki, A., Lassotta, R., Kobayashi, T., Rizzi, L., & Franck, J. (2012). Delay of word order in Japanese? Evidence from a preferential looking study with 19 and 30-month-old children. In N. Miyake, D. Peebles & R. P. Cooper (Eds.), *Proceedings of 34th Annual Meeting of the Cognitive Science Society* (pp. 2826). Austin, TX: Cognitive Science Society.
- R Development Core Team (2015). The R project for statistical computing. R Foundation for Statistical Computing, Vienna, Australia. Retrieved from <http://www.r-project.org>.
- Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, 4(3), 423–436.
- Sun, C. F., & Givón, T. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language*, 61, 329–351.

- Tomasello, M.** (2000). Do young children have adult syntactic competence? *Cognition*, 74, 209–253.
- Tomasello, M.** (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Travis, L.** (1984). *Parameters and the effects of word order variation*. Doctoral dissertation, Massachusetts Institute of Technology.
- Tsai, D. W.-T.** (2008). Object specificity in Chinese: A view from the vP periphery. *The Linguistic Review*, 25, 479–502.
- Van Bergen, G.** (2006). *To ba or not to ba: Differential object marking in Chinese*. M.A. Thesis, Radboud University Nijmegen.
- Wexler, K.** (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23–79.
- Yang, X. L., Shi, R.-S., & Xu, K. L.** (2018). Grammatical aspect in early child Mandarin: Evidence from a preferential looking experiment. *Journal of Psycholinguistic Research*, 47(4), 1–20. DOI: <https://doi.org/10.1007/s10936-018-9590-7>
- Yeh, M. Y. C.** (2015). *The role of maternal input in early word order acquisition: The case of Mandarin Chinese*. Doctoral dissertation, University of Connecticut.
- Yokoyama, M., & Miyata, S.** (2017). *Yokoyama Corpus*. Pittsburgh, PA: TalkBank. Retrieved from <https://childes.talkbank.org/access/Japanese/Yokoyama.html>
- Yu, S. Y., & Tamaoka, K.** (2018). Age-related differences in the acceptability of non-canonical word orders in Mandarin Chinese. *Lingua Sinica*, 4(3). DOI: <https://doi.org/10.1186/s40655-018-0035-x>
- Zhang, Z., Shi, R. S., & Li, A. J.** (2015). Grammatical categorization in Mandarin-Chinese-learning infants. *Language Acquisition*, 22(1), 104–115.
- Zhou, J.** (2001). *Pragmatic development of Mandarin speaking young children: from 14 months to 32 months*. Doctoral dissertation, The university of Hong Kong. Available from HKU Thesis Online (HKUTO).
- Zhu, J.T., & Gavarró, A.** (2019). Testing language acquisition models: Null and overt topics in Mandarin. *Journal of Child Language*, 46(4), 707–732. DOI: <https://doi.org/10.1017/S0305000919000114>

Cite this article: Zhu J, Franck J, Rizzi L, Gavarró A (2022). Do infants have abstract grammatical knowledge of word order at 17 months? Evidence from Mandarin Chinese. *Journal of Child Language* 49, 60–79. <https://doi.org/10.1017/S0305000920000756>