Cognitive Development 25 (2010) 30–45

Contents lists available at ScienceDirect

Cognitive Development

Young children's comprehension of English SVO word order revisited: Testing the same children in act-out and intermodal preferential looking tasks

Angel Chan[a,b,*], Kerstin Meints[c], Elena Lieven[a], Michael Tomasello[a]

[a] Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
[b] The Hong Kong Polytechnic University, Hong Kong, China
[c] University of Lincoln, Lincoln, UK

ARTICLE INFO

Keywords:
Early linguistic/cognitive representations
Word order comprehension
Act-out
Intermodal preferential looking

ABSTRACT

Act-out and intermodal preferential looking (IPL) tasks were administered to 67 English children aged 2–0, 2–9 and 3–5 to assess their comprehension of canonical SVO transitive word order with both familiar and novel verbs. Children at 3–5 and at 2–9 showed evidence of comprehending word order in both verb conditions and both tasks, although children at 2–9 performed better with familiar than with novel verbs in the act-out task. Children at 2–0 showed no evidence of comprehending word order in either task with novel verbs; with familiar verbs they showed competence in the IPL task but not in the act-out task. The difference in performance for familiar and novel verbs from the same children at 2–0, on the IPL task, and at 2–9, on the act-out task, is consistent with the hypothesis that early linguistic/cognitive representations are graded in strength, with early representations still weak and very task dependent. However, these representations also become more abstract with development, as indicated by the familiarity effect even in the more sensitive IPL task.

In their spontaneous speech English-speaking children use the normal subject–verb–object (SVO) word order in transitive constructions fairly consistently from quite early in development (Brown, 1973). Similarly, in comprehension tasks, children as young as 2 years of age appropriately act-out

* Corresponding author at: Room GH632, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hunghom, Hong Kong. Tel.: +852 2766 7439; fax: +852 2334 0185.
E-mail address: ctaachan@polyu.edu.hk (A. Chan).

0885-2014/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2009.10.002

"The doggie bites the cat" (reversible transitives) that depend exclusively on a knowledge of canonical word order (DeVilliers & DeVilliers, 1973). We cannot conclude from these facts, however, that very young children have full productive control of English SVO word order. Unless we know what children have and have not heard, it might be that they are simply reproducing the ordering of the particular words they have heard adults using, or they might be marking SVO relations syntactically, but only locally for some highly familiar verbs (Tomasello, 1992).

One source of evidence for productivity is overgeneralization errors such as *She falled me down* or *Don't giggle me*, in which intransitive verbs are used productively in the SVO transitive frame (Bowerman, 1982; Pinker, 1989). However children seldom make such errors prior to about 3 years of age. Another source of evidence comes from experiments with novel verbs, enabling researchers to control what the child has and has not heard. In the case of English transitive word order, the experiments often involve elicited production. In these experiments, children are exposed to a novel verb in a syntactic construction, such as an intransitive or passive sentence, and then tested to see if they can use that verb productively in the canonical SVO transitive construction, with cues to syntactic roles other than word order (e.g., animacy of the S and O participants, use of case marked pronouns) carefully controlled. Experiments of this type have clearly demonstrated that by 3–6 to 4 years of age most English-speaking children can readily assimilate novel verbs to an abstract SVO schema (Maratsos, Gudeman, Gerard-Ngo, & DeHart, 1987; Pinker, Lebeaux, & Frost, 1987). That is, the vast majority of children from by this age can produce a canonical transitive SVO utterance with a completely novel verb (e.g., "The cat's meeking the dog") that they have never heard before.

However, this is not true of younger children. In a series of studies by Tomasello and colleagues, 2- to 3-year-old children readily used a novel verb in the construction in which they heard it modeled, but they did not use it in constructions, including the SVO transitive construction, in which they had not heard it modeled (see Tomasello, 2000, 2003, for reviews). Similar findings appear based on other elicited production methods with different response demands, such as the weird word order method of Akhtar (1999; see also Abbot-Smith, Lieven, & Tomasello, 2001, and Matthews, Lieven, Theakston, & Tomasello, 2005), and the priming methodology of Savage, Lieven, and Tomasello (2003; see also Huttenlocher, Vasilyeva, & Shimpi, 2004). In all cases, children below about 3 years of age are generally poor at producing newly learned verbs in constructions in which they have not heard them used previously.

Despite these consistent findings using different production methodologies, some researchers argue that production experiments are too demanding from a performance point of view (although this criticism has not been leveled against the priming methodology). The classic test of comprehension is the act-out task. For example, Akhtar and Tomasello (1997, Study 2) taught children the novel verb *dacking* (as a single word) for a canonical transitive action. They then, using novel characters, instructed children, "Show me. Cookie Monster is dacking Big Bird." All 10 children aged 3–8 were excellent at this task, whereas only 3 of the 10 children aged 2–9 performed above chance, even though most did well on a control task using familiar verbs. However, this task is still somewhat demanding because children must learn the novel verb. Therefore, Akhtar and Tomasello (1997, Study 3) created a modified act-out task for use with novel verbs. Children first learned to act-out a novel action on a novel apparatus with two toy characters; only then (their first introduction to the novel verb in a transitive frame) did the adult hand them two new characters and request (while pushing the apparatus in front of them), "Show me. Cookie Monster is meeking Big Bird." In this case children's only exposure to the novel verb was in a transitive sentence frame used for an action they already knew how to perform. (They did not have to learn the verb.) Since every child knew the names of the novel characters and on every trial attempted to make one of them act on the other in the appropriate way, the only question was which character should play which role. The under-3-year-old children, as a group, performed at chance level on this task, with only 3 of the 12 children performing above chance as individuals.

The other major technique used to assess children's comprehension of English word order is intermodal preferential looking (IPL). In this technique, a child is shown two displays (often on two television screens) and hears a single utterance (through a centrally located loudspeaker) that describes only one of the displays felicitously. The question is which display the child will prefer to look at. The IPL task places minimal performance demands on children, involving essentially looking

behaviors as the required response mode. Although there are findings of children's comprehension of English word order using familiar verbs (Hirsch-Pasek & Golinkoff, 1996), and findings concerning verb semantics using novel verbs (e.g., the difference between causative and inchoative meanings; Naigles, 1990), there is only one study testing SVO comprehension with novel verbs. Gertner, Fisher, and Eisengart (2006) used a modified IPL methodology to investigate early word order comprehension with novel verbs and found that children as young as 21 and 25 months old preferred watching the causative event that matched the reversible SVO sentence they heard. This study provides the first evidence of verb-general, abstract comprehension of early word order in young 2-year-old English-speaking children. However, this study had a training phase that used the same characters and nouns in SVO sentences with familiar verbs as in the test phase with the novel verbs, and this may have influenced the children's looking behavior, a point we return to.

Thus, the IPL methodology seems to suggest an earlier age for comprehension of verb-general SVO word order in English than does the act-out methodology. It is possible that this discrepancy simply reflects the different performance demands of the two tasks. But it is also possible that it reflects different kinds or different strengths of underlying syntactic representations. That is, if one takes an exemplar-based approach to the development of syntactic representations, it could be that the abstraction process is a gradual one and that the different methods enable children to access representations of different strengths (Abbot-Smith, Lieven, & Tomasello, 2008; Abbot-Smith & Tomasello, 2006; Tomasello, 2003; Tomasello & Abbot-Smith, 2002). Thus, Chang, Dell, and Bock (2006) argue that there are different requirements involved in looking to one of two visual events that match a sentence (as in IPL), as opposed to creating a matching event to the sentence heard (as in act-out), and that these interact with the time course of learning.

In the current study, therefore, we attempted to resolve the discrepancy between these different findings by using the two different comprehension methodologies with the same children—at three different ages. First, we administered the improved act-out task of Akhtar and Tomasello (1997) to children at a younger age than in any previous study (coupled with some further methodological improvements) using both familiar and novel verbs. Second, we administered an IPL task to the same children, again using both familiar and novel verbs. This is the first time to our knowledge that the same children's performance has been compared using these two methodologies. By using two converging methods aimed at the same underlying competence, we attempted to resolve the discrepancy in the literature and thus to determine more precisely what exactly young children's performance in comprehension tasks reveals about their knowledge of canonical English word order.

1. Method

1.1. Participants

Sixty-seven British children participated in the study: 23 aged 2–0 (M = 2–0.10, range 1–11 to 2–1); 22 aged 2–9 (M = 2–8.24, range 2–7 to 2–10); and 22 aged 3–5 (M = 3–5, range 3–4.7 to 3–7), with approximately equal numbers of boys and girls at each age. All children came from middle-class families, were born full-term and healthy, and were exposed to English as their only language.

An additional 31 children were tested but excluded for the following reasons: child could not sit through tasks (n = 11); child looked only to one side of the screen during the salience trial in two out of three verb trials in the IPL task (n = 2); child had difficulty identifying all the animal characters (n = 6); child did not arrive for the second task within 2 week interval (n = 7); experimenter or video error (n = 2); child later found to be bilingual (n = 2) or born prematurely (n = 1).

1.2. Design and procedures

A British English version of the Communicative Development Inventory (Fenson et al., 1993; Hamilton, Plunkett, & Schafer, 2000) was first administered to check specific words in children's vocabulary that they needed to know for testing purposes. Each child then participated in two comprehension tasks: act-out and intermodal preferential looking (IPL), counterbalanced for order across children.

The act-out task was based on that of Akhtar and Tomasello (1997: study 3, 'no highlighted' condition), although with several improvements. First, for best fit with the transitive, only verbs of direct causation were used. Second, when requesting the child to act-out the action we used a simpler directive (instead of "Can you make the X VERB the Y?" we said "Look! The X is VERBing the Y. Can you do it?"). Third, there were fewer trials per child.

The preferential looking methodology followed that of Meints, Plunkett, and Harris (1999) and Meints, Plunkett, Harris, and Dimmock (2002). Children were seated on the parent's lap facing a large back projection screen at a distance of 90 cm apart in a darkened testing cubicle. Images or video displays were shown at the child's eye-level. A loudspeaker presenting the auditory-stimulus was positioned centrally above the screen. The light from the displayed images allowed subsequent analysis of eye fixations. Two miniature cameras were placed immediately above the large screen adjusted to an angle to record the infant's eye- and head-movements to the left and right side of the screen. Parents were asked to close their eyes during testing and listened to instructions played over headphones. All parents were shown all images and displays following all testing.

During testing, the experimenter (E) was in an adjacent, but separate control room to operate the computer program for running IPL. Cameras were connected to a video and monitor via a video-mixer that permitted video recording of a split-screen 'twin image' of the child during testing. This split-screen image was displayed on the monitor in the control room so E could observe the child on the monitor. A new trial was not launched until the child had centered his or her attention either spontaneously or by being attracted to a flashing red light at the center of screen. Inter-trial intervals lasted for a minimum of 0.5 s and varied according to the child's attention to the task.

Both tasks, act-out and IPL, consisted of three phases: (1) the demonstration phase, in which the objects/actions were introduced; (2) the training phase, in which the child learned the verb-action mapping; and (3) the test phase, in which the child was tested for comprehension of word order. In the demonstration and training phases of both tasks, the verbs were presented without any syntactic marking of agent or patient roles of the causative events (Look! VERB-ing!). In the test phase, the verbs were used in transitive SVO sentence frames. Both testing sessions were videotaped for later coding (see below).

Verb stimuli consisted of two sets of verbs, one used for each task (counterbalanced). There were three verbs in each set, two novel and one familiar. The verbs in one set were 'meek', 'tam', and 'push', and in the other set were 'gorp', 'pilk', and 'kick'. The actions corresponding to these verbs, together with the accompanying apparatuses, are listed in Table 1. The order of verb presentation among the

Table 1
Novel and familiar verbs, their meanings and the accompanying apparatuses.

Verb	Description
Meek	Resting position: Both animals stand on the protruded part of a spherical toy which can move up and down. Causative action: Agent pushes patient down. Upright patient moves up and down.
Pilk	Resting position: Both animals each stand on one spinning handle on a rectangular platform, one on each side. Causative action: Agent spins patient around.
Gorp	Resting position: Both animals stand on the circular platform of a rocking chair apparatus comprising of a large circular platform on top and 'rocking' legs at the bottom. Causative action: Agent rocks patient back and forth.
Tam	Resting position: Both animals are hung onto a tripod set-up having two strings hanging down and a piece of Velcro attached to each of the strings. Each animal is attached to a piece of Velcro. Causative action: Agent makes patient swing.
Push	Resting position: Both animals stand at the start of a track with a large circular hole on the other end. Causative action: Agent pushes the patient along the track and down into the hole.
Kick	Resting position: Both animals stand on the top of a slide. Causative action: Agent pushes patient down the slide.

three verbs within a set (within a task) was also counterbalanced across children. The four novel verbs ('meek', 'tam', 'gorp', 'pilk') were chosen to meet the following criteria: monosyllabic in form, feasible in both tasks (child can act-out action, action readily filmable for IPL), involving direct causation with change of state in patient, and reversible between the two participants. The familiar verbs 'push' and 'kick' were chosen from the action words on the Infant CDI to meet these same criteria. We established that these two verbs were also in our children's comprehension repertoire from their parental reports on the Infant CDI.

For noun stimuli, 10 pairs of plastic toy animals were used to enact the actions. (All 20 were among the most frequent animal words in the comprehension repertoire of British English children up to 27 months of age in the Oxford CDI Database; Hamilton, Plunkett, & Schafer, 2000.) In addition, children could identify all the characters by their names, as reported by their mothers on the Infant CDI; we also administered an animal identification task before each of the two main tasks (see below). The 10 pairs were used as follows: One pair (monkey–bee) was used for demonstrating use of the apparatus in the act-out task, and 6 pairs for counterbalancing between the act-out (the first test trials) and the IPL tasks. These 6 pairs were divided into two casts—duck–bunny, bird–fish, dog–pig; and sheep–cat, horse–cow, bear–lion; the use of these two casts was counterbalanced across task. The final 3 pairs were used for the act-out second trials (elephant–tiger, spider–frog, chicken–mouse), as each child received 2 test trials per verb in act-out and 1 test trial per verb in IPL. A typical trial set for IPL and Act-Out was: IPL (Verb 1: duck–bunny; Verb 2: bird–fish; Verb 3: dog–pig); Act-Out first trials (Verb 1: sheep–cat; Verb 2: horse–cow; Verb 3: bear–lion) and Act-Out second trials (Verb 1: elephant–tiger; Verb 2: spider–frog; Verb 3: chicken–mouse). Which member of the pair was the agent of the action was also counterbalanced across children.

1.2.1. Act-out task

The experimenter (E) first introduced the animal toys in groups of four by their names and then encouraged the child to point to or to name the relevant animal toy. Clarifications and corrections were made when necessary. As reported above, children included in our data analyses responded appropriately for virtually every animal.

For the demonstration phase, E took out an apparatus and said "Look! We are going to play a game. I'll show you first and then it's your turn. Watch carefully; you have to do it exactly like me." The child saw two animate toys (the demonstration animal pair bee and monkey) presented side-by-side on the apparatus while watching the left (or the right) animate toy being picked up by the experimenter to act causatively on the other one twice with the accompanying description "Look! this is VERB-ing!."

For the training phase, E then said "Now it's your turn, can you do it?" placing the bee and the monkey on the apparatus in resting position and pushing the apparatus forward toward the child. The child then performed the action with the demonstration animal pair, while E commented "Yes, this is VERB-ing!". Demonstration and/or correction by physical manipulation were used when the child had difficulty in imitating the action properly.

E then reversed the agent–patient roles and demonstrated the action twice again with the appropriate description. A second training phase in which the child was again given a turn ensured that the child attempted to make one character act on the other in a causative manner on two consecutive trials involving role reversals before proceeding to the test phase.

In the test phase, E began by holding the apparatus back toward her side and said "Now you know how to VERB, so I have two new animals for you." E took out two new animal characters (from the 'shared' set assigned for that child), placed them side-by-side on the apparatus in resting position and said "*Look! The A is V-ing the B! Can you do that?*"[1]. E then pushed the apparatus toward the child and repeated "*Look! The A is V-ing the B! Can you do that?*" and looked expectantly to the child. (This was the first time the child had heard the test verb in a transitive frame.) Following the child's response, a second identical trial with a new animal pair was administered. Across trials for a given child, the

[1] We tried "Show me: The A is V-ing the B!" in the pilot, but found that "Look! The A is V-ing the B! Can you do that?" was a better prompt in getting and maintaining children's attention in the act-out task, especially for the youngest children at 2–0.

Table 2
The structure of each verb trial series in the IPL task.

Trial	Left side of the screen	Center of the screen	Right side of the screen	Auditory stimuli
Trial 1 (demonstration)		1st video presented showing, e.g., X VERB-ing Y		Look!
Trial 2 (training)		1st video presented again showing, e.g., X VERB-ing Y		Look! VERB-ing!
Trial 3 (demonstrtion)		2nd video presented showing, e.g., Y VERB-ing X		Look!
Trial 4 (training)		2nd video presented again showing, e.g., Y VERB-ing X		Look! VERB-ing!
Trial 5 (baseline trial)	Showing X VERB-ing Y		Showing Y VERB-ing X	Look!
Trial 6 (test trial)	Showing again X VERB-ing Y		Showing again Y VERB-ing X	Look! X is VERB-ing Y! *or* Look! Y is VERB-ing X!

resting position of the agent was left or right in alternating fashion. Each child experienced six trial sets, four trial sets with novel verbs and two trial sets with familiar verbs.

The child's response was coded into one of four categories correct; role reversal error; irrelevant action (including responses in which the child acted as agent); and no attempt. E praised the child verbally (and sometimes gave stickers) for engaging with the task regardless of the correctness of response. Responses of 4 children from each age group (18% of the data) were coded independently by a second scorer. Inter-rater agreement was 100%.

1.2.2. Intermodal preferential looking task (IPL)

The visual stimuli used the same animal toys and apparatuses as the act-out task. The causative actions between the animal toys on the apparatuses were filmed to create a video-version for the IPL task. (Animal toys were manipulated by invisible fishing lines, so no hands were visible.)

For initial recognition of the animal toys in IPL, digital images of the animal toys were used. The six animals involved in the IPL task (the 3 pairs from the shared set not used in the act-out task) were first introduced with the images appearing one by one at the center of the screen, with the accompanying auditory-stimulus "Look, look at the [ANIMAL NAME]!." Following this initial round of character introduction, E presented the two paired animals one on the left side the other on the right side of the screen and named only one of them: "Look, look at the [ANIMAL NAME]!' There were six trials in this phase, using six animal characters (in three pairs), each trial naming one animal. Therefore, each pair appeared twice on the screen in this phase, each time naming one of the two animals. The Left/Right sidedness of the matching image was alternated across trials within child.

As in the act-out task, there were three phases in the testing of each verb—demonstration, training, and test. Table 2 shows the structure of this series. Each series consisted of six trials. The first trial served to introduce the action to the child. The second trial repeated the same animation but with the accompanying "Look! VERB-ing!" auditory-stimulus, thus serving to train the child to map the verb to the action, with no information relevant to word order marking. The third and fourth trials were again the same except that the images presented involved role reversal of agent and patient. All the animated images in these four trials (for demonstration and training) were shown at the center of the screen. Each video clip lasted for 7000 ms during which 3–4 of the same actions were repeatedly performed (Meints, Plunkett, & Harris, 2008). During the fifth trial, children were shown simultaneously two animations on each side of the screen depicting pairs of transitive events. These differed only in role reversal between agent and patient. The pair was edited so that the animation was synchronized throughout. The accompanying auditory-stimulus was "Look!" without any indication to draw the child's attention to a particular image. This served as the baseline trial (called salience trial by Hirsch-Pasek & Golinkoff, 1996), based on which we determined whether preference to one image existed before (and independent of) the target auditory-stimulus in the test trial. The sixth trial was the test trial, during which E repeated the visual stimulus in exactly the same way as in the fifth trial, but while playing the auditory-stimulus "Look! The X is VERB-ing Y!" Similar to the act-out task, the test trial was the child's first introduction to the novel verb in a transitive sentence frame. Children heard no prior syntactic marking of agent and patient roles prior to the test trial. In addition, children were not subject to the other potential confounds that we identified in the 2006 study by Gertner et al. Our children were tested using multiple characters and had not seen the same character being the actor in both videos while hearing that character's name mentioned first in an NVN frame. There were three verb trials per child in IPL, one with a familiar verb and two with novel verbs.

The sound files for the still animal images in the schema "Look! Look at the [ANIMAL NAME]!" were matched for overall length (5500 ms) as well as for the time to onset of the target word (2750 ms) (Meints, Plunkett, Harris, & Dimmock, 2002). The sound files for the action videos in the schema "Look! The AGENT is VERB-ing the PATIENT!" were matched for overall length (7000 ms), the time to the onset of the AGENT word (1400 ms), and the time to the sentence offset (3800 ms).

To ensure that the change in looking preference was driven by the linguistic stimulus, both animals within an animation appeared in continuous motion from the beginning, before the agent initiated the causative action on the patient, so that the child could not locate the agent by merely attending to which character moved first (Hirsch-Pasek & Golinkoff, 1996). Also counterbalanced were the first video presented (whether E showed the child the matching image or the non-matching image first

during the demonstration and training); the Left/Right sidedness of the agent within a video; and the Left/Right sidedness of the matching image.

The video recording of children's looking behavior was used to assess direction of eye gaze during an experimental trial. This assessment of the video recording (played back at standard speed) was carried out off-line after data collection was completed. The coder used a button-box to trigger a data-registration program that synchronized with the video recording. Each trial was scored twice by the same coder for the child's looks to the left side and twice for the child's looks to the right side. The coder was blind to the left–right location of the target image. The two scores for each side (length of looking in milliseconds) were then averaged, yielding two measures (one for left and one for right) for each trial. The coding was accepted only when the intra-rater correlations between the two scorings on each side were both at least .90. Correlations reached as high as .98 on average for the first author, who scored all the trials. An additional coder scored 12% of the data, with 91% agreement.

Similarly to Gertner, Fisher, and Eisengart (2006), we report the proportion of looking time to the matching video ('matching' with respect to the auditory-stimulus heard) as the measure of target preference for our data analyses, referred to as Target looking (TL). In the case of the single word animal identification task, this measure was calculated both before and after the onset of the target word (at 2750 ms). For the verb trials, the onset of the agent word (at 1400 ms) in the template, "Look! The AGENT is VERB-ing the PATIENT!" was used as the starting point for the analysis. The proportion of looking time to the matching video was calculated from the start point at 1400 ms till the end of the trial from both the baseline trial (without directive auditory-stimulus) and the test trial (with directive auditory-stimulus describing only one of the two images). For ease of distinction, we call them pre-auditory-stimulus and post-auditory-stimulus phases here. Following the logic of existing IPL studies, we expected that if children comprehended the word order of the heard sentence, their proportion of looking time to the matching image (TL) would be significantly higher in the post-auditory-stimulus phase than in the pre-auditory-stimulus phase.

2. Results

2.1. Act-out task

Results indicated that older children acted out the appropriate action with both characters in their appropriate roles more often than younger children, with the children at 3–5 being correct more often than the children at 2–9, who were in turn correct more often than the children at 2–0. This is indicated by a main effect of Age $F(1,64) = 56.69$, $p < .001$, $\eta_p^2 = .64$ from a 3×2 ANOVA of Age group (2–0, 2–9, 3–5) $\times$ Verb condition (novel vs. familiar verb)[2]; and significant differences in the percentage accuracy between children aged 3–5 and aged 2–9, and between children aged 2–9 and aged 2–0 ($p < .001$ in both cases), from post hoc pair-wise comparisons using the Fischer's LSD procedure. Also, over all age groups children produced correct answers more often with familiar verbs than with novel verbs, indicated by a main effect of Verb condition $F(1,64) = 4.33$, $p = .041$, $\eta_p^2 = .063$.

As we were interested in the behavior of each age group for the two different types of verbs (familiar vs. novel), findings were also analyzed separately for each age group. As expected, the children at 3–5 performed above chance for both the novel and familiar verbs, $t(21) = 21$, $p < .001$ in the familiar verb condition, $t(21) = 43$, $p < .001$ in the novel verb (one-sample t-tests), with no difference between them, $t(21) = -1$, $p = .329$. In contrast, the youngest children at 2–0 were not above chance for either of the two verb types and were actually below chance (50%) with the novel verbs, $t(22) = -3.27$, $p = .003$, one-sample t-test; they were at chance if trials with no attempt or an irrelevant action are excluded—with no reliable difference in performance between the two verb conditions, $t(22) = .843$, $p = .408$. The children at 2–9 were clearly above chance for the familiar verbs, $t(21) = 5.14$, $p < .001$, one-sample t-test, and marginally so for the novel verbs, $t(21) = 2.03$, $p = .056$, one-sample t-test. Children at this age

[2] The initial analysis was a $3 \times 2 \times 2$ ANOVA of Age group (2–0, 2–9, 3–5) $\times$ Verb condition (novel vs. familiar verb) $\times$ Task order (Act-out first vs. IPL first), with the number of correct responses out of the total number of trials as dependent measure. Since task order produced no significant main effects or interactions, it was removed from further analyses.

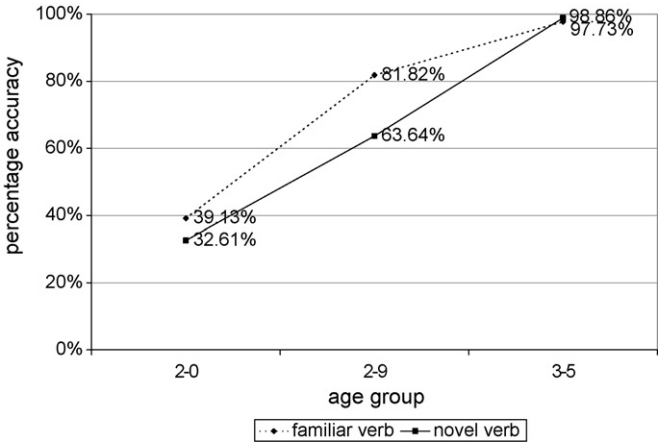

Fig. 1. Results for different ages in the familiar and novel verb conditions of the act-out word order task.

Table 3
Mean proportion of looking time to the matching video in the IPL animal character identification task.

Age	Pre-auditory-stimulus Trial Phase Mean (standard error)	Post-auditory-stimulus Trial Phase Mean (standard error)	Difference between pre- and post-auditory-stimulus looking *t* and *p* values
2–0	.48 (.01)	.54 (.02)	$t(22) = -1.62$, $p = .060$ (see note below)
2–9	.52 (.01)	.57 (.02)	$t(20) = -2.42$, $p = .013$
3–5	.47 (.01)	.60 (.02)	$t(19) = -5.13$, $p = .000$

Note: Non-parametric Wilcoxon test yielded significant result at age 2–0: $Z = -1.73$, $p = .043$ (exact sig., one-tailed).

were correct more often with the familiar verbs than the novel verbs, $t(21) = 2.25$, $p = .035$, as shown in Fig. 1.

Analysis using as a dependent measure the total number of attempted trials only (i.e., excluding those in which the child made no response), yielded an identical pattern of significant results.

2.2. Intermodal preferential looking task (IPL)

We first checked children's understanding of the animal names; and children's *baseline* looking preference in the familiar and novel verb conditions.

2.2.1. Animal names

Children at 2–0, 2–9 and 3–5 localized the referents of the animal names reliably, a precondition for the IPL word order task. This was shown by a significant main effect for Trial Phase, $F(1,61) = 25.14$, $p = .000$, $\eta_p^2 = .29$, in a 3×2 ANOVA including Age (2–0, 2–9, 3–5) and Trial Phase (pre- vs. post-auditory-stimulus phase). Over all age groups children looked significantly more at the "matching" animal image upon hearing it named, compared to their baseline looking behavior before hearing the auditory-stimulus. Also, there were significant or tendency-toward-significant effects between pre- and post-auditory-stimulus looking times for all three age groups, based on both one-tailed planned *t*-tests and non-parametric Wilcoxon tests. Task order and gender had no effects. Means appear in Table 3.

2.2.2. Baseline looking preference

Importantly, results suggested that children did not have any biased preference to one of the two videos when hearing only a non-directive "Look!" in the pre-auditory-stimulus phase (the base-

line trial). That is, the proportion of looking time to one image vs. the other did not differ from chance (0.5) for the familiar verb condition, $t(18) = -1.18$, $p = .255$ at age 2–0; $t(18) = .059$, $p = .954$ at age 2–9; $t(16) = .881$, $p = .392$ at age 3–5 (one-sample t-tests), nor for the novel verb condition, $t(20) = -.719$, $p = .481$ at age 2–0; $t(21) = -.663$, $p = .514$ at age 2–9; $t(21) = .646$, $p = .525$ at age 3–5 (one-sample t-tests). In both cases looks to each image were also not significantly different from each other. For familiar verbs, $F(2,54) = 1.18$, $p = .317$; for novel verbs, $F(2,64) = .637$, $p = .532$). In addition, the corresponding effect sizes (indicated by the eta square values) were all small (familiar verb condition $\eta^2 < .13$ in all cases; novel verb condition $\eta^2 < .05$ in all cases) and unlikely to reach significance.

2.2.3. Familiar verb condition

Results comparing pre-naming baseline looking with post-naming in the familiar verb condition indicated that over all age groups, children looked significantly longer to the matching video (pre-auditory-stimulus baseline mean TL = .490, post-auditory-stimulus mean TL = .604). This is evidenced by a highly significant main effect of Trial Phase, $F(1,52) = 8.312$, $p = .006$, $\eta_p^2 = .138$, based on a 3×2 ANOVA involving Age (2–0, 2–9, 3–5) and Trial Phase (pre- and post-auditory-stimulus). In addition, the difference between pre- and post-auditory-stimulus looking (indicated by the highly significant main effect of Trial Phase above) was consistent across the three age groups. This is indicated by a non-significant interaction between Trial Phase and Age, $F(2,52) = .38$, $p = .69$, $\eta_p^2 = .014$. Task order yielded neither main effect nor interactions with other variables. No other effects reached significance.

2.2.4. Novel verb condition

The major findings are based on the second novel verb trial, since initial analyses revealed no significant effects for the first novel verb trial[3] (not uncommon among IPL studies). For the second trial, two dependent measures were used. First, we compared children's proportion of looking time to the "matching" video before vs. after hearing the auditory-stimulus for six subgroups (the three age groups split by whether they received IPL as the first or second task). These post hoc comparisons were licensed by an interaction between Trial Phase, Age and Task Order based on a $3 \times 2 \times 2$ ANOVA of Age (2–0, 2–9, 3–5) $\times$ Trial Phase (pre- and post-auditory-stimulus) $\times$ Task order (IP first vs. IPL second).[4] The post hoc comparisons were carried out using the Fischer's LSD procedure (with p values adjusted due to multiple testing). Means appear in Table 4.

We also looked at the number of children who watched the matching video more than the distractor video after hearing the auditory-stimulus (called the 'target' lookers here), and determined the probability of this number of 'target' lookers occurring simply by chance, as done in other IPL studies (Hirsch-Pasek & Golinkoff, 1996; Naigles, Bavin, & Smith, 2005). We not only looked at the number of 'target' lookers based on the entire post-auditory-stimulus interval, but also each of the "sub-intervals". This was done by dividing the entire post-auditory-stimulus interval (i.e., from the onset of the agent word at 1400 ms to the end of the test trial) into three sub-intervals: first 2 s, second 2 s, and last 1600 ms, similarly to Gertner, Fisher, and Eisengart (2006). Doing so allowed us to observe in greater detail the time course of the stimulus effect along the trial.

[3] The initial analysis was a $3 \times 2 \times 2 \times 2$ ANOVA of Age (2–0, 2–9, 3–5), $\times$ Trial (the first trial and the second trial) $\times$ Trial Phase (pre- and post-auditory stimulus) $\times$ Task order (IP first vs. IPL second). There were a highly significant interaction between Trial and Age, $F(2,40) = 5.81$, $p = .006$, $\eta_p^2 = .225$, a tendency-toward-significant interaction between Trial, Trial phase and Age, $F(2,40) = 2.61$, $p = .086$, $\eta_p^2 = .115$, as well as a tendency-toward-significant interaction between Age and Task order, $F(2,40) = 2.73$, $p = .077$, $\eta_p^2 = .120$). Since the initial analysis revealed different results between the two novel verb trials, a $3 \times 2 \times 2$ ANOVA of Age (2–0, 2–9, 3–5), $\times$ Trial Phase (pre- and post-auditory stimulus) $\times$ Task order (IP first vs. IPL second) was then carried out for each trial. There were neither significant nor even tending-toward-significant effects registered for the first novel verb trial.

[4] In addition, there were a series of significant results registered for the second novel verb trial, including a tending-toward-significant interaction between Trial phase and Age, $F(2,49) = 2.71$, $p = .077$, $\eta_p^2 = .099$, a tending-toward-significant interaction between Trial phase, Age and Task order, $F(2,49) = 2.61$, $p = .084$, $\eta_p^2 = .096$, a significant effect of Age, $F(2,49) = 3.98$, $p = .025$, $\eta_p^2 = .14$, as well as a highly significant interaction between Age and Task order, $F(2,49) = 5.6$, $p = .006$, $\eta_p^2 = .186$. Since Trial phase interacted with Age and Task order in the second novel verb trial, we reported the results for the two task-order-subgroups at each age group separately.

Table 4
Mean proportion of looking time to the matching video in the IPL word order task (based on the second novel verb test trial).

Age		Pre-auditory-stimulus baseline trial Mean (standard error)	Post-auditory-stimulus Trial Phase Mean (standard error)	Difference between pre- and post-auditory-stimulus looking *p* values
2–0	IPL first	.48 (.06)	.38 (.09)	.30
	IPL second	.51 (.05)	.41 (.07)	.14
2–9	IPL first	.51 (.05)	.49 (.07)	.71
	IPL second	.60 (.05)	.77 (.07)	.02**
3–5	IPL first	.53 (.05)	.62 (.07)	.20
	IPL second	.45 (.05)	.43 (.07)	.68

** Significant at .05 level (*p* values have been adjusted for multiple comparisons.

The two dependent measures yielded a consistent pattern of significant findings, especially for the youngest and the middle age groups. Taken together, the essential results are that children at 2–0 displayed no reliable sign of word order competence with novel verbs, while children at 2–9 and age 3–5 displayed some evidence of word order productivity in IPL. Specifically, no significant results indicating word order productivity were obtained from the youngest 2–0 children regardless of task order and regardless of the dependent measure. Robust significant findings were registered from the middle age group (aged 2–9) on both dependent measures, especially from those who received IPL as the second task. With the first dependent measure (see Table 4), these children looked significantly longer to the matching video *after* vs. *before* hearing the SVO sentence (pre-auditory-stimulus baseline mean TL = .60, post-auditory-stimulus mean TL = .77, $p = .02$). In the interval analyses, 100% of these children (9/9) watched the matching video more than the distracter video during as early as the first 2 s after hearing the sentence stimulus (0–2 s post-auditory-stimulus interval). The probability that this number of "target lookers" occurring simply by chance was highly significant ($p = .002$). As for the children who received IPL as the first task, 70% of these children (7/10) were target lookers, again representing a preponderance of the children, despite the fact that the probability of this number of "target lookers" occurring by chance was non-significant. In the subsequent 2-s post-auditory-stimulus interval (2–4 s), significant results were registered from *both* subgroups: 80% of the children receiving IPL as the first task (8/10) were target lookers ($p = .055$), and 88.9% of the children receiving IPL as the second task (8/9) were target lookers ($p = .02$).

Some significant findings appeared for the oldest children (aged 3–5) who received IPL as the first task. In the interval analysis, 80% of these children (8/10) watched the matching video more than the distractor video during the first 2 s after hearing the sentence stimulus (0–2 s post-auditory-stimulus interval), and the probability that this number of "target lookers" occurred simply by chance was close-to-significantly small ($p = .055$). In the subsequent 2-s post-auditory-stimulus interval (2–4 s), 70% of these children (7/10) were target lookers, again representing a preponderance of the children, despite the fact that with only 1 target looker less, the probability of this number of "target lookers" occurring by chance became non-significant. Note also that on the other dependent measure (see Table 4), these children also looked longer to the matching video upon hearing the directive sentence stimulus, compared to their baseline looking preference (mean proportion of looking time to the matching video *prior to* hearing the sentence (.53) and mean proportion of looking time to the matching video *after* hearing the sentence (.62), although the difference is not significant ($p = .20$).

2.3. Correlation between tasks

We attempted to correlate children's performance on the two tasks in several different ways. In none of these analyses were there any significant correlations. Only in the novel verb condition, with children aged 2–0, who participated second in IPL, during the second 2 s post-stimulus interval of the first verb trial, was there a near-significant positive correlation between act-out percentage accuracy and the mean proportion of looking time to the matching video ($r_s = .578$, $N = 11$, $p = .063$). Overall, the finding is that the two tasks do not correlate.

3. Discussion

Two comprehension methods, act-out and intermodal preferential looking (IPL), were used to assess the same children's understanding of SVO word order as a marker of the agent–patient relations in English transitive construction. We used not only familiar but also novel verbs to assess children's linguistic productivity. We also compared children at different points of development. The main findings are: (a) children at 2–9 and 3–5 showed some evidence of word order competence in both verb conditions across both tasks; (b) with novel verbs, children at 2–0 did not show any reliable sign of word order productivity in either task; (c) with familiar verbs, the same 2–0 aged children comprehended word order in IPL but not in act-out; (d) no significant correlations appeared between the acting-out and preferential looking measures.

To interpret these results, we must be clear what the two tasks are measuring. The act-out task measures the child's accuracy in acting out a sentence according to his or her interpretation of the sentence. It is thus an accuracy measure. In the present study, since the child knew the names of the animal characters and also knew how to perform the causative actions, the only question was which character should play which role as indicated by word order. In contrast, the IPL task measures the child's looking preference to the matching image vs. the non-matching image, that is, 'matching' with respect to the linguistic stimulus heard. It is thus a preference measure. In this technique, a child is shown two images and hears a single utterance that describes only one of the images felicitously. The question is which display the child prefers to look at, in particular, in the present case, whether the child prefers watching the causative event that matches the reversible SVO sentence heard.

It is thus possible that the lack of correlation between the two tasks in this study reflects simply task-specific processing demands—they are measuring somewhat different things. However, one can also look at the discrepancy in performance between the IPL and the act-out methods as reflecting something of how the child cognitively represents linguistic constructions. In a "graded representations" model of cognitive representation, linguistic constructions develop in their abstractness along a graded continuum of strength (Abbot-Smith & Tomasello, 2006). Different methods tap into different strengths of representation. The act-out task needs stronger representations than the IPL task. This would account for why within the same verb condition, the same children showed word order competence in the IPL but not in the act-out tasks.

A related perspective is that the IPL task needs not only weaker representations but only partial representations or knowledge to succeed (Chang et al., 2006). Recall that IPL is a forced-choice matching task. In the case of understanding SVO word order in the active reversible transitive construction, the visual-paired comparison is between an iconic event (the matching event) and a non-iconic event (the role reversal non-matching event) with respect to the sentence heard. In addition, current IPL studies (including Gertner, Fisher, & Eisengart, 2006, as well as the current study) have only used highly prototypical action verbs: the agent role imparts a force on the patient causing the patient to undergo a change of location or motion. In this context, matching in the IPL task could be put down to lower-level association between the order of mention of the two character names (in an active transitive sentence) and the highly salient directionality of force transmission in the causal chain of an iconic event (Croft, 1998: 23; Langacker, 1990). The act-out task, in contrast, involves acting out (creating) the whole matching causative event. Matching the iconic parallels between word sequence and the directionality of energy flow within a causative event in IPL is obviously part of the word order knowledge. In fact, especially in the most prototypical cases, recognizing the salient correspondence between word order and the directionality of energy flow within an iconic event is likely to constitute an important initial part of the form-meaning mapping as children gradually acquire knowledge about how word order marks the agent–patient relations. However, this lower-level association is not sufficient for a full blown knowledge of the syntactic function of word order either in less prototypical cases or when a complete acting out of the sentence is required. Thus we would argue that sensitivity to word order and a full productive control of the English SVO word order are two different things. There are interesting parallels here to debates in other parts of the developmental psychology literature as to theoretical and methodological issues raised by comparing performance in 'discrimination-type' tasks to tasks requiring other, action-based responses (see Hood, Carey, & Prasada, 2000).

Our findings suggest that even when the "more sensitive" IPL method is used to tap early competence, the youngest children, at 2–0, did not exhibit productive understanding of SVO word order that goes beyond their experience. The 2–0 s comprehended transitive sentences like 'The duck kicks the rabbit' with familiar verbs like 'kick' and 'push', but the same children did not comprehend transitive sentences like 'The pig meeks the sheep' when they had not experienced the newly learnt causative verb 'meek' in an SVO transitive frame. From a methodological perspective, they had no general problem acting out the causative action expressed by the novel verb, and the IPL task demanded a minimal looking response. Despite the fact that the sentences are all instantiations of the transitive schema, what differed between the two types of sentences was whether the children had experienced that verb in an SVO transitive frame before. These findings indicate that at early stages of development children show differential and restricted competence in comprehension and that children's linguistic productivity is restricted and tied closely to their linguistic experience.

Both types of comprehension data portray a developmental stage at which there is a dissociation in performance between familiar and novel verb conditions from the same children—at 2–0, captured by IPL, and at 2–9, captured by act-out. The findings are consistent with the usage-based theory of language acquisition: The acquisition of the transitive construction initially involves the acquisition of low-scope patterns based on frequent items in children's linguistic experience. Based on these exemplars children subsequently abstract a more general transitive construction schema and become linguistically more productive. Not only do linguistic representations become less item-based and more abstract/schematic as development proceeds; complexity also develops in terms of emerging patterns of relationships between constructions and their constituents (Lieven & Tomasello, 2008; Tomasello, 2003). However, much theoretical work remains to be done to characterize the precise nature of early linguistic/cognitive representations and how they develop (instance, Goldberg, 2006; Ibbotson & Tomasello, 2009). In this respect two points should be noted in terms of children's grasp of SVO word order more broadly. First, existing studies, whether IPL or act-out, have only tried action verbs, but the notion of syntactic transitivity goes beyond the prototypical action scene of a volitional agent physically acting on a patient (for instance 'The key opened the door' c.f. Dowty, 1991). Second, children acquiring different languages show productive understanding of SVO word order with novel verbs at different ages in the same act-out task and this is closely related to how informative the cues to this are in the language they are learning (Chan, Lieven, & Tomasello, 2009, for children learning English, German and Cantonese).

These results raise several important methodological issues. First, one might ask whether our negative results with novel verbs from the youngest, 2–0 children arose from methodological problems with our IPL task. The fact that the same 2–0 children were successful with familiar verbs (and animal names) in the same IPL paradigm excludes this possibility. Moreover, children at 2–9 and 3–5 displayed some reliable evidence of novel verb word order productivity, lending further credibility to the IPL procedure. Gertner, Fisher, and Eisengart (2006) found earlier competence with novel verbs; however, as noted earlier, their positive results might be due to the training children received in the initial practice phase. Thus, Dittmar, Abbot-Smith, Lieven, and Tomasello (2008) could only replicate Gertner, Fisher, and Eisengart (2006)'s IPL when children received a practice phase containing full transitive sentences with familiar verbs and the same nouns as those used in the test phase; with a more neutral practice phase children were unable to succeed. These findings suggest that for children this young to succeed in a word order test with novel verbs in IPL, they must first go through some kind of learning or priming period in which the additional linguistic experience somehow prepares them for the test (Dittmar et al., 2008). The contrast between the methods in these two studies again raises the question of the nature of children's linguistic representations at this early point of development.

The second possibility is that our IPL task is more demanding than that in Gertner, Fisher, and Eisengart's (2006) study. In their study, the two video stimuli actually involved different actions, whereas ours involved the same action. Having two videos with the same action but differing only in role reversal might be a more demanding word order test for young children. Although our positive IPL results with familiar verbs at all ages exclude the possibility that the video pairs were not discriminable to our children, it is still possible that in the novel verb condition when the similar video pair of the novel action was coupled with the novel verb auditory-stimulus, all of this became too demanding in an on-line preferential looking task for the youngest children.

How robust, then, is the IPL method? In the present study we found trial order effects, task order effects, and lack of positive results with some of the 3–5 s. Regarding trial order effect, this is not uncommon among IPL studies. Previous IPL studies such as Hirsch-Pasek and Golinkoff (1996) and Dittmar et al.'s (2008) testing the early comprehension of SVO word order (with familiar verbs in Hirsh-Pasek & Golinkoff and with novel verbs in Dittmar et al.) also found negative results with the first verb trial, and positive results in subsequent trial(s), suggesting that children might need "a bit more exposure to the stimuli to warm to the task" (Hirsh-Pasek & Golinkoff, 1996, p. 113).

Regarding the task order effect with the children aged 2–9, here the act-out task might have worked as a kind of pre-training. In the act-out task, children had to create, and therefore attend to, a visual event that matched what they had heard. Thus the act-out task might have trained the children to attend to a visual event that matched what they heard, which in turn might have helped them orient faster toward the matching image in IPL than their peers who received the IPL first.

As for the lack of positive results with some of the oldest group (those who received IPL as the second task), it is important to note that this happened in the IPL task but not in the act-out task. Recall from Fig. 1 that the 3–5 children displayed robust evidence of word order productivity with novel verbs in the more demanding act-out task, scoring above 95% accuracy as a group with no task order effect. It could be that the apparent task order effect observed among the 3-year-olds in the IPL novel verb task reflects the use of a visual fixation (preference) measure that may no longer be as sensitive for older children at 3–5 as for the younger children at 2–9. Findings from other IPL studies have also indicated fragility in using the mean proportion of looking time to the matching video as a dependent measure with children at this age. For instance, Kidd, Bavin, and Rhodes (2001) tested a group of English-speaking children between 28 and 34 months old with an IPL paradigm and found that while the children were capable of mapping transitive argument structure frames with causative events and intransitive argument structure frames with non-causative events—even with novel verbs, the same children showed negative results with familiar verbs, which should be a less demanding condition. This pattern (positive IPL results in a more demanding condition, but negative IPL results in a less demanding condition) was also found in IPL studies by Meints, Plunkett, and Harris (1999), and Meints, Plunkett, Harris, and Dimmock (2002, 2004), and is similar to our findings for the 3–5 age group, who displayed robust evidence of word order productivity with novel verbs in the more demanding act-out task, scoring above 95% accuracy, but those receiving the IPL task second showed no positive results in the less demanding IPL task with novel verbs. In addition, other IPL studies such as Gertner et al.'s (2006) that tested more than one age group found that an older group of children actually looked less to the matching video than a younger age group, suggesting that the mean proportion of looking time to the matching video, as an IPL measure, could become less sensitive with older children. Clearly, more studies that focus on IPL methodologies need to be done, in which different dependent measures such as longest look, first look, latency to initiate a shift in fixation from the distracter to the target visual stimulus (see, for example, Fernald, Zangl, Portillo, & Marchman, 2008), in addition to the mean proportion of looking time to the target video, should be systematically investigated and optimized.

To conclude, no prior studies to our knowledge have tested the same children's grammatical competence using two methods. In this study, we compared the act-out and IPL measures from the same children, enabling us to directly address the discrepancy between the act-out and IPL word order findings in the current literature. Our youngest 24-month-olds failed to act-out the causative events accurately according to word order with familiar verbs, while the same children preferred watching the target video in the IPL task. Given the current controversies concerning when English learners begin to show evidence of grammatical knowledge, this is an important finding. It shows, for the first time with very similar stimuli and the very same children, that the IPL method reveals earlier competence than act-out tasks. Another important finding is the lack of robust correlations between the two tasks. This finding is consistent with the idea that early linguistic/cognitive representations are graded in strength, with early representations still weak and very task dependent. In addition, the finding is also compatible with the related perspective that some tasks such as the IPL task need not only weaker representations but only partial representations or knowledge for children to succeed. Together, the results from our two methods, in combination with previous research, suggest that the

time from about the middle of the third year of life is especially crucial for English children's productive understanding of word order in transitive constructions.

Acknowledgements

We thank Elena Rossi, Charlotte Hall, Sophie Hall and Stephanie Yarwood for help with piloting and data collection; Petra Jahn, George Rodis and Alan Woodward for technical help; Daniel Stahl and Roger Mundry for statistical advice; and Kirsten Abbot Smith, Franklin Chang, Evan Kidd and Edith Bavin for helpful input. Parts of the findings were presented at the Xth International Congress for the Study of Child Language in Berlin, July 2005, and at the Child Language Seminar in Newcastle, July 2006. We thank especially Anne Fernald, Cindy Fisher, Yael Gertner, Julian Pine and Renate Zangl for their useful comments. We also thank all the children and parents who came in to the Infant Lab at the University of Lincoln.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What pre-school children do and do not do with ungrammatical word orders. *Cognitive Development*, *16*, 679–692.

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, *23*., 48–66.

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, *23*, 275–290.

Akhtar, N. (1999). Acquiring basic word order: evidence for data-driven learning of syntactic structure. *Journal of Child Language*, *26*, 339–356.

Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology*, *33*(6), 952–965.

Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: implications of developmental errors with causative verbs. *Quaderni di Semantica*, *3*(1), 5–67.

Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.

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, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272.

Croft, W. (1998). Event structure in argument linking. In M. Butt, & W. Geuder (Eds.), *The projection of arguments* (pp. 21–63). CSLI Publications.

De Villiers, J., & De Villiers, P. (1973). Development of the use of word order in comprehension. *Journal of Psycholinguistic Research*, *2*(4), 331–341.

Dittmar, M., Abbot-Smith, K., Lieven, E. V. M., & Tomasello, M. (2008). Young German children's early syntactic competence: a preferential-looking study. *Developmental Science*, *11*(4), 575–582.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*, 547–619.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (1993). *The MacArthur communicative development inventories: User's guide and technical manual*. Baltimore: Paul H. Brokes Publishing Co.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing*. Benjamins Publishing.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, *17*, 684–691.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than the USA. *Journal of Child Language*, *27*, 689–705.

Hirsch-Pasek, K., & Golinkoff, R. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development*, *71*(6), 1540–1554.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*(2), 182–195.

Ibbotson, P., & Tomasello, M. (2009). Prototype constructions in early language acquisition. *Language and Cognition, 1*(1), 59–85.

Kidd, E., Bavin, E. L., & Rhodes, B. (2001). Two-year olds' knowledge of verbs and argument structure. In M. Almgren (Ed.), *Research on child language acquisition*. Somerville, MA: Cascadilla Press.

Langacker, R. (1990). *Concept, image and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Lieven, E., & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson, & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maratsos, M., Gudeman, R., Gerard-Ngo, P., & DeHart, G. (1987). A study in novel word learning: the productivity of the causative. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Matthews, D., Lieven, A., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, *20*(1), 121–136.

Meints, K., Plunkett, K., & Harris, P. L. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology*, *35*(4), 1072–1078.
Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. K. (2002). What is "on" and "under" to 15-, 18- and 24-month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology*, *20*, 113–130.
Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. (2004). The Cow on the High Street: effects of background context on early naming. *Cognitive Development*, *19*(3), 275–290.
Meints, K., Plunkett, K., & Harris, P. L. (2008). Eating apples or houseplants? The role of typicality in early comprehension of verbs. *Language and Cognitive Processes*, *23*(3), 434–463.
Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.
Naigles, L., Bavin, E. L., & Smith, M. (2005). Toddlers recognize verbs in novel situations and sentences. *Developmental Science*, *8*, 424–431.
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
Pinker, S., Lebeaux, D., & Frost, L. (1987). Productivity and constraints in the acquisition of the passive. *Cognition*, *26*, 195–267.
Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of young children's linguistic representations: Lexical and structural priming of syntactic constructions. *Developmental Science*, *6*, 557–567.
Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
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*. Harvard University Press.
Tomasello, M., & Abbot-Smith, K. (2002). A tale of two theories: Response to Fisher. *Cognition*, *83*, 207–214.