Toddlers’ Use of a Third Party’s Gaze Information in Verb-Action Mapping

Yuriko Oshima-Takane

1. Introduction

Research has shown that children are sensitive to social cues such as eye gaze, gestures, and head postures when inferring speaker’s referential intentions in word learning (Aktar & Tomasello, 1996; Baldwin, 1993). Understanding speaker’s referential intention is particularly important for verb learning for several reasons (Golinkoff & Hirsh-Pasek, 2006; McGuire, Hirsh-Pasek, & Golinkoff, 2006; Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009). For instance, concepts denoted by verbs (e.g., actions, events, relations) are often less salient than other elements of the visual scene (e.g., the actor or the object). Furthermore, verbs refer to one of many possible elements of an action (i.e., manner of motion, direction relative to speaker, instrument involved, results achieved) and the meaning is cross-linguistically variable (Gentner, 1982). In addition, verbs present a perspective problem: it often happens that two distinct verbs can be used to describe the same event, but the one that is used can be contingent upon the speaker’s point of view (e.g., chasing vs. running away) (Gleitman, 1990). Thus, simply observing an action or an event is insufficient; rather, one must understand the event from the speaker’s point of view (Nappa et al., 2009).

Due to these difficulties, children’s capacity to decipher social cues may be particularly important to assist them in the process of verb learning. For instance, verbs that present the perspective problem (e.g., I took a cookie vs. I gave a cookie) require the child to understand the perspective of the speaker in order to learn the correct meaning. Thus, it is important that children are able to decipher the speaker’s social cues, such as line-of-regard, gestures and head postures, to determine the event to which she/he is attending in order to understand this event from the perspective of the speaker. Furthermore, social cues may also be important for learning verbs that are contingent upon the actor’s intentions (e.g., to spill vs. to pour). Learning the meaning of these verbs

* Yuriko Oshima-Takane (yuriko.oshima-takane@mcgill.ca), McGill University. This research was supported by SSHRC and FQRSC grants. I would like to thank Kristen Boyle, Leila Cohen, Erica Heiden, Vivian Sue, and Erica Cheng for assistance in collecting and coding the data, and all parents and children for participating in this study.

requires the child to understand the perspective of another person, that of the actor, in order to understand his/her intentions. Akhtar and Tomasello (2000) claim that if a child were able to successfully decipher the socio-pragmatic cues of the speaker, they would be able to understand if the speaker were referring to an action, even if this action were not taking place at the time of the utterance. This is important given the finding that in child-directed speech, verbs are not often uttered during an action, but rather right before the action takes place (Golinkoff & Hirsh-Pasek, 2006; Tomasello & Kruger, 1992) or to comment on an action already completed (Akhtar & Tomasello, 2000). In fact, Tomasello (1995) argues that children must develop an intentional perception, that is, understanding of another person’s intentions that will motivate the child to direct her/his attention to that which the other person is attending.

Theories of language acquisition have incorporated the importance of social cues in their explanation of verb learning (Akhtar & Tomasello, 2000; Golinkoff & Hirsh-Pasek, 2006, 2008; Maguire et al., 2006). For instance, the Emergentist Coalition Model claims that children may only be able to start learning verbs once they have fully developed the capabilities to utilize perceptual, social, and grammatical cues, especially for learning less perceptually salient actions (Golinkoff & Hirsh-Pasek, 2008). In support of this theory, Brandone, Pence, Golinkoff, and Hirsh-Pasek (2007) reported that 21 to 24-month-old English-speaking children were capable of learning a new verb when perceptual salience of the action and social cues coincided; however, they were incapable of accomplishing this task when perceptual salience and social cues conflicted. Furthermore, when two actions were equalized in terms of perceptual salience, 22-month-olds were incapable of using solely social and linguistic information to learn a new verb, despite various forms of social cues being provided, including head direction, eye gaze, and physically handling the referent. However, by 34-months of age, children were able to override the more perceptually salient cue to use solely the speaker’s social and linguistic cues when learning a verb. Nappa et al. (2009) investigated the role of speaker’s gaze and syntactic cues in verb learning with older children using an eye tracker and found that 3-year-old English-speaking children use speaker’s gaze cues to interpret the meaning of novel perspective verb pairs (e.g., chasing vs. running away) when syntactic information is uninformative but do not use them when syntactic information is informative.

Based on these studies, it is clear that children are able to use a variety of ostensive social information to determine the correct referents of novel verbs; however, it remains unclear if children are able to use more subtle social cues, such as eye gaze without head direction and gestures, to make correct verb-action mappings. Eye gaze is a particularly important social cue to be able to use in verb learning, since caregivers do not always accompany their linguistic output with ostensive gestures to aid the child. It is common for caregivers to merely look at the event that they are describing. In addition, a factor not addressed in the previous studies is children’s ability to use social cues provided by a third party onlooker, who is neither the speaker nor the actor. As a child
observes a scene, it could be helpful for the child to use the attentional focus of third party onlookers, particularly their eye gaze, to find the referred event. However, no studies have investigated whether a third party onlooker’s eye gaze actually helps toddlers learn novel verbs, although there is evidence for 15-month-olds to use third party onlooker’s gaze cues for noun learning (Houston-Price et al., 2006). Third party onlooker’s gaze cues may be more difficult to use than speaker’s gaze cues, because children need to understand that third party onlookers generally look at the event to which a speaker is referring. The present study investigated from what age French-speaking toddlers are able to use a third party onlooker’s eye gaze cue for identifying the referent of novel verbs.

2. Experiment 1: 30 month-olds

2.1. Participants

Twenty-eight 30-month-old (16 male, 12 female) French-speaking toddlers with a mean of 30 months and 14 days (range: 30 months to 30 months and 28 days) participated in this study. They were randomly assigned to either the Social and the Non-Social condition. All participants were recruited from the university database of mothers interested in participating in developmental studies with their children.

2.2. Verb learning task

A modified version of the Habituation method with a switch design was used for this verb-learning task (Lippeveld, Samuel, & Oshima-Takane, 2009). The habituation method was used to familiarize children with novel objects, actions, and verbs that were used throughout the reminder of the verb learning task in order to control for novelty effects of these visual and linguistic stimuli while children were taught novel verb-action pairings with morphosyntactic cues alone. It incorporated an Intermodal Preferential Looking Paradigm to assess children’s ability to learn novel verb-action mappings more accurately than the habituation method with a switch design.

In this verb learning task, children watched a short movie composed of two phases: the Morphosyntactic and the Test Phase. The Morphosyntactic Phase contained morphosyntactic cues only and all participants watched the same movie clips. It was aimed to familiarize them with the given task. The Test Phase varied according to condition. In the Social condition, children were presented a movie with both morphosyntactic cues and a social cue, in the form of eye gaze, to assist them in learning the verb-action mapping, whereas those in the Non-Social condition were presented another movie with only morphosyntactic cues.

Two novel object sets were used in the movie in which a female actor, whose face was not visible, performed two different causative actions on each object. One set of objects comprised of a red pyramid structure with a blue and silver stick through it and a green cubic structure with a pink and silver stick through
it. With both objects the actor either slid the stick back and forth through the object or used the stick to lift the upper portion of the object open and close. The other set of objects were a blue triangular shaped object on top of a blue and orange-checkered rectangular base and a green rectangular shape on top of a black circular base with red spots. The actor either pushed the top part of the object down or turned the top part around in a circular fashion.

A transitive verb sentence frame, “Je __ le jouet”, with a novel verb was presented while the child was shown a movie clip in which an actor performed a causative action. The novel verbs used were “laf”, “nim”, “vop” and “dax.” For example, the child heard “je vop le jouet” (i.e., I vop the toy) as the actor slid a stick in and out of an object and heard “je dax le jouet” (i.e., I dax the toy) as she lifted the upper portion of the object open and close. The linguistic stimuli were recorded by a native female Quebecois-French speaker using child-directed speech.

Each phase of the verb learning task consisted of four trials: familiarization, control, teaching and assessment. The familiarization trial consisted of two movie clips alternately presented three times each, for a total of six times. The movie clips alternately presented one of the objects in the center of an otherwise all-black television screen. For example, a clip of the actor pushing a stick through the red pyramid while “je vop le jouet” was uttered would alternate with that of the actor using the stick to lift the top portion of the green cube while “je dax le jouet” was uttered. During each movie clip, the verb sentence was repeated four times. During the familiarization trials, the children learn to associate two different novel verbs with two different actions. However, since they received only 6 trials (i.e., 3 trials for each word-action pairing), they were not expected to fully habituate to these pairings (i.e., the habituation rate should be higher than 50% reduction criterion) at this time. This is important because children tend to rely on eye gaze cues when they are not sure about the verb meaning (Nappa et al., 2008).

The control trial consisted of two movie clips. The first control clip (Control 1) determined if the child had a preference towards any particular object, action, or side of the screen, while the second control clip (Control 2) allowed the child to become accustomed to a novel female onlooker’s face. In the Control 1, the same two objects from the familiarization trial were presented on either side of the screen; however, the actions that had been performed on each object during the familiarization trials were now being performed on the other object in order to test whether children map the novel verb onto the action or the object in the assessment trial. Continuing with the previous example, on one side of the screen, the actor was now lifting the top portion of the red pyramid using the stick to open, while on the other side of the screen, the actor was sliding the stick in and out of the green cube. The auditory stimuli, “Maintenant ils sont différents. Comme ils sont amusants” (i.e., Now they are different. How fun they are!), accompanied this visual scene. The Control 2 was identical to the Control 1 with the exception of a female onlooker’s face in the center of the screen looking forward.
The teaching trial of the Morphosyntactic Phase was the same for all children. The visual scene was identical to that of the Control 2; however, the female onlooker’s eyes were now closed. The auditory stimuli also differed from the control clip in that the children heard one of the two verb sentences uttered in the familiarization trial (continuing with the previous example, either “je vop le jouet” or “je dax le jouet”). In the Test Phase, this teaching trial varied according to condition. In both conditions, the children heard one of the verb sentences from the familiarization trials and, once again, the visual scene was identical to the Control 2 with the exception of the female onlooker’s eyes in the center of the screen. In the Non-Social condition, the female onlooker in center of the screen had her eyes closed, whereas in the Social condition, the female onlooker in the center of the screen was gazing at the target action being performed. Therefore, the children in the Social condition received a social cue to assist them in finding the correct verb-action pair.

The visual scene of the assessment trial was identical to that of Control 1. Following the above example, the assessment trial of the Morphosyntactic Phase consisted of the actor opening the top half of the red pyramid using the stick on one side of the screen and the actor sliding the stick in and out of the green cube on the other side. The children were then asked to identify the correct action-referent of the novel verb using two different questions: “Sais-tu lequel je vop” (i.e., *Do you know which one I vop?*), “Trouve lequel je vop” (i.e., *Find the one I vop*).

Movies were counterbalanced according to the side of the screen with the correct verb-action pairing and the object set used. As such, half of the movies had the correct verb-action pairing on the left side of the screen in the Morphosyntactic Phase and on the right side of the screen in the Test Phase and half had the opposite. In addition, half of the movies presented the red pyramid/green cube object set during the Morphosyntactic Phase and the blue triangle shape atop a rectangular base/rectangle shape atop a circular base object set during the Test Phase and half had the opposite.

A pre-test/pre-post/post-test trial consisting of a duck and a dog hand puppet dancing behind a desk was shown before the Morphosyntactic Phase (i.e., pre-test), after the Test phase (i.e., post-test), and in between the two phases (i.e., pre-post). These trials lasted a total of 16 seconds. The pre-test trial allowed us to ensure that the child was attentive at the beginning of the Morphosyntactic Phase. The pre-post test trial allowed us to ensure that the child was still attentive before the Test Phase. The post-test trial informed us of whether or not the child had remained attentive during the Test Phase.

With the exception of the pre-/pre-post/post-test, all trials were 20 seconds in length. A 1-second attention grabber that consisted of an expanding white star and a “bing” sound was presented in between each trial. This was used to bring the child’s attention back to the television screen in case the child had stopped looking at the screen and to re-orient his/her gaze to the center of the screen for the following clip.
2.3. Data Analysis

The child’s looking times were coded offline using SuperCoder (Hollich, 2003) by a coder who was unaware of both the contents of the verb learning task movie and condition (i.e., Social or Non-Social) of the participant. Each second was broken down into 30 frames and each frame was coded for the entire duration of the movie to determine if the child had been looking left, right, center, or away. A proportion score of looking times during the Control 1 and during the teaching trial was calculated by dividing the number of frames a child spent looking at the side of the screen with the correct referent of the novel verb (i.e., the matching screen) by the sum of the number of frames spent looking at the side of the screen with the incorrect verb referent (i.e. the non-matching screen) and the number of frames looking at the matching screen. The proportion score of the Control 1 indicated if the child had a preference for a particular side of the screen. A proportion score greater than 0.5 indicated that the child preferred the matching screen, whereas a proportion score less than 0.5 showed that the child preferred the non-matching screen. A proportion equal to 0.5 signified that the child had no preference for a particular side of the screen. The proportion score of the teaching trial was used to determine if the conditions differed in their looking behaviour during this trial.

A proportion score of looking times was also calculated for each question in the assessment trial to determine whether or not the child was able to map the novel verb onto the correct action referent by using the above equation. Children in the Morphosyntactic Phase would look at the correct referent in the assessment trial significantly longer than in the Control 1 if they could use morphosyntactic cues provided during the familiarization and teaching trials. Because children in both conditions were given the same familiarization, control, and teaching trials, they should show no significant difference in their looking times in these trials. In the Test Phase, children in the Social condition were expected to look at the correct referent in the assessment trial significantly longer than in the Control 1 if they could use a third party onlooker’s gaze cue provided during the teaching trial. Children in the Non-Social condition were expected to look at the correct referent in the assessment trial significantly longer than in the Control 1 if they could use morphosyntactic cues alone provided during the familiarization and teaching trials. If the children in the Social condition look significantly longer at the matching screen in the assessment trial than in the Control 1, whereas those in the Non-Social condition do not, this would indicate that they are able to use a third party onlooker’s eye gaze cue provided during the teaching trial in order to learn the novel verb-action mapping. The significance level for all statistical tests used in the present study was set at $p=.05$. 
2.4. Results

2.4.1. Morphosyntactic Phase

A 2 (Social, Non-Social) x 2 (first two, last two familiarization trials) mixed ANOVA was performed to determine the extent to which the children had habituated to the verb-action pairings presented during the familiarization trials. The results revealed a main effect of trial in looking times during the first two familiarization trials ($M=18.54\text{sec}, SD=2.05$) and the last two familiarization trials ($M=16.06\text{sec}, SD=3.05$), $F(1, 25^1)=18.06, p<.001$. The significant decrease in looking times suggests that the children had habituated to the verb-action pairings to a significant degree during the familiarization trials. There was neither a main effect of the condition nor an interaction effect, indicating that there was no significant difference in the decrease in looking times between the two conditions. In addition, the comparison of the mean looking times for the first two familiarization trials to the last two familiarization trials indicates that children in neither condition habituated to the word-action pairings fully using 50% reduction criterion ($M=87\%$).

A one-sample t-test revealed that neither the mean proportion score of the Control 1 (C1) in the Social condition ($M=.54, SD=.14$) nor that of the Non-Social condition ($M=.49, SD=.09$) differed significantly from chance (0.5), indicating that the children in both conditions did not have a preference for a particular side of the screen.

Consistent with our prediction that the teaching trial in the Morphosyntactic Phase should not differ in children’s looking patterns between the two conditions, an independent sample t-test showed that the Social ($M=.50, SD=.16$) and Non-Social ($M=.54, SD=.17$) conditions did not differ significantly in their proportion scores.

Fig. 1 presents the mean proportion scores for Control 1 (C1), assessment Question 1 (Q1), and Question 2 (Q2) by condition. A 2 (Social, Non-Social) x3 (C1, Q1, Q2) mixed ANOVA showed neither main effects nor an interaction effect. Consistent with our predictions, the results confirmed that there were no differences in looking patterns between the two conditions for any of the trials during the Morphosyntactic Phase. Furthermore, children in both conditions were not able to map the novel verb onto the correct referent using the morphosyntactic cues alone.

---

1 The statistical tests were performed on the familiarization data from 27 participants because one participant looked away most of the time in the first two trials (i.e., only 10 sec looking times out of 40 sec).
2.4.2. Test Phase

A 2 (Social, Non-Social) x 2 (first two, last two familiarization trials) mixed ANOVA revealed a main effect of looking times between the first two familiarization trials ($M=18.48\text{sec, }SD=1.77$) and the last two familiarization trials ($M=14.23\text{sec, }SD=5.08$), $F(1,26)=19.79, p < .001$, but no other effects. This indicates that the children in both conditions habituated to the verb-action events of the familiarization trials to a significant degree from the first two trials to the last two trials similarly. The comparison of the mean looking time for the first two familiarization trials to the last two familiarization trials indicates that children in neither condition habituated to the word-action pairings fully using 50% reduction criterion ($M=77\%$).

A one-sample $t$-test revealed that neither the proportion scores of the Control 1 (C1) in the Social condition ($M=.47, SD=.14$) nor those in the Non-Social condition ($M=.52, SD=.10$) differed significantly from chance. This indicates that the children in neither condition had a preference for a particular side of the screen during the control trial.

The teaching trial of the Test Phase varied according to condition with the Social condition receiving an additional social cue; therefore, it is possible that proportion scores during the teaching trial would differ according to condition. However, the Social ($M=.51, SD=.19$) and Non-Social ($M=.48, SD=.14$) conditions did not differ significantly in their proportion scores.

The mean proportion scores for Control 1 (C1), assessment Question 1 (Q1), and Question 2 (Q2) by condition are presented in Fig.1. A 2(Social, Non-Social) x 3(C1, Q1, Q2) mixed ANOVA showed a main effect of trials, $F(2, 52)=3.88, p=.027$, but no other effects. Paired $t$-tests showed that the main effect of trials was due to the fact that the proportion score of Q2 in the Social condition ($M=.68, SD=.23$) was significantly higher than that of the C1 ($M=.47, SD=.14$), $t(13)=-3.00, p =.01$, although no significant difference between C1 and Q1 ($M=.53, SD=.30$). In contrast, the proportion scores of Q1 and Q2 in the Non-Social condition (Q1: $M=.46, SD=.21$; Q2: $M=.55, SD=.26$) were not significantly different from that of the Control 1 ($M=.52, SD=.10$).

The results of the assessment trial in the Morphosyntactic and the Test Phase were not due to fatigue because there was no main effect of looking times during the pre-test ($M=13.96, SD=4.85$), the pre-post ($M=14.84, SD=3.54$) and the post-test ($M=13.83, SD=3.47$). There was neither a main effect of condition nor an interaction effect. This indicates that the conditions did not differ in their attentiveness. In addition, toddlers in both conditions did not differ significantly in their total productive vocabulary scores, as measured by the French-equivalent of the MCDI (Social: $M=410, SD=163$; Non-Social: $M=436, SD=122$), indicating that the difference found in the Test Phase was not due to the difference in vocabulary development between two conditions (Frank, Poulin-Dubois & Trudeau, 1997).
3. Experiment 2: 26 month-olds
3.1. Method

Thirty-two 26-month-old (16 male, 16 female) French-learning toddlers with a mean of 26 months and 18 days (range: 26 months and 3 days to 26 months and 30 days) participated in this experiment. The same verb learning task and data analysis as those used in Experiment 1 were used in Experiment 2 in order to investigate whether 26-month-olds were able to use a third party onlooker’s eye gaze information to learn a novel verb-action mapping.

3.2. Results
3.2.1. Morphosyntactic phase

A 2 (Social, Non-Social) \( \times \) 2 (first two, last two familiarization trials) mixed ANOVA was performed to determine the extent to which the participants had habituated to the verb-action pairings presented during the familiarization trials. The results revealed a main effect of trial in looking times during the first two familiarization trials \( (M=18.30\text{sec}, SD=1.94) \) and the last two familiarization trials \( (M=16.61\text{sec}, SD=2.40) \), \( F(1,30) = 16.49, p < .001 \). The significant decrease in looking times suggests that the children had habituated to the verb-action pairings to a significant degree during the familiarization trials. There was a main effect of the condition, \( F(1,30)=4.58, p=.041 \) (Social first two trials: \( M=17.71\text{sec}, SD=2.17 \); Social last two trials: \( M=15.88\text{sec}, SD=2.64 \); Non-Social first two trials: \( M=18.87\text{sec}, SD=1.53 \); Non-Social last two trials: \( M=17.35\text{sec}, SD=1.95 \) ) but no interaction effect. However, an independent sample t-test performed on the familiarization difference scores (i.e., a difference between the mean looking times of the first two trials and that of the last two trials) showed no significant difference between the Social \( (M=1.83\text{sec}, SD=2.92) \) and Non-Social \( (M=1.55\text{sec}, SD=2.05) \) conditions. Furthermore,
children in neither condition did not reach the 50% habituation criterion (M=90–92%).

Since Control 1 and the teaching trial in the Morphosyntactic Phase did not vary according to condition, the two conditions should not differ in children’s looking patterns in these trials. Indeed, the conditions did not differ significantly in the proportion scores of the Control 1 (Social: M=.52, SD=.12; Non-Social: M=.49, SD=.15) and those of the teaching trial (Social: M=.52, SD=.12; Non-Social: M=.49, SD=.20). In addition, neither the proportion score of the Control 1 in the Social condition nor in the Non-Social condition differed significantly from chance.

The proportion scores for Control 1 (C1), assessment Question 1 (Q1), and Question 2 (Q2) by condition are presented in Fig. 2. As previously stated, we did not expect the proportion scores of the assessment trial in the Morphosyntactic Phase to differ significantly between conditions since children were exposed to the same stimuli leading up to this trial. A two-way mixed ANOVA with condition (Social, Non-Social) as a between-group factor and trial (C1, Q1, Q2) as a within-group factor showed no main effects nor an interaction effect. Consistent with our predictions, these results confirmed that there were no significant differences in looking patterns between the two conditions in the Morphosyntactic Phase. Furthermore, children in both conditions were not able to map the novel verb onto the correct referent using the morphosyntactic cues alone.

3.2.2. Test Phase

A 2 (Social, Non-Social) x 2 (first two, last two familiarization trials) mixed ANOVA was performed to determine the extent to which the children had habituated to the verb-action pairings presented during the familiarization trials. The results revealed a main effect of looking times between the first two (M=17.65sec, SD=2.18) and the last two (M=14.64sec, SD=3.13) familiarization trials, F(1, 292)=29.01, p<.001 and a main effect of condition, F(1,29)=9.21, p=.005 (Social first two: M=17.17sec, SD=2.40; Social last two: M=13.00 sec, SD=2.18; Non-Social first two: M=18.10sec, SD=1.91; Non-Social last two: M=16.18 sec, SD=2.59) but no interaction effect. This suggests that the children had habituated to the verb-action events of the familiarization trials to a significant degree from the first two trials to the last two trails but the degree of the habituation depends on the condition. An independent sample t-test performed on the familiarization difference scores showed a marginally significant difference between the Social (M=4.18sec, SD=2.92) and Non-Social (M=1.92sec, SD=3.34) condition, t(29) =2.00, p=.056. This suggests that the children in the Non-Social condition might have habituated to the verb-action

---

2 The statistical tests were performed on the familiarization data from 31 participants because one participant looked away for the entire period of the first two trials.
pairings during the familiarization trials to a relatively lesser degree than their counterparts in the Social condition. However, children in neither condition habituated to the word-action pairings fully using 50% reduction criterion (M=76–92%).

A one-sample *t*-test revealed that neither the proportion score of the Control 1 in the Social condition (M=.53, SD=.13) nor that of the Non-Social condition (M=.46, SD=.11) differed significantly from chance, indicating that the children in neither condition had a preference for a particular side of the screen. In addition, an independent sample *t*-test showed no significant difference between two conditions.

The teaching trial of the Test Phase varied according to condition with the Social condition receiving an additional social cue; therefore, it is possible that proportion scores during this trial would differ according to condition. However, the Social (M=.49, SD=.17) and Non-Social (M=.56, SD=.26) conditions did not differ significantly in their proportion scores.

A two-way mixed ANOVA with condition (Social, Non-Social) as a between-group factor and trial (C1, Q1, Q2) as a within-group factor performed on the proportion scores showed neither main effects nor interaction effect. These results indicate that 26-month-olds were not able to use the third party’s eye gaze information in order to map the novel verb onto the correct referent.

A 2(Social Condition, Non-Social) x 3(pre-test vs. pre-post vs. post-test) mixed ANOVA showed no main effect of looking times during the pre-test (M = 15.57 sec, SD = 2.35), the pre-post (M=15.80 sec, SD=3.10) and the post-test (M = 16.22 sec, SD=1.84). There was neither a main effect of condition nor an interaction effect. This indicates that the conditions did not differ in their looking times before, during, and after the verb learning task, suggesting that the results of the assessment trial of the Morphosyntactic and the Test Phases were not due to fatigue. Furthermore, the children in both conditions did not differ

![Fig. 2. The 26-month-olds’ mean proportions scores and the standard errors in the control and the assessment trial by condition.](image-url)
significantly in their total productive vocabulary scores (Social: \(M=373, SD=143\); Non-Social: \(M=349, SD=133\)).

4. Discussion

The present results showed that 30-month-olds, but not 26-month-olds, in the Social condition were able to learn the novel verb-action mapping, whereas those in the Non-Social condition were not able to do so. This finding suggests that the toddlers understand that third party onlookers generally look at the event to which a speaker is referring by 30 months of age. The results of the assessment trial were not due to the differences in language development between the two conditions, because they did not differ significantly in their total productive vocabulary scores. Furthermore, there was no significant decrease in looking times from the pre-test to the post-test for either condition; therefore the present results are not due to fatigue. In addition, since there was no significant differences in looking times between the two conditions in the Morphosyntactic Phase, toddlers in both conditions were equally familiar with the task at hand upon presentation of the different teaching trial of the Test Phase. Finally, the proportion scores during the assessment trial were not a result of a preference for a particular side of the screen, since the toddlers in neither conditions looked significantly longer at a particular side of the screen during the Control 1 trial. The difference in proportion scores between the assessment and the Control 1 trial suggests that the 30-month-olds in the Social condition were able to take advantage of the eye gaze cue provided during the teaching trial of the Test Phase allowing them to more readily learn the verb-action mapping.

Brandone et al. (2007) have found that 22-month-old children are unable to use a combination of social cues, including head direction, eye gaze, and physically handling the referent, to learn a novel verb-action mapping. By contrast, 34-month-olds are not only able to use social cues, but are also able to override conflicting perceptual salience to use social and linguistic cues to learn a verb-action mapping. Furthermore, Akhtar and Tomasello (1996) have shown that, at 24-months, children are able to use a combination of social cues, consisting of head direction, eye gaze, and physically handling a referent, to map a novel verb to an action. The present study demonstrated that children younger than 34-months of age are able to use eye gaze, a much less perceptually salient cue in comparison to the combination of various social cues used in the two aforementioned studies, to learn a novel verb-action mapping. Moreover, the present study is unique in that it also shows that 30-month-olds are able to interpret the eye gaze of a third party onlooker, who is neither the speaker nor the actor, to form a verb-action mapping. The capacity to decipher eye gaze cues to learn a novel verb is particularly important because people do not always pair their speech with salient gestures that were often the types of social cues provided in previous studies on verb learning (Akhtar & Tomasello, 1996; Brandone et al., 2007).
Being able to use the eye gaze information from a third party onlooker is particularly useful for word learning because it allows children to learn words when merely observing a scene of interacting people who are not necessarily purposefully attempting to communicate with nor teach the observing child (O’Doherty, Troseth, Shimpi, Goldenberg, Akhtar, & Saylor, 2011). Despite children’s ability to follow eye gaze from a young age, Beier and Spelke (2012) found that it was not until 10-months of age that children understand the communicative intentions of mutual eye gaze, i.e. that two people generally make eye contact when talking with each other. In order to use eye gaze to learn new words, it is necessary for children to understand the communicative intentions of this social cue, specifically that people generally look at what someone is talking about. The present results show that, by the age of 30-months, toddlers are able to use the eye gaze of a third party onlooker to form a novel verb-action pairing, suggesting that the toddlers understand that third party onlookers generally gaze at the event to which a speaker is referring. It should be noted, however, that the social cue provided in the present study was perceptually subtle and less salient than the combination of various social cues used in previous studies (Akhtar & Tomasello, 1996; Brandone et al., 2007; Houston-Price et al., 2006). It is possible that toddlers younger than 30-month-olds are able to use third-party onlooker’s gaze cue to identify the referent of novel verbs if it is provided in combination with perceptually more salient cues such as head postures and gestures. Further research is necessary to investigate this possibility.

References


