Conflicting Social Cues: Fourteen- and 24-Month-Old Infants' Reliance on Gaze and Pointing Cues in Word Learning

Markus Paulus & Paula Fikkert

To cite this article: Markus Paulus & Paula Fikkert (2014) Conflicting Social Cues: Fourteen- and 24-Month-Old Infants' Reliance on Gaze and Pointing Cues in Word Learning, Journal of Cognition and Development, 15:1, 43-59, DOI: 10.1080/15248372.2012.698435

To link to this article: http://dx.doi.org/10.1080/15248372.2012.698435

Accepted author version posted online: 31 Jul 2012.
Published online: 12 Nov 2013.

Submit your article to this journal

Article views: 258

View related articles

View Crossmark data

Citing articles: 2 View citing articles
Conflicting Social Cues: Fourteen- and 24-Month-Old Infants’ Reliance on Gaze and Pointing Cues in Word Learning

Markus Paulus
Ludwig Maximilian University, Germany

Paula Fikkert
Radboud University Nijmegen, The Netherlands

Language acquisition is a process embedded in social routines. Despite considerable attention in research to its social nature, little is known about developmental differences in the relative priority of certain social cues over others during early word learning. Employing an eye-tracking paradigm, we presented 14-month-old infants, 24-month-old infants, and adults with movies in which an actor repeatedly gazed at one and pointed to the other of two objects while presenting them with a novel word. The results show that the 14-month-old infants pay more attention to a model’s eye gaze when learning to map a novel word to a referent, whereas 24-month-old infants and adults rely more on pointing cues. Our results provide evidence for a developmental change in the relative priority of pointing versus eye-gazing cues in language acquisition.

One of the most important steps in human development is the acquisition of language, and in particular, the construction of a mental lexicon (e.g., Swingley, 2009). An important question is how infants map words to referents and, more specifically, which cues infants use to determine the referent of a speaker’s utterance. The importance of this question is intriguingly demonstrated by Quine (1960), who stated that there is an infinite number of possible word-to-world mappings to choose from when confronted with a novel word form. Philosophers of language and social theorists have repeatedly emphasized that language is best understood as a social phenomenon (Brandom, 1994; Wittgenstein, 1953), which is deeply rooted in the pragmatics and interactions of our social lives (Carpendale & Lewis, 2004; Racine & Carpendale, 2007). In line with these considerations, previous research has shown that infants rely heavily on social cues. Two of the most prominent and widely studied social cues concern gaze and pointing (e.g., Butterworth, 2001; Carpenter, Nagell, & Tomasello, 1998; Flom, Lee, & Muir, 2007; Moore, 2008; Paulus, 2011; Thoermer & Sodian, 2001). Infants have been shown to use gaze and pointing gestures in joint attention contexts to acquire novel words (e.g., Baldwin, 1993; Hirotani, Stets, Striano, & Friederici, 2009; Hollich et al., 2000; Houston-Price, Plunkett,

Correspondence should be sent to Markus Paulus, Department of Psychology, Ludwig Maximilian University, Leopoldstr. 13, 80802 Munich, Germany. E-mail: markus.paulus@lmu.de
& Duffy, 2006; for a comprehensive review, see Waxman & Lidz, 2006). In line with these findings, longitudinal studies have reported positive relations between infants’ emergent ability to follow gaze and productive vocabulary size at a later age (Brooks & Meltzoff, 2005; Morales, Mundy, & Rojas, 1998). Based on these and similar studies, developmental theories have stressed the social pragmatics of early language acquisition (e.g., Tomasello, 2003).

Research has shown that at least by 12 months of age, infants start to rely on social cues to disambiguate between multiple possible referents (e.g., Hennon, Chung, & Brown, 2000). However, infants’ reliance on social cues for word learning in the first half of the 2nd year of life is rather fragile. For example, by 13 months of age, infants need clear behavioral cues that relate a speaker’s utterance to an object, as they have problems mapping words to referents if the behavioral connection is less clear (Woodward, 2000). At 18 months, word learning can still be disturbed when, while a model is gaze-cueing one object, a second object is more salient and thus competing for the infants’ attention (Houston-Price et al., 2006; Moore, Angelopoulos, & Bennett, 1999). By 2 years of age, the social pragmatics of word learning are more consolidated and advanced. For example, it has been shown that children at this age are able to learn novel words from overhearing (Akhtar, Jipson, & Callanan, 2001), even when the teacher is engaged in a distracting activity and the object is not explicitly labeled (Akhtar, 2005). Besides evidence that infants rely on social cues, research has also shown that infants actively search for speakers’ cues and their usefulness in disambiguating between several possible referents (e.g., Akhtar, Carpenter, & Tomasello, 1996; Baldwin, 1993; Bloom, 2000).

Taken together, the reviewed literature provides ample evidence for the claim that social pragmatics play an important role in early language acquisition. Furthermore, it shows developmental progress in infants’ reliance on social cues and interpretation of others’ behavior in the context of word learning. However, one question that has not been addressed in the literature concerns developmental changes in relative priority of certain social cues over others; and in particular, little is known about infants’ reactions in the case of conflicting social cues. Most of the studies using pointing and gaze cues have either presented them alone or coupled them. Yet, there are daily life situations in which hand and gaze cues are decoupled (e.g., Flanagan, Terao, & Johansson, 2008). For example, while looking at a partner one is talking to, one can be simultaneously pointing at the referent of the ongoing conversation; hence, pointing and gazing are decoupled events (cf. Hayhoe & Ballard, 2005). Another example is children’s observation of third-party teaching interactions (e.g., Gräfenhain, Behne, Carpenter, & Tomasello, 2009), in which from the observer’s point of view, gaze and pointing can appear to be decoupled. Studying infants’ word learning in such situations of conflicting cues could reveal how these cues interact with each other and on which of the cues infants preferentially rely in word learning. Infants’ relative weighting of social cues and developmental changes in this weighting would provide important information concerning the developmental pathways of relying on these cues and the psychological mechanisms subserving their development.

Concerning the processing of gaze cues, research has provided evidence that the perception of eye-gaze directions already trigger shifts of attention in 3-month-old infants (Hood, Willen, & Driver, 1998). Furthermore, it has been shown that the perception of eye gaze facilitates object processing (e.g., Hoehl, Reid, Mooney, & Striano, 2008) and, as already mentioned, learning about object names (e.g., Hennon et al., 2000). At about 1 year of age, infants are more likely to engage in gaze follow, when the other has open eyes compared with closed eyes (Brooks & Meltzoff, 2002; D’Entremont & Morgan, 2006). Around the same time, their gaze following
becomes more sophisticated as they tend to follow someone’s gaze when it previously revealed an interesting object compared with someone who merely looked in an empty container (Chow, Poulin-Dubois, & Lewis, 2008). Additionally, infants expect someone to grasp a novel object when the previously grasped object is out of the actor’s sight (Sodian, Thoermer, & Metz, 2007), revealing a progressing understanding of the intentional nature of human gaze. Given the social and evolutionary relevance of the processing of others’ eye gaze, it has been proposed that eye gaze plays a dominant role in social learning and social understanding (Frischen, Bayliss, & Tipper, 2007). Notwithstanding the claim that humans—and to some extent already, infants—are able to appreciate the intentional nature of gaze, it has been suggested that people often rely on eye gaze—already early in development—in a highly automatic, reflex-like routine (e.g., Baron-Cohen, 1995; Emery, 2000).

Studies with adults have shown that even though another person’s gaze provides important information in conversations (e.g., Hanna & Brennan, 2007), in the case of direct conflict, the pointing gesture is perceived as being more relevant: They orient along another person’s pointing cue rather than his gaze (Tomoko, Dairoku, Naoki, & Susumu, 2004). That is, adults interpret pointing as a more reliable cue to reference than they interpret eye gaze. This is probably due to the fact that the pointing gesture is a more specific expression of a speaker’s intention than is his eye gaze. It is obvious that people gaze at various objects throughout the day: Sometimes, looking behavior serves the monitoring and control of goal-directed actions, such as when looking at the target of one’s action (e.g., Neggers & Bekkering, 2001) or the anticipation of one’s own action effects (e.g., Kenward, 2010). Sometimes, gaze behavior is automatically triggered by other processes such as language comprehension (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) or the orienting reflex (e.g., Sokolov, 1963). As a consequence, gazing is not always a communicative action, directed toward others. In contrast, people only rely on pointing when they emphatically want to communicate about something specific (see Deák, Triesch, 2006). It has therefore been suggested that pointing forms a more reliable cue for reference than does gaze (Deák, Flom, & Pick, 2000), and it has been characterized as a prototypical way to anchor communication (Bangerter, 2004; Clark, 2003).

Developmental evidence for the relevance of pointing cues comes from a number of investigations. In a study by Grassmann and Tomasello (2010), an experimenter pointed to a novel object while verbally requesting another well-known object (e.g., a car), which was also accessible. Two- and 4-year-old children followed the pointing gesture over the verbal information in the model’s request, stressing the importance of social cues (and especially of pointing) as indicating a speaker’s intention. Further evidence for the impact of pointing comes from a recent study of Palmquist and Jaswal (2012). They presented 3- to 4-year-old children with two actors. One of the actors hid a ball under one of four cups, whereas the other could not see the baiting event. Subsequently, both actors pointed to different cups, grasped different cups, or indicated none of the cups (i.e., baseline condition). When the children were asked to indicate which actor knew the location of the ball, they indicated the correct actor only in the baseline and the grasping condition above chance. Yet, in the pointing condition, they indicated both actors to an equal extent. The authors suggested that “seeing an informant point can lead preschoolers to attribute knowledge to that informant even if he or she did not have perceptual access” (Palmquist & Jaswal, p. 231). That is, regardless of whether the informant actually saw the baiting event or not, children relied to the same extent on both persons when they were pointing to convey information. Palmquist, Burns, and Jaswal (2012) also showed that this effect is unique for pointing
and does not occur for other referential forms. Taken together, these considerations and studies provide evidence that for preschool children, the pointing gesture is an important indicator of another person’s referential intention and his knowledge state, regardless of the person’s visual access, which has previously been considered the strongest indicator of a person’s knowledge (cf. the seeing-is-knowing rule; Ruffman, 1996).

Nothing is known, however, about the early development of young children’s preferential reliance on pointing cues over other forms of reference such as gaze following. This question is of particular interest given the earlier-reviewed empirical findings and theoretical claims about the automaticity of early gaze following. More concretely speaking, information about this issue would be of interest for two reasons: First, it would be informative with respect to the developmental origins of selective reliance on the pointing cue. It would reveal whether it is the product of learning processes or whether children exhibit such a preference from the moment they engage in point following (that is, around or before their 1st birthday; cf. Behne, Liszkowski, Carpenter, & Tomasello, 2012; Gredebäck, Melinder, & Daum, 2010). A shift in the preferential reliance from gaze to pointing cues could indicate a shift from often relatively automatic and reflex-like processes (cf. Baron-Cohen, 1995; Emery, 2000) to the controlled use of the more informative pointing gesture. Second, given that we examine infants’ reliance on social cues in a word-learning paradigm, such knowledge would help us to understand the development and developmental changes in the early social pragmatics of word learning.

To this end, we examined infants’ behavior during word learning in a situation of conflicting gaze and pointing cues (for similar conflict paradigms, see Graham, Nilsen, Collins, & Olineck, 2010; Grassmann & Tomasello, 2010). Participants watched a video on an eye-tracking screen in which a model repeatedly gazed at one object (i.e., had direct visual contact to one object) and pointed at another while teaching the participants a new word. Subsequently, they were asked in several test trials, without the model present, to look at the target (see “Stimuli” section). Based on findings that language comprehension modulates looking behavior in children (e.g., Bergmann, Paulus, & Fikkert, 2012; Swingley, Pinto, & Fernald, 1998) and adults (Tanenhaus et al., 1995), we analyzed participants’ looking behavior during both the teaching phase and the test phase of the experiment.

Our main question was to examine developmental changes in infants’ reliance on social cues for word learning and the developmental origins of their strong reliance on the pointing cue. In particular, we examined whether this preferential reliance on pointing is already to be found in the early origins of point following. Alternatively, it could be the consequence of learning processes as speculated by Palmquist et al. (2012). In particular, given the automaticity of eye-gaze processing, one could argue that younger infants predominantly rely on a person’s eye gaze and that their learning about the higher reliability of pointing cues proceeds during the 2nd year of life, supported by their engagement in social activities (Bibok, Carpendale, & Lewis, 2008) and their developing ability for attentional flexibility and control (Colombo, 2001; Johnson, Posner, & Rothbart, 1991). In other words, given that the eyes attract attention and trigger automatic shifts of attention, it might require considerable attentional control to focus and process the information of the pointing cue when it is in conflict with the gaze cue. As infants’ ability for the intentional control of attention develops considerably during the first 2 years of life (Ruff & Rothbart, 1996), we expected that younger infants would preferentially rely on the gaze cue and that only 2-year-old children and adults would be able to appreciate the pointing cue. Because it has been shown that infants’ reliance on the pointing gesture develops around their
1st birthday (e.g., Behne et al., 2012; Butterworth & Grover, 1990; Woodward & Guajardo, 2002), we decided to investigate 14- and 24-month-old infants as well as a group of adults.

To examine how infants extract information during the teaching trials, we were additionally interested in exploring participants’ looking behavior during the teaching phase. Most of the studies investigating infants’ reliance on social cues during word learning have focused only on infants’ behavior in test phases (i.e., after they have been taught the novel word), but little is known about how infants extract information during the teaching situation. Given this lack of knowledge, our expectations about infants’ behavior in the teaching phase was less specific. If it were true that the reliance on pointing cues is the result of learning processes, one might expect that the younger infants (who use early gaze information to learn about objects; e.g., Hoehl et al., 2008) would more strongly rely on the actor’s gaze and would therefore show a greater preference for the teaching person’s face compared with older infants. Based on our considerations that pointing gains increasing relevance during the course of the 2nd year of life, we reasoned that the 24-month-old infants and adults would show a stronger preference for the model’s hand and the object indicated by the pointing gesture. Furthermore, based on findings that young infants are eager to explore their environment to acquire knowledge (e.g., von Hofsten, 2007), we expect that participants at all ages would show a shift of attention to one of the two objects when the model explicitly teaches them the novel word, compared with a phase directly before the teaching event.

METHOD

Participants

The infant sample consisted of 32 healthy full-term Dutch infants without any complications who were recruited through the Baby Research Center in Nijmegen, The Netherlands. The sample contained sixteen 14-month-old infants (age range = 1;2–1;3; \textit{M} \text{age} = 1;3; 8 boys) and sixteen 24-month-old infants (age range = 1;12–2;1; \textit{M} \text{age} = 2;1; 8 boys). Three additional 14-month-old infants and 2 additional 24-month-old infants were not included in the final sample because of general inactivity (\textit{n} = 1), refusal to remain seated (\textit{n} = 1), or not providing enough data for analysis (\textit{n} = 3; see “Data Analysis” section). The infants’ parents gave informed consent for participation of their child and were given a baby book or monetary compensation for their visit. Additionally, we tested 12 adults (age range = 18–30 years). The adult participants were students and were credited with course points for their participation in the study.

Stimuli

The stimulus material consisted of two types of short movies—learning movies and test movies (see Figure 1 for screenshots of the key phases). The learning movies displayed a female model sitting at a table. On the left and right side of the table, there were two novel toy objects. The learning movies showed the model initially looking at her upright-oriented index finger. After exactly 1.5 seconds, she pointed with her index finger to the object located contralateral to her pointing hand while she gazed at the same time at the other object. The model’s pointing
behavior was clearly voluntary as the movement was directed at an object and the arm was voluntarily held in the air. Following this, she said in Dutch, ‘‘Look! Isn’t it nice? This is a taan!’’ The word taan has been chosen as a nonword because its properties are characteristic of typical Dutch words without being similar to known words (Fikkert & Levelt, 2008). Thereupon, the model moved the index finger and gazed back to the point of departure (see the first three pictures of Figure 1 for screen shots from key phases of the movies). The whole action sequence lasted 12 seconds. To control for direction and object biases, we balanced the position of the objects at the left and right side of the table (two possibilities) and which object the model was pointing to and gazing at (two possibilities) so that four different learning movies were constructed.

In addition, there were test movies in which only the two objects were shown. Approximately 2 seconds after stimulus onset (baseline phase; slightly varying between 1,000 ms and 2,000 ms to reduce stimulus-response automaticity; cf. Paulus, Lindemann, & Bekkering, 2009), the infants were addressed by the model’s voice, ‘‘Oh! Look!’’ (attention phase), which was followed by ‘‘A taan!’’ (naming phase); approximately 8 seconds after stimulus onset, infants were again addressed, ‘‘There! Look!’’ (attention phase), which was followed by ‘‘A taan!’’ (naming phase). The whole test movie was a duration of 12 seconds (see Figure 2B for a visualization of the time course). There were two versions of the test movies to balance which object was presented at the left- and right-hand side of the table.

We included the baseline phase and attention phase in the test movies to control for saliency and general attention-enhancing effects. If a possible effect in the naming phase was to be due to
saliency and general attention-enhancing effects, we expect the same effect to be present in the baseline phase and the attention phase. If the effect is time-locked to the naming of the word, this would support our claim that infants have used the model’s cues to map a novel word form to a referent.

Ten learning movies (serving as learning trials) and four test movies (serving as test trials) were combined and presented to the participants (see Figure 2A). This number of learning trials was chosen because previous studies have provided evidence for such rapid word learning (e.g., Woodward, Markman, & Fitzsimmons, 1994). We counterbalanced between participants the object the model was pointing to and gazing at and within participants the location of the two objects on the left and right side of the table (to prevent infants associating cues and objects with a particular side). To prevent monotony during presentation of the movies, we first presented six learning trials, then two test trials, followed by four additional learning trials and finally the last two test trials.

**Experimental Setup and Procedure**

Infants were seated in an infant seat on the lap of the caregiver, who sat on a chair that was 60 cm away from the monitor of a corneal reflection eye tracker (Tobii 1750, Tobii Technology, Sweden). Adult participants were seated directly at the chair. The stimuli were shown on a 17-inch TFT flat-screen monitor, and gaze was recorded at 50 Hz with an average accuracy of 0.5° visual angle. A nine-point calibration procedure was used to calibrate the gaze of each participant. If only seven or fewer points were calibrated successfully, the calibration of the missing points was repeated. An attention getter was presented to attract participants’ attention to the screen. Then, the experimenter started the experiment with a button press. The stimuli were automatically presented by Clear View (Tobii Technology, Sweden) in a predefined order (see “Stimuli” section).

**Data Analysis**

Participants’ looking behavior was recorded using Clear View (Tobii Technology, Stockholm, Sweden). A fixation filter was activated so that only a constant gaze within a 30-pixel diameter area for at least 200 ms was counted as a fixation (cf. Gredebäck, Theuring, Hauf, & Kenward, 2008).
Learning trials. Data from the learning trials were separated into a preteaching phase (approximately 2,000 ms), during which the children’s attention was directed to the scene by saying, “Look! Isn’t it nice?” (see “Stimuli” section), and a teaching phase (approximately 2,000 ms), during which the model taught the novel word by saying, “This is a taan!” Four areas of interest (AOIs) were defined around the two objects, the model’s face, and the model’s hand (see third picture of Figure 1 for an example), and infants’ fixations to the four AOIs and the background (i.e., counting as fifth AOI) were analyzed with the help of Clear View. Trials in which no eye gaze was registered at all were omitted from further analysis (out of the 10 trials, on average 1.3 had to be omitted). The relative proportion of looking times to each of the five regions was first computed separately for each trial and subsequently averaged across trials for every child. To statistically analyze participants’ fixation behavior during the teaching phase, the data were entered into a two-way analysis of variance (ANOVA), with object (head, hand, gaze object, point object, background) as the within-subject factor and age group (14 months, 24 months, adults) as the between-subject factor. Finally, as we were interested in age-related gaze changes across phase, we calculated for every participant and every AOI the mapping changes in gaze by subtracting the proportion of looking in the teaching phase from the preteaching phase (see Figure 3B). Data were entered into a multivariate analysis of variance (MANOVA), with age as the independent factor (14 months, 24 months, adults). Post-hoc comparisons were employed to examine which AOI mapping changes differed from 0.

Test trials. Data from the test stimuli were separated into the baseline phase, which included the time before the model’s voice could be perceived (varying between approximately 1,000 ms and 2,000 ms), an attention-directing phase (e.g., “There! Look!” varying between 2,000 and 3,000 ms), and a naming phase (“A taan!”) of exactly 1,000 ms. The naming phase included data from 360 ms to 1,360 ms after the onset of the target word (cf. Swingley & Aslin, 2002). Two equally sized AOIs were defined around the two objects (see fourth screenshot of Figure 1 for an example), and participants’ fixations to both objects were analyzed with the help of Clear View. Trials in which no eye gaze at all was registered were omitted from further analysis as were infants from whom data were obtained in fewer than two test trials. The relative proportion of looking times to each target was first computed separately for each test trial and subsequently averaged across trials for every participant. We decided to calculate and statistically analyze the proportion of looking times because we were interested in the relative preference for one object over the other, independent of the overall amount of looking at the objects (cf. Houston-Price et al., 2006; Wu & Kirkham, 2010). Data were statistically analyzed using a three-way ANOVA, with the within-subjects factors of cued object (gaze object, point object) and phase (baseline, attention, naming) as well as the between-subjects factor of age group (14 months, 24 months, adults).

RESULTS

Learning Trials

Figure 3A shows the means for participants’ fixations on the different AOIs. The ANOVA yielded a main effect of object, $F(4, 164) = 18.525$, $p < .001$, $\eta_p^2 = .311$, and a significant interaction between object and age Group, $F(8, 164) = 3.776$, $p < .001$, $\eta_p^2 = .156$. To further
FIGURE 3 (A): The average proportion of looking times directed to the areas of interest by the 14-month-olds, the 24-month-olds, and the adult participants in the teaching phase. Error bars indicate the standard errors of the means. (B): Mapping of gaze changes from the preteaching phase to the teaching phase, depicted as differences in the average proportion of looking in the teaching phase and the preteaching phase. Error bars indicate the standard errors of the means. (C): The average proportion of looking times directed to the gaze object (left bars) and to the point object (right bars) by the 14-month-olds, the 24-month-olds, and the adult participants in the naming phase. The black horizontal line marks 50% (i.e., chance performance). Error bars indicate the standard errors of the means.
examine the effects, we first performed post-hoc t-tests to examine looking preferences within each age group. These showed that the 14-month-old infants spent more time looking at the head than they did at any other areas (all ps < .01) and that they spent less time looking at the hand than at the gaze object and point object (all ps < .05; all other ps > .24), suggesting that the head area strongly attracted infants’ visual attention.

The 24-month-old infants looked more at the head than at the hand and gaze object (all ps < .01), less at the head than at the point object and the background (all ps < .05), and more at the point object than at the gaze object, t(15) = 4.068, p = .001, and more at the background than at the gaze object, t(15) = 2.315, p < .05 (all other ps > .24). This shows an emergent preference to orient along the pointing cue than along the gaze cue.

The adults looked more at the head than at the hand and the gaze object (all ps < .05), less at the hand than at the point object and the background (all ps < .05), and more at the point object than at the gaze object, the hand, and the background (all ps < .01; all other ps > .15). The pattern of results indicates that the point object strongly attracted adults’ visual attention.

The MANOVA on participants’ mapping changes in gaze revealed no significant effect, F < 1. We subsequently dropped age as factor from the analysis and performed one-sample t-tests to examine whether mapping changes differed significantly from 0. The t-tests showed that only mapping changes to the point object differed from 0, t(43) = 2.198, p < .05 (all other ps > .27), suggesting that all three age groups showed a shift to the point object in the transition from the preteaching phase to the teaching phase.

Test Trials

A Levene’s test showed that the error variance was equal across groups (all ps > .14) so that the data could be analyzed by an ANOVA. The three-way ANOVA revealed a significant effect of cued object, F(1, 41) = 11.352, p < .01, η² = .217, a significant interaction between the factors cued object and age group, F(2, 40) = 5.114, p = .01, η² = .2, which was further qualified by a three-way interaction between all factors, F(4, 82) = 2.985, p < .05, η² = .127 (all other ps > .32). To explore this effect in greater detail, we conducted post-hoc t-tests for every phase and age group. For the baseline phase, there was no preference for one of the objects in any of the age groups (on average, point object = 53.6%; gaze object = 46.4%; all ps > .30). Likewise, there was no preference during the attention phase for the infants (on average, point object = 54.4%; gaze object = 45.6%; all ps > .25). For the adult group, the difference in looking times to both objects (point object = 63.3%; gaze object = 36.7%) approached significance, t(11) = 1.851, p = .09, probably as they partly already anticipated the utterance of the target word. The post-hoc tests for the naming phase confirmed that the 14-month-old infants looked on average longer at the gaze object than at the point object, t(15) = 2.392, p < .05 (see Figure 3C). The 24-month-old infants looked significantly longer at the point object than at the gaze object, t(15) = 2.337, p < .05, as did the adult participants, t(11) = 6.366, p < .001.

Additionally, to examine whether the effect was driven by a few subjects only, we conducted an individual-level analysis for each age group. To this end, we decided for each participant whether he looked on average more at the gaze object or the point object during the test phase. Twelve of the sixteen 14-month-old infants showed a preference for the gaze object. A chi-square test revealed that this was different from chance, χ²(1, 16) = 4.000, p < .05.
Numerically similar results were obtained for the 2-year-olds; 12 of the 16 infants had a preference for the point object, $\chi^2(1, 16) = 4.000, p < .05$. For the adults, 11 of the 12 participants showed a preference for the point object, $\chi^2(1, 12) = 8.333, p < .01$.

DISCUSSION

We investigated which of two social cues 14- and 24-month-old infants as well as adults rely more on when learning to match a novel word form to a referent. In particular, we examined on which cue participants would preferentially rely, when both gaze and point cues are available but in conflict with each other. Participants were presented with movies in which an actor repeatedly gazed at one and pointed to the other of two objects while teaching the participants a novel word. In a subsequent test phase, participants’ looking times to the named object revealed that the 14-month-olds relied to a greater extent on the actor’s gaze cues, whereas the 24-month-olds and the adults relied more on the actor’s pointing behavior when relating words to referents. This result supports the hypothesis that there are developmental differences in the priority of certain social cues over others in the early social pragmatics of language acquisition.

It is important to note that the visual preferences for one over the other object were only found in the naming phase of the test trials. That means that only when the novel name was uttered would participants look longer at one of the objects. The fact that no such preference was found in the baseline phases of the test trials excludes the possibility that the naming effect could be due to a general object enhancement effect. Furthermore, no preference for one of the objects was found during the attention phase (e.g., “Look!”; “There!”). This strongly suggests that it is the participants’ perception of the novel word that guided their attention to one of the objects and not an unspecific attention-enhancing mechanism like the onset of the actor’s voice. Together, these results suggest that it was the naming of the novel word that directed the participants’ attention to one of the objects and that participants thus successfully matched a novel word form to a referent.

The present findings advance current theoretical discussions on two points. First, they provide empirical evidence that the relative impact of social cues of infants’ word learning changes over time. In particular, research has suggested several biases and mechanisms that underlie young children’s acquisition of novel words such as a mutual exclusivity bias (Markman, 1989), a shape bias (Landau, Smith, & Jones, 1988), pragmatic inferences (Diesendruck, Hall, & Graham, 2006), or their ability for statistical learning (Yu & Smith, 2007). Furthermore, it has been suggested that social cues play an important role in early lexical learning (e.g., Baldwin, 1993; Tomasello, 2003). However, current models of word learning have primarily examined whether and when infants use social cues and how these cues interact with other mechanisms (e.g., Hollich et al., 2000; Yu & Ballard, 2007). For example, Yu and Ballard’s model integrated social and statistical cues in an attempt to unify accounts on word learning. Yet, they treat social information as a unitary concept by stating that infants use social cues (from a particular age onward). Our results therefore advance the current theoretical debate on early word learning by the finding that there are multiple social cues whose relative impact may change in the course of development (cf. Deák, Walden, Yale Kaiser, & Lewis, 2008) as they probably follow different routes of acquisition (Sodian & Thoermer, 2004). That is, the main question for future research is not whether infants use social cues, but to which extent they rely on which cues in the course of development.
Second, the results contribute to recent discussions about the ontogeny and relevance of young children’s preferential reliance on the pointing gesture (Grassmann & Tomasello, 2010; Palmquist et al., 2012; Palmquist & Jaswal, 2012) and the relevance of gaze cues (Emery, 2000). In particular, it has been shown that preschool children rely on a person’s pointing gesture to an object rather than to the verbal information given by this person (Grassmann & Tomasello). Furthermore, it has been shown that 3- to 4-year-old children attribute knowledge to a pointing person even when that person does not have visual access to the occlusion event (e.g., Palmquist & Jaswal). Concerning the origins of this tendency, Palmquist and colleagues speculated that it is the result of learning processes. In other words, it was hypothesized that children in the course of the 1st years of life learn that the pointing gesture is a valid indicator of another persons’ referential intention and reliability, regardless of whether the informant had visual access to an event or not. Our study extends these findings by showing that 2-year-old children, but not 14-month-old children, show a similar effect. Despite a person’s visual attention to one object while teaching the children a novel word, the older children relied on the actor’s pointing cue to another object. In contrast, the younger children relied on the actor’s gaze cue, supporting claims about the priority of eye gaze in young infants’ social perception and social learning (Baron-Cohen, 1995; Emery, 2000; Moore, 2006). Our results therefore support the hypothesis that young children’s strong reliance on the pointing cue is the result of learning (i.e., experience with the pointing cue).

The current findings of developmental changes in infants’ preferential reliance on gaze and pointing cues in early word learning leads to the question of the nature of the psychological mechanisms subserving this change. On the one hand, simply learning the relevance of pointing cues based on experience could play an essential role as suggested by Palmquist et al. (2012). This suggestion is in line with theoretical assumptions stating that infants become socialized in human communities. Social signals may differ in different communities, and hence, their learning may well be a product of enculturation and socialization (Rogoff, 2003). It has been reported that the penchant to pointing as well as the form of pointing varies across human cultures. For example, some Australian and African tribes use their lips to point to an object (see Wilkins, 2003). In other words, people in Western societies rely heavily on index-finger pointing to refer to objects or events, and infants have to learn to engage in these activities (see also Bibok et al., 2008).

Furthermore, this development might be related to the development of attentional shifting in situations of information conflict. In particular, research during the past decades has shown that the ability to voluntary shift attention develops in the course of the 1st years of life (Colombo, 2001; Johnson et al., 1991), and it has been argued that the ability to flexibly shift attention and resolve conflicts shows a marked increase from the age of 2 years onward (for a review, see Garon, Bryson, & Smith, 2008). Accordingly, younger infants’ preferential reliance on the gaze cue could be partly due to the fact that they are particularly attracted to another person’s face (as evident in their looking behavior during the learning trials), to their reflex-like processing of gaze cues (Emery, 2000), and to the inability to shift attention to the more informative pointing cue. This interpretation is in line with findings that gaze cueing of attention is an often highly automatic process, which even adults cannot consciously inhibit (Driver et al., 1999; for a current overview, see Frischen et al., 2007). That is, even though infants’ gaze following might under some circumstances reveal a progressing understanding of the intentional nature of human gaze (Chow et al., 2008; Sodian et al., 2007), it does not necessarily rely on a reflected estimation of another person’s intention (Driver et al.; Moore, 2006).
Both explanations are not mutually exclusive. It is possible that to learn about the relevance of pointing cues in social interactions (cf. Bibok et al., 2008; Palmquist et al., 2012), infants first need to develop a higher attentional flexibility. If this interpretation is true, then our results provide evidence for the broader theoretical notion that the development of attentional control plays an important role in infants’ progressing ability to employ social cues (for similar considerations about the grounding of joint attention abilities in perceptual processes, see Deák & Triesch, 2006).

Our study shows, upon a closer look at the learning trials, that the different age groups scanned the actor and her behavior differently. To summarize the most interesting differences: Whereas younger infants spent more time looking at the actor’s face than at any other AOI and did so more than any other age group (extending studies demonstrating strong face preferences in young infants; e.g., Blass & Camp, 2003), they showed no difference in looking times to one of the two objects during the learning trials. In contrast, the 24-month-olds and adults already displayed a preference for the point object in the learning phase; moreover, all three age groups spent less time looking at the pointing hand compared with everything else. The finding that the 24-month-olds and adults focused much more on the point object and only marginally on the pointing cue itself shows that they were able to quickly extract the relevant social cue and direct their attention to the relevant information. This is in line with the results from our test trials that both age groups attached the novel label to the object to which the actor was pointing. Interestingly, it should be noted that the 24-month-old children, even though preferentially attaching the novel word to the pointed object, still showed a greater sensitivity to the actor’s face as indicated by their looking behavior in the teaching phase. This indicates that by this age, sensitivity and reliance were separable, probably based on their advanced understanding of other peoples’ intentions and motivations (Moore, 2006). Finally, it should be noted that the fact that all age groups focused least at the actor’s pointing cue in the learning phase renders it implausible that the difference in 14- and 24-month-olds’ reliance on gaze and pointing cues was due to a differential sensitivity for the saliency of the hand between both age groups.

Already in the learning phase, the 24-month-old children and adults spent more time looking at the point object than at the gaze object. This suggests that during the observation of conflicting gaze and pointing cues, these two age groups predominantly oriented along the pointing cue. This confirms earlier studies with adults (Tomoko et al., 2004) and extends these findings to 2-year-old children. However, no such preference was found for the 14-month-olds. Nevertheless, a closer look at participants’ mapping changes of their gazes across the different phases of the study (i.e., from preteaching to the teaching phase) revealed for all age groups an increase in focus on the point object. How can this be explained, given that the 14-month-olds eventually relied more on the gaze cue than on the pointing cue? There are two different possible interpretations—a cognitive explanation and one based on attentional processes. On the one hand, given the preference for the point object among the older participants, the results could indicate a nascent sensitivity for the relevance of pointing already in 14-month-old infants, which, however, is still given less weight than the gaze cue. On the other hand, the effect could also be due to short-term saliency effects, as the hand moved during the execution of the pointing gesture and could thus have directed participants’ attention to the object that was close to the hand (in the sense of reflexive orienting; cf. Ristic & Kingstone, 2009). Interestingly, it should be noted that similar effects have also been reported by a study of Houston-Price and colleagues (2006, Experiment 2). The authors presented 15-month-old children with two objects while...
teaching them a novel word. One of the objects was socially cued, while the other object was moving and thus attracting infants’ attention. The authors reported that even though in the training phase the moving object captured infants’ attention to a greater extent, infants nevertheless attached the novel label to the socially cued word. Further research is needed to investigate these possible explanations in more detail.

From an applied point of view, our results suggest that whenever one wants to support infants’ language acquisition and, in particular, their word learning, one needs to take into account that the impact of the different social cues changes over time. The present findings show that gaze cues are more important around 1 year of life, whereas pointing cues become a powerful tool for word learning in the course of the 2nd year of life. Further research is necessary to examine in greater detail how and under which circumstances infants and toddlers benefit from the different social cues provided by their environment and how this knowledge could be used to facilitate language acquisition in educational settings. Taken together, the present study provides evidence for developmental changes in priority of gaze and pointing cues in infants’ word learning. It shows that 14-month-old infants preferentially rely on gaze cues, whereas 24-month-old children and adults preferentially rely on pointing, even when in conflict with gaze.

ACKNOWLEDGMENTS

We thank the parents and infants who participated in the study. Furthermore, we are grateful to Marlies Vissers for her help preparing the stimuli and Angela Khadar for her support with data acquisition as well as for proofreading the manuscript. Finally, we would like to thank two anonymous reviewers and Susan Graham for valuable feedback.

REFERENCES

Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see: Infants prefer to follow the gaze of a reliable
Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). One-year-olds’ understanding of nonverbal gestures
targeted to a third person.
developing phonological grammar. In P. Avery, B. Elan Dresher & K. Rice (Eds.), Contrast in phonology: Theory,
perception, acquisition (pp. 231–268). Berlin, Germany: Mouton.
Neurophysiology, 100, 1533–1543.
Thomson Associates.
Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). One-year-olds’ understanding of nonverbal gestures
directed to a third person. Cognitive Development, 24, 23–33.
Developmental Science, 13, 252–263.
Social Neuroscience, 5, 441–450.


