The Development of Pointing Perception in Infancy: Effects of Communicative Signals on Covert Shifts of Attention

Moritz M. Daum  
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, and University of Zurich

Julia Ulber  
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, and Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Gustaf Gredebäck  
Uppsala University

The present study aims to investigate the interplay of verbal and nonverbal communication with respect to infants’ perception of pointing gestures. Infants were presented with still images of pointing hands (cue) in combination with an acoustic stimulus. The communicative content of this acoustic stimulus was varied from being human and communicative to artificial. Saccadic reaction times (SRTs) from the cue to a peripheral target were measured as an indicator of the modulation of covert attention. A significant cueing effect (facilitated SRTs for congruent compared with incongruent trials) was only present in a condition with additional communicative and referential speech. In addition, the size of the cueing effect increased the more human and communicative the acoustic stimulus was. This indicates a beneficial effect of verbal communication on the perception of nonverbal communicative pointing gestures, emphasizing the important role of verbal communication in facilitating social understanding across domains. These findings additionally suggest that human and communicative (ostensive) signals are not qualitatively different from other less social signals but just quantitatively the most attention grabbing among a number of other signals.

Keywords: infants, action perception, pointing

A large amount of everyday interaction involves the communication of information to others by either verbal or nonverbal means. If, for example, you want to direct the attention of another person to a certain aspect in the environment, you can not only tell him or her to look there but you can achieve the same goal by pointing at the respective location. To be understood as directed toward a location in space, these nonverbal pointing gestures need to be embedded into a communicative context (Tomasello, Carpenter, & Liszkowski, 2007). In the present study, we aimed to investigate the development of infants’ perception of pointing gestures and the nature of its communicative context, that is, whether it needs to have human and communicative qualities or whether it may involve a broader range of attention-grabbing signals.

At around their first birthday, infants start to point to others (Legerstee & Barillas, 2003; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, & Tomasello, 2007; Meltzoff & Brooks, 2008) and follow other people’s pointing gestures (Carpenter, Nagell, & Tomasello, 1998; Deák, Flom, & Pick, 2000; Leung & Rheingold, 1981; Morissette, Ricard, & D’ecarie, 1995; von Hofsten, Dahlström, & Fredriksson, 2005). At this age, pointing already serves different functions such as informing another person about an interesting object (i.e., declarative pointing) or requesting something the infant wants to have (i.e., imperative pointing; Bates, Camaioni, & Volterra, 1975).

These findings indicate that pointing is used as a means of nonverbal communication (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Infants use pointing gestures to communicate something to somebody else at an age when they are not yet able to communicate the same content verbally. However, the pointing gesture itself can only be understood in a communicative context. Tomasello and colleagues (Tomasello et al., 2007) nicely describe this fact by saying that “by itself, pointing is nothing” (p. 706). When infants observe an actor who looks at his or her own wrist while pointing at a hidden target infants do not follow the pointing direction in order to retrieve the hidden object (Behne, Carpenter, & Tomasello, 2005). To perceive pointing as goal directed, the observer and the pointer need to establish a common ground for their interaction in order to make sure that both attend to the same thing and both know about this (Clark, 1996; Tomasello et al., 2007).
This common ground is often established by communicative signals (also referred to as ostensive signals; e.g., Csibra & Gergely, 2009). These communicative signals convey information about the communicative intent, preparing an observer to process forthcoming information. They can be presented in the visual or the auditory modality. Infants have been shown to prefer mutual gaze over averted gaze from birth (Farroni, Csibra, Simion, & Johnson, 2002). Similarly, infants and even newborns prefer infant-directed speech (so called motherese) over adult-directed speech (Cooper & Aslin, 1990; Werker & McLeod, 1989).

Infants furthermore expect that concurrently occurring communicative signals corefer to the same object (Gliga & Csibra, 2009). Gliga and Csibra reported that 1-year-old infants were surprised when an object that was previously named by the experimenter did not appear at the location where the experimenter was pointing to. Besides, infants learn words more easily in a context of communicative and deictic signals (Woodward, Markman, & Fitzsimmons, 1994), and this beneficial effect has been interpreted in the sense that infants expect verbal labels and deictic gestures to co-occur and to be co-referenced (Baldwin et al., 1996; Tomasello, 2001).

The phenomenon of the beneficial influence of communicative signals on various tasks was nicely captured by Csibra and Gergely (2009), who introduced the term “natural pedagogy,” a specific communication system that helps to process relevant information by marking it via the use of communicative signals, which indicate that the communicated information is specifically relevant for the receiver. The framework of natural pedagogy is partially supported by Senju and Csibra (2008), who demonstrated that 6-month-olds followed an adult model’s gaze only when the gaze shift was preceded by a communicative signal (either direct gaze in comparison to averted gaze or infant-directed speech as compared with adult-directed speech).

Following the logic of Senju and Csibra (2008), the aim of the present study was to further investigate how communicative signals modulate infants’ perception of others’ referential actions, in our case pointing gestures. In addition, going beyond the question of whether or not a communicative signal influences the perception of pointing gestures, we were interested in the extent to which the communicative signal needs to be human and referential or not. When you think of a ringing doorbell, this artificial sound naturally entails a communicative function that is to inform somebody to open the door. Accordingly, we furthermore wanted to study whether communicative signals that are not human and referential (e.g., an artificial sound) are already effectual to focus attention to relevant aspects of the observed action.

To investigate these issues, different acoustic stimuli were presented along with the still picture of a pointing hand, followed by a peripheral target. The acoustic stimuli varied in the degree to which they included human and referential information. Infants’ perception of the pointing gestures was measured via the modulation of their covert attention in the direction of the pointing gesture using a spatial cueing paradigm. Before we describe the present study in more detail, we provide a brief introduction to the spatial cueing paradigm in the following.

An observer’s covert attention can be modulated toward a specific location without explicitly looking at it (Posner, 1980). When presented with a directional cue, for example an arrow, followed by a peripheral target, adults’ reaction times are reduced if the target appears at a location that is congruent with the direction of the cue compared with a target that appears at a location that is incongruent with cue direction. We refer to this facilitation of reaction times as cueing effect. Similar cueing effects as in adults have been demonstrated to be already present in 4-month-old infants (Johnson, Posner, & Rothbart, 1991).

Cueing effects have been demonstrated not only for abstract cues like arrows but also in response to more socially relevant cues. Observed gaze shifts modulate covert attention in adults (Driver et al., 1999; Friesen & Kingstone, 1998; Langdon & Smith, 2005), infants (Farroni, Johnson, Brockbank, & Simion, 2000; Hood, Douglas, & Driver, 1998), and newborns (Farroni, Massaccesi, Pividori, & Johnson, 2004). The same is true for observed grasping hands both in adults (Daum & Gredebäck, 2011b; Fischer, Prinz, & Lotz, 2008) and infants (Daum & Grédebäck, 2011a). With regard to pointing, few studies have reported cueing effects based on observed pointing gestures. These studies have shown that pointing gestures cause a cueing effect in adults that is similar to arrows and grasping hands (Daum & Gredebäck, 2011b). Pointing gestures made with the index finger produce a significantly larger cueing effect than pointing with other fingers (Arita & Watanabe, 2009), and pointing direction interferes with head or gaze direction (Langton & Bruce, 2000). In infants, only one study has explicitly investigated the connection between covert attention and pointing gestures (Gredebäck, Melinder, & Daum, 2010). This event-related potential (ERP) study demonstrated that congruent and incongruent pointing gestures elicit different ERP components already in 8-month-old infants, which is much earlier than the usually reported developmental onset of pointing at around 12 months of age. Besides this study, nothing is known about whether infants’ attention can be modulated by still pictures of manual pointing gestures and the degree to which communicative signals support or interfere with the perception of these gestures.

The use of a spatial cueing paradigm is specifically appropriate to measure the perception of pointing gestures. It measures the modulation of attention in the direction of the observed pointing gesture. In order to shift attention, an observer has first to perceive the directionality of the pointing gesture. Only then is attention modulated in the pointing direction. The modulation of covert attention can be seen as an implicit measure for pointing perception. An observer does not have to explicitly follow another person’s pointing gesture. He or she does not have to decide when and toward which location in space a gaze shift has to be initiated. For this reason, one could expect that infants are found to perceive the directedness of a pointing gesture earlier (as reported by Gredebäck et al., 2010) when measured via a spatial cueing paradigm as compared with when pointing perception is measured via more explicit measures.

To summarize, the aim of the present study was to investigate the development of infants’ perception of pointing gestures and its modulation by communicative signals by using an implicit spatial cueing paradigm. Three main research questions guided this research.

1 In the present study, we use the terminology communicative signals for a broad range of acoustic signals of human and nonhuman nature as artificial sounds, such as the ringing of a doorbell, as well as to entail communicative information.
The first question was based on the fact that, at least for infants, a pointing gesture by itself has no specific referential meaning (Tomasello et al., 2007). We were specifically interested whether infants could infer the directionality from an observed pointing gesture per se or whether a referential communicative context was necessary to perceive a pointing gesture as being directed toward a location in space. Previous research has shown that infants can infer the directionality from a grasping hand without any additional communicative context given (Daum & Gredebäck, 2011a). Furthermore, adults have been shown to shift their attention in the direction of an observed pointing gesture without an additional communicative signal (Daum & Gredebäck, 2011b). To explore the necessity of a communicative context, we measured 12-month-old infants’ perception of pointing gestures in both a communicative speech condition, in which the visually displayed pointing gestures were accompanied by an acoustic stimulus of human and communicative nature (a female voice saying “Look! – There!”) and a no-sound condition with no additional acoustic stimulus (Experiment 1). The short answer to this research question is that a communicative signal is necessary to perceive the directionality of a pointing gesture.

Second, we were interested in the extent to which the acoustic stimulus presented along with the pointing gesture needs to feature human and referential qualities or whether any acoustic stimulus is sufficient for infants to focus their attention toward the pointing gesture and perceive its directionality. For this reason, we ran two additional conditions in Experiment 1. In the reversed speech condition, the acoustic stimulus was modified to keep some human features but was no longer referential (therefore, the acoustic stimulus from the communicative speech condition was played backward). In the artificial sound condition, a computer-generated sound was played along with the pointing gesture. Via these two additional conditions, our aim was to investigate whether the influence of an acoustic stimulus follows an all-or-none rule or whether it is graded in nature. Via these four conditions, we tested four possible assumptions: (a) In order to perceive the direction of a pointing gesture, the required communicative context might be exclusively generated by human and referential acoustic signals. Accordingly, a cueing effect would only occur in the communicative speech condition. (b) A communicative context might be created by any acoustic signal as long as it is of human nature. Accordingly, a cueing effect should be present in the two human conditions (communicative and reversed speech). (c) Any acoustic signal might put an observed pointing gesture into a communicative context. In this case, all conditions with an acoustic signal would result in a cueing effect. (d) The influence of the acoustic signal might be graded depending on the communicative nature of the acoustic stimulus. This would result in an increased size of the cueing effect over the four conditions with the following order (from weakest to strongest): no sound–artificial sound–reversed speech–communicative speech.

Third, we were interested in the development of the perception of pointing gestures. The common finding is that infants usually start to point and to follow others’ pointing around their first birthday (e.g., Carpenter et al., 1998). This finding is in line with a recent spatial cueing study that showed that infants shift their attention in the direction of an observed grasping gesture at a similar age as they start to intentionally grasp for objects (Daum & Gredebäck, 2011a). In contrast, the only study so far that used a spatial cueing paradigm to study the neurophysiological basis of infants’ perception of pointing gestures using electroencephalography (EEG) revealed that 8-month-old infants are already sensitive to the directionality of a pointing hand (Gredebäck et al., 2010). Accordingly, to explore the extent to which the modulation of attention is related to infants’ own pointing skills or not, we tested infants at an age when they are usually about to start pointing and to follow other’s pointing, that is, at the age of 12 months (Experiment 1). Additionally, to further explore whether the reported sensitivity to pointing direction at an earlier age can also be transferred to a behavioral version of the spatial cueing paradigm, we additionally tested a group of younger infants at the age of 10 months (Experiment 2).

Experiment 1

In Experiment 1, infants’ modulation of covert attention based on the perception of a goal-directed pointing gesture was investigated using a spatial cueing paradigm (Daum & Gredebäck, 2011a; Hood et al., 1998). The basic paradigm consisted of the presentation of a central attention grabber, followed by a centrally presented image of a hand pointing either to the left or the right (cue) and a toy (target) subsequently being presented in the periphery of the screen at a location that was either congruent or incongruent with the pointing direction. The presentation of the cue was accompanied by an acoustic stimulus that was varied in the above-mentioned conditions between subjects. Infants’ saccadic reaction times (SRTs; i.e., the time before the infants fixated the target) were assessed using eye-tracking technology.

Method

Participants. A total of 96 infants at the age of 12 months participated successfully in Experiment 1; 24 in each of the four conditions (51 girls, 45 boys, mean age = 12 months 3 days, range = 11 years 13 months–12 years 15 months). Twenty-three additional infants (11 girls, 12 boys) were tested but not included in the final sample due to fussiness (n = 2), technical problems (n = 2), fewer than 10 valid trials (n = 8), or mean RTs of three standard deviations above the mean (evaluated separately per each condition, n = 11). Contact information for the infants was obtained from public birth records. All infants were born full term (37–42 weeks gestation) and with normal birth weight (> 2.500 g). All parents gave informed consent before the study. The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki.

Test environment, stimuli, and apparatus. The laboratory was unfurnished except for the test equipment. The infants were seated in a car safety seat (Maxi Cosi Cabrio), which was placed in front of the eye tracker. The stimuli were presented, and gaze was measured using a Tobii 1750 near infrared eye tracker (Tobii Technology AB, Danderyd, Sweden) with an infant add-on (precision: 1 deg, accuracy: 0.5 deg, sampling rate: 50 Hz) and the software Clearview (Version 2.7.1; Tobii, Danderyd, Sweden). A 9-point infant calibration was used. During calibration, a blue and white square expanded and contracted (extended diameter = 3.3 visual degrees) in synchrony with a sound. Viewing distance was approximately 60 cm.

An exemplary trial is depicted in Figure 1. Each trial started with an attention catcher that consisted of a looming stimulus (a...
multicolored wooden tower, a yellow rubber duck, a multicolored soft textile cube, or a multicolored soft textile cone; horizontal and vertical dimensions: maximum 4.5 deg, minimum 2.3 deg) presented at the center of the 17-in. monitor (24.8 \times 20.7 deg) in combination with an attention-catch sound. This sound was constant for each stimulus but varied between the stimuli.

As soon as the infant fixated the central stimulus, the experimenter started each trial manually. A trial consisted of the presentation of a cue that was presented at the center of the screen and a target that was presented peripherally. The cue consisted of the presentation of a human hand closed to a fist presented for 600 ms and the same hand pointing to the left or the right for 720 ms. The central cue then disappeared and was followed by a renewed presentation of the initial stimulus (now referred to as the target). The target appeared at a peripheral location on the screen that was either congruent or incongruent with the direction of the pointing hand. Target appearance was accompanied by a short version of the initial attention-grabbing sound played for 800 ms. The distance from the nearest part of the hand to the nearest part of the target was 9.3°. The target remained visible until the infant looked at it for approximately 1,000 ms or until 5,000 ms had elapsed. Then a new trial began with a different centrally presented looming stimulus.

The visual presentation of the cue was presented in combination with an additional acoustic stimulus that was manipulated in the following four conditions. In the communicative speech condition, the additional acoustic stimulus consisted of a female human voice saying “Look!” (German original: “Schau mal!”) as the first part of the cue (i.e., the fist) appeared, followed by “There!” (German original: “Da!”) as the second part of the cue (i.e., pointing hand) appeared. In the reversed speech condition, the first part of the cue was accompanied by the word Look played backward (which sounded like “Lam uasch” in the German original). The second part of the cue was accompanied by the word There played backward (which sounded like “Ad” in the German original). The manipulation of the acoustic stimuli was realized using the software Praat (Boersma & Weenink, 2010). In the artificial sound condition, the presentation of the cue was accompanied by two different computer-generated artificial sounds (first part of the cue: Windows XP Hardware Insert.wav; second part of the cue: Windows XP Hardware Remove.wav). These sounds were different from the four sounds used in combination with the looming attention grabbers and overall lower in pitch. In the no-sound condition, no additional acoustic stimulus was presented.

The order of the targets as well as the relation between the direction of the pointing hand and the location of the target was pseudorandomized. In order to avoid adaptation effects, not more than three repetitions of cue direction, target location, or cue–target relation were allowed. This reduced the possibility that the latency of target-oriented gaze shifts on any given trial was influenced by one of these factors of the previous trial. The maximum number of trials presented was 64. The infants were randomly assigned to one of four different orders that were created using the software MATLAB (MathWorks, Natick, Massachusetts).

Procedure. Infants were tested in the laboratory at a time of day when they were likely to be alert and in a good mood. All infants were tested individually with one parent present. Each participant and his or her parents were first escorted to a reception room. For approximately 10 min, the infant was allowed to explore the room while the research assistant described the test procedure.
to the parents and one of the parents signed a consent form. Then, the infant and one parent were brought to the test room. The research assistant helped the parent to position the infant in the car seat. During stimulus presentation, the parent sat on a chair behind the infant. Parents were instructed not to interact with their children during testing. They were encouraged, however, to put both hands symmetrically close to the child if it appeared necessary to comfort the infant. Once the infant and the parent seemed comfortable, the research assistant left the room and the stimulus presentation was started.

Data analysis. For the analysis of gaze, three square areas of interest (AOI) were defined on the screen. The cue AOI covered the cueing hand (horizontal and vertical dimension: 7.5°), whereas the target AOIs covered each of the targets (horizontal and vertical dimension: 4.7°). A trial was considered to be valid if the infant fixated the central cue for at least 200 ms (Gredebäck, Örnikloo, & von Hofsten, 2006) prior to making a gaze shift to the target. The SRT was defined as the duration between the appearance of the target and the arrival of the infant’s gaze in the respective target AOI (Gredebäck, Johnson, & von Hofsten, 2009). Individual RTs of less than 100 ms and greater than two standard deviations of each individual mean were excluded from analysis. Infants had to provide a minimum number of 10 trials to be included in the final analysis. The analysis was performed in five steps. (a) The mean SRTs were first analyzed using an overall 2 × 4 × 2 (congruency [congruent, incongruent], condition [communicative speech, reversed speech, artificial sound, no speech], sex [male, female]) analysis of variance (ANOVA). (b) The mean SRTs of the different conditions were further analyzed discretely by separate t tests for each condition comparing the SRTs of congruent and incongruent trials. (c) The number of infants who shifted their gaze faster toward the congruent target was compared with the number of infants who shifted their gaze faster to the incongruent target separately for each condition using a nonparametric Sign tests for each condition. (d) Both the parametric and the nonparametric data were rank ordered on the basis of the degree of communicative nature of the cue (i.e., communicative speech–reversed speech–artificial sound–no speech), and correlations were calculated with the size of the cueing effect (SRT_{incongruent} − SRT_{congruent}) using a Spearman’s rank correlation and with the number of infants showing a cueing effect using Somer’s d coefficient. (e) Finally, in order to test the four assumptions about the nature of the acoustic stimulus (see Test Environment, Stimulus, and Apparatus section above), linear regressions were calculated to test the following four assumptions: The communicative context (a) needs to have human and communicative qualities (human communicative cue model), (b) needs to only have human qualities (human cue model), (c) can be created by any acoustic signal (acoustic cue model), or (d) is graded in nature based on the human and referential nature of the cue (graded model).

Results

Mean SRTs are displayed in Figure 2, and the number of infants who shifted their gaze faster toward the congruent versus the incongruent target is displayed in Figure 3. P values are reported two-tailed throughout. The average number of trials accomplished was 42.3 (SD = 12.8; range = 16–61) in the communicative speech condition, 35.2 (SD = 13.2; range = 13–60) in the reversed speech condition, 39.9 (SD = 13.0; range = 14–61) in the artificial sound condition, and 37.3 (SD = 10.5; range = 19–57) in the no-sound condition. A comparison of the number of trials using separate t tests revealed no significant differences between the four conditions (all ps > .20; Tukey’s honestly significant difference test corrected for multiple comparisons).

Overall analysis. The omnibus analysis of the SRTs revealed a significant Congruency × Condition interaction, F(3, 88) = 3.35, p = .02, η² = .10, indicating an increase of the cueing effect with increasing communicative nature of the acoustic stimulus (see Figure 2). There were two marginally significant effects: a main effect of congruency, F(1, 88) = 3.72, p = .06, η² = .04, and a Congruency × Sex interaction, F(1, 88) = 3.09, p = .08, η² = .03. Although only marginally significant, the SRTs tended to be faster overall in congruent compared with incongruent trials, and this difference appeared to be larger for girls relative to boys. No further effect or interaction was significant (with all ps > .26).

The results of the parametric analysis were further supported by two nonparametric analyses of the number of infants showing a cueing effect (i.e., shifting their gaze faster to the target in congruent compared with incongruent trials) and the number of infants showing an opposite cueing effect. First, the above-reported marginal effect of congruency was supported by a Sign test that showed that over all conditions, 61 infants showed a cueing effect compared with 35 who showed an opposite cueing effect, and that this difference was significant (p = .01).

Second, when the data were ranked according to the communicative nature of the acoustic stimulus presented with the cue (ranking order: communicative speech, reversed speech, artificial sound, no sound; see separate analyses below and Figure 3), the size of the cueing effect was significantly correlated with the communicative grade of the cue (r = −.24, p < .05; Spearman’s
rank correlation), and the number of infants showing the cueing effect decreased across the different experimental manipulations ($d = .20, p < .05$; Somer’s $d$ coefficient). This result indicates that the more human and the more referential the additional acoustic stimulus, the stronger the cueing effect. In order to provide a more detailed picture of the data, separate analyses for each condition are reported in the following.

**Separate analyses per condition.** In the communicative speech condition, the 12-month-olds showed significantly faster SRTs in congruent trials ($M = 395.29$ ms, $SD = 54.02$ ms) compared with incongruent trials ($M = 421.91$ ms, $SD = 51.52$ ms), $t(23) = 2.76$, $p = .01$. This result was confirmed by a nonparametric Sign test, revealing that 19 infants showed a cueing effect, whereas only five showed an opposite cueing effect ($p = .007$). In the reversed speech condition, no SRT differences between the congruent ($M = 422.74$ ms, $SD = 15.97$ ms) and incongruent trials ($M = 436.06$ ms, $SD = 15.08$ ms) was found, $t(23) = 1.41$, $p = .17$. Sixteen infants shifted their gaze faster to congruent targets, and eight did the opposite ($p = .15$; Sign test). In the artificial sound condition, again no significant effect of congruency was found; the SRTs did not differ between congruent ($M = 420.43$ ms, $SD = 10.51$ ms) and incongruent ($M = 428.95$ ms, $SD = 11.17$ ms) trials, $t(23) = 0.90$, $p = .38$. Fourteen infants shifted their gaze faster to congruent targets, and 10 did the opposite ($p = .54$; Sign test). Finally, in the no-sound condition, again no SRT difference between congruent ($M = 444.38$ ms, $SD = 14.06$ ms) and incongruent ($M = 430.62$ ms, $SD = 11.32$ ms) trials was found, $t(23) = 1.48$, $p = .15$. Twelve infants shifted their gaze faster to congruent targets, and 12 did the opposite ($p = 1$; Sign test).

**Regression analysis.** In order to get a clearer picture about the relation between the different nature of the acoustic cues and the size of the congruency effect, we tested the four models, as introduced above, using linear regressions. Although the amount of variance that was explained was not high, the graded model of the congruency effect significantly predicted the size of the congruency effect (standardized $\beta = -0.30$), $t(95) = 3.00$, $p = .003$, adjusted $R^2 = 0.08$; $F(1, 94) = 8.99$, $p = .003$, and did so better than the acoustic cue model ($\beta = 0.27$), $t(95) = 2.74$, $p = .007$, $R^2 = 0.06$, $F(1, 94) = 7.50$, $p = .007$; the human cue model ($\beta = 0.24$), $t(95) = 2.37$, $p = .020$, $R^2 = 0.05$, $F(1, 94) = 5.60$, $p = .020$; and the communicative human cue model ($\beta = 0.22$), $t(95) = 2.16$, $p = .033$, $R^2 = 0.04$, $F(1, 94) = 4.66$, $p = .033$.

**Discussion.** The findings of Experiment 1 showed that infants at the age of 12 months shift their covert attention in the direction of an observed pointing gesture. However, this effect was less robust or even absent in the conditions in which the acoustic stimulus was less human, less referential, or not present. In the following, in order to further explore the developmental trajectory of the cueing effect found in the communicative speech condition, we additionally tested infants at a younger age to see whether adding a referential language cue helps to perceive the directionality of a pointing gesture already at a younger age.

**Experiment 2.** In Experiment 2, a group of 10-month-old infants was tested with the same paradigm that was used in the communicative speech condition of Experiment 1.
Method

Participants. Twenty-four 10-month-old infants participated successfully in Experiment 2 (nine girls, 15 boys, mean age = 10 years 6 months, range = 10 years 0 months–10 years 13 months). Three additional 10-month-olds (two girls, one boy) were tested but not included in the final sample due to fussiness (n = 1), technical problems (n = 1), or mean RTs of three standard deviations above the overall mean (n = 1). As in Experiment 1, contact information for the infants was obtained from public birth records.

Apparatus, procedure, and data analysis. The same apparatus and design were used as in the communicative speech condition of Experiment 1. The SRTs were analyzed using a 2 × 2 (congruency [congruent, incongruent] × sex [male, female]) repeated measures ANOVA in combination with a nonparametric Sign test. The data of Experiment 2 were further compared with the data of the 12-month-olds in the communicative speech condition of Experiment 1 using a 2 × 2 (congruency [congruent, incongruent] × experiment [Experiment 1, Experiment 2]) repeated measures ANOVA.

Results

The average number of trials was 41.5 (SD = 12.4; range 13–59). A t test revealed no difference between the average number of trials between Experiment 2 and the communicative speech condition in Experiment 1 (M = 42.3, SD = 12.8), r(46) = 0.82, p = .76. An overall data analysis yielded no significant effect of congruency, sex, or their interaction (all Fs < 1). The SRTs of the 10-month-olds did not differ between congruent (M = 409.39 ms, SD = 67.38 ms) and incongruent trials (M = 406.44 ms, SD = 68.34 ms). A nonparametric Sign test showed that 11 infants shifted their gaze faster toward the congruent target, whereas 13 did the opposite (p = .84).

The comparison of the data of the 10-month-olds from Experiment 2 with the 12-month-olds from Experiment 1 revealed a significant Congruency × Age interaction, F(1, 46) = 5.43, p = .02, η2 = .11, reflecting the fact that in the communicative speech condition, the 12-month-olds showed a significant cueing effect, whereas the 10-month-olds did not yet do so. Furthermore, we found a marginally significant effect of congruency, F(1, 46) = 3.48, p = .07, η2 = .07, and no effect of age (F < 1).

Discussion

In contrast to Experiment 1, where the 12-month-olds did show a reliable cueing effect in the condition in which the pointing gesture was accompanied by communicative speech, the 10-month-olds in Experiment 2 did not show a cueing effect in the same condition. This is in line with a number of findings mentioned in the introduction that showed that at 12 months of age, infants start to independently point to others (Brooks & Meltzoff, 2008; Liszkowski et al., 2004, 2007) and to follow others’ pointing gestures (Deák et al., 2000; Legerstee & Barillas, 2003; Leung & Rheingold, 1981; Morissette et al., 1995; von Hofsten et al., 2005). Furthermore, this is in line with numerous studies showing that there is a very close relationship between perception and production of goal-directed actions present in early infancy (Daum & Gredebäck, 2011a; Daum, Prinz, & Aschersleben, 2011; Kanakogi & Itakura, 2011; Loucks & Sommerville, 2012; Sommerville & Woodward, 2005; Woodward & Guajardo, 2002).

General Discussion

The present study was designed to investigate 10- and 12-month-olds’ perception of pointing gestures and the nature of the communicative context needed. Three main questions were guiding this research. First, we intended to explore whether a communicative context, provided by a verbal and referential communicative signal is or is not necessary for infants to perceive the directionality of a pointing gesture. Second, we were interested in the extent to which the acoustic stimulus presented along with the pointing gesture need to be human and referential or whether an acoustic stimulus per se is sufficient for the infants to focus their attention at the pointing gesture and perceive its directionality. Third, we intended to explore the extent to which the modulation of attention is related to the age at which infants usually start to point and to follow others pointing.

With respect to the first question of whether a communicative context is necessary, the results showed that 12-month-olds shifted their gaze faster toward a target that appeared at a location that was congruent with the direction of the pointing gesture compared with a target that appeared at a location that was incongruent with the direction of the pointing gesture. This cueing effect was only significant when the pointing gesture was accompanied by an acoustic stimulus that included communicative, human speech (communicative speech condition). In contrast, no cueing effect was present when no acoustic stimulus was presented with the pointing gesture (no-sound condition). These results are in accordance with previous findings indicating that covert attention can be modulated by still pictures of human manual actions and gestures both in adults (Ariga & Watanabe, 2009; Daum & Gredebäck, 2011b; Sato, Koshiyama, Uono, & Yoshikawa, 2010) and infants (Daum & Gredebäck, 2011a; Rohlfing, Longo, & Bertenthal, 2012). With respect to grasping and pointing, the age at which a respective cueing effect is first reported coincides with the age at which infants usually start to actively produce the respective actions, that is, around the age of 5 months for grasping (e.g., von Hofsten, Vishton, Spekel, Feng, & Rosander, 1998) and around the age of 12 months for pointing (e.g., Liszkowski et al., 2007). The present findings are furthermore in line with a wide variety of findings that indicate a beneficial effect of communicative human signals on the processing of observed actions and events (Csibra & Volein, 2008; Senju & Csibra, 2008; Topál, Gergely, Miklósi, Erdöhegyi, & Csibra, 2008).

With respect to the second question about the nature of the acoustic stimulus, the interpretation of the results of the four conditions in Experiment 1 is challenging. A conservative interpretation of the present findings is that a cueing effect was only present in the communicative speech condition, but not in the other conditions. Accordingly, at the age of 12 months, infants need additional referential and human information that serves as communicative context in which a pointing hand is perceived as directed toward a location in space. However, we prefer a more speculative interpretation of the results that is based on the correlations reported between the conditions (when ranked according to their referential and human nature) and the cueing effect and on the results of the linear regression analysis.
These analyses showed that when the acoustic stimulus was less referential and less human, the cueing effect was gradually decreased and that a linearly graded model was best in predicting the size of the cueing effect. One can, thus, speculate that at the age of 12 months, the cueing effect caused by an observed pointing gesture was larger the more human and the more communicative the quality of the accompanying acoustic stimulus was. With this, the present findings go beyond indicating that communicative signals put the observed scene in a pedagogical context. They furthermore provide evidence suggesting that the impact of communicative signals does gradually depend on the nature of the signal. Accordingly, a communicative human cue seems to be not qualitatively but rather quantitatively different from other acoustic signals. The influence of pedagogical and ostensive cues as reported (Csibra & Gergely, 2009) might therefore just be the most attention-grabbing signals among other less attention-grabbing signals. This further suggests that the perception of the direction of a pointing gesture might not follow an all-or-nothing rule with respect to the presented communicative context. On the contrary, the results indicate that adding any acoustic stimulus to a visual stimulus already helps infants to better focus their attention toward the relevant part of a visually presented stimulus. This interpretation is, to some extent, speculative, as a significant congruency effect was only found in the communicative speech condition and the correlation found might be driven by the qualitative difference between either the communicative speech or the no-sound condition and the respective other three conditions. Further studies are needed to provide additional evidence for this hypothesis.

In any case, the optimal communicative context in which the infant’s attention is best modulated in the direction of the pointing gesture was created with an acoustic stimulus that was human and communicative, that is, when the content of the acoustic stimulus did support the content of the observed nonverbal gesture. In the communicative speech condition in Experiment 1, the referential content of the verbal stimulus (“Look! There!”) coincided with the content of the nonverbal stimulus, that is, the pointing gesture, due to the fact that both referential and emphasize that something is going to be communicated and that there is a certain location that is of increased interest.

Similar beneficial effects of co-occurring stimulus dimensions were reported by Bahrick and Lickliter (2000). They showed that when 5-month-old infants were habituated to a rhythm that was presented in two modalities (auditory and visual), they were able to discriminate a novel rhythm, but not when they were habituated to the rhythm in only one modality. On the basis of these findings, Bahrick and colleagues (Bahrick & Lickliter, 2000; Gogate & Bahrick, 1998) formulated the intersensory redundancy hypothesis. This hypothesis assumes that in infancy, information that is presented redundantly across different modalities selectively recruits attention and facilitates perceptual differentiation compared with when the same information is presented in only one modality. In line with a number of related findings (Lewkowicz, 2000; Morrongiello, Fenwick, & Chance, 1998; Zmyj, Jank, Schütz-Bosbach, & Daum, 2011), this suggests that infants are able to integrate and use the information coming from different modalities and that redundant information improves their perception of events. In the present study, the redundant information of a verbal and a nonverbal referential cue was beneficial for perceiving a pointing gesture as goal directed. This suggests that infants at 12 months are able to integrate referential verbal cues and referential action cues, which further suggests that verbal and nonverbal communication are already linked at this age. This suggests that with 12 months of age, not only are action and perception closely related, but action and language are similarly related, as it is the case in adults (Kelly, Özyürek, & Maris, 2010). It remains, however, to be tested whether it is the actual linguistic content that supports the action content or whether the effect is driven by nonlinguistic characteristics of the voice, such as, for example, pitch contour.

Concerning the third question about development of pointing perception, the results of the 10-month-olds tested in Experiment 2 indicate that, in contrast to recent neurophysiological findings (Gredebäck et al., 2010), in the present behavioral task, younger infants did not yet modulate their attention in the direction of the observed pointing gesture.

There are two open issues concerning the development of pointing perception that need to be discussed in more detail. First, it has previously been reported that infants shift their covert attention on the basis of observed grasping gestures at a similar age as when they develop grasping skills (Daum & Gredebäck, 2011a). The conclusion of this study was that the modulation of covert attention during the observation of a grasping hand is based on the infants’ own grasping experience. The same is true for the present findings. The shift of covert attention during the perception of a pointing hand seems likewise to be based on the infant’s own pointing experience, as the age of 12 months coincides with the onset of both skills. However, in the grasping study, no communicative signals were needed in addition to the presentation of the grasping hand in order to cause a shift of covert attention. In contrast, in the present study, the presentation of a communicative and referential signal was essential in order to trigger a respective shift of covert attention.

One reason for this difference might lie in the nature of the respective action. Grasping is performed by and for oneself, for example, in order to get an object that is needed or desired. Grasping does not contain any communicative meaning. Pointing, in contrast, is performed in order to inform somebody else about something. Its primary purpose is communication (Tomasello et al., 2007). The still picture of a disembodied pointing hand used in the present study seems indeed to be “nothing,” and by itself might not entail enough communicative information for a 12-month-old to trigger shifts of covert attention in the same way as in adults. Only when the pointing gesture was integrated into a communicative context, infants did perceive the pointing gesture as a means for communication. For adults, this additional integration into a communicative context is not needed (Daum & Gredebäck, 2011b), as they are able to infer the communicative intention from the mere observation of a pointing hand.

Second, the present results are in contrast to recent findings showing that infants at the age of 8 months already demonstrate differential activation in their EEG signals when presented with congruent and incongruent pointing gestures (Gredebäck et al., 2010). In their study, the authors reported an increased P400 component over posterior temporal areas for congruent compared with incongruent pointing gestures. This component was suggested to have a common source with the increased N200 component found in adults for incongruent compared with congruent pointing gestures over the same areas (for similar findings and interpreta-
tions, see Csibra, Kushnirenko, & Grossmann, 2008; Leppänen, Mouson, Vogel-Farley, & Nelson, 2007; Senju, Johnson, & Csibra, 2006). In contrast, in the present study, only infants at the age of 12 months (thus 4 months later) showed a differential perception of congruent and incongruent pointing gestures.

Two factors might account for this discrepancy. (a) Neurophysiological measures have repeatedly been shown to demonstrate a higher sensitivity of measurement, indicating that information is already processed by the infant brain while it cannot yet be transferred to an overt behavioral response (e.g., in language processing, see Männel & Friederici, 2008). From this perspective, in combination with the reported EEG findings (Gredebäck et al., 2010), the present data provide evidence for a timeline of pointing comprehension going from maturation of cortical networks at 8 months to behavioral expression at 12 months of age. (b) The two studies differed also with respect to their experimental design. In the present study, infants were presented with a pointing hand followed by a peripheral target, and we measured whether they shifted their covert attention in the direction of the pointing gesture. This very/particular experimental design was used to measure shifts of covert attention and involves the prediction of a future aspect of the perceived action, that is, the location of a potential target object. In contrast, in the EEG study, the order of the cue–target presentation was reversed. There, the peripheral target was presented first, followed by the central pointing gesture. This was done in order to measure the infants’ brain responses to the congruency of the pointing hand and not to the congruency of the location of the target that appeared. In this case, the infants’ evaluation of the pointing hand with respect to the previously presented location of the target was measured post hoc.

Previous studies that compared different experimental designs assessing action prediction (e.g., via anticipatory gaze shifts) with experimental designs assessing post hoc action evaluation (e.g., via looking times or pupil dilation) found that the results of the two designs can be dissociated early in life and become associated not until a later point in lifetime (Daum, Attig, Gunawan, Prinz, & Gredebäck, 2012; Gredebäck et al., 2010). In the study by Gredebäck and Melinder (2010), infants were able to predict the outcome of an observed manual feeding action measured via anticipatory gaze shifts beginning with the age of 12 months. However, their pupil responses were already different during the observation of a rational and a non-rational outcome of the action at the age of 6 months. Similarly, Daum and colleagues (2012) used a paradigm following the logic of Woodward (1998) and presented infants with an animated agent moving toward one of two goal objects. Results showed that the infants’ action evaluation measured via looking times reflected the encoding of the goal-directedness of the observed action already at the age of 9 months, whereas at the same age, their action anticipation, measured via predictive gaze, did not. Furthermore, it was not until infants’ third birthday that the predictive gaze pattern reflected the same encoding strategies as the younger infants’ looking times. Similar findings were reported from the domain of physical reasoning (Hood, Cole-Davies, & Dias, 2003; Mash, Novak, Berthier, & Keen, 2006). On the basis of the findings of these studies, the disparity between the findings of the present study and the findings of the EEG study (Gredebäck et al., 2010) might be explained by a potential discrepancy between the different experimental paradigms, and therefore mechanisms used and their respective different dependent variables.

To conclude, in the present study, we investigated infants’ perception of pointing gestures and its interrelation with additional communicative signals. Infants do shift their covert attention in the direction of a perceived pointing hand, but this shift of attention is modulated by the communicative signal in the sense that the more human and the more referential the communicative signal is, the more it helps to focus the attention on the relevant part of the observed stimulus.

References


Call for Brief Reports: Teaching and Training in Psychotherapy Integration

Journal of Psychotherapy Integration will start publishing a new recurring brief reports section titled, “Teaching and Training in Psychotherapy Integration.” This section solicits reports on innovative ways to teach psychotherapy integration or to train novice psychotherapists in integrating therapeutic techniques.

Submitted manuscripts should not exceed 2500 words, including references. Manuscript should be submitted electronically through the journal’s submission portal under the Instructions to Authors section at www.apa.org/pubs/journals/int. The deadline for paper submission will be November 1, 2013.

Please note in your cover letter that you are submitting for this brief reports section. We look forward to your submissions!