PAPER

Young children follow pointing over words in interpreting acts of reference

Susanne Grassmann and Michael Tomasello

Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Abstract

Adults refer young children's attention to things in two basic ways: through the use of pointing (and other deictic gestures) and words (and other linguistic conventions). In the current studies, we referred young children (2- and 4-year-olds) to things in conflicting ways, that is, by pointing to one object while indicating linguistically (in some way) a different object. In Study 1, a novel word was put into competition with a pointing gesture in a mutual exclusivity paradigm; that is, with a known and a novel object in front of the child, the adult pointed to the known object (e.g. a cup) while simultaneously requesting 'the modi'. In contrast to the findings of Jaswal and Hansen (2006), children followed almost exclusively the pointing gesture. In Study 2, when a known word was put into competition with a pointing gesture – the adult pointed to the novel object but requested 'the car' – children still followed the pointing gesture. In Study 3, the referent of the pointing gesture was doubly contradicted by the lexical information – the adult pointed to a known object (e.g. a cup) but requested 'the car' – in which case children considered pointing and lexical information equally strong. Together, these findings suggest that in disambiguating acts of reference, young children at both 2 and 4 years of age rely most heavily on pragmatic information (e.g. in a pointing gesture), and only secondarily on lexical conventions and principles.

Introduction

Adults refer young children's attention to things in two basic ways: through the use of pointing (and other deictic gestures) and through the use of words (and other linguistic conventions). These two means of reference work in somewhat different ways.

Comprehending pointing would seem to be more basic and 'natural'. It is based on infants' (and other primates') natural tendency to follow the gaze direction of others to external targets, which they do from at least 6 months of age (Moore & D'Entremont, 2001). Infants themselves also typically begin pointing before they begin using language (Carpenter, Nagel & Tomasello, 1998). Pointing and other deictic gestures are natural and pragmatically based in the sense that they simply direct the visual attention of another person to an object or event in the immediate here and now. Words and other linguistic expressions, in contrast, are conventional in that they become effective only through the social learning of a convention in which we all use an 'arbitrary' sound in the same way to direct attention, typically, to a particular kind of referent. Indeed, unlike pointing, common nouns and most verbs are not used to refer to particular objects or events at all – that is, not without some kinds of grounding devices such as

determiners or tense markers – but rather to classes of particular kinds of objects or events.

The psychological primacy of the pointing gesture is also apparent in the fact that adults – at least in Western culture – often teach children words by pointing to referents, but they do not teach them pointing at all, certainly not by using words. In the language learning process, when children hear a new word simply emanating from a loudspeaker outside of any social-pragmatic context, they do not learn it (Baldwin, Markman, Bill, Desjardins & Irwin, 1996). To learn the new word they need some kind of independent social-pragmatic information about what the adult is referring to with the new word – and pointing is an especially effective source of such independent information for doing this (Tomasello, 2001, 2008). In general, a growing body of research suggests that children's word learning rests fundamentally on their social-pragmatic skills, in which an understanding of the pointing gesture plays an important role (Baldwin & Moses, 2001; Hollich, Hirsh-Pasek & Golinkoff, 2000; Saylor, Sabbagh & Baldwin, 2002; Saylor, Baldwin & Sabbagh, 2004; Tomasello, 1992, 2003).

Children quite often experience gestures and language together (Masur, 1997), and indeed when pointing and a word are used together, infants expect them to refer to

Address for correspondence: Susanne Grassmann, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103, Leipzig, Germany; e-mail: grassmann@eva.mpg.de

the same thing (Gliga & Csibra, 2009). An interesting question, then, is what children might think and do when a word and pointing gesture conflict, that is, when they suggest different referents. In a recent study, Jaswal and Hansen (2006) provided conflicting information of this type for 3- to 4-year-old children. The children were simply supposed to give the adult the object he asked for – out of two that were on the table in front of them, with one being a familiar object and one being novel. The problem was that the adult pointed (or looked) to one of two objects on the table in front of them while indicating the other with language. Actually, the linguistic reference in this study was more indirect, as the idea of the study was to pit against one another the pointing gesture, as social-pragmatic information, and mutual exclusivity, as a general lexical principle. Thus, specifically, what Jaswal and Hansen did was to ask children for the 'modi' – which the principle of mutual exclusivity would suggest is the novel object – while pointing to the familiar object. Perhaps surprisingly, children tended to hand over to adults the novel object, as indicated by the language, not the familiar object, as indicated by the pointing gesture. This is surprising because, based on the analysis above, one might expect that pointing would be a more basic and compelling act of reference than would language. It is even more surprising given that the linguistic reference was not direct, but was indirect through the principle of mutual exclusivity – meaning that the child could determine that the novel word referred to the novel object only by exclusion. One explanation for this result might be that Jaswal and Hansen (2006) produced their pointing gesture without the kinds of ostensive accompaniments (e.g. eye contact with child, gaze alternation to referent) that young children typically experience (see below).

In the current study we attempted to replicate the findings of Jaswal and Hansen with some important modifications and extensions. In a first study, we replicated Jaswal and Hansen directly in one condition, but in addition we had another condition in which the pointing gesture was more clearly ostensively marked for the child (mainly by gaze alternation). This study thus pitted the pointing gesture against reference by exclusion (as described by the principle of mutual exclusivity), which we might call indirect linguistic reference. The adult, for example, pointed to a car (with a novel object also on the table) and told the child to 'Give me the modi'. In a second study, we addressed the possibility that competition between pointing and language as acts of reference might work differently if the linguistic reference was effected not by exclusion but rather directly. Specifically, we pitted an ostensively marked pointing gesture against direct linguistic reference. The adult, for example (with a car and a novel object on the table in front of the child), pointed to the novel object and said 'Give me the car'. Finally, in a third study, the child faced two familiar objects (e.g. a car and a cup), and the adult pointed to the car while asking the child to

'Give me the cup'; in this case there were two familiar labels working against the pointing gesture.

If pragmatic information is somehow more fundamental than lexical information for young children in resolving acts of reference, then children in all studies should choose objects based on the pointing gesture rather than linguistic reference. Because the amount of experience children have with the linguistic conventions might play some role in all of this, the studies were run with 2-year-olds as well as 4-year-olds to investigate possible developmental changes in children's weighting of pragmatic and lexical information.

Study 1

Jaswal and Hansen (2006) addressed the issue of the relative strength of social-pragmatic information and lexical principles by pitting mutual exclusivity against a pointing or looking gesture in a disambiguation task. Children were presented with two objects (one novel and one familiar), and then the experimenter (E) requested, 'Give me the blicket', while he pointed to (Study 1) or looked at (Study 2) the familiar object. Thus, according to mutual exclusivity, E intended to refer to the novel object (since the familiar object already had a label) but according to social-pragmatic information E intended to refer to the familiar object because he pointed to it. It was found that 3- to 4-year-old children selected the novel object as referent for 'blicket' more often than the familiar, pointed-to object. However, the problem is that Jaswal and Hansen either pointed to or looked at the familiar object without gaze alternation, that is, without clear ostensive cues that this was a communicative act intended *for the child* (Csibra, 2003). Specifically, in their study pointing was performed such that E 'extended the index finger … toward the familiar object … and placed it on the table about halfway between his body and the object. As his finger touched the table, he tapped the table twice to draw the child's attention to the pointing gesture … throughout the trial; E looked at the child …' (Jaswal & Hansen, 2006, p. 160). In the condition involving looking, the adult simply stared at one of the objects. In neither the pointing nor the looking condition did the adult alternate gaze between child and referent or in any other way express the communicative intention that this is *for you*. In control conditions, Jaswal and Hansen established that in a forced choice situation the children could follow these cues to the intended object. But nevertheless it remains possible that these are relatively weak social-pragmatic cues as compared with their more highly ostensively marked counterparts – pointing and/or looking with gaze alternation between child and object – which children experience most often in the natural world (Bates, 1979; Masur, 1997).

In the current study, therefore, we used a pointing gesture in one condition non-ostensively without gaze alternation (as in Jaswal and Hansen), and in

one condition ostensively with gaze alternation. Our prediction was that a pointing gesture with clear ostensive marking would lead children in this paradigm to pick the familiar, not the novel, object. Because previous research has shown that 2-year-olds rely less on mutual exclusivity than do 4-year-olds (cf. Merriman & Bowman, 1989; Merriman & Schuster, 1991; Graham, Poulin-Dubois & Baker, 1998; Mervis & Bertrand, 1994), we tested two age groups; 4-year-olds (similar to Jaswal & Hansen, 2006) and also 2-year-olds.

Method

Participants

Seventy-two monolingual German-speaking children participated in the study. Half of the children were 2-year-olds (range 1;11,16–2;0,13, M = 1;11,29); the other half were 4 years old (range 3;10,3–4;1,20, M = 3;11,28). Children were randomly assigned to one of the three between-subjects conditions. The mean ages of the children in each condition were: Baseline 1;11,26 and 3;11,20; Ostensive Pointing 2;0,1 and 4;0,4; and Non-ostensive Pointing 2;0,1 and 4;0,0.

Children's parents had previously volunteered to participate in studies of child development. Fourteen additional children participated but were not included in the final sample: Six 4-year-olds (two turned out to be bilingual, two labeled too many novel objects, and two avoided touching the objects) and eight 2-year-olds (four were fussy, one bilingual, and three did not know the labels of at least three of the familiar objects). Children were randomly assigned to one of the experimental conditions, for a total of 12 children in each condition in each age group. The proportion of boys and girls in each condition was 50%.

Materials and design

Each child received four trials in the condition he or she was assigned to. In each trial children were presented with two objects (one novel and one familiar) in an object choice task. E said, 'Let's play with the baffe. Give me the baffe (toma, dofu, modi)'. According to the condition, E just looked at the child (Novel Label Baseline Condition) or, in the conflict conditions with pointing, pointed simultaneously with the verbal request either ostensively or non-ostensively to the familiar object (see Procedure). As novel objects we used an unusual yoyo, a U-shaped door stopper, a bottle cap, and a cone-shaped plastic piece. These objects were selected from a pretested pool of novel objects that are unlikely to be labeled by 2- to 4-year-old children. If children had labeled the novel object during the trial, this trial would have been excluded (but it never happened). The familiar objects were: a toy car, a shoe, a comb, and a pen. Parents of the 2-year-olds were asked to fill in a short questionnaire and report whether their children *produced* a label for each of the familiar objects. If

a child did not know the label for one of the familiar objects, then the corresponding trial was later excluded from analyses (six cases). If more than one trial had to be excluded for a child, he or she was dropped from the study and replaced by another child (three children – see above). We counterbalanced toy pairings, their occurrence with each of the novel words, the order of presentation of the toy pairs, and the left–right position of the novel and familiar objects within and across children.

Procedure

Children were tested individually in a quiet room in their daycare center (4-year-olds) or came for an approximately 30 minute visit to our lab (2-year-olds). During the study the child sat at a table in a child chair or on the parent's lap, with E directly across the table. All sessions were videotaped.

Novel Label Baseline. At the beginning of each trial, E held one novel and one familiar object approximately at children's eye level and said 'Look, here'. After the child had looked at each object, E put the objects on the table – centered in front of the child approximately shoulder-width apart and 30 cm away from the child. Then E said a German equivalent of 'Let's play with the baffe. Give me the baffe' or 'Let's play with the baffe. Take the baffe'. E looked straight at the child during the whole trial until he or she made a choice. When the child had selected an object, E responded in a neutral manner independently of the choice the child made. E and the child played for approximately 1 minute with the selected toy (e.g. sliding down a chute). Three trials followed (for a total of four) in exactly the same manner, with a counterbalanced left–right position of the novel toy.

Non-Ostensive Pointing. The procedure of this condition was identical to the Novel Label Baseline except that E pointed non-ostensively to the familiar toy while using the novel label (as in Jaswal and Hansen): During the verbal request ('Let's play with the baffe') E's hand was positioned on the table halfway between E and the objects with an extended index finger towards the familiar toy. E tapped with her index finger on the table twice to draw attention to the pointing. The hand remained on the table until the child made a choice and E looked straight to the child during the whole trial.

Ostensive Pointing. The procedure of this condition was identical to the Non-Ostensive Pointing Condition except that the pointing gesture was performed naturally in an ostensive manner with gaze alternation: While E said, 'Give me the baffe', she pointed to the familiar object with the whole forearm (and index finger) and repeatedly alternated her gaze between the child and the familiar object until the child made a choice.

Coding and reliability

Children were scored as having chosen an object if they picked it up and handed it to E, held it up, or pushed it

toward E. When children selected both toys, first touch was scored. When children selected both objects simultaneously, their response was coded as 'no choice'. We did not exclude trials in which children made no choice since we judged them as informative in the conflict conditions in which pointing and labeling indicated different objects. Two additional independent coders scored a randomly selected sample of nine 2-year-olds (coder A) and eight 4-year-olds (coder B) from videotape. As estimated by Cohen's Kappa, inter-observer reliability was 0.88 for the 2-year-olds and 1.0 for the 4-year-olds.

Results

Figures 1 and 2 show the mean proportions of novel and familiar object selection in each of the three conditions. For purposes of statistical analysis, the dependent variable was how often the child selected the novel object (consistent with mutual exclusivity) in response to E's request ('Give me the modi').

Comparisons against chance (both ages collapsed, using one-sample *t*-tests) revealed that in the Novel Label Baseline children selected the novel object as the referent of the novel word significantly above chance ($M = 78.8\%$, $SD = 17.7$, $t(23) = 7.98$, $p < .001$, Cohen's $d = 1.627$). Thus, we replicated the disambiguation effect in the traditional mutual exclusivity paradigm. If, however, a non-ostensive pointing gesture contradicted mutual exclusivity, we found that children chose objects at chance level and picked the novel object only 48.7% of the time ($SD = 37.2$). Most importantly, when the contradictory pointing was enriched with gaze alternation and performed ostensively, children hardly ever chose the novel object ($M = 2.1\%$, $SD = 0.07$); instead, children selected the pointed-to familiar object as the speaker's intended referent 97.9% of the time ($SD = 0.07$, $t(23) = 33.26$, $p < .001$, Cohen's $d = 6.746$). Thus, children overrode the assumption that novel words
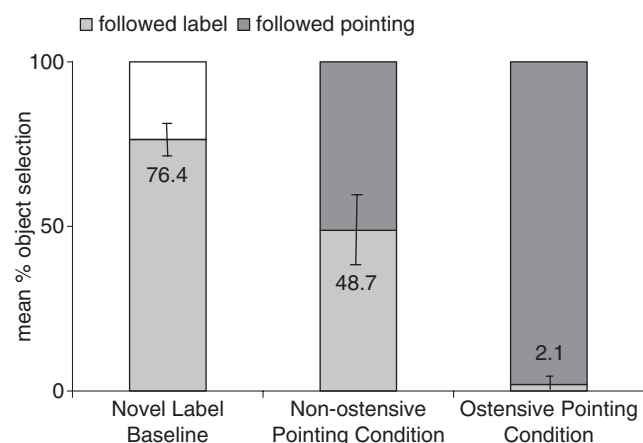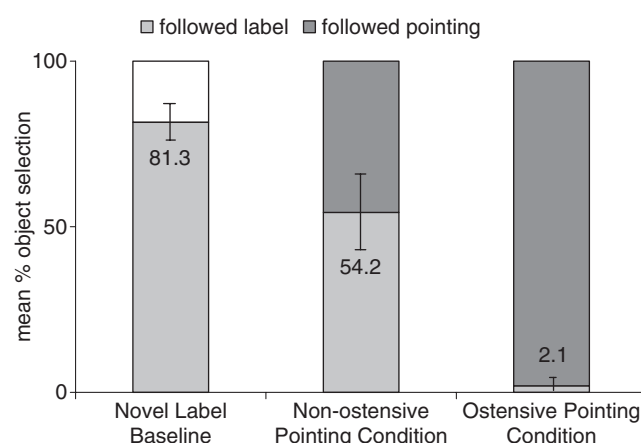

Figure 2 *Mean proportions of 4-year-olds' object selection in each condition in Study 1 (n = 12 per condition).*

refer to novel objects in the face of an ostensive pointing gesture indicating otherwise. There were no major age differences in these patterns, except that examination of individual responses revealed for 4-year-olds in the Non-Ostensive Pointing Condition a bimodal distribution (five had a clear preference for pointing and five for the novel label); no such distribution was found in the younger children, who tended to show mixed responses across trials.

A 2 (age) × 3 (condition) ANOVA indicated that children selected the novel object differently across conditions ($F(2, 66) = 60.18$, $p < .001$, $\eta^2 = 0.646$). Bonferroni post-hoc pair-wise comparisons revealed that children chose the novel object more in the Novel Label Baseline condition (with no pointing) than in the Non-Ostensive Pointing condition ($p = .001$) and more in the Novel Label Baseline than in the Ostensive Pointing Condition ($p < .001$). Furthermore, children chose the novel object more in the Non-Ostensive Pointing Condition than in the Ostensive Pointing Condition ($p < .001$). The two age groups did not perform differently.

Analyses of children's response in the first trial revealed exactly the same results. Table 1 shows children's object selection in the first trial. A Fischer's

Figure 1 *Mean proportions of 2-year-olds' object selection in each condition in Study 1 (n = 12 per condition).*

Table 1 *Children's object selection in the first trial in each condition in Study 1*

	Object selection	
	Familiar object	Novel object
	2-year-olds	
Novel Label Baseline	3	9
Non-ostensive Pointing Condition	6	6
Ostensive Pointing Condition	11	1
	4-year-olds	
Novel Label Baseline	2	10
Non-ostensive Pointing Condition	6	6
Ostensive Pointing Condition	12	0

Exact chi-square test (collapsed over age) revealed that children's object selection differed among conditions ($\chi^2(2, N = 72) = 30.77$, $p < .001$). Pair-wise comparisons revealed that children selected the novel object more often in the Novel Label Baseline than in the Non-Ostensive Pointing Condition ($\chi^2(1, n = 48) = 4.46$, $p = .034$ (one-tailed)) and more often in the Novel Label Baseline than in the Ostensive Pointing Condition ($\chi^2(1, n = 48) = 5.63$, $p = .024$). Furthermore, children chose the novel object more in the Non-Ostensive Pointing Condition than in the Ostensive Pointing Condition ($\chi^2 (1, n = 48) = 12.77$, $p < .001$.

Discussion

In this study we examined whether and how strongly 2- and 4-year-olds rely on mutual exclusivity when it conflicts with different types of pointing. In a baseline condition (with no pointing) we found that both 2- and 4-year-olds used mutual exclusivity to identify the novel object as the referent of a novel word when they were presented with a novel and a familiar object and E asked them to select the 'modi' – the classic exclusion effect. In two conflict conditions, mutual exclusivity was pitted against two different pointing gestures. When mutual exclusivity was in conflict with a non-ostensive pointing gesture, children chose objects at chance level. But when the pointing was ostensive (more salient articulation and with gaze alternation), all of a sudden children followed this cue – and forgot about mutual exclusivity – almost 100% of the time. The current study has thus identified a common, everyday social-pragmatic cue that overrides mutual exclusivity almost every time.

Interestingly, we found no difference between 2- and 4-year-olds. In all conditions children at both ages performed nearly identically. This is somewhat surprising because other studies have typically found a stronger exclusion effect in older or linguistically more advanced children (Graham *et al*., 1998; Merriman & Schuster, 1991). The only explanation we have for this discrepancy is that the specific materials we used in our study helped younger children to exclude familiar objects as likely referents of novel words. We selected four familiar objects which we assumed to be in the *productive* vocabulary even of 2-year-olds (a car, a pen, a shoe, and a comb) – all parents of the 2-year-olds filled in a questionnaire to confirm this for each child individually. Previous studies on young children's use of mutual exclusivity have only controlled for children's *comprehension* of the familiar objects' labels (e.g. Graham *et al*., 1998). Interestingly, we also found no age effect in the two conflict conditions. This suggests either that the power of the lexical principles is not increasing with age or the importance of social-pragmatic information is not decreasing with age. This is interesting because it suggests that pointing is a strong cue not only for children who are still novice word learners, but also for relative experts.

We are not sure how to explain why we found only equal selection of the familiar and novel object in the Non-Ostensive Pointing condition, whereas Jaswal and Hansen (2006) found a preference for the novel object based on mutual exclusivity. But it is worth noting that Yow and Markman (2007) also did not replicate this finding in a control group. Also a bit mysterious is why, in Jaswal and Hansen's (2006) control conditions, children used non-ostensive pointing and looking at above-chance levels. One suggestion is that the verbal information in their control conditions provided cues to the relevance of the accompanying pointing/looking. The phrasing 'give it to me' in the pointing control condition, for example, indicated that there is some non-verbal cue to the reference of the otherwise unlicensed pronoun (and a similar argument can be made about the looking control condition). In terms of non-ostensive pointing, Enfied, Kita and de Ruiter (2007) differentiate what they call small points and big points (which may bear some similarity to non-ostensive and ostensive pointing as defined here). Big points carry primary information indicating the great importance of the indicated referent while small points indicate to the speaker 'you probably don't need this extra information, because you probably know what I mean' (p. 1735). Thus, when a speaker produces a small point he indicates that the verbal information is primary. With this background, Jaswal and Hansen's (2006) finding that 3- to 4-year-old children disregarded a non-ostensive pointing gesture when it was accompanied by a novel word might actually indicate sophisticated knowledge about when to rely on speech.

Our finding that children at both ages overrode the lexical principle does not contradict the Constraints Approach to word learning – which accords a primary role to lexical principles in word learning (e.g. Markman, 1989, 1990). The Constraints Approach recognizes some role for pragmatic cues in children's word learning, especially when mutual exclusivity must be overridden to learn second labels for already familiar objects (e.g. Markman & Wachtel, 1988; Liittschwager & Markman, 1994; Mervis, Golinkoff & Bertrand, 1994). However, we favor the Social-Pragmatic Account of word learning since our results are not only consistent with this approach but even can be explained by it. In the Social-Pragmatic Account the key is always the child's understanding of the intentional, rational actions of others, including their communicative intentions (cf. Tomasello, 2000; Akhtar & Tomasello, 2000). Pointing and other deictic gestures are both phylogenetically and ontogenetically primary because they can be interpreted without any social learning of conventions. Indeed, people can communicate very effectively in a foreign culture whose language they do not speak by using pointing and other gestures. Moreover, an argument can be made that arbitrary conventions such as words can only be

learned if the child has some non-arbitrary way of knowing what the person is using the novel word to refer their attention to in the first place (Tomasello, 1992, 2001, 2008). That is to say, social-pragmatic information plays a foundational role in word learning (e.g. Baldwin *et al.*, 1996; Baldwin & Moses, 2001; Saylor *et al.*, 2004). Finally, in this view, lexical principles such as mutual exclusivity are not *a priori* conditions for language acquisition, but rather are derived generalizations from particular communicative actions on particular occasions (Saylor *et al.*, 2004; Diesendruck & Markson, 2001; Mervis & Bertrand, 1994; Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992). The reasoning is thus that the more fundamental social-pragmatic information (pointing) should trump the derived general principle (mutual exclusivity).

Despite these general theoretical considerations, there are currently very few concrete predictions – from either theoretical perspective – about precisely when social-pragmatic cues should override mutual exclusivity. The one exception is Saylor *et al.* (2002, 2004), who suggest that children apply mutual exclusivity only when there is some pragmatic evidence to do so (in their study gestural or verbal whole–part juxtaposition). However, children's object selection in the Non-Ostensive Pointing Condition and in Jaswal and Hansen's (2006) study contradict this suggestion: children at least sometimes applied mutual exclusivity with no pragmatic evidence for it (indeed there was some evidence against it). The important difference between Saylor *et al.* (2002, 2004) and the current study (and Jaswal & Hansen, 2006) is that in our study children did not have to override the Whole Object Assumption to apply mutual exclusivity, whereas in their study they did have to. So perhaps children do not need pragmatic information to apply mutual exclusivity but only to override it (as in our study) or to override the Whole Object Assumption (as in Saylor *et al.*'s (2002, 2004) studies).

In our Study 1, the pointing gesture singled out its intended referent directly, whereas children had to infer the referent of the novel word via exclusion. This difference in directness of reference may thus help to explain our finding that children find ostensive pointing a more powerful indicator of intended referents than language. Therefore, in a second study we sought to provide a 'fairer' comparison of pointing and lexical information by making the lexical cue totally direct; that is, rather than pointing competing against a novel word (relying on exclusion learning), we had it compete with a familiar word. Specifically, in a set-up similar to Study 1, we pitted the use of a familiar word against the pointing gesture by pointing to the novel object while at the same time asking for the familiar object by name (e.g. 'car'). This tests directly which cue children follow in disambiguating an ambiguous request when both indicate their intended referents directly.

Study 2

In Study 2 the question was whether children would also rely on pointing when it conflicted with a direct lexical reference using a familiar word. Study 2 is thus the reverse of the Ostensive Pointing Condition from Study 1 (see Figures 3a and 3b). Whereas in Study 1 a novel word was used while pointing to a familiar object, in Study 2 a familiar word was used while pointing to a novel object. That is, with a familiar (e.g. a car) and a novel object on the table, we asked children to give us 'the car' while pointing to the novel object. We assumed that familiar words provide more direct and reliable lexical information than novel words since familiar words identify their referent directly while novel words (in an exclusion paradigm such as that of Study 1) do so only indirectly. Assuming that the directness and reliability of pointing and familiar words are very similar as referential acts, the current study represents a more or
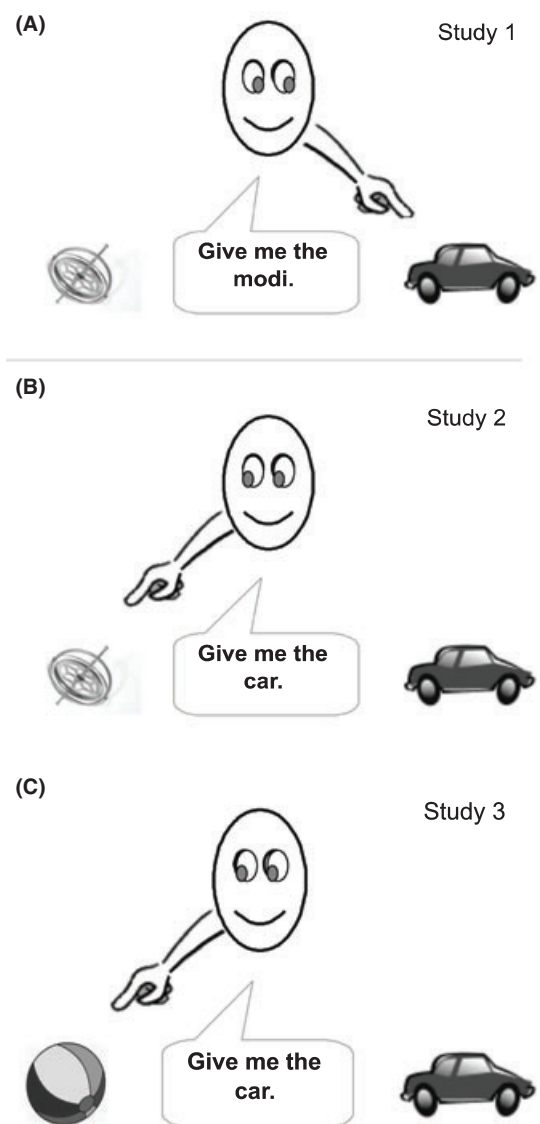

Figure 3 *Basic cues given: (A) Study 1 (Ostensive Pointing Condition), (B) Study 2, and (C) Study 3.*

less direct test of the relative strength of pointing and lexical reference (not principles) in children's reference resolution. Our expectation was that pointing, as the more fundamental act of reference, would win out over direct lexical reference in this study as well.

Method

Participants

Twenty-four monolingual German-speaking children participated in the study. Twelve children were 2 years old (range 1;11,6–2;0,23, *M* = 2;0,2) and 12 children were 4 years old (range 3;10,12–4;2,0, *M* = 4;0,13). One additional 2-year-old participated but was not included in the final sample because she was fussy. The proportion of boys and girls was 50% in each age group.

Materials and design

The same materials were used as in Study 1. The design and counterbalancing was identical to Study 1. As the lexical cue, E used the label of the familiar object (car, comb, pen, and shoe) that was presented to the child in a given trial. Pointing always indicated the novel object. Parents of the 2-year-olds were asked to judge whether their children *produced* a label for each of the familiar objects. If a child did not know the label for one of the familiar objects the corresponding trial was later excluded from the analyses (three cases).

Procedure

The procedure in Study 2 is the reverse of the Ostensive Pointing Condition in Study 1. With one familiar and one novel toy placed on the table, E said, 'Let's play with the car. Take the car', and simultaneously pointed ostensively to the novel object (with gaze alternation between the chil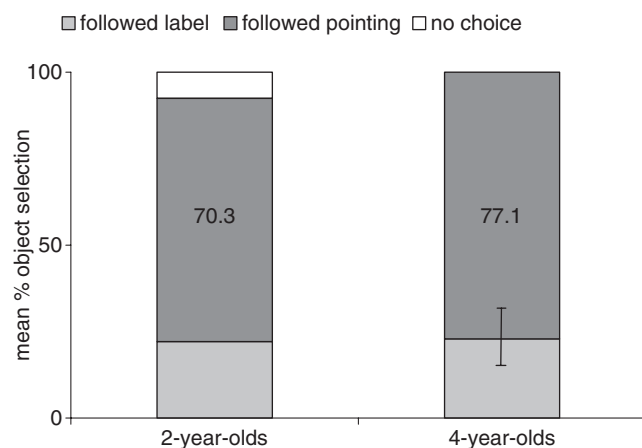d and the novel object) until the child made a choice. As in Study 1, E and the child played some time with the selected object. The procedure was repeated for a total of four trials.

Coding and reliability

Coding was identical to Study 1. Two independent coders scored a randomly selected sample of two 2-year-olds (coder A) and three 4-year-olds (coder B) from videotape. As estimated by Cohen's Kappa, interobserver reliability was 1.0 for the 2-year-olds as well as for the 4-year-olds.

Results

Figure 4 shows the mean proportions of familiar and novel object selection. Familiar object selection indicates that children relied on the spoken familiar word and novel object selection indicates that children relied on pointing. Two-year-olds (but not 4-year-olds) sometimes did not make a choice and selected both objects (*M* = 7.6%, *SD* = 13.9). For purposes of statistical analysis, the dependent variable was how often the child selected the pointed-to object in response to E's request.

A one-sample *t*-test collapsed over age revealed that children chose the pointed-to object more than expected by chance ($t(23)$ = 4.56, $p < .001$, Cohen's d = 1.122). An independent samples *t*-test revealed no difference among the age groups in their object selection: 2-year-olds chose the pointed-to object 70.3% and the labeled familiar object 22.2% of the time (*SD* = 21.4). Four-year-olds chose the pointed-to object 77.1% and the labeled familiar object 22.9% of the time (*SD* = 29.1).

In a further independent samples *t*-test we compared children's object selection in Study 2 with their performance in its reverse condition in Study 1 (Ostensive Pointing Condition). This analysis revealed that although even in Study 2 children preferably relied on pointing over the word, children relied less on pointing in Study 2 (*M* = 73.7%, *SD* = 25.4) than in Study 1 (*M* = 97.9%, *SD* = 7.1, $t(46)$ = 4.50, $p < .001$, Cohen's d = 1.327). To put it another way, children relied more on the familiar word in Study 2 (*M* = 22.5%, *SD* = 24.99) than on the novel word in the Ostensive Pointing Condition in Study 1 (*M* = 2.1%, *SD* = 7.1).

Figure 4   *Mean proportions of object selection in Study 2 (n = 12 per age group).*

Table 2 *Children's object selection in the first trial in Study 2 and Study 3*

	Object selection	
	Pointed-to object	Labeled object
	2-year-olds	
Study 2	7	5
Study 3	10	2
	4-year-olds	
Study 2	9	3
Study 3	6	6

Analyses of children's response in the first trial revealed similar results. Table 2 shows children's object selection in the first trial in Study 2. A binominal test (collapsed over age) revealed that children chose the pointed-to object more often ($n = 16$) than the labeled object ($n = 8$) ($p = .076$). A Fischer's Exact chi-square test (collapsed over age) revealed that children's object selection differed among studies – indicating that children in Study 2 relied less on pointing (or more on the label) than children in Study 1 ($\chi^2(1, n = 48) = 6.70, p = .023$).

Discussion

In this study we examined whether and how strongly 2- and 4-year-olds relied on a speaker's pointing gesture when it conflicted with a known word. We found that 2-year-olds as well as 4-year-olds preferred pointing over familiar words to determine the speaker's intended referent. This thus shows that pointing is a very powerful indicator of referential intent, even overriding children's established word knowledge.

The current findings mean that children's strong reliance on pointing in Study 1 was not due just to the indirectness and ambiguity of the lexical reference in the exclusion learning paradigm – since in the current study pointing and familiar words are equally direct acts of reference. In our view, the explanation for this finding is again that pointing is, in some sense, the primordial act of reference. Even pre-linguistic children rely on pointing to determine someone's intended referent (Tomasello, Carpenter & Lizkowski, 2007; Goldin-Meadow, 2007). Nevertheless, the comparison of children's object selection in Studies 1 and 2 revealed that they chose the pointed-to object less when it conflicted with a familiar word (Study 2) than when it conflicted with a novel word (Study 1). This indicates that children do not just follow pointing blindly in the sense that when it is available they follow it no matter what. Rather, children had both sources of information available (pragmatic/gestural and conventional/lexical) in both studies, and they used and integrated them both into one coherent interpretation, albeit with different weightings on different occasions.

Usually, speech and gesture are not in conflict as in our studies, but agree in their reference. The assumption of co-reference of gesture and word is thus crucial if children are to integrate information to learn new words and extend the meanings of known words (cf. Hollich *et al.*, 2000; Golinkoff *et al.*, 1992). Indeed, Gliga and Csibra (2009) found that children as young as 13 months assume co-reference of pointing and words. In our first study, even though the adult used the novel word seemingly for an unexpected referent (the familiar object as indicated by her pointing gesture) children resolved this potential ambiguity by, apparently, assuming that the adult intended the novel word to indicate in some new way some aspect of the familiar object (e.g. a superordinate category of some kind). They thus simply had to build an

initial lexical entry for the novel word and its referent – while the meaning of the familiar word (as the known label of the pointed-to object) was not affected. In Study 2, in contrast, to build a coherent integration of the speaker's use of a familiar word for a novel object, children had to modify in an apparently difficult way their already established lexical entry for a highly familiar word whose referential extension they thought they knew well. This difference between creating a first lexical entry for a novel word and modifying in an unusual way the lexical entry for a well-established word may thus account for the different effects of pointing and lexical information in the two studies.

Overall, then, we think that children's strong reliance on pointing in both studies, together with their slightly reduced reliance on pointing in Study 2, indicates that children tried to integrate pointing and words into one coherent act of reference, with pointing being primary. But still in this study, the novel object gave them an opening – they 'only' had to assume that this novel object was an unusual exemplar of a familiar word. In Study 3 we attempted to bias things totally in favor of lexical information by using two familiar objects, so that integrating pointing and word contradicted established lexical knowledge in two ways.

Study 3

In Study 3 we made the lexical information which contradicted pointing even stronger. Specifically, with two familiar objects (e.g. a ball and a car) on the table we asked children to give us the 'car' while pointing to the ball (see Figure 3c). In this case, compared to Study 2, children additionally knew the label of the pointed-to object – which might lead them to rely less on the pointing gesture than in Study 2 since a single coherent interpretation of this situation requires a re-interpretation of *both* familiar object labels. We were not sure *a priori* what children would do in this situation.

Method

Participants

Twenty-four monolingual German-speaking children participated in the study. Twelve children were 2 years old (range 1;11,5–2;0,28, $M = 2;0,0$) and 12 children were 4 years old (range 3;10,4 – 4;1,25, $M = 4;0,11$). Three additional 2-year-olds participated but were not included in the final sample: two were fussy and two did not know enough familiar words. The proportion of boys and girls was 50% in each age group.

Materials and design

The same familiar objects as in Study 1 and Study 2 were used. A second set of familiar objects (toy tree, mini-

bucket, toy handbag, and child-necklace) were used instead of the novel objects in Studies 1 and 2. Parents of the 2-year-olds were asked to judge whether their children *produced* a label for each of the familiar objects. Two children were dropped from the study and replaced by another child because they did not know enough of the familiar objects' labels. E verbally requested the same objects as in Study 2. Design and counterbalancing were identical to Study 2 except that two familiar objects were on the table and E labeled one but pointed to the other.

Procedure

The procedure was identical to the procedure in Study 2 with the exception that two familiar objects were placed on the table and E pointed to one of them but verbally requested the other one (the same familiar labels were used as in Study 2).

Coding and reliability

Coding was identical to Study 2. Two independent coders scored a randomly selected three 2-year-olds (coder A) and two 4-year-olds (coder B) from videotape. As estimated by Cohen's Kappa, interobserver reliability was 0.86 for the 2-year-olds and 1.0 for the 4-year-olds.

Results

Figure 5 shows the mean proportions of children's object selection. Children could choose either the labeled or the pointed-to object. Two-year-olds (but not 4-year-olds) sometimes did not make a choice and selected both objects ($M = 2.1\%$, $SD = 7.2$). A one-sample *t*-test collapsed over age revealed that children chose objects at chance level. An independent samples *t*-test revealed no difference among the age groups in their object
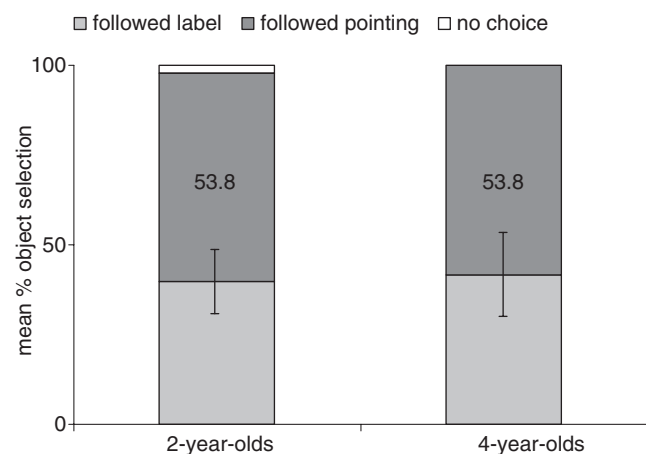

Figure 5 *Mean proportions of object selection in Study 3 (n = 12 per age group).*

selection. Two-year-olds chose the labeled familiar object 39.6% of the time and the pointed-to familiar object 58.3% of the time ($SD = 31.0$). Four-year-olds chose the labeled familiar object 41.7% and the pointed-to familiar object 58.3% of the time ($SD = 40.6$). However, examination of individual performance revealed that 4-year-old children's chance performance is due to the fact that half of them had a preference for pointing and the other half of the children had a preference for the label (in addition, six children followed their non-preferred cue in one trial). No such bimodal distribution was found in the 2-year-olds.

In an independent samples *t*-test we compared children's object selection in Study 3 with children's performance in Study 2. The studies differed with respect to the familiarity of the pointed-to object: while in Study 2 the pointing gesture referred to a novel object, in Study 3 it referred to a false familiar object. The dependent variable was how often children selected the pointed-to object. This analysis revealed that children relied equally on pointing in both studies. They chose the pointed-to novel object in Study 2 73.7% of the time ($SD = 25.4$) and they chose the pointed-to familiar object in Study 3 58.3% of the time ($SD = 35.1$) ($p < .10$).

We also compared children's object selection in Study 3 with their object selection in the Ostensive Pointing Condition in Study 1. In both studies the adult pointed to a familiar object while she said a novel (Study 1) or a false familiar word (Study 3). The most likely referent of the spoken word was present in both studies: a novel object in Study 1 and the labeled familiar object in Study 3 (see Figure 3). An independent samples *t*-test on how much children relied on pointing revealed that children did so more in the Ostensive Pointing Condition in Study 1 than in Study 3 ($t(46) = 5.42$, $p < .001$, Cohen's $d = 1.597$).

Analyses of children's response in the first trial revealed similar results. Table 2 shows children's object selection in the first trial in Study 3. A binominal test (collapsed over age) revealed that children chose the pointed-to object marginally more often ($n = 16$) than the labeled object ($n = 8$) ($p = .076$). This preference for pointing is due to 2-year-olds' object selection; 4-year-olds displayed no such preference. Fischer's Exact chi-square test (collapsed over age) revealed that children's object selection in the first trials did not differ among Study 2 and Study 3 but among Study 1 and Study 3, such that they relied more on pointing in Study 1 ($\chi^2(1, n = 48) = 6.701$, $p = .023$).

Discussion

In Study 3 we examined children's disambiguation of a two-sidedly false referential expression, when a speaker pointed to a familiar object but said a different familiar word (with its referent also in front of the child). We found that children followed pointing and lexical information equally often in this situation. We see this

again as evidence that children attempted to integrate the contradictory cues to the speaker's intended reference. Integrating pointing and lexical information into a single coherent interpretation was especially difficult in this case, as children's inherent reliance on pointing (see Studies 1 and 2) was challenged by the fact that the lexical information provided by the speaker and their lexical knowledge suggested that the pointing gesture was wrong in two ways.

In this study children chose the pointed-to object less than in the one-sidedly false situation in Study 1 in which a novel word was used. That is, in Study 1 children readily accepted a second label for a familiar object (calling a car 'modi'), but in Study 3 they did not (calling a car 'ball'). This is presumably because in Study 3 both of the objects involved (car and ball) already had highly familiar labels whose referential extensions the child knew well – whereas in Study 1 the word 'modi' had no previous associations. But children did not choose the pointed-to object more often in Study 3 than in Study 2 in which the adult asked for 'the ball' but pointed to a novel object (with a ball also on the table). This means that when children extend a familiar label, it is just as easy to extend it to a familiar as an unfamiliar object. The asymmetry is thus that it seems to be harder for children to modify the lexical entry for a highly familiar word in opaque ways (i.e. they are not sure of the basis for the extension) than to understand reference to a familiar object with a novel word.

It should be noted that in a previous study on children's integration of familiar words and pointing, Thompson and Massaro (1994) found that when verbal and gestural cues contradicted one another, 5- and 9-year-olds' object selection was mainly driven by the familiar word, and the pointing gesture was usually disregarded. This may indicate that the balance of language and gesture changes in strength over development. However, Thompson and Massaro used only the initial syllables of the familiar words as verbal information in this study, and so further research using comparable methods is needed to establish possible developmental changes after the ages studied here.

General discussion

In three studies we examined how children weigh a pointing gesture versus lexical information in determining a speaker's intended referent. Typically, adult pointing and lexical information work redundantly for children to establish reference, and indeed children as young as 13 months expect this (Gliga & Csibra, 2009). Our finding was that when these two sources of information conflict, children most often trust the pointing gesture over the lexical information – unless the lexical information is multiply contradicted (in which case they still trust them equally). This was the same for both 2-year-olds and 4-year-olds.

The pattern of children's object selection across studies suggests that they tried to interpret the speaker's contradictory acts of reference coherently. In Study 1 children integrated both cues easily and accepted a novel word as a second label (of some kind) for a familiar object. In Study 2 they accepted and may have extended a familiar word to cover reference to a novel object. Children's somewhat reduced reliance on pointing in Study 2 might indicate that the integration of pointing and words is harder when pointing conflicts with a familiar word (that has to be reinterpreted) than when pointing conflicts with a novel word (which only gets its first bit of meaning). In Study 3 where a familiar word conflicted with pointing to another familiar object children followed the gestural and lexical cues equally. This suggests that it is hardest to make sense of such a two-sidedly false reference – perhaps because it makes the reinterpretation of two familiar words necessary: first, the spoken familiar word has to be extended to cover a new (itself familiar) object, and second, this object's label has to be reinterpreted as being, for example, a subordinate term of the spoken word. Thus, to accept 'car' for a ball children might extend their car-category to entail balls but then 'car' is a superordinate term and balls form a subcategory of cars. Thus, a complex relationship between two familiar words results from the attempt to integrate the use of a familiar word for a different familiar object.

The main point is thus that children relied on ostensive pointing to different degrees in the different studies, which shows that they did not just always go automatically with the pointing gesture but also considered the lexical information – differently depending on their confidence in it – to try to come up with a single coherent interpretation of the adult's act of reference. The asymmetry we found by comparing across studies is that it seems to be harder for children to modify their lexical entry for a highly familiar word in opaque ways (i.e. they are not sure of the basis for the extension) than to understand reference to a familiar object with a novel word.

An alternative explanation for all of these findings might be that children simply thought that the adult had misspoken and meant to say another word. We do not think that this is what happened, however, since one would then expect that (at least the older children) would frequently try to correct the adult or make comments such as 'this one?' or 'that's not a car!' – but children at both ages rarely produced such comments. And furthermore, even if this is what children thought, it still raises the question of why children should preferentially assume that the adult misspoke rather than mis-pointed – which would again argue for the primacy and naturalness of the pointing gesture.

Our results are also relevant for the question why children in some experiments avoid looking at wrongly labeled objects and instead search for an appropriate referent of the spoken word (Koenig & Echols, 2003;

Swingley & Fernald, 2002). In Koenig and Echols (2003), for example, children saw a picture of a familiar object on a screen (e.g. a car) and E mislabeled it with a false familiar word (e.g. 'That's a ball'). Thus, the mislabeling was two-sidedly false, as in our Study 3. Koenig and Echols noted that it remains open which component of the two-sidedly false mislabeling caused children's avoidance of the wrongly labeled object – the expectation of a conventional referent of the familiar word or the expectation of a conventional label for the familiar object. Although in our studies we examined object choice rather than children's looking, our results help to answer this question. In our studies children's expectations about how familiar words should be used referentially were more important than their expectations about how familiar objects should be labeled. Children accepted a novel word for a familiar object nearly 100% of the time (presumably because there might be a corresponding novel referent, for example, a superordinate category or the like), but they clearly did not like using a familiar label for a 'wrong' object, especially a familiar one (presumably because it was not easy to see any basis for the surprising extension). Our findings thus suggest that children's avoidance of two-sidedly false reference in Koenig and Echol's (2003) studies is caused by the word (which is wrongly used) more than by the object (for which an unexpected label was used).

The current findings thus provide new evidence for the special importance of social-pragmatic information not only in word learning but in language use in general. Specifically, these findings suggest that in reference resolution, social-pragmatic experience is primary for young children, and lexical information and general lexical principles are derivative. This is not to deny the importance of conventional linguistic labels or lexical principles. It is only to argue that conventional linguistic communication is grounded in naturally meaningful social interactions and gestural communication (Tomasello, 2008), and that in fact such social-pragmatic grounding makes language acquisition possible in the first place.

References

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R.M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford: Oxford University Press.

Baldwin, D.A., Markman, E.M., Bill, B., Desjardins, N., Irwin, J.M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word–object relations. *Child Development*, **67** (6), 3135–3153.

Baldwin, D.A., & Moses, L.J. (2001). Links between social understanding and early word learning: challenges to current accounts. *Social Development*, **10** (3), 309–329.

Bates, E. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York: Academic Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, **63** (4, Serial No. 176).

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society, London B*, **358**, 447–458.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Developmental Psychology*, **37** (5), 630–641.

Enfield, N.J., Kita, S., & de Ruiter, J.P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, **39** (10), 1722–1741.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Pyschological Science*, **20**, 347–353.

Goldin-Meadow, S. (2007). Pointing sets the stage for learning language – and creating language. *Child Development*, **78** (3), 741–745.

Golinkoff, R.M., Hirsh-Pasek, K., Bailey, L.M., & Wenger, N.R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, **28**, 99–108.

Graham, S.A., Poulin-Dubois, D., & Baker, R.K. (1998). Infants' disambiguation of novel object words. *First Language*, **18**, 149–164.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: an emergentist coalition model of word learning. *Monographs of the Society for Research in Child Development*, **65** (3, Serial No. 262).

Jaswal, V.K., & Hansen, M.B. (2006). Learning words: children disregard some pragmatic information that conflicts with mutual exclusivity. *Developmental Science*, **9**, 158–165.

Koenig, M.A., & Echols, C.H. (2003). Infants' understanding of false labeling events: the referential roles of words and the speakers who use them. *Cognition*, **87** (3), 179–208.

Liittschwager, J., & Markman, E. (1994). Sixteen- and 24-month-olds' use of mutual exclusivity as a default assumption in second label learning. *Developmental Psychology*, **30**, 955–968.

Markman, E. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science*, **14**, 57–77.

Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, **20** (2), 121–157.

Masur, E.F. (1997). Maternal labelling of novel and familiar objects: implications for children's development of lexical constraints. *Journal of Child Language*, **24** (2), 427–439.

Merriman, W.E., & Bowman, L.L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, **54**, 1–55.

Merriman, W.E., & Schuster, J.M. (1991). Young children's disambiguation of object name reference. *Child Development*, **62**, 1288–1301.

Mervis, C.B., & Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Development*, **65** (6), 1646–1662.

Mervis, C.B., Golinkoff, R.M., & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic-level category. *Child Development*, **65** (4), 1163–1177.

Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, **2** (2), 109–129.

Saylor, M., Baldwin, D., & Sabbagh, M. (2004) Converging on word meaning. In D.G. Hall & S.R. Waxman (Eds.), *Weaving a lexicon* (pp. 509–531). Cambridge, MA: MIT Press.

Saylor, M.M., Sabbagh, M.A., & Baldwin, D.A. (2002). Children use whole–part juxtaposition as a pragmatic cue to word meaning. *Developmental Psychology*, **38** (6), 993–1003.

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, **46** (1), 39–56.

Thompson, L.A., & Massaro, D.W. (1994). Children's integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology*, **57** (3), 327–354.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, **10**, 401–414.

Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 132–158). Cambridge: Cambridge University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, **78** (3), 705–722.

Yow, Q.W., & Markman, E.M. (2007, April). Monolingual and bilingual children's use of the mutual exclusivity assumption and pragmatic cues in word learning. Poster presented at 2007 SRCD Biennial Meeting, Boston, MA.