Monolingual and Bilingual Children's Resolution of Referential Conflicts: Effects of Bilingualism and Relative Language Proficiency

Josje Verhagen[1], Susanne Grassmann[2], Aylin C. Küntay[1,3]

[1]Utrecht University, Department of Special Education, Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands

[2] Institute for Educational Research and Development, School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Brugg-Windisch, Switzerland

[3]Koç University, Department of Psychology, Rumelifeneri Yolu, Sarıyer 34450, Istanbul, Turkey

Email addresses: J.Verhagen@uu.nl, susanne.grassmann@fhnw.ch, akuntay@ku.edu.tr

Corresponding author:

Dr. Josje Verhagen,

Department of Special Education: Cognitive and Motor Disabilities, Utrecht University,

Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

Phone number: +31-(0)30-2531511, Email address: J.Verhagen@uu.nl

Abstract

Monolingual children follow pointing over labeling when these are in conflict in object selection tasks. Specifically, when a speaker labels one object, but points at another object, monolinguals select the object pointed at. Here, we ask whether (i) bilingual children show the same behavior as monolinguals and (ii) relative language proficiency affects bilinguals' conflict resolution. 35 monolingual and 32 bilingual two- to four-year-olds performed an experiment involving a conflict between pointing and labeling. The bilinguals were tested in Dutch and in English. The bilinguals had a stronger preference for pointing over labeling and selected both objects less often than the monolinguals. Point following was stronger in the bilinguals' weaker language than in their stronger language. These results support earlier findings on bilinguals' increased sensitivity to socio-pragmatic cues and weaker reliance on mutual exclusivity, and show that previously acquired language knowledge affects how children weigh socio-pragmatic and lexical cues.

Monolingual and Bilingual Children's Resolution of Referential Conflicts:

Effects of Bilingualism and Relative Language Proficiency

Young children learn new words from very early on in life. Previous research has shown that both lexical principles and socio-pragmatic principles guide word learning, but studies differ regarding the importance they attach to either of these principles. One line of research focuses on lexical principles such as mutual exclusivity (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Hansen & Markman, 2009; Markman & Wachtel, 1988; Markman, Wasow, & Hansen, 2003), showing, for example, that children tend to avoid having two labels referring to the same object. In another line of research, socio-pragmatic principles such as joint attention, eye gaze and pointing are considered the key determinants of early word learning (Baldwin, 1991; Baldwin, 1993; Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996; Bannard & Tomasello, 2012). Baldwin et al. (1996) found, for example, that 18- to 20-month-old children were able to learn a novel label for a novel object when they saw a speaker attending to this object, but not when they only heard the speaker's voice, suggesting that socio-pragmatic cues are necessary for word learning (cf. Bannard & Tomasello, 2012). Hirotani, Stets and Friederici (2009) found in an ERP study that joint attention was needed to enable word learning over mere associative learning in 18- to 21-month-olds.

These latter studies show that socio-pragmatic cues are important for word learning. Likewise, studies looking at children's reference resolution have found that children rely on socio-pragmatic information when resolving a conflict between socio-pragmatic and lexical referential cues (Ateş, 2016; Grassmann & Tomasello, 2010). Specifically, in these studies, children's task was to disambiguate the reference of a speaker using conflicting referential cues,

namely, labeling one of two objects on the table while pointing or gazing at the other one. Children could resolve this conflict by following either the lexical (labeling) or socio-pragmatic (pointing/gazing) cue. The results of these studies indicated that children relied on both cues, with the relative strength of these cues being dependent, amongst others, on whether pointing occurred in an ostensive manner. In the current study, we investigate children's resolution of conflicts between labeling and pointing to find out (i) whether bilingual children rely on pointing versus labeling to the same extent as monolingual children, and (ii) whether bilingual children's reliance on pointing versus labeling is affected by whether the referential conflict is presented in their weaker or stronger language.

Jaswal and Hansen (2006) were the first to investigate how children weigh pointing and the mutual exclusivity principle when resolving a referential conflict. These authors administered a disambiguation task in which three- and four-year-old children were shown a novel and a familiar object. The experimenter then asked for the novel object ("Can you give me the blicket?"), while she pointed or looked at the familiar object. This study showed that English monolingual three- and four-year-old children overwhelmingly followed labeling over pointing and eye gaze when these cues were in conflict. The authors concluded that children expect words to be mutually exclusive even in the presence of a conflicting socio-pragmatic cue (cf. Graham, Nilsen, Collins, & Olineck, 2010).

More recently, Grassmann and Tomasello (2010) repeated Jaswal and Hansen's study with German monolingual two- and four-year-olds. However, these authors used ostensive pointing, that is, pointing combined with gaze alternation between the child and the object pointed at. In Jaswal and Hansen (2006), the pointing gesture was directed unambiguously at the object, while gazing was directed toward the child. Ostensive pointing, as used by Grassmann

and Tomasello, may be considered a more natural, ecologically valid way of pointing, which expresses more clearly the communicative intention that the object pointed (and looked) at is *for you* (Grassmann & Tomasello, p. 253). Grassmann and Tomasello found that children followed (ostensive) pointing over labeling, but not non-ostensive pointing as used by Jaswal and Hansen (2006). Grassmann and Tomasello also found that children's preference for pointing was weaker when the experimenter used a familiar label (e.g., 'car') than when the experimenter used a novel label (e.g., modi). Subsequent studies have replicated these findings for Turkish-learning two- and four-year-olds (Ateş, 2016) and German-learning four-year-olds (Grassmann, Magister, & Tomasello, 2011).

To date, studies on children's resolution of a conflict between pointing and labeling have exclusively looked at monolingual children. Worldwide, however, the number of children acquiring more than one language outnumbers the number of monolingual children (Associated Press, 2016; Grosjean, 2010; Tucker, 1998) and, in many parts of the world, the number of bilinguals is still increasing (e.g., Eurobarometer, 2012; Shin & Kominski, 2007). There are two main differences between bilingual and monolingual children that make it worthwhile to investigate bilingual children's resolution of a conflict between socio-pragmatic and lexical principles. First, bilingual children are known to rely less on lexical principles such as mutual exclusivity than their monolingual peers, most likely because they are used to knowing two labels for the same object, from their two languages (Byers-Heinlein & Werker, 2009; Davidson, Jergovic, Imami, & Theodos, 1997; Houston-Price, Caloghiris, & Raviglione, 2010). Second, bilingual children have been found to rely more strongly on non-verbal referential cues, such as gaze direction, than monolingual children (Brojde, Ahmed, & Colunga, 2012; Yow & Markman, 2011; Yow, 2014). This has been attributed to the fact that children growing up in a bilingual

environment need to constantly monitor what language a speaker is using and how to respond appropriately to avoid communication breakdowns, which would enhance their attentional sensitivity to the socio-pragmatic contexts surrounding verbal communication (Hung, Patrycia, & Yow, 2015; Yow & Markman, 2011). Also, the fact that bilingual children often have comparably smaller lexicons in each of their languages makes it more likely for them to encounter unknown words, which may be related to their enhanced attention to non-verbal information to extract linguistic meaning when they are uncertain (Siegal, Iozzi, & Surian, 2009; Yow, 2010).

In the current paper, we investigate whether bilingual children weigh lexical and socio-pragmatic referential cues differently than monolingual children during the resolution of a conflict between the two types of cues. We also investigate if bilingual children's previously acquired language knowledge affects their reliance on lexical versus socio-pragmatic cues such that they resolve referential conflicts differently depending on whether they are tested in their weaker or stronger language. Equal language proficiency levels are rare, even in bilinguals exposed who are exposed to two languages from birth (Paradis, 2007) and a variety of measures have been used to assess language proficiencies in bilinguals' two languages, including spontaneous speech measures and standardized vocabulary tests (Bedore et al., 2012).

To investigate effects of bilingualism and relative language proficiency on children's resolution of referential conflicts, we replicate the ostensive pointing experiment reported in Grassmann and Tomasello (2010) with monolingual and bilingual two- to four-year-olds. This age group is very similar to the age groups studied in earlier work, which also looked at preschoolers (Grassman & Tomasello, 2010; Jaswal & Hansen, 2006). Our aim is twofold. First, we investigate if bilingual children – due to their weaker reliance on mutual exclusivity

(Davidson et al., 1997; Byers-Heinlein & Werker, 2009) and/or stronger sensitivity to socio-pragmatic cues (Brojde et al., 2012; Yow & Markman, 2011) – show an even stronger preference for pointing over labeling than monolingual peers. Second, we examine effects of previously acquired language knowledge in the bilingual children. Specifically, we compare bilinguals' reference resolution across two datasets that were collected within the same experiment, but in two different languages (Dutch and English), one of which was their weaker and the other their stronger language. To date, previous studies have not looked at effects of relative language proficiency on bilingual children's resolution of referential conflicts, so it is an open question if children's existing language knowledge impacts on disambiguation behavior. However, if bilinguals' increased sensitivity to socio-pragmatic information is indeed related to their experience with lexical gaps, as was argued above, we predict that bilingual children may show a stronger preference for pointing over labeling in their weaker language than in their stronger language.

Method

Participants

Thirty-five Dutch monolingual children and 32 bilingual children participated in the study. These children had been selected out of a larger sample of 57 monolingual children and 59 bilingual children if they had completed all four trials of the experiment (monolinguals and bilinguals) as well as Dutch and English receptive vocabulary tasks (bilinguals), and if they attended Dutch-English bilingual daycare (bilinguals). Children had been recruited via daycare centers through an information letter asking parents to participate in the study. In this letter, parents indicated which languages were spoken at home, and, in the case of the bilingual

children, they also indicated which language out of English and Dutch their child knew best (answering options: English, Dutch, equally well).

The monolingual children ranged between 2;6 and 3;11 years in age (M = 3;2, SD = 0;5) and the bilingual children ranged between 2;6 and 4;4 years (M = 3;3, SD = 0;7). There were 14 (40%) boys in the monolingual group and 17 (53%) boys in the bilingual group. In both groups, the vast majority of children came from families in which at least one parent had obtained a bachelor degree (90% of the monolinguals, 97% of the bilinguals). Differences in age, gender and parental education were not significant between the groups (all ps > .10).

A subset of 28 bilingual children performed the experiment twice, once in Dutch and once in English. In this subgroup, the children ranged between 2;6 and 4;4 years of age (M = 3;3, SD = 0;6) and there were 12 boys (43%). The remaining four bilingual children did not have (complete) sessions in both languages due to experiment error (n = 2), or were excluded because they had the same scores on Dutch and English receptive vocabulary tasks (n = 2), rendering these children's data uninformative for our research question comparing performance in bilinguals' weaker versus stronger language.

All bilingual children attended one of two bilingual daycare centers in the Netherlands where they were exposed to both Dutch and English through (near-)native speakers. Although they were all exposed to Dutch and English at daycare, they formed a heterogeneous group in terms of the language(s) used at home. Specifically, some children were exposed to Dutch and English at home (n = 6), but other children were exposed to Dutch and another language (n = 7), or to English and another language than Dutch (n = 7). Others had parents who spoke neither Dutch nor English, but one other language (e.g., Hungarian) (n = 5), or each a different language

(e.g., Portuguese and Italian) (n = 1). Finally, there were six bilingual children who came from monolingual Dutch families and who were only exposed to their second language (English) at daycare. Thus, the bilingual group consisted of bilingual children who were exposed to Dutch and English at daycare and Dutch and/or English at home, as well as trilingual and quatrolingual children who were exposed to Dutch and English at daycare and one or two other languages at home. In the present study, all these children are termed 'bilingual'.

Materials

Main experiment. Children were given a disambiguation task that was a replication of the ostensive pointing condition in Grassmann and Tomasello (2010), in which a pointing cue was pitted against a labeling cue. Children performed this task while sitting in a child chair opposite to the experimenter. They were first introduced to a shoe-box-sized chute and invited to play with it by sliding objects through the chute in order to engage them in the task and familiarize them with the task materials. Subsequently, they were given four object selection trials in the condition they were randomly assigned to. In each trial, children were presented with two objects – one novel and one familiar – that were simultaneously shown to the child and placed in front of the child on a table. In the 'familiar label' condition, the experimenter said (the Dutch equivalent of) the following instruction: "Let's play with the car. Take the car". While producing this instruction, the experimenter pointed at the other (novel) object. In the 'novel label' condition, the experimenter said (the Dutch equivalent of): "Let's play with the modi. Take the modi". While saying this, the experimenter pointed at the familiar object (e.g., car) (see Figure 1). Pointing was ostensive in both conditions, such that pointing co-occurred with gaze alternation, in a natural manner. Specifically, while the experimenter pointed to the object with her forearm, and using her index finger, she repeatedly alternated her gaze between the child and

the object until the child produced a response. After children had made a choice, they were allowed to slide the selected object down the chute and play with it for approximately one minute.

The same familiar objects were used as in Grassmann and Tomasello (i.e., a toy car, a baby shoe, a comb, and a pen). As novel objects, we used a wheeled piece of metal, a round red label, a piece of a garden hose, and a closing clip. The novel words that were used were the same as in Grassmann and Tomasello (i.e., *modi*, *toma*, *bafo*, and *dofu*). For the bilingual children, who performed the experiment in both Dutch and English, the same novel toys were used across the two experiments, but with different labels in the familiar label condition, that is, either English or Dutch words. In the novel label condition, the same toys and novel labels were used across experiments, but with the phonology of the labels adapted to English/Dutch.

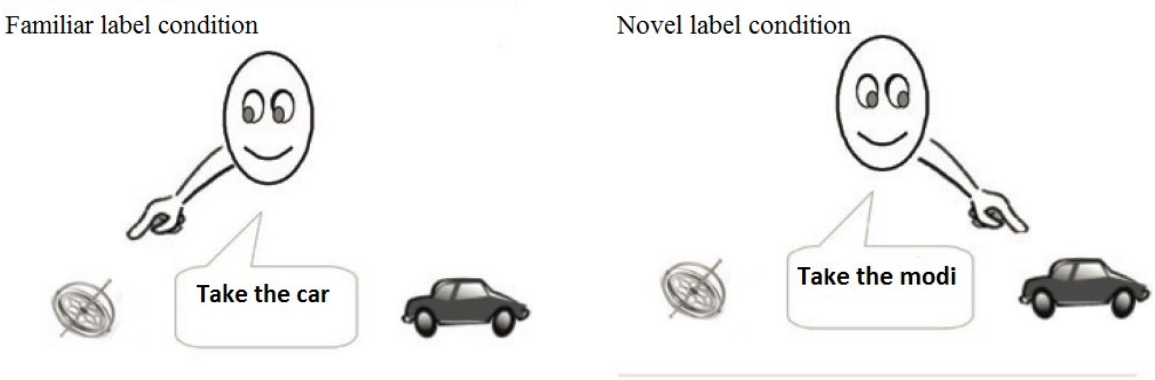

Figure 1. Familiar label and novel label conditions of the experiment (adapted from Grassmann and Tomasello, 2010, p. 257).

Design

The experiment contained a between-subjects factor 'condition' (familiar label vs. novel label). Furthermore, for the bilingual children, there also was a within-subjects factor 'relative

language proficiency' (weaker language vs. stronger language). The dependent variables were the number of trials in which the children showed a certain response type: (a) selecting the object pointed at ('point following'), (b) selecting the labeled object ('label following'), and (c) selecting both objects ('both objects')

We counterbalanced the pairings of novel and familiar objects, order of presentation of the object pairs, the left-right positioning of the novel and familiar objects on the table, and the assignments of the novel words to the novel objects. Furthermore, for the bilingual children, the order in which they participated in the two language versions of the experiment was varied. For the latter, strict counterbalancing was not possible for logistic reasons (i.e., limited number of days on which children/experimenters were available, and frequent absences of children due to illness, holidays etc.).

Scoring

Children's responses were scored on the basis of video recordings by trained assistants. Responses of a subset of 10 monolingual and 10 bilingual children were scored by an additional coder, showing100% agreement in scores for both groups. Following Grassmann and Tomasello (2010), children were scored as having chosen an object if they picked it up and handed it to the experimenter, held it up, or pushed it toward the experimenter. When children selected both objects, either simultaneously or consecutively, this was scored as 'both objects', unlike in Grassmann and Tomasello's study, who classified such responses on the basis of first touch, or – in the case of simultaneous selections – coded these as 'no response'. Thus, in our study, the following three response categories were used: 'followed pointing', 'followed labeling', and 'both objects'.

Vocabulary

Dutch receptive vocabulary. All children performed the Dutch version of the Peabody Picture Vocabulary Test (PPVT) (PPVT-III-NL, Dunn, Dunn, & Schlichting, 2005) which assesses receptive vocabulary. In this test, children choose one out of four pictures after an orally presented word. The test is adaptive such that testing is stopped when children make a fixed number of errors. Standard scores were computed.

English receptive vocabulary. The bilingual children also performed the English Peabody Picture Vocabulary Test (PPVT-4, Dunn & Dunn, 2007). The procedure was the same as in the Dutch PPVT. Standard scores were computed.

Procedure

Children were tested individually by trained research assistants in a quiet room at their daycare centers. The experiment formed part of a larger battery of language and executive function tasks, which are not reported on in the present paper. The tasks were administered in a fixed order within sessions in which the PPVT preceded the experiment.

The monolingual children performed the experiment in Dutch only. The bilingual children performed the experiment twice, once in Dutch and once in English. The English and Dutch versions of the experiment were administered in separate sessions on two different days that were between one week and two weeks apart. Different research assistants administered the two experiments, and assistants only spoke the relevant language to the children during the test session. The Dutch experiment was part of a session containing Dutch language and executive function tasks, which were administered by a Dutch native speaker. This session included the Dutch PPVT. The English experiment was part of a session containing English language and executive function tasks, administered by a (near-)native speaker of English. This session

included the English PPVT. The task procedure in the English experiment was exactly the same as in the Dutch experiment, and children were presented with the same condition in both experiments.

The order of presentation of both experiments varied across children: 16 out of the 28 children who did both experiment versions conducted the experiment in their stronger language first and the remaining 12 children conducted the experiment in their weaker language first. All sessions were videotaped for scoring purposes. At the end of each session, children received a small gift.

Analyses

To check for possible differences in disambiguation behavior between monolinguals and bilinguals (Question 1), we ran linear mixed-effect logistic regression models (or 'mixed models') in the statistical package R with 'subjects' and 'items' as random factors, 'group' and 'condition' as the independent variables, and 'point following' (0 = no point following, 1 = point following), 'label following (0 = no label following, 1 = label following), and 'both objects' (0 = one object only, 1 = both objects) as the dependent variables. In these analyses, we used for the bilingual group the data from the experiment conducted in children's stronger language, operationalized as the language in which they obtained the highest PPVT score. Comparing this group split on the basis of PPVT scores with a group split on the basis of parents' judgments of children's proficiency in both languages, we found that 24 of the 32 children (75%) had obtained the highest PPVT standard score in the language that their parents had indicated as their child's strongest language. For six of the 32 children, their parents had indicated that they were equally proficient in both languages, which was reflected in very similar cross-language PPVT scores for

these children. For the remaining two children, the language in which children obtained the highest PPVT score differed from the language judged by parents as their children's best language differences, but for these children, the differences in English and Dutch vocabulary scores were very small (i.e., 4- and 5-point differences in standard scores, respectively). This suggests that, overall, children received the highest PPVT score in the language that was indicated as their best language by their parents.

To examine possible effects of language proficiency on the bilingual children's disambiguation behavior (Question 2), we performed two analyses. First, we conducted mixed-effect logistic regression models on children's responses (e.g., point following vs. no point following), with 'subjects' and 'items' as random factors and 'relative language proficiency' (weaker vs. stronger language) and 'condition' as the predictor variables. As above, children's 'weaker language' and 'stronger language' were determined on the basis of PPVT scores, and checked against parents' judgments, showing that, in this subsample of 28 children who had completed both language versions of the experiment as well as both PPVTs, 24 out of 28 (86%) children obtained the highest PPVT standard score in the language characterized by their parents as their stronger language. We also tested for effects of administration order of the two experiments in the bilingual group by adding 'experiment order' (weaker language vs. stronger language in first session) as an independent variable to the analyses. Lastly, in order to investigate whether vocabulary scores correlated with children's disambiguation behavior, we calculated correlations between children's vocabulary scores in Dutch and point following in the experiment conducted in Dutch, as well as between children's vocabulary scores in English and point following in the experiment conducted in English.

Results

Question 1: Comparing Monolingual and Bilingual Children

Table 1 shows children's mean scores and standard deviations for the English and Dutch receptive vocabulary tasks (PPVTs). Standard scores are provided. A t-test showed that mean performance on the Dutch PPVT was higher in the monolingual group than in the bilingual group ($t(1,65) = 5.06$, $p < .001$, $d = 1.22$). In the bilingual group, there was no significant difference in mean scores across languages, but, on average, children obtained the highest score in Dutch ($t(1,31) = 1.57$, $p = .13$, $d = 0.41$). Importantly, none of the children scored below the minimum standard score of the tasks. Specifically, for the Dutch PPVT, the lowest standard score is 55 and the lowest score observed in the current sample was 57. For the English PPVT, the lowest standard score is 20 and the lowest score observed in the current sample was 51.

Table 1.

Descriptive Statistics for Dutch and English Receptive Vocabulary Tasks (PPVTs) for the Monolinguals and Bilinguals (Standard Scores)

	Monolinguals				Bilinguals			
	M	SD	min-max	N	M	SD	min-max	N
Dutch PPVT	107.17	9.53	81-124	35	89.06	18.68	57-118	32
English PPVT	-	-	-	-	81.75	17.28	51-112	32

Table 2 presents the descriptive statistics for the experiment conducted in children's stronger language, operationalized as the language in which they had obtained the highest PPVT

standard score out of English and Dutch. In this table, mean numbers per response type and standard deviations are provided for the two conditions for the monolingual and bilingual children separately.

Table 2.

Mean Number (and Standard Deviations) per Response Type per Condition for the Monolinguals and Bilinguals

		Followed pointing	Followed labeling	Both objects
Monolinguals	Familiar label	2.61 (1.50)	0.50 (0.79)	0.72 (1.18)
	Novel label	3.65 (0.60)	0.18 (0.53)	0.18 (0.39)
Bilinguals	Familiar label	3.41 (1.00)	0.53 (0.87)	0.06 (0.24)
	Novel label	3.33 (1.45)	0.67 (1.45)	0

Note. Monolinguals: n = 18 in familiar label condition, n = 17 in novel label condition. Bilinguals: n = 17 in familiar label condition, n = 15 in novel label condition. For the monolinguals in the familiar label condition, the data do not add up to four, as there were a few additional responses (n = 3) in which children did not select any of the objects. For the bilingual children, the data are based on a combination of English and Dutch (i.e., 13 English, 19 Dutch), as children differed as to in which language they had obtained the highest PPVT score.

Linear mixed-effect logistic regression models run in the statistical package R with 'subjects' and 'items' as random factors, 'group' and 'condition' as the independent variables and 'point following' (0 = no point following, 1 = point following) as the dependent variable showed an effect of 'group' (b = 2.57, SE = 1.29, z = 2.00, p = .04). 'Condition' was marginally

significant (b = 6.03, SE = 3.11, z = 1.94, p = .05), and the interaction between 'group' and 'condition' was not significant (b = -2.94, SE = 2.06, z = -1.43, p = .15). Taken together, these results show that the bilinguals followed pointing more often than the monolinguals and that there was an overall tendency to follow pointing more often in the novel label condition than in the familiar label condition.

As for the children's label following responses, there were no effects of group (b = -0.34, SE = 2.02, z = -0.17, p = .87) nor condition (b = -3.58, SE = 5.95, z = -0.60, p = .55), and no significant interaction between group and condition (b = -1.73, SE = 3.61, z = -0.48, p = .63). Finally, regarding the children's selection of both objects, there was a main effect of group (b = -2.69, SE = 1.05, z = 2.56, p = .01), but no effect of condition (b = -1.96, SE = 1.23, z = -1.59, p = .12), and no interaction between group and condition (b = -0.31, SE = 2.25, z = -0.19, p = .85). This shows that the monolingual children chose both objects significantly more often than the bilingual children.

Question 2: The Role of Language Proficiency

To address our second question on the role of relative language proficiency on children's reference resolution, we compared the bilingual children's behavior across the experiments conducted in their weaker versus stronger language. 'Weaker language' here refers to the language (out of English and Dutch) for which children obtained the lowest PPVT standard score. Table 3 shows mean numbers and standard deviations per response type and per condition for the experiment conducted in bilingual children's weaker versus stronger language. Note that the sample was somewhat smaller, as it consisted of children who had performed both PPVTs and for whom data in both language versions of the experiment were available (28 out of 32).

Table 3.

Mean Number (and Standard Deviations) per Response Type per Condition for the Bilinguals in their Weaker Versus Stronger Language

		Followed pointing	Followed labeling	Both objects
Weaker language	Familiar label	3.93 (0.27)	0.07 (0.27)	0
	Novel label	3.79 (0.80)	0.21 (0.80)	0
Stronger language	Familiar label	3.29 (1.07)	0.64 (0.93)	0.07 (0.27)
	Novel label	3.29 (1.49)	0.71 (1.49)	0

Note. n = 14 in the familiar label condition, n = 14 in the novel label condition. Note that these data are based on a combination of English and Dutch, because children differed in which language was their stronger language (stronger language: 11 English, 17 Dutch; weaker language: 17 English, 11 Dutch).

Linear mixed-effect logistic regressions on children's point following responses (0 = 'no point following', 1 = 'point following') with 'subjects' and 'items' as random factors, and 'relative language proficiency' and 'condition' as the predictor variables showed a main effect of 'relative language proficiency' (b = 3.18, SE = 1.29, z = 2.46, p = .01) indicating that the bilinguals followed pointing significantly more often in their weaker language than in their stronger language. There was no effect of 'condition' (b = 0.32, SE = 1.67, z = 0.19, p = .85) and no interaction between 'language proficiency' and 'condition' (b = -1.14, SE = 1.56, z = -0.73, p = .46). A model in which 'experiment order' (weaker language in first session vs. stronger language in first session) was added as a predictor did not show an effect of experiment order (b

= 1.06, *SE* = 1.66, *z* = 0.64, *p* = .52) and very similar results for the other factors as in the earlier model without 'experiment order' (i.e., relative language proficiency: *b* = 3.19, *SE* = 1.29, *z* = 2.47, *p* = .01; condition: *b* = 0.21, *SE* = 1.68, *z* = 0.13, *p* = .90; relative language proficiency*condition: *b* = -1.16, *SE* = 1.55, *z* = -0.74, *p* = .46).

While the above analyses show that the children followed pointing more often in their weaker language than in their stronger language, these results do not present direct evidence for an effect of language proficiency as measured by vocabulary knowledge on the children's disambiguation behavior. Specifically, one child might have a score of 70 in Dutch and 120 in English, while, for another child, these scores are 90 and 100 respectively. To test more directly how vocabulary scores related to children's disambiguation behavior, we investigated how children's PPVT scores in Dutch related to their disambiguation behavior in the experiment conducted in Dutch, and, likewise, how children's PPVT scores in English related to their disambiguation behavior in the experiment conducted in English. Correlation analyses showed moderate correlations for both languages for the novel language condition such that children with higher vocabulary scores followed pointing less often (*r*(14)= -.38, *p* = .19 for Dutch; *r*(14) = -.59, *p* = .03 for English), but only the correlation for English reached significance. For the familiar label condition, both correlations were close to zero and non-significant (*r*(14)= -.01, *p* = .98 for Dutch; *r*(14) = -.04, *p* = .89). So, when presented with a novel label in the English experiment, bilingual children with high vocabulary scores in English followed pointing less often than children with lower English vocabulary scores.

Discussion

This study replicates a previous experiment on children' resolution of referential conflicts (Grassmann & Tomasello, 2010) which showed that German monolingual children overwhelmingly followed pointing over labeling when both cues were pitted against each other in a disambiguation experiment. In our study, Dutch monolingual children and bilingual children were tested. The aim of our study was twofold: (i) compare monolingual and bilingual children's preferences for pointing versus labeling and (ii) investigate possible effects of relative language proficiency on the bilingual children's disambiguation behavior.

For the monolingual children, our results closely resembled those of Grassmann and Tomasello (2010): Dutch monolingual children aged two to four years largely followed pointing over labeling, and did so more often when presented with a novel label than with a familiar label. The main finding of the study is that the reference resolution patterns of the bilingual children differed from those of the monolingual children in two ways. First, the bilinguals followed pointing over labeling significantly more often overall. Second, the bilingual children hardly ever selected both objects in response to the experimenter's seemingly contradictory reference, while such responses did sometimes occur in the monolingual children.

A further finding was the bilingual children showed an effect of relative language proficiency such that they followed pointing more frequently in their weaker language than in their stronger language. Furthermore, we found some tentative evidence that children with higher vocabulary scores in a given language followed pointing less often in that language. However, only the correlation between English vocabulary scores and children's responses in the English experiment reached significance, so this finding needs to be interpreted with caution.

Taken together, our results suggest that bilingual children, just as monolingual children, rely on ostensively cued pointing rather than lexical labels to resolve reference when referential cues are contradictory, and in fact, do so even more than monolinguals. This is in line with previous research demonstrating that young bilingual children rely on pointing more than monolingual peers (e.g., Yow & Markman, 2011) and that bilingual children rely less on mutual exclusivity than monolingual children (e.g., Byers-Heinlein & Werker, 2009) in resolving referential acts. As such, our findings add to the growing body of evidence that bilingual children are more sensitive to socio-pragmatic information than monolingual children (Brojde et al., 2012; Yow & Markman, 2011; Yow, 2014). In addition, our finding that the bilingual children chose both objects significantly less often than the monolingual children suggests that the strategies to cope with the conflicting two cues may differ between monolingual and bilingual children. More research is needed to explain this difference, but a tentative explanation would be that the monolingual children were more hesitant to rely on either one of the cues and therefore resorted to a conflict resolution strategy in which both cues were equally weighted.

The finding that the bilingual children relied on pointing more strongly in their weaker language than in their stronger language suggests that children who know a language less well – lexically speaking – trust the corresponding information less, and instead show a stronger reliance on nonverbal cues such as pointing. This could also explain why young bilinguals generally rely on non-verbal information more strongly than monolingual peers: as they often know fewer words in each of their languages (Pearson, Fernández, & Oller, 1993), they may trust non-verbal information more. This idea is in line with earlier proposals suggesting that bilinguals' greater familiarity with encountering unknown words may make them attend more to non-verbal cues as a compensatory strategy (cf., Siegal, Iozzi, & Surian, 2009; Yow, 2010). It

Another possible explanation is that bilinguals' enhanced sensitivity to non-verbal information is related to their more intensive training in avoiding communication breakdowns (cf., Hung, Patrycia, & Yow, 2015; Yow & Markman, 2011) Future work could investigate in more detail if bilingual children indeed show increased sensitivity to non-verbal cues depending on specific experiences bilingual children may or may not have. For instance, some bilingual children may experience communication breakdowns more often than others, for example, because one or both of their parent(s) are not fluent in at least one of the children's languages. It would be interesting to investigate how such experiences relate to the development of communicative skills. We would predict that those bilingual children who have intensive experience with challenging verbal communication rely more on non-verbal cues than other bilingual children.

However, a recent study by Hung, Patrycia, and Yow (2015) seems to contradict such a hypothesis. This study showed that 3- and 4-year-old bilingual children who were presented with a communication breakdown were less likely to rely on pointing (versus labeling) than children who were not presented with such a communication breakdown. Specifically, in this study, bilingual children were presented with a story prior to a disambiguation experiment similar to the one used in the current study but with non-ostensive pointing, as in Jaswal and Hansen (2006). Crucially, the story contained either a switch from a familiar to another familiar language, or a switch from a familiar to an unfamiliar language, or no code-switching. The authors found that the children who had heard a story containing a switch to an unfamiliar language relied less on pointing when resolving the reference of a novel label that was used in conflict with a pointing gesture than children who heard either no switch or a switch to a familiar language. The authors suggest that the communication breakdown induced by the unfamiliar switch may have led the children to look at the experimenter more, perhaps because they expected the experimenter to

clarify the switch, but the inaction of the experimenter to repair the switch induced them to not rely on pointing in the disambiguation experiment. This study suggests that there may be complex interplays between specific properties of bilingual children's language environments, for example with respect to language switching and language proficiency, their interlocutors' communicative behaviors and children's reliance on non-verbal cues.

There are a few limitations to our study. First, even though only highly-frequent words were used, no pre-test was used to check whether children actually knew the words used in the familiar label condition (and conversely, did not know the words used in the novel label condition). However, we used the Dutch equivalents of the items used in Grassmann and Tomasello (2010), who did check whether the participants in their study actually understood those words. As German and Dutch are very similar, we think it is unlikely that there was a major problem with our stimuli. Second, one may wonder whether children's strong reliance on pointing in the current study as well as in earlier work (Ateş, 2016; Grassmann & Tomasello, 2010) is at least in part due to children's assumption that the aim of the adult's gesture is to help them solve the task. As no explicit instruction to the children was given prior to the experiment as to tell them that the adult was not going to help them, we cannot exclude that children relied on the pointing gesture because they did not want to counteract a cooperative looking adult. One way to test this would be to have an 'untrustworthy adult' or young child administer the test, and see if children follow pointing less often under such conditions. Finally, relative language proficiency in the bilingual children's languages was determined on the basis of receptive vocabulary scores on standardized tests in both Dutch and English. A more careful assessment of language proficiency in bilingual children would have entailed the measurement of other language skills as well, including grammar and language production skills. In our study, such a

detailed investigation was not possible, but checking our PPVT-based group split against parents' judgments of children's language proficiency showed a very similar pattern.

Despite these limitations, we think the results of our study provide at least initial evidence that bilingual children weigh referential cues differently from monolingual children. The current findings support earlier results that bilingual children are more sensitive to non-verbal cues (Brojde et al., 2012; Yow & Markman, 2011; Yow, 2014) and rely less on principles of mutual exclusivity than monolingual children (Byers-Heinlein & Werker, 2009), and expand these to situations of conflict. The current results also show that bilingual children's reliance on pointing versus labeling is related to existing language knowledge. Children are more inclined to follow an adult speaker's point in the language which they know less well – lexically speaking – than in the language they know more words in. A question of particular interest for future research would be whether the children indeed learn from their resolution of the contradictory reference and retain and integrate the information to their lexical knowledge (as suggested by Grassmann & Tomasello, 2010) – or whether the effect observed in the current paper is a short-lived effect of reference resolution.

Acknowledgements

This research was financed by a VENI grant from the Netherlands Organisation for Scientific Research (NWO) awarded to XXX, grants from the Prince Claus Chair for Development and Equity and NWO awarded to XXX, and a grant from the Education and Learning Sciences focus area at Utrecht University awarded to XXX, XXX, and XXX. We would like to thank Jelle Fremerij for his help with data collection and setting up the experiment, and Sharon Unsworth and Elise de Bree for allowing us to include the experiment as part of a larger test battery. Finally, we would like to thank Willemijn Doedens for recoding part of the data, and Marleen Berkhout, Sascha Couvee, Mandy Dorée, Marlies van Hombergen, and Eline van Knijff for collecting the data.

References

Associated Press. (2016). Some facts about the world's 6,800 tongues. September 1, 2016, http://edition.cnn.com/2001/US/06/19/language.glance/index.html?_%2520s=PM:US

Ateş, S. B. (2016). Developing verbal and non-verbal referential communication skills in young children's naturalistic and experimental interactions with adults. Unpublished doctoral dissertation. Koç University, Turkey.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*, 875–890.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*, 395–418.

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*, 3135–3153.

Bannard, C., & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PLoS ONE, 7:* e49881.

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., and Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition, 15*, 616-629.

Brojde, C. L., Ahmed, S., & Colunga, E. (2012). Bilingual and monolingual children attend

to different cues when learning new words. *Frontiers in Psychology, 3,* 1–11.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word learning heuristic. *Developmental Science, 12*, 815–823.

Davidson, D., Jergovic, D., Imami, Z., & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language, 24,* 3–24.

Dunn, L. M, & Dunn, D. M. (2007). *Peabody picture vocabulary test (PPVT-4)*. Minneapolis, USA: Pearson.

Dunn, L. M., Dunn, L. M., & Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Amsterdam: Pearson.

Eurobarometer (2012). Europeans and their languages. European Commission. Retrieved 23 September 2016: http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf

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., Nilsen, E. S., Collins, S., & Olineck, K. (2010). The role of gaze direction and mutual exclusivity in guiding 24-month-olds' word mappings. *The British Journal of Developmental Psychology*, *28*, 449–465.

Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science, 13*, 252–263.

Grassmann, S., Magister, C., & Tomasello, M. (2011). What children do when pointing and naming conflict. Paper presented at the SRCD Biennial Meeting, Montreal, Canada.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Hirotani, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence. *Neuroreport, 20*, 600–605.

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy, 15*, 125–150.

Hung, W-Y., Patrycia, F., & Yow, W. Q. (2015). Bilingual children weigh speaker's referential cues and word-learning heuristics differently in different language contexts when interpreting a speaker's intent. *Frontiers in Psychology, 6*: 796.

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

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

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*, 241–275.

Paradis, J. (2007). Early bilingual and multilingual acquisition. In P. Auer & L. Wei (Eds.), *Multilingualism and multilingual communication. Handbook of applied linguistics*, Vol.5 (pp. 15-44). Berlin: Mouton de Gruyter.

Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*, 93–120.

Siegal, M., Iozzi, L., & Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition, 110,* 115–122.

Shin, H.B., & Kominski, R.A. (2010). *Language Use in the United States: 2007*. American Community Survey Reports, ACS-12. U.S. Census Bureau, Washington, DC.

Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Clevedon: Multilingual Matters.

Yow, W. Q. (2010). Monolingual and bilingual children's integration of multiple cues to understand a speaker's referential intent – the role of experience in cognitive development. Unpublished doctoral dissertation. Stanford University, US.

Yow, W. Q. (2014). Monolingual and bilingual preschoolers' use of gestures to interpret ambiguous pronouns. *Journal of Child Language*, *42*, 1394–1407.

Yow, W. Q., & Markman, E. M. (2011). Young bilingual children's heightened sensitivity to referential cues. *Journal of Cognition and Development, 12,* 12–31.

Yow, W. Q., & Markman, E. M. (2014). A bilingual advantage in how children integrate multiple cues to understand a speaker's referential intent. *Bilingualism: Language and Cognition, 18*, 391–199.