SPECIAL ISSUE ARTICLE

A bilingual advantage in 54-month-olds’ use of referential cues in fast mapping

W. Quin Yow,1 Xiaoqian Li,1 Sarah Lam,2 Teodora Gliga,3 Yap Seng Chong,2,4,5 Kenneth Kwek6 and Birit F.P. Broekman2,7,8

1. Singapore University of Technology and Design, Singapore
2. Singapore Institute for Clinical Sciences, Singapore
3. Centre for Brain and Cognitive Development, Birkbeck, University of London, UK
4. Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
5. Department of Obstetrics and Gynaecology, National University Hospital, National University Health System, Singapore
6. KK Women’s and Children’s Hospital, Singapore
7. Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore,
8. Department of Psychological Medicine, National University Hospital, National University Health System, Singapore

Abstract

Research has demonstrated a bilingual advantage in how young children use referential cues such as eye gaze and pointing gesture to locate an object or to categorize objects. This study investigated the use of referential cues (i.e. eye gaze) in fast mapping in three groups of children that differed in their language exposure. One hundred and seven 54-month-old children who were English monolinguals (n = 28), English-Mandarin bilinguals (n = 48), and English-Mandarin bilinguals with exposure to a third language (i.e. trilinguals, n = 31) were assessed with a word learning task using two types of test – a referent test and a mutual exclusivity test. During the task, following the gaze of an adult speaker was needed to be able to indicate the correct referent of a novel word at test. All three groups of children demonstrated successful word learning in explicit selection of and implicit looking time toward the target object during testing. However, bilingual and trilingual children outperformed their monolingual peers in both types of test when they were asked to explicitly select the correct objects. These findings suggest positive effects of bilingualism on children’s use of referential cues in fast mapping.

Research highlights

• Bilingual and trilingual children were significantly better able to use the speaker’s eye gaze to learn object–label associations and further infer a new label for an unlabeled object than monolingual children.
• Multilingual children’s greater sensitivity to a speaker’s communicative intent and perspective extends to fast mapping of words in a social context.
• The ability to quickly associate a novel word with a novel object can be enhanced by diverse experiences in a language-learning environment.
• Word learning may require more than paying attention to the co-occurrence of words, social referential cues and objects – the ability to understand the referential intent of the interlocutor is critical to any successful form of learning in a social context.

Introduction

Children are able to quickly derive the meaning of a novel word even when exposed to it only briefly. This strategy, known as fast mapping, occurs when children encounter a novel word and use the linguistic and nonlinguistic contexts to rapidly acquire information
about its meaning (Carey, 1978; Carey & Bartlett, 1978; Heibeck & Markman, 1987; for a review, see Woodward & Markman, 1998). Fast mapping can be assessed with a typical novel-object-word pairing task, in which participants see a novel object and hear its corresponding novel label. Participants are then asked to either name the new object or identify the target object that corresponds to the new word after minimum exposure to the word–referent mapping.

Research shows that children as young as 2 years old are able to use novel-object-word pairings to fast map novel words after a brief single exposure (Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992; Spiegel & Halberda, 2011). Children are also able to use other sources of information to rapidly narrow down a word’s meaning, including grammatical properties (e.g. nouns and adjectives; Gelman & Markman, 1985; Hall, Quantz & Persoage, 2000), object properties (e.g. shape or color of objects; Colunga & Smith, 2005), social contexts (e.g. having access to location for search; Akhtar & Tomasello, 1996), and referential cues (e.g. eye gaze and pointing gestures; Baldwin, 1991, 1993; Diesendruck & Markson, 2001; Tomasello, 1998; Woodward, 2000).

Referential cues are believed to be one of the most reliable heuristics that children use to build word–referent mappings from the early stages of language development. It is well documented that young children are sensitive to a speaker’s communicative cues that indicate referential intent (e.g. Berman, Chambers & Graham, 2010; Jaswal, 2004). They are able to monitor the speaker for cues, such as gaze direction, to determine meanings of novel words. By 9 to 10 months of age, infants can successfully follow speakers’ referential gestures and they are more likely to map a novel word to an object which the speaker is attending to (e.g. Brooks & Meltzoff, 2002; Woodward, 2003).

Children are also able to follow a speaker’s referential cues in spite of the presence of a salient distractor to arrive at a novel-object-word mapping (Gliga, Elsabagh, Hudry, Charman & Johnson, 2012; Houston-Price, Plunkett & Duffy, 2006, Experiment 4). In Gliga et al.’s (2012) novel word learning paradigm, 3-year-old children were shown video clips of an ambiguous word learning situation, where they saw two novel objects presented together. One of the objects was physically salient (such as having moving parts or it changed color), and the other was less salient (no physical change in state). A speaker turned, looked at the less salient object, and labeled it with a novel word. Children had to follow the speaker’s referential cue (i.e. eye gaze) to the less salient of the two objects in order to learn the novel word. In the following test trial, children were asked to look at and point to the referent of the novel word (i.e. referent test). Children’s eye movements were recorded with an eye tracker, which allows researchers to explore the relation between word learning and gaze-following. The researchers found that children were able to follow the speaker’s gaze toward the less salient object, and looked at and pointed to the less salient object above chance during the referent test trials.

Research suggests that different language exposure may foster the development of different word learning heuristics (Bialystok, Barac, Blaye & Poulin-Dubois, 2010; Brojde, Ahmed & Colunga, 2012; Merriman & Kutlesic, 1993; Smith, Colunga & Yoshida, 2010). In particular, children tend to use the mutual exclusivity principle in word learning to assume that different object categories have different names; thus one object should only have one label (ME; Markman, 1991; Markman & Wachtel, 1988). Children from a multilingual environment are typically more relaxed with the ME constraint when mapping novel words to novel objects compared to monolingual children (e.g. Byers-Heinlein & Werker, 2009; Davidson, Jergovic, Imami & Theodos, 1997; Houston-Price, Caloghiris & Raviglione, 2010; but cf. Au & Glusman, 1990; Frank & Poulin-Dubois, 2002). For example, Davidson et al. (1997) showed that bilingual children aged 3.5 to 6 years were less likely to show an ME bias and also less likely to reject a new name for a familiar object than monolingual children. Studies also found that when mutual exclusivity was in conflict with referential cues such as pointing, bilingual children were more likely to favor the referential cues over mutual exclusivity, compared to monolingual children who showed a more robust use of the word constraint (Brojde et al., 2012; Diesendruck, 2005; Healey & Skarabela, 2008; Jaswal & Hansen, 2006; Yow & Markman, 2007; but cf. Grassmann & Tomasello, 2010).

Research comparing monolingual and bilingual children in a word learning task suggests that the two groups of children may attend to different sources of information when learning words. This difference is likely because monolingual and bilingual children grow up in linguistic environments that have different characteristics and have different learning demands (Brojde et al., 2012; Hung, Patrycia & Yow, 2015). Bilingual children frequently encounter situations where an object has two different names in two different languages. Bilingual children also need to adapt and adopt different linguistic perspectives unique to individual languages. In addition, bilingual children are faced with the additional task of discovering what language a speaker is using. As such, they have to develop some form of self-generated monitoring strategies to manage their more demanding communicative environment, which then lead to greater attention to the speaker for cues in word learning.
Indeed, bilingual children are found to be more sensitive to a speaker’s referential cues and have better perspective-taking skills compared to monolingual children (Fan, Liberman, Keysar & Kinzler, 2015; Yow & Markman, 2011, 2016). In contrast, monolingual children have the tendency to attend more to object properties when learning new words (Brojde et al., 2012; Haryu & Imai, 1999). For example, Brojde et al. (2012) found that monolingual children relied more on object property cues (i.e. physical similarity between the objects), whereas bilingual children relied more on pragmatic cues (i.e. experimenter’s eye gaze) to extend the just-learned labels to other objects.

Although studies found that monolingual and bilingual preschoolers do succeed in fast mapping tasks (e.g. Kan & Kohnert, 2008; Merriman & Kutlesic, 1993; Rohde & Tiefenthal, 2000), it is unknown whether differences in language exposure would affect children’s use of referential cues in fast mapping. This is especially so when faced with conflicting information such as saliency of a distractor. In this study, we seek to examine how differences in language exposure might affect children’s ‘monolingual’ if the child was exposed to English at least 90% of the time, ‘bilingual’ if the child was exposed to a second language (i.e. Mandarin) regularly besides English at least 25% of the time, and ‘trilingual’ if the child was exposed to a third language besides English and Mandarin at least 10% of the time (see a similar criterion used in Byers-Heinlein & Werker, 2009). The final sample included 28 English monolinguals (14 males, 14 females), 48 English-Mandarin bilinguals (20 males, 28 females) and 31 English-Mandarin-Language X trilinguals (17 males, 14 females). The distribution of the monolinguals, bilinguals, and trilinguals in this sample is similar to the distribution of the population (Singapore Department of Statistics, 2015). The average English exposure for monolinguals was 92.3%. The average language exposure for bilinguals was 54.9% English and 45.5% Mandarin. The average language exposure for trilinguals was 53.5% English, 21.1% Mandarin, and 25.3% third language (Bengali = 1, Cantonese = 3, Dialect = 2, Dutch = 1, Hindi = 2, Hokkien = 5, Indonesian = 1, Malay = 13, Tamil = 1, Teochew = 2). Table 1 shows the descriptive statistics of the three language groups.

Information on maternal education and monthly household income was also collected. Children in the three language groups were comparable in demographic characteristics. Chi-square test showed that there were no significant differences in the number of children whose mothers completed at least high school between the language groups, $\chi^2 (2) = 2.21, p = .33$ (monolingual: 78.6%, bilingual: 62.5%, trilingual: 71.0%). The average monthly household income, on a scale from 1 ($< $1000) to 5 ($> $6000), was also not different across the language groups, $F(2, 90) = 1.01, p = .37$, $\eta^2 = .022$ (monolingual: 4.1, bilingual: 3.9, trilingual: 3.7). All children were of Asian descent and were Singaporean residents born in Singapore.

**Method**

**Participants**

The sample comprised 107 children (51 males, 56 females). The mean age of participants was 54.95 months (range = 52.53–56.15 months). Participants were recruited as part of the Growing Up in Singapore Towards Healthy Outcomes (GUSTO) birth cohort. Parents were asked to list the language(s) and estimate the percentage of each language spoken to their children at 48 months. A child was considered to be a ‘monolingual’ if the child was exposed to English at least 90% of the time, ‘bilingual’ if the child was exposed to a second language (i.e. Mandarin) regularly besides English at least 25% of the time, and ‘trilingual’ if the child was exposed to a third language besides English and Mandarin at least 10% of the time (see a similar criterion used in Byers-Heinlein & Werker, 2009). The final sample included 28 English monolinguals (14 males, 14 females), 48 English-Mandarin bilinguals (20 males, 28 females) and 31 English-Mandarin-Language X trilinguals (17 males, 14 females). The distribution of the monolinguals, bilinguals, and trilinguals in this sample is similar to the distribution of the population (Singapore Department of Statistics, 2015). The average English exposure for monolinguals was 92.3%. The average language exposure for bilinguals was 54.9% English and 45.5% Mandarin. The average language exposure for trilinguals was 53.5% English, 21.1% Mandarin, and 25.3% third language (Bengali = 1, Cantonese = 3, Dialect = 2, Dutch = 1, Hindi = 2, Hokkien = 5, Indonesian = 1, Malay = 13, Tamil = 1, Teochew = 2). Table 1 shows the descriptive statistics of the three language groups.

**Design and materials**

**Novel word learning paradigm**

This paradigm was adapted from Gliga et al. (2012) and was administered as part of the GUSTO 54-month neurodevelopment visit. An experimenter introduced the study to the child in the language most often used by the child, as reported by the child’s mother (English = 85,
Table 1: Means (and standard deviations) of demographic information and scores on KBIT-2 and flanker task by language group

<table>
<thead>
<tr>
<th></th>
<th>Monolingual</th>
<th>Bilingual</th>
<th>Trilingual</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>28</td>
<td>48</td>
<td>31</td>
</tr>
<tr>
<td>Male:Female</td>
<td>14:14</td>
<td>20:28</td>
<td>17:14</td>
</tr>
<tr>
<td>Age (months)</td>
<td>54.8 (0.8)</td>
<td>55.0 (0.6)</td>
<td>55.0 (0.7)</td>
</tr>
<tr>
<td>Maternal education (%)</td>
<td>78.6%</td>
<td>62.5%</td>
<td>71.0%</td>
</tr>
<tr>
<td>Monthly household income $ (1–5)</td>
<td>4.1 (1.0)</td>
<td>3.9 (1.0)</td>
<td>3.7 (1.1)</td>
</tr>
<tr>
<td>Exposure to English (%)</td>
<td>92.3 (4.0)</td>
<td>54.5 (15.3)</td>
<td>53.5 (15.9)</td>
</tr>
<tr>
<td>Exposure to Mandarin (%)</td>
<td>–</td>
<td>45.5 (15.3)</td>
<td>21.1 (13.7)</td>
</tr>
<tr>
<td>Exposure to other language (%)</td>
<td>–</td>
<td>–</td>
<td>25.3 (16.3)</td>
</tr>
<tr>
<td>KBIT-2 (Overall IQ)</td>
<td>95.6 (13.5)</td>
<td>98.5 (14.3)</td>
<td>96.4 (12.8)</td>
</tr>
<tr>
<td>KBIT-2 (Verbal)</td>
<td>92.6 (14.5)</td>
<td>93.0 (18.8)</td>
<td>86.7 (15.7)</td>
</tr>
<tr>
<td>KBIT-2 (Nonverbal)</td>
<td>99.4 (15.1)</td>
<td>103.8 (11.6)</td>
<td>106.8 (11.0)</td>
</tr>
<tr>
<td>Flanker RT (ms) – Congruent</td>
<td>1358 (260)</td>
<td>1307 (303)</td>
<td>1417 (233)</td>
</tr>
<tr>
<td>Flanker RT (ms) – Incongruent</td>
<td>1507 (320)</td>
<td>1464 (341)</td>
<td>1491 (322)</td>
</tr>
<tr>
<td>Flanker effect (ms)</td>
<td>149 (227)</td>
<td>157 (212)</td>
<td>101 (306)</td>
</tr>
</tbody>
</table>

Notes: KBIT-2 = Kaufman Brief Intelligence Test, 2nd edn. aNumber of children whose mother had completed at least high school. bMonthly household income on a scale of from 1 (< $1000) to 5 (> $6000). cFor KBIT-2, n = 26, 41, and 28 for monolingual, bilingual, and trilingual, respectively. dFor flanker task, n = 26, 46, and 30 for monolingual, bilingual and trilingual, respectively.

Chinese = 21, Malay = 1), but the stimuli were presented in English to all children. Preliminary analysis showed no significant differences in task performance between the different instruction languages.

The stimuli were eight video clips each consisting of a familiarization phase (two intervals: baseline and teaching) and a test phase (two types of test: a referent test and a mutual exclusivity test). In each video clip, a static novel object (non-changing and non-moving) was paired with a salient novel distractor (changed color or had moving parts). At the beginning of each familiarization phase, a female speaker in the video (different from the experimenter) was seated behind a table with two objects placed on it. In a 4-second baseline interval, the speaker looked straight at the camera, maintained a still face for 3s, and then said in a cheerful voice, ‘Hello!’ This was followed by a 4-second teaching interval, where the speaker turned her gaze toward the static object and labeled it (see Figure 1). There were two types of trials in the teaching interval. In four of the video clips (short-length), the speaker looked at the static novel object and labeled it using a novel word twice. For example, ‘Look at this! It’s a neem.’ Looking back at the camera she then said, ‘Do you want to play with the neem?’ The other three novel words used in the short teaching clips were zoop, kobe, and teri. In the remaining four video clips (long-length), the speaker labeled the static object three times. For example, ‘Look at this! It’s a blicket! Wow! Look at the blicket.’ She then looked back toward the camera and said, ‘Do you want to play with the blicket?’ The other three novel words used in the long teaching clips were dax, sefo, and toma (the two types of teaching trials were meant to explore whether the amount of exposure to the novel word would affect the word learning performance; see Appendix A for a list of objects and novel words used in the study). The placement of the static and salient objects on either side of the screen was counterbalanced across trials. The order of short and long teaching trials was also counterbalanced across trials and participants.

In the test phase, a central animation was first shown, followed by a referent test and a mutual exclusivity test, or vice versa. For both tests in each video clip, a still image of two objects appeared on the table with their positions reversed compared to the familiarization phase.

In the referent test, the speaker drew attention to the static object, e.g. ‘Look at the blicket! Look at the blicket!’ and then requested it, ‘Can you show me the blicket? Can you show me the blicket?’ In the mutual exclusivity (ME) test, the speaker referred to the salient distractor with a new novel label, e.g. ‘Look at the plume! Look at the plume!’ followed by ‘Can you show me the plume?’ Can you show me the plume?’ (see Figure 1).

Each test ended when children pointed to one of the two objects in response to the last question of each test. Children’s eye gaze data were recorded throughout the test phase with an eye tracker but only the looks during the first 4 seconds of each test (i.e. the test interval) were analyzed, during which children were requested to look at the target object. The order of the two types of test was counterbalanced across trials, such that the referent test came before the ME test on four video clips and vice versa for the other four video clips. There were two additional familiar word test clips, where a pair of familiar objects (a rubber duck and a child’s shoe) was presented together, and the speaker requested one of the two objects. These two familiar word trials were used to help children understand the task.

Kaufman Brief Intelligence Test, 2nd edition (KBIT-2)

The KBIT-2 was administered to measure children’s verbal and nonverbal abilities (Kaufman & Kaufman, 2004). The test consists of three subsets with a total of 154 items. On each item, the experimenter said a word or asked a question and the child was to provide one-word responses or point to one of five or six pictures that illustrated the meaning of the word or the answer to the
question. The test was administered and scored following the standard procedures. Each child obtained a standardized score with a mean of 100 for overall IQ Composite Score, as well as a Verbal Score and a Nonverbal Score. Data for 12 children (monolingual = 2, bilingual = 7, trilingual = 3) were missing; thus they were not included in the analysis of this task.

Flanker task

The flanker task was adapted from Diamond, Barnett, Thomas and Munro (2007), which included a standard and a reverse version of the task. In the standard flanker task, children saw one to five blue fish in the center of the screen pointing to (i.e. facing) either the left or the right. They were to indicate the direction of the center fish by pressing one of two response pads placed on the left or right side of the display (i.e. press the left pad if the center fish pointed to the left and press the right pad if it pointed to the right). There were four conditions that differed in the type of flankers. In the standard-control condition, a single fish appeared in the center of the screen. In the standard-neutral condition, two flankers appeared on each side of the center fish and both flankers pointed downward. In the standard-congruent condition, all five fish pointed in the same direction, whereas in the standard-incongruent condition, the four flankers pointed in the opposite direction of the center fish. In the reverse flanker task, children saw four or five pink fish on each trial and they were to indicate the direction of the flankers. That is, they were to focus on the flankers and inhibit attention to the center stimulus. In the reverse-control condition, only four flankers (with no center fish) were presented. In the reverse-neutral condition, four flankers pointed to either the left or the right while the center fish pointed downward. The stimuli for the reverse-congruent and reverse-incongruent trials were the same as those used in the standard task except for the color of the fish.

Each child completed both a block of the standard flanker task and a block of the reverse flanker task each consisting of 16 trials (i.e. four trials of each type of flankers), with the order of blocks counterbalanced across participants. Each block was preceded by practice with feedback. Trials began with a blank screen for 250 ms, followed by the stimulus for 2000 ms or until the child responded. Accuracy and response time (RT) were recorded for each trial. Five children (monolingual = 2, bilingual = 2, trilingual = 1) did not complete this task and thus were excluded from analysis of this task.
Preliminary analyses indicated no effects of task version (standard vs. reverse), so they were combined in subsequent analyses. The flanker effect was calculated as the difference between the mean RT of incongruent and congruent trials, which reflects the cognitive cost of inhibiting attention to the distractors (see Bialystok & Barac, 2012).

Procedure

Children were seated on a chair in front of a computer monitor mounted with a Tobii T60 eye tracker. The children were told that a ‘teacher’ was going to teach them some new words in English that they might not have heard before. At the beginning of the experiment, a 5-point calibration sequence was run until at least 4 points were properly calibrated for each eye.

Each child was then presented with a video clip that consisted of a pair of familiar objects (a rubber duck and a shoe) and a speaker’s voice from the video clip requesting one of the two objects (e.g. ‘Look at the duck! Look at the duck! Can you show me the duck? Can you show me the duck?’). The child was then verbally prompted by the experimenter to point to one of the two objects (e.g. the experimenter asked ‘Where is it?’) and was reminded that he/she should point only when asked. The child was then shown two video clips that each consisted of a familiarization phase and a test phase and was verbally prompted to point to one of the two objects at each test trial. The same pair of familiar objects was shown again, but this time the speaker’s voice requested the other object. The child was then shown the remaining six video clips. There were two fixed orders of trial presentation and children were randomly assigned to either one of the orders. The experimenter recorded children’s pointing responses to the test trials while the eye tracker recorded the children’s eye movements.

Children who completed the novel word learning task were tested on the KBIT-2 and flanker task in a separate session on the same day of the visit to assess their verbal and nonverbal intelligence as well as their inhibitory control ability.

Results

Based on previous literature (Gliga et al., 2012; Senju & Csibra, 2008), we derived two measures of word learning: the percentage of correct pointing during the referent and the mutual exclusivity tests, and the proportion of looking time directed to the correct object during the test intervals (i.e. the static object in the referent test and the salient object in the mutual exclusivity test). Trials with no pointing were considered incorrect responses. There were a total of 13 trials (0.76% out of all test trials) where children did not point to either one of the objects. These 13 trials were distributed across children from the three language groups. Three monolingual, one bilingual and one trilingual child failed to point in one of the eight referent test trials. In addition, two monolingual, three bilingual and three trilingual children did not point in one of the eight ME test trials.

Eye gaze data were processed and analyzed using eyetrackingR (Dink & Ferguson, 2015). As the objects of each pair differed in size, rectangular areas of interests (AOIs) were defined around each object with 0.6 degree of visual angle margin as suggested by Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka et al. (2015) and Orquin, Ashby and Clarke (2016). The average proportional looking time across groups was examined in a 3-second window of analysis after the onset of the first label during the test interval (i.e. 1 second after the onset of the trial). The proportional looking time toward the target object was calculated with respect to the total amount of time spent looking at the two objects. All data were normalized for area sizes. Of the 107 participants, nine were excluded from the eye gaze data analysis because they did not accumulate enough looking time due to either no gaze data at all (one bilingual and one trilingual), a technical problem (one monolingual, one bilingual, and one trilingual), or no accumulated looking time toward both objects in more than half of the test trials (two bilingual and two trilingual).

Preliminary analyses found no significant main effects of gender, teaching type (short-length vs. long-length), or interaction with language group, so they were combined in subsequent analyses. One-way ANOVAs comparing the language groups on their performance in the KBIT-2 and the flanker task revealed no significant differences across groups, KBIT-2 overall IQ: \( F(2, 92) = 0.40, p = .67, \eta^2_p = .009 \), KBIT-2 verbal score: \( F(2, 92) = 1.32, p = .27, \eta^2_p = .028 \), KBIT-2 nonverbal score: \( F(2, 92) = 2.39, p = .097, \eta^2 = .049 \), Flanker: \( F(2, 99) = 0.50, p = .61, \eta^2 = .010 \) (see Table 1). Thus, the monolingual, bilingual, and trilingual groups were comparable in their English verbal knowledge and inhibitory control ability. Initial analyses on the looking behavior during the familiarization phase found no significant group differences in the amount of time spent looking at each of the AOIs (i.e. static object, salient object, and face) during the baseline interval, \( F(2, 95) < 1.10, \text{all } ps > .34, \eta^2 < .023 \), and the teaching interval, \( F(2, 95) < 2.40, \text{all } ps > .097, \eta^2 < .048 \). All groups of children demonstrated similar distribution of looking amongst the various AOIs during the familiarization phase (see Table 2).
Children’s performance on the familiar word test trials was also analyzed. For pointing, all but four children pointed to the duck or shoe correctly in the two familiar word test trials. Two monolingual and one trilingual child failed to point in the first familiar word test trial when prompted to do so; however, they all pointed and chose the correct referent in the second familiar word test trial. One bilingual child chose duck in both of the two familiar word test trials. There were no significant differences between language groups in terms of the number of children who succeeded in both of the two familiar word test trials, \( \chi^2(2) = 1.29, p = .53 \). The results did not change after removing these four children, so the children were included in the remaining analyses. For looking times, there were also no significant differences between language group in the proportional looking toward the correct referent, \( F(2, 95) = 1.54, p = .22, \eta^2 = .031 \). All groups looked toward the target object at above chance levels – monolingual: \( t(26) = 4.80, \) bilingual: \( t(42) = 6.43, \) and trilingual: \( t(27) = 8.29, \) all ps < .001, \( ds > 0.92 \).

**Pointing during test**

We first analyzed the performance during the test based on children’s explicit pointing behavior using repeated measures ANOVA with language group (monolingual, bilingual and trilingual) as the between-subjects variable and test type (referent vs. mutual exclusivity test) as the within-subject variable. There was a significant main effect of language group in the mean percentage of correct response on these first test trials was found, \( F(2, 104) = 3.19, p = .045, \eta^2 = .032 \). There was no significant main effect of test type, \( F(1, 104) = 2.67, p = .11, \eta^2 = .011, \) or interaction of language group and test type, \( F(2, 104) = 0.11, p = .90, \eta^2 = .001 \). Bilinguals and trilinguals were significantly better than monolinguals in establishing first referent based on the speaker’s eye gaze in the referent tests, as well as using the speaker’s eye gaze to constrain the second novel label in the ME tests.

Separate one-sample \( t \)-tests were conducted to evaluate children’s overall word learning performance in each of the language groups. All groups performed above chance level in pointing to the correct object, monolinguals in referent test: \( t(27) = 6.84, p < .001, d = 1.25 \), monolinguals in ME test: \( t(27) = 5.79, p < .001, d = 1.09 \), bilinguals in referent test: \( t(47) = 15.19, p < .001, d = 2.19 \), bilinguals in ME test: \( t(47) = 13.26, p < .001, d = 1.94 \), trilinguals in referent test: \( t(30) = 11.21, p < .001, d = 2.00 \), and trilinguals in ME test: \( t(30) = 11.64, p < .001, d = 2.19 \). This suggests that children were able to correctly point, above chance, to

\[
\begin{array}{l|c|c|c}
\text{Table 2} & \text{Proportion looking duration toward each area of interest during baseline and teaching intervals of each trial} \\
\hline
\text{Group} & \text{Monolingual} & \text{Bilingual} & \text{Trilingual} \\
\hline
\text{N} & 27 & 44 & 27 \\
\text{Baseline Interval} & & & \\
\text{Static object} & 0.10 (0.08) & 0.09 (0.07) & 0.09 (0.07) \\
\text{Salient object} & 0.34 (0.13) & 0.38 (0.13) & 0.40 (0.14) \\
\text{Face} & 0.55 (0.15) & 0.52 (0.13) & 0.52 (0.16) \\
\text{Teaching Interval} & & & \\
\text{Static object} & 0.31 (0.10) & 0.35 (0.09) & 0.31 (0.09) \\
\text{Salient object} & 0.15 (0.08) & 0.17 (0.10) & 0.21 (0.14) \\
\text{Face} & 0.54 (0.12) & 0.48 (0.12) & 0.48 (0.16) \\
\end{array}
\]

\textit{Notes:} \( N = 98 \). An additional 9 participants (1 monolingual, 4 bilingual and 4 trilingual) were excluded from this analysis.

**Figure 2** Task performance indicated by children’s pointing behavior during testing. In the referent test trial, correct pointing was pointing toward the static object; in the mutual exclusivity test trial, correct pointing was pointing toward the salient object. Error bar = standardized error of the mean. * \( p < .05 \).
the static object in the referent test (the one that the speaker looked at while labeling during the familiarization phase) and the salient object in the ME test (ability to use mutual exclusivity principle to map a second novel label onto the object not labeled during the familiarization phase).

**Looking during test**

Looking time directed to the correct object during the test intervals was analyzed in a similar way. Repeated measures ANOVA found no significant main effect of language group on the proportional looking time toward the correct object, $F(2, 95) = 0.78, p = .46, \eta^2 = .009$. The main effect of test type and the language group x test type interaction were not significant, both $F$s < 1.26, $p$s > .29, $\eta^2$s < .012.

We also conducted repeated measures ANOVA on the proportional looking time toward the correct object on the first test trials of the referent and ME tests. There was also no significant difference between language groups in the looking time toward the target object on these first test trials, $F(2, 95) = 0.98, p = .38, \eta^2 = .010$.

One-sample $t$-tests showed that all three groups of children performed at above-chance levels for both the referent test and the ME test: referent test – monolingual: $t(26) = 3.50, p = .002, d = 0.67$, bilingual: $t(43) = 4.39, p < .001, d = 0.66$, and trilingual: $t(26) = 5.55, p < .001, d = 1.07$; ME test – monolingual: $t(26) = 4.27, p < .001, d = 0.82$, bilingual: $t(43) = 7.50, p < .001, d = 1.13$, and trilingual: $t(26) = 3.87, p < .001, d = 0.74$ (Figure 3).

**Looking during teaching**

We investigated whether this multilingual advantage in word learning during the pointing test trials might be due to the children spending a longer time looking at the target static object during the teaching intervals compared to the monolinguals. Children who were better at following the speaker’s gaze to look at the target object more during the teaching interval might also be better at mapping the novel label onto the static object during the test trials.

There was a significant positive correlation between children’s proportional gaze toward the static object during the teaching interval and their overall pointing performance during the referent test trials for the sample as a whole (Pearson’s $r = 0.23, p = .025$). There was no significant language group difference in the proportional looking time toward the static object during the teaching interval, $F(2, 95) = 1.26, p = .29, \eta^2 = .026$. Thus, children who looked longer at the same static object the speaker was looking at when she labeled it during the teaching interval also performed better when asked to explicitly point to the target (static) object in the subsequent referent test trials. However, no significant correlations were found between children’s proportional gaze toward the static object during the teaching interval and their proportional gaze toward the static object during the referent test trials, $r$s < 0.15, $p$s > .45.

All groups also looked longer than expected by chance toward the static object during the teaching interval, monolingual: $M = 0.68, SD = 0.14$, $t(26) = 6.72, p < .001, d = 1.29$, bilingual: $M = 0.69, SD = 0.15$, $t(43) = 8.32, p < .001, d = 1.27$, and trilingual: $M = 0.63, SD = 0.20$, $t(26) = 3.44, p = .002, d = 0.65$.

In summary, although all three groups of children demonstrated successful word learning in terms of both explicit pointing behavior and implicit looking time, 54-month-old children who were exposed to more than one language regularly outperformed those who were predominantly exposed to a single language in tasks requiring referential word learning abilities. However, this effect is demonstrated in a task that requires explicit pointing to the target objects but not so with the implicit looking measures.

**Discussion**

The current study is the first to compare monolingual and multilingual children’s performance in a fast
mapping task where pragmatic cues contradict object properties. This study also extends Brojde and colleagues’ (2012) work on bilingual children’s sensitivity to pragmatic cues in a generalization task, where children were required to use pragmatic cues to generalize a just learned novel word to other novel objects rather than learning a new mapping. In this study, we examined whether differences in language exposure would affect children’s use of referential cues in fast mapping during a novel word learning paradigm, especially when children are faced with conflicting information, i.e. the saliency of a distractor. Fifty-four-month-old English monolinguals, English-Mandarin bilinguals, and English-Mandarin-Language X trilinguals were shown pairs of static and salient objects. A speaker looked at the static object and provided a novel label. The same pair of objects was then shown to the children and the speaker requested either the static object (with the same novel label – referent test) or the salient object (with a new novel label – mutual exclusivity test). Results suggested that while all children were able to use the speaker’s eye gaze to learn the object-label association and further infer a new label for the unlabeled object above chance levels, bilingual and trilingual children performed significantly better than monolingual children in the referential fast mapping tasks requiring explicit pointing.

This study provided novel evidence that bilingual children with or without exposure to a third language were better at quickly associating the label with the object referred to by an adult’s eye gaze than monolingual children. The bilinguals in this study were also better able than the monolinguals in using the speaker’s referential cues to establish ME constraint in a novel-object-word mapping task. This result seems to contradict existing literature describing bilingual children as either having a more relaxed ME constraint than monolingual children (e.g. Davidson et al., 1997), or that there was no difference between monolingual and bilingual children in their propensity to honor the ME constraint (e.g. Au & Glusman, 1990; Frank & Poulin-Dubois, 2002). These studies typically present children with either a simple disambiguation task (given a novel word to choose one referent from a pair of one familiar and one novel object), or explicitly teach children a novel label on a novel object and then provide a second novel label to see if children would avoid applying both novel labels to the same object (and saying that the second novel label was from a different language). The ME test in our study is different in that children are required to first use the speaker’s gaze to establish the word mapping of a novel object (no explicit teaching), and then use this association to constrain a second novel label on this newly named object. Therefore, the ability to honor ME in our study depends critically on whether the children were successful in using the speaker’s gaze to establish the prior novel-object-word mapping.

The bilingual advantage we found in our referential fast mapping task may be a general reflection of bilingual children’s greater sensitivity to the communicative context than monolingual children. Bilingual children’s daily experience interacting with people who speak different languages demands greater attention to the speaker for cues in word learning as well as in other communicative contexts. Research comparing bilingual and monolingual children in word learning task suggests that when bilingual children are faced with an additional task of discovering what language a speaker is using, they attend more to the speaker for cues (e.g. Brojde et al., 2012; Hung et al., 2015). Thus, bilingual children may have adopted some form of self-generated, active engagement strategies to manage their more demanding everyday communicative environment, which then lead to their heightened sensitivity to a speaker’s referential cues (Yow & Markman, 2015). Our findings provide converging support to accounts that suggest children learn to tune attention efficiently to the most relevant dimensions of a task so as to achieve the most optimal learning (Smith et al., 2010).

Previous research suggests that bilingualism may contribute to better inhibitory control skills (e.g. Bialystok, 2009; Poulin-Dubois, Blaye, Coutya & Bialystok, 2011). It may be argued that the presence of the salient object might have distracted the children from attending to the speaker’s gaze toward the static object during the teaching interval, which was critical to establishing the correct referent. Bilingual children may have an advantage in such a task due to their better inhibitory control skills compared to monolingual children. However, eye gaze data revealed that children from the three language groups did not differ in their gaze following toward the static object during the teaching interval. Similarly, no significant differences were found in inhibitory control ability as measured by the flanker task between the language groups (consistent with other studies such as Fan et al., 2015; Yow & Markman, 2015). That said, previous studies that examined bilingual advantage in children’s cognitive abilities suggest that bilingual experience changes children’s cognitive processes, such as executive functions and attentional control (Bialystok, 1999; Bialystok, 2007; Bialystok, Craik, Grady, Chau, Ishii et al., 2005), problem solving (Bialystok & Majumder, 1998), cognitive flexibility (Bialystok & Shapero, 2005), and episodic and semantic memory (Kormi-Nouri, Shojaei, Moniri, Gholami, Moradi et al., 2008; Thorn, Gathercole & Frankish, 2002). As such, it
remains plausible that both bilingual children’s unique cognitive processes and their adaptation of linguistic perspectives contribute to the bilingual advantage in the referential fast mapping tasks used in our study.

It should be noted that while strong effects of language experience in referential fast mapping tasks were found based on explicit behavioral responses (i.e. children’s point to the correct objects), bilingual and trilingual children did not show greater gaze following to the speaker during the teaching phase than monolingual children. Gaze following may be a prerequisite for establishing joint attention and learning from social cues, but it may not guarantee the success of word learning in a social context. In other words, while gaze following is necessary to direct attention to a referent, it is not sufficient to learn a word successfully. A child needs to understand the referential intent behind the association between the word he/she hears and the object referred to by the speaker’s cues (Waxman & Gelman, 2010). Gliga et al. (2012) found that children who had poor social and communicative skills could also follow gaze to the correct object but did not learn the word associated with that object. Indeed, past studies found that bilingual children were better than monolinguals at understanding the referential intent of the speaker based on the speaker’s communicative cues (e.g. Yow & Markman, 2011, 2015, 2016). Furthermore, to understand a speaker’s intention, one must take the speaker’s perspective. A recent study by Fan and colleagues (2015) found that bilingual children and children who were exposed to a multilingual environment could take the perspective of a speaker when interpreting the speaker’s intended meaning, but not the monolingual children. Thus, while gaze following is required to establish joint referents, the ability to understand the referential intent of an interlocutor is essential to any successful form of learning in a social context.

Similarly, we found that bilingual and trilingual children outperformed monolingual children during the test phase in a task that requires explicit pointing but not with implicit looking measures. Some developmental research suggests that implicit looking time measure is often a more sensitive measure than reaching or pointing (e.g. Allopenna, Magnuson & Tanenhaus, 1998; Harberda, 2003; Trueswell & Gleitman, 2004). Preferential looking to a novel object after hearing a novel label is taken as a measure of success at mapping the novel label to the novel object. This is particularly important for very young children whose overt behavioral responses are not yet reliable and well developed, or whose limited vocabulary masks their understanding and expression of the tasks (e.g. theory-of-mind tasks; see Low & Perner, 2012, for a recent review). However, there are studies, similar to our study, where looking time dissociates from overt measures of knowledge, both in infancy and in early and mid-childhood (e.g. Garnham & Ruffman, 2001; Hood, Cole-Davies & Dias, 2003; Lee & Kuhlmeier, 2013; Shinskey & Munakata, 2005). For example, Brandt-Kobe and Höhle (2010) found that implicit and explicit measures of sentence comprehension can give diverging results within the same age group of 3- and 4-year-olds using identical materials. In our study, it could be reasonably assumed that the 54-month-old children are not constrained by their physical and cognitive abilities in this simple overt behavioral task. Most importantly, explicit pointing to the target objects may in fact represent a stronger representation of the learned mapping than implicit looking measures as it requires additional processing steps (Adani & Fritzschke, 2015). In other words, it is possible that a child needs to fully or almost fully represent the object–label mapping before being able to make a correct overt response.

Our study defined participants’ language groups (i.e. whether they were monolingual, bilingual, or trilingual) based on the percentage of parent-to-child language use, as reported by parents. While language exposure is important in bilingual language development (e.g. Thordardottir, 2015), language use by the child, which was not obtained in our study, is also important and may be different from that used by their caregiver (e.g. Branum-Martin, Mehta, Carlson, Francis & Goldenberg, 2014). Recent studies have shown that other factors (e.g. child’s use of both languages, interactions between two languages, cultural contexts) might contribute to the variability of bilinguals’ language skills in each language (e.g. Branum-Martin et al., 2014; Chan, Brandon & Tardif, 2009). Thus, future studies should consider obtaining such information in order to have a better understanding of the language profile of the bilingual/trilingual children.

In sum, word learning may require more than paying attention to the co-occurrence of words, social referential cues and objects (e.g. Gliga et al., 2012; Waxman & Gelman, 2009). It is important to understand the underlying communicative intent of a speaker’s cues and not just merely notice and follow communicative cues. Children in a multilingual environment regularly experience demands that require assimilation and accommodation of different linguistic perspectives, which in turn require taking the perspective of the interlocutors and understanding their referential intents. Our study suggests that multilingual children’s greater sensitivity to a speaker’s communicative intent and perspective extends to fast mapping of words in a social context. The ability to quickly associate a novel word with a novel object can be enhanced by diverse experiences in a language-learning environment.
In conclusion, our study found that children’s exposure to more than one language has positive effects on their ability to fast map word–object pairings based on referential cues. Bilingual and trilingual children were better able than monolingual children to map a novel label with the novel object that the speaker is looking at when she provided the label as well as to use the mutual exclusivity principle to associate a new label with another novel object not referred to earlier. Diversity in language experiences thus contributes to the development of children’s understanding of referential intent in the word learning process.

Acknowledgements

We wish to thank the GUSTO study participants who participated in this study and the research coordinators at the Neurodevelopment Research Centre, who helped with data collection. We also thank Elaine Tham, Jun Pei Lim, Adam Bin Abdul Malik, Ranjani D/O Nadarajan and Mumtaz Baker for their help with the task design and data extractions. This research is supported by the Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Programme and administered by the Singapore Ministry of Health’s National Medical Research Council (NMRC), Singapore- NMRC/TCR/004-NUS/2008; NMRC/TCR/012-NUHS/2014.

References


Markman, E.M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in...*


Merriman, W.E., & Kutlesic, V. (1993). Bilingual and mono-


Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E.


Singapore Department of Statistics (2015). General household


Y ow, W .Q., & Markman, E.M. (2007). Monolingual and


Y ow, W .Q., & Markman, E.M. (2010). Different kinds of concepts and different kinds of words: what words do for


Received: 30 December 2015
Accepted: 11 July 2016

© 2016 John Wiley & Sons Ltd
### Appendix A

**Novel objects and novel words used in the experiment**

<table>
<thead>
<tr>
<th>Pair</th>
<th>Static object</th>
<th>Salient object</th>
<th>Word in referent test</th>
<th>Word in ME test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td>Blicket</td>
<td>Plume</td>
</tr>
<tr>
<td>2</td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td>Dax</td>
<td>Kellar</td>
</tr>
<tr>
<td>3</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td>Neem</td>
<td>Gombie</td>
</tr>
<tr>
<td>4</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td>Zoop</td>
<td>Timbo</td>
</tr>
<tr>
<td>5</td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td>Sefo</td>
<td>Moxie</td>
</tr>
<tr>
<td>6</td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td>Toma</td>
<td>Togo</td>
</tr>
<tr>
<td>7</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td>Kobe</td>
<td>Fimmet</td>
</tr>
<tr>
<td>Pair</td>
<td>Static object</td>
<td>Salient object</td>
<td>Word in referent test</td>
<td>Word in ME test</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>----------------</td>
<td>----------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>8</td>
<td><img src="image1.png" alt="Static object" /></td>
<td><img src="image2.png" alt="Salient object" /></td>
<td>Teri</td>
<td>Wug</td>
</tr>
</tbody>
</table>

© 2016 John Wiley & Sons Ltd