Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants

Krista Byers-Heinlein a,⁎, Janet F. Werker b

a Concordia University, Department of Psychology, 7141 Sherbrooke St. West, Montreal, QC, Canada H4B 1R6
b University of British Columbia, Department of Psychology, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4

A R T I C L E   I N F O
Article history:
Received 11 July 2012
Revised 22 March 2013
Accepted 18 May 2013
Available online 15 June 2013

Keywords:
Infant development
Mutual exclusivity
Disambiguation
Bilingualism
Word learning

A B S T R A C T
In ambiguous word learning situations, infants can use systematic strategies to determine the referent of a novel word. One such heuristic is disambiguation. By age 16–18 months, monolinguals infer that a novel noun refers to a novel object rather than a familiar one (Halberda, 2003), while at the same age bilinguals and trilinguals do not reliably show disambiguation (Byers-Heinlein & Werker, 2009; Houston-Price, Caloghiris, & Raviglione, 2010). It has been hypothesized that these results reflect a unique aspect of the bilingual lexicon: bilinguals often know many translation equivalents, cross-language synonyms such as English dog and Mandarin gǒu. We studied the role of vocabulary knowledge in the development of disambiguation by relating 17–18 month-old English–Chinese bilingual infants’ performance on a disambiguation task to the percentage of translation equivalents in their comprehension vocabularies. Those bilingual infants who understood translation equivalents for more than half the words in their vocabularies did not show disambiguation, while infants who knew a smaller proportion of translation equivalents showed disambiguation just as same-aged monolinguals do. These results demonstrate that the structure of the developing lexicon plays a key role in infants’ use of disambiguation.

Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.

1. Introduction
Young children are highly skilled word learners. They are able to take advantage of a wealth of information to figure out a word’s meaning, including associative (Smith & Yu, 2008), social (Baldwin & Moses, 2001), and referential (Waxman & Gelman, 2009) information. Children are also able to infer a word’s meaning even in contexts where information is very sparse. In a disambiguation task (sometimes referred to as a mutual exclusivity task; Markman & Wachtel, 1988), children are presented with a pair of objects, one familiar with a known name (such as a cup) and another unfamiliar without a known name (such as a spatula). Although they have no prior knowledge of the word spatula, children as young as 16–18 months tend to infer that the novel word refers to the long-handled utensil, rather than the round drinking vessel (Byers-Heinlein & Werker, 2009; Halberda, 2003; Markman, Wasow, & Hansen, 2003). This fast-mapping of words onto referents via disambiguation is thought to be a first step towards the long-term retention of the meanings of novel words (Bion, Borovsky, & Fernald, 2013; Carey & Bartlett, 1978; Horst, McMurray, & Samuelson, 2006; Horst & Samuelson, 2008; Swingley, 2010).

What is the origin of infants’ ability to infer the referent of a novel word via disambiguation? One possibility is that one-to-one mappings are inherently preferred by the child’s learning system. For example, under Markman & Wachtel’s (1988) mutual exclusivity account, children operate under the default assumption that object labels denote mutually exclusive categories, and use this assumption to infer that a novel label could not go with
an object that already has a label. It has also been suggested that mutual exclusivity is an emergent property of computational processes that support word learning (Frank, Goodman, & Tenenbaum, 2009; McMurray, Horst, & Samuelson, 2012; Merriman, 1999; Regier, 1996, 2003), or that it might be founded in children's preference for novelty (Horst, Samuelson, Kucker, & McMurray, 2011).

An alternative view is that the development of disambiguation is driven by experience, and that it emerges only once a child has established that each object should have a basic-level label (Mervis & Bertrand, 1994), or has ascertained that adults use different words to refer to different kinds of objects (Diesendruck & Markson, 2001).

There is mounting evidence that early language experience does play a role in the development of disambiguation. Key support for this view is that bilingual and trilingual infants do not show disambiguation from the same age as monolinguals (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010), and preschool and school-age bilinguals use disambiguation less robustly than monolinguals (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Davidson, Jergovic, Imami, & Theodos, 1997; Davidson & Tell, 2005; Diesendruck, 2005; but see also Byers-Heinlein, Chen, & Xu, in preparation; Frank & Poulin-Dubois, 2002; Merriman & Kutlesic, 1993 for studies that showed similar performance by monolinguals and bilinguals). These group differences amongst children from varying language backgrounds indicate that early language experience can influence the development of disambiguation. However, because studies have yet to investigate exactly what aspect of monolinguals' and bilinguals' experience contributes to observed differences, the specific role that experience plays remains unclear.

One explanation for the difference between monolinguals, and bilinguals' development of disambiguation (e.g. Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010; Merriman & Kutlesic, 1993) is that bilinguals regularly learn translation equivalents, pairs of words that are synonyms across the two languages such as English dog and Mandarin góu. Bilinguals' knowledge of translation equivalents results in infants' lexicons having a many-to-one (or at least two-to-one) mapping structure between words and concepts. In contrast, each new word that a monolingual child learns is usually another one-to-one mapping between a word and a concept. The notion that children's knowledge of one-to-one mappings between words and concepts contributes to the early development of disambiguation in both monolingual and bilingual infants has been called the lexicon structure hypothesis (Byers-Heinlein & Werker, 2009; see also McMurray et al., 2012, for a discussion of the impact of a one-to-many lexicon structure on disambiguation).

The lexicon structure hypothesis predicts that individual children's use of disambiguation will vary according to the extent to which their lexicons have a one-to-one versus a many-to-one mapping structure. As one-to-one mapping between words and concepts is the typical structure of a single language (Clark, 1987, 1990, 1993), monolingual children's use of disambiguation should become more robust with age as their vocabularies grow. Indeed, several studies have found that monolingual children's use of disambiguation increases with age (Bion et al., 2013; Halberda, 2003) and other studies report positive correlations between monolinguals' vocabulary size and their use of disambiguation or related behaviors (Bion et al., 2013; Graham, Poulin-Dubois, & Baker, 1998; Houston-Price et al., 2010; Mervis & Bertrand, 1994). However, for monolingual children, the extent to which their lexicons are characterized by one-to-one mappings is conflated with age and overall vocabulary size, making such results a weak test of the lexicon structure hypothesis.

Studies of individual differences amongst children's use of disambiguation can disentangle the effects of age and vocabulary size from the extent to which children's vocabularies have a one-to-one mapping structure. When a bilingual learns a new word, it might be for a previously unlabeled concept and thus be consistent with a one-to-one lexicon structure, or it might be a translation equivalent of an already known word, and thus contribute to a many-to-one lexicon structure. Bilingual children vary widely in the proportion of words for which they know a translation equivalent (David & Wei, 2008; De Houwer, Bornstein, & De Coster, 2006; Pearson, Fernández, & Oller, 1995; Sheng, Lu, & Kan, 2011; Umbel, Pearson, Fernández, & Oller, 1992). However, bilingual children do not seem to systematically learn or avoid translation equivalents, as the proportion of translation equivalents known by bilingual children is equivalent to what would be expected by chance (i.e. the average overlap found in the single-language lexicons of two randomly selected children; Pearson et al., 1995).

Because bilingual children's knowledge of translation equivalents is both highly variable and unsystematic, bilingual children provide a natural experiment for investigating the role of a one-to-one lexicon structure in the development of disambiguation. The lexicon structure hypothesis predicts that individual bilingual children who know many translation equivalents will not show disambiguation at the same age as those individual bilingual children who know few translation equivalents. Further, it predicts that bilingual children who know few translation equivalents will show disambiguation from the same age as monolinguals.

Only two studies to date have reported data that pertain to the relationship between bilinguals' use of disambiguation and their knowledge of translation equivalents, with somewhat equivocal results. Houston-Price et al. (2010) reported a non-significant negative correlation between the number of translation equivalents bilinguals knew and their performance on a disambiguation task ($r(19) = -0.29, p = .23$). Frank and Poulin-Dubois (2002) tested children on a task related to disambiguation, their willingness to learn two labels for the same object. They also found a non-significant trend for the children who knew the fewest translation equivalents to be the most likely to avoid mapping two labels to the same object (27-month-olds: $r(24) = -0.27, p = .18$; 35-month-olds: $r(26) = -0.20, p = .30$). While the correlations reported in these studies were in the predicted direction, there is thus far no statistically robust evidence that bilinguals' knowledge of translation equivalents influences their use of disambiguation.
The current study sought to examine whether the development of disambiguation is influenced by children’s experience with language as a one-to-one versus a many-to-one mapping system by investigating the relationship between bilinguals’ knowledge of translation equivalents and their use of disambiguation. We studied Chinese–English bilingual infants’ disambiguation in a preferential looking task, in which infants heard a novel word and their looking at a novel picture over a familiar distracter was measured (see Byers-Heinlein & Werker, 2009; Halberda, 2003; Houston-Price et al., 2010 for studies using a similar method). To provide a more sensitive test than previous investigations, we studied bilingual infants at the lower limit of the age when monolingual infants reliably show disambiguation in this paradigm, 17–18 months.

We predicted that if experience with one-to-one mappings between words and their referents contributes to the development of disambiguation then 17–18 month-old bilinguals should show systematic variation in their use of disambiguation as a function of their knowledge of translation equivalents. We predicted that bilingual infants whose lexicons contained a large proportion of translation equivalents would not show disambiguation, replicating previous results that averaged across bilingual infants (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010). Importantly, we also predicted the novel result that those bilingual infants whose lexicons contained only a small proportion of translation equivalents would disambiguate a novel word, just like same-aged monolinguals. Finally we predicted specificity in this relationship, that experience with language as a one-to-one versus a many-to-one mapping system by investigating the relationship between bilinguals’ knowledge of translation equivalents and their use of disambiguation. We studied Chinese–English bilingual infants’ disambiguation in a preferential looking task, in which infants heard a novel word and their looking at a novel picture over a familiar distracter was measured (see Byers-Heinlein & Werker, 2009; Halberda, 2003; Houston-Price et al., 2010 for studies using a similar method). To provide a more sensitive test than previous investigations, we studied bilingual infants at the lower limit of the age when monolingual infants reliably show disambiguation in this paradigm, 17–18 months.

2. Methods

2.1. Participants

A total of 20 (12 female) bilingual infants learning English and Chinese (either Cantonese or Mandarin) were included in the study, with a mean age of 17m27d (range: 17m17d–18m12d). Eight additional infants were tested but excluded from the analyses because the infant was too restless or inattentive to complete the study (4), because the infant had a major health concern (2), or because the infant was not reported to understand any of the familiar English words used in the study (2). Infants were primarily recruited by approaching new parents at a maternity hospital in Vancouver, Canada, at the time of the infant’s birth.

2.1.1. Language background

All infants came from homes where English and Chinese had been spoken regularly since the infant was born. Fifteen bilinguals were hearing English and Cantonese, and five were hearing English and Mandarin. Most parents were ethnically Chinese, and typically one or both parents was Canadian-born or had been born elsewhere but primarily educated in English. The Language Exposure Questionnaire was used to assess how often infants heard each language (Bosch & Sebastián-Gallés, 1997). On average, infants heard English 49% of the time (range: 27–75%) and Chinese 49% of the time (range: 25–73%). One infant heard a small amount of Korean (7%). To gauge infants’ relative exposure to each language, a balance score was calculated for each infant as 50 minus the difference in exposure between the infants’ two languages (e.g. infants with perfectly balanced exposure of 50/50 had a score of 50–50 = 0, while those with the largest imbalance of 75/25 had a score of 75–25 = 50). The mean balance score was 21 (SD = 12; range: 0–47), meaning that the average participant heard their dominant language 60.5% of the time and the other language 39.5% of the time.

2.2. Materials

2.2.1. Vocabulary and translation equivalents measures

Estimates of infants’ English vocabulary size were obtained by asking parents to complete the Words and Gestures form of the MacArthur–Bates Communicative Development Inventory (CDI; Fenson et al., 2007) which has shown high validity for bilinguals (Marchman & Martinez-Sussman, 2002). Estimates of infants’ Chinese vocabularies were obtained via the Mandarin or Cantonese adaptation of the form as appropriate (Tardif & Fletcher, 2008). The total number of items across the forms was similar: 396 English, 410 Mandarin, and 388 Cantonese. The use of the Words and Gestures form allowed measurement of both infants’ comprehension and their production vocabularies. One parent per infant completed the form in each language. Where possible, this was the parent who was most familiar with the infant’s vocabulary in that language, thus infants varied as to whether the same parent or different parents completed the two CDIs. In four cases, parents failed to return complete CDI forms for one or both languages. These infants were excluded from the analyses that required CDI data, thus reducing the sample size to 16 infants in these analyses.

A trilingual research assistant who was a native Cantonese–Mandarin bilingual and had moved to an English-speaking country in childhood, identified word pairs in the English and Chinese CDIs that were translation equivalents (e.g. English dog and Mandarin gūi/Cantonese gūi). She verified all English–Cantonese pairings with an English–Cantonese bilingual, and all English–Mandarin pairings with an English–Mandarin bilingual. In some cases, English and Chinese lexicalize the same concept differently. For example, whereas English has the single word brother, Mandarin has separate words gēge for older brother and didi for younger brother. Thus, if the child knew the English word brother and the Mandarin words gēge and didi, these three words counted as two pairs because they
lexicalize two different concepts. A total of 297 English–Cantonese pairs and 294 English–Mandarin pairs were identified.

2.2.2. Stimuli

Visual stimuli consisted of four brightly-colored objects, three familiar (ball, car, and shoe) and one novel. The novel object was a slightly modified version of a phototube from the TarrLab Object DataBank (1996). The objects were presented on a black background in yoked pairs: car–ball and phototube–shoe. The yoked presentation of the objects followed the procedure used by several previous studies of disambiguation in infants of this age (Byers-Heinlein & Werker, 2009; Halberda, 2003). The objects appeared in different colors on different trials to maintain infant interest, and to ensure that infants generalized across different-colored exemplars of the same category.

All auditory stimuli were recorded in English by a female native speaker. Conventional English labels were recorded for the familiar objects – ball, car, and shoe. The nonce word nil was recorded as the novel label for the phototube object because it is not a possible Mandarin or Cantonese word, and although it is a real word for English-speaking adults, it is unlikely that infants would be familiar with this word. On each trial, the label was presented in a carrier phrase and then repeated in isolation (e.g. “Look at the ball! Ball!”). Three different carrier phrases were used for each word, presented quasi-randomly on different trials. “Look at the ___”, “Find the ___”, and “Where is the ___”.

We examined infants’ reported comprehension of the target words on the CDI to ensure that they were indeed familiar. Comprehension of ball, car, and shoe ranged from 80% to 100%. Although infants were never tested using Chinese stimuli, we also examined their understanding of the Chinese translations of these words. Reported comprehension of the translations ranged from 94% to 100% of infants. Therefore, the vast majority of infants were reported to understand the test words in both English (the language of testing) and in Chinese.

2.2.3. Apparatus

Data were collected using a Tobii 1750 eye tracking system, consisting of a 17” LCD monitor for the presentation of visual stimuli with a built-in eye tracking camera that sampled at a rate of 50 Hz using a corneal reflection technique. Auditory stimuli were presented via computer speakers located behind a black cardboard panel, on either side of the eye tracker. A PC computer running the Tobii Clearview software program both controlled the stimulus presentation and collected the eye tracking data.

2.3. Procedure

The experimental procedure followed Halberda (2003), and was identical to Byers-Heinlein and Werker (2009). Infants were tested in a dimly lit, sound-attenuated room. The experimenter controlled the study from a computer and a closed-circuit TV monitor, located in a screened-off area of the same room. Before commencing the study, the experimenter settled the infant on the parent’s lap, and positioned the eye tracking monitor approximately 60 cm away from the infant’s eyes. Parents wore a blindfold or closed their eyes for the duration of the study to avoid influencing the infant. The experimenter first carried out a built-in five-point infant calibration routine, and then began the experimental procedure.

At the beginning of testing, each infant saw a warm-up trial during which a spinning waterwheel appeared sequentially on each side of the monitor. Test trials began immediately following the warm-up. On each trial, infants’ baseline preference for the object pair was measured in silence for 3 s. Next, an auditory stimulus was played that named one of the objects (e.g. “Look at the ball! Ball!”). The objects then remained in silence on the monitor, yielding a total duration of 9.5 s. After the test phase of the trial was completed, the unlabeled object disappeared, while the labeled object moved around on the monitor for 2 s with accompanying music. Previous studies of word comprehension have suggested that such visual feedback is effective to maintain infant engagement in the task (Killing & Bishop, 2008). The results of past studies using the same paradigm have found no evidence that this reinforcement drives infants’ performance on the task (Byers-Heinlein & Werker, 2009; Halberda, 2003).

Infants were presented with 24 test trials, in four blocks of six trials per block. The first and third blocks consisted of known versus known trials (ball–car), while the second and fourth blocks consisted of known versus novel trials (shoe–nil). Each object was labeled on half of the trials in which it appeared, thus a total of six times. Trials where the familiar object was named were called familiar label trials and trials where the novel object was named were called disambiguation trials. Eight stimulus orders were created to counterbalance side and order of presentation across infants, however, the configuration of objects (e.g. ball on left, car on right) was constant for each infant. A central circular attention-getter was presented between trials to reorient infants to the screen. The total duration of the study was approximately 7 min.

Infant eye-gaze data were collected by the eye tracker, and each data point was classified as a look towards the left side object, a look towards the right side object, or failure to look towards either object. Parents completed the questionnaires following the experimental session.

3. Results

3.1. Vocabulary measures

Detailed descriptive statistics of infants’ comprehension and production vocabularies in both English and Chinese are presented in Table 1. English and Chinese vocabularies were calculated by counting the number of words the child was reported to understand and say on the corresponding forms. Total vocabulary was calculated by summing the words the child knew in English and in Chinese. The percentage of infants’ vocabularies that constituted translation equivalents was calculated by taking the total number of words for which the infant knew a translation equivalent (each word in a pair such as dog and gǒu was
counted separately), and dividing by the total vocabulary size (see also De Houwer et al., 2006; Pearson et al., 1995 for alternate methods of quantifying translation equivalents). All infants had at least some translation equivalents in their comprehension vocabularies, and all but two had translation equivalents in their productive vocabularies. The total conceptual vocabulary was calculated as the total vocabulary minus the number of redundant words whose meaning was captured by its translation equivalent, providing an index of the number of meanings lexicalized by the infant (Pearson, Fernández, & Oller, 1993). Thus if the infant knew both dog and its translation gōu, this would count for only one concept.

Relationships amongst vocabulary measures were examined. There was a significant correlation between English comprehension and English production, \( r(14) = .69, p = .003 \), but not between Chinese comprehension and Chinese production, \( r(14) = .10, p = .70 \). Children were reported to understand on average 54 (SD = 95) more words in Chinese than in English, which was a statistically significant difference \( t(15) = 2.249, p = .04, d = .56 \). Infants were reported to produce 10 (SD = 29) more words in English than in Chinese, but this difference was not significant, \( t(14) = 1.30, p = .21, d = .33 \). No significant correlation was found between comprehension and production for infants’ total vocabularies, \( r(14) = .25, p = .36 \), or for their total conceptual vocabularies, \( r(14) = .27, p = .32 \). Further, there was no significant correlation between the number of translation equivalents infants were reported to understand and the number they could produce, \( r(14) = .21, p = .44 \).

### 3.2. Experimental task

As in previous studies (e.g., Byers-Heinlein & Werker, 2009), an analysis window of 360–2000 ms after the onset of the target word was identified based on the minimum time needed to process a word and initiate an eye movement (Dahan, Swingley, Tanenhaus, & Magnuson, 2000), and the typical length of time that infants sustain fixation after responding to a word (Dahan et al., 2000; Fernald, Perfors, & Marchman, 2006; Swingley, Pinto, & Fernald, 1999). Infants’ responses on familiar label and disambiguation trials were analyzed only in this time window. Any trial where the infant looked 750 ms or less at the objects was excluded from the analyses.

An individual baseline score was calculated for each infant, as the proportion of time the infant looked at a particular object during the 3 s silent baseline period, averaged across all trials in which that object was onscreen. Trials during which the infant looked less than 1 out of the 3 s were excluded from the calculation. Infants tended to show more interest during baseline to the familiar objects as opposed to the novel object, \( t(19) = 3.00, p = .007, d = .67 \), a pattern reported in numerous previous studies that have used similar paradigms (e.g., Byers-Heinlein & Werker, 2009; Mather & Plunkett, 2009; Schafer, Plunkett, & Harris, 1999; White & Morgan, 2008). Thus, to control for inherent baseline preferences as in previous work (e.g., Byers-Heinlein & Werker, 2009; Halberda, 2003; White & Morgan, 2008), all subsequent analyses were conducted with difference scores, which subtracted infants’ individual baseline preference from the proportion of time they looked at the target object after labeling. A positive difference score therefore indicated increased looking at the target object after labeling.

Familiar label trials were analyzed first. A preliminary between-subjects ANOVA showed that infants’ performance on familiar label trials did not vary as a function of object, \( F(2,38) = 1.86, p = .17, \eta^2_p = .089 \), therefore data were collapsed across the three familiar objects. A one-tailed \( t \)-test indicated that infants significantly increased their looking towards the familiar target upon hearing its label, \( M = .072, SD = .15, t(19) = 2.15, p = .023, d = .48 \). Next, a one-tailed analysis was performed on disambiguation trials to examine whether, as a group, infants oriented to the novel object upon hearing the novel label. Replicating previous results, as a group, Chinese–English bilinguals did not show disambiguation, as there was no significant increase in looking at the novel object upon hearing the novel label, \( M = .060, SD = .30, t(19) = .89, p = .19, d = .20 \). Finally, a correlation was computed between infants’ performance on familiar label and disambiguation trials. No significant correlation was found, \( r(18) = -.25, p = .29 \).

### 3.3. Relationship of performance to knowledge of translation equivalents

The main hypothesis was that infants’ use of disambiguation would be related to their knowledge of translation equivalents. We focused on comprehension of translation equivalents, rather than production of translation equivalents, because comprehension indexes all of the words for which infants have a lexical representation. Three predictions were tested: 1) infants who knew the smallest proportion of translation equivalents would show the best performance on disambiguation trials, 2) knowledge of translation equivalents would not be related to performance on familiar label trials, and 3) performance on novel...
label trials would not be related to any other individual variable (e.g. vocabulary size, percent exposure to English).

Pearson correlations were computed between infants’ performance on both disambiguation and familiar label trials, and all individual variables related to vocabulary size and language exposure (see Table 2 for the correlations discussed below; a complete table of correlations between these variables is available in the online Supplementary material). Correlations were tested for statistical significance both using the traditional parametric technique for computing $p$-values, and also using an empirical bootstrapping technique for computing 95% confidence intervals. The latter were computed using the boot function in the R statistical package (Canty & Ripley, 2011; Davison & Hinkley, 1997) with 1000 bootstrap replicates per correlation. As predicted, a significant negative correlation indicated that bilingual infants who knew a smaller percentage of translation equivalents showed significantly better performance on disambiguation trials, $r(14) = -.55$, $p = .026$. Further, as predicted, there was no significant correlation between infants’ knowledge of translation equivalents and their performance on familiar label trials, $r(14) = .24$, $p = .37$. Their performance on disambiguation trials was also not significantly correlated with any other measure, including vocabulary size in English or Chinese, total vocabulary size, total conceptual vocabulary size, percent of exposure to each language, or balance of exposure to each language. Bootstrapped confidence intervals are reported in Table 2, and they were consistent with the parametric analysis. Namely, those correlations that had a parametrically-estimated $p$-value less than .05 had an empirically-estimated confidence interval that did not contain zero, while those correlations that had parametrically-estimated $p$-values greater than .05 had empirically-estimated confidence intervals that did contain zero.

As a further test of the main hypothesis, data were analyzed following a median split of the proportion of translation equivalents known by each infant. A low-overlap group of eight infants was reported to know translation equivalents known by each infant. A low-overlap group of eight infants was reported to know translation equivalents for less than 50% of their vocabularies ($M = 34\%$), and a high-overlap group of eight infants knew translation equivalents for 50% or more of their vocabularies ($M = 58\%$). As predicted, on disambiguation trials, the low-overlap group was significantly more likely to disambiguate the novel label than the high-overlap group, $t(14) = 3.52$, $p = .003$, $d = .18$, one-tailed (these data are plotted in Fig. 1). Infants in the low-overlap group successfully disambiguated the novel label by increasing looking at the novel object (higher proportion) suggests use of disambiguation. T-tests are one-tailed. Error bars represent the standard error of the mean.

### Table 2

<table>
<thead>
<tr>
<th>Familiar label trial performance</th>
<th>Disambiguation trial performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson’s $r$</td>
<td>Bootstrapped 95% CI</td>
</tr>
<tr>
<td>% Translation equivalents</td>
<td>.24 [−.46, .66]</td>
</tr>
<tr>
<td>English comprehension</td>
<td>.36 [−.31, .76]</td>
</tr>
<tr>
<td>Chinese comprehension</td>
<td>.11 [−.31, .48]</td>
</tr>
<tr>
<td>Total comprehension</td>
<td>.31 [−.14, .67]</td>
</tr>
<tr>
<td>Total conceptual comprehension</td>
<td>.32 [−.10, .66]</td>
</tr>
<tr>
<td>% English exposure</td>
<td>.33 [−.02, .64]</td>
</tr>
<tr>
<td>Balance score</td>
<td>.24 [−.30, .70]</td>
</tr>
</tbody>
</table>

Note: Bootstrapped 95% confidence intervals showed the same results as parametrically-derived $p$-values.

$^* p < .05$.

Fig. 1. Bilinguals’ average performance on disambiguation label trials, and as a function of overlap group. Longer looking to the novel object (higher proportion) suggests use of disambiguation. T-tests are one-tailed. Error bars represent the standard error of the mean.

4. Discussion

The current study investigated how early language knowledge influences the development of the disambiguation word learning heuristic, infants’ tendency to infer that

---

1 The exact value of the median score was 47%. However, as there were no infants with scores between 47% and 50%, 50% was used as the nominal cutoff.
a novel noun refers to a novel rather than to a familiar object. We predicted that bilingual infants’ use of disambiguation would be related to the extent to which their lexicons had a one-to-one versus a many-to-one mapping structure as operationalized by their knowledge of translation equivalents. Seventeen and 18-month-old Chinese–English bilingual infants participated in a preferential looking study conducted in English. Familiar label control trials asked infants to find the referent of a known noun (e.g. “Look at the ball!”), and infants were successful on these trials. Disambiguation trials asked infants to find the referent of a novel noun (e.g. “Look at the nil!”) when presented with a novel object and a familiar distracter. As a group, bilingual infants did not show a significant tendency to disambiguate the novel noun, replicating previous results (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010). However, as predicted, bilingual infants’ performance was not homogeneous. Those infants whose vocabularies contained a smaller proportion of translation equivalents (<50%; low-overlap group whose lexicons had primarily a one-to-one structure) disambiguated the referent of the novel noun, while those whose vocabularies contained a larger proportion of translation equivalents (>50%; high-overlap group whose lexicons had primarily a many-to-one structure) did not. This pattern of difference was robust across numerous analyses, including correlation, median-split, and chi-squared tests. These results support the position that infants’ knowledge of one-to-one mappings between words and their referents contributes to their use of disambiguation.

As discussed in the Introduction, several previous studies did not find a significant correlation between infants’ knowledge of translation equivalents and their performance in disambiguation and related tasks (i.e. Frank & Poulin-Dubois, 2002; Houston-Price et al., 2010). However, all three correlations reported in these previous studies were in the predicted negative direction: children who knew more translation equivalents showed less disambiguation or related behaviors. Thus, from an effects size perspective, the current results are largely consistent with previous work, albeit with a larger effect. One reason our results reached statistical significance and those from the previous studies did not, could be that our infants were younger on average than those in the previous studies. As discussed later in this paper, by older ages bilingual children do eventually develop disambiguation. Testing infants in a narrow age window during which monolinguals are known to develop disambiguation could have increased our power to detect an effect.

It is important for the current interpretation that the data showed specificity in the relationship between infants’ knowledge of translation equivalents and their use of disambiguation. We found no significant correlation between disambiguation and any other language measure including amount of exposure to each language, balance between exposure between the two languages, total vocabulary size, total conceptual vocabulary size, and vocabulary size in each language. Further, infants’ ability to perform a preferential looking task did not account for the main finding, as infants successfully looked to the familiar object in response to its label, and performance on disambiguation trials was not correlated with performance on familiar label trials. Performance on familiar label trials also did not vary as a function of the proportion of translation equivalents infants knew, suggesting that knowledge of translation equivalents affected performance on disambiguation trials specifically, and not on the task in general. Thus, while it is impossible to entirely rule out a third-variable explanation, particularly given our sample size, we do not find support for such an explanation in the current data.

Another alternate explanation is that disambiguation and knowledge of translation equivalents are correlated because bilinguals who adhere strongly to disambiguation avoid learning translation equivalents, while those who have not developed the disambiguation heuristic do not. However, this explanation does not predict the overall differences between monolingual and bilingual infants’ use of disambiguation at this age (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010). If disambiguation develops independently of language experience, one would expect just as much individual variation amongst monolinguals as bilinguals, and thus no group differences. The most parsimonious explanation for both results is that bilingual infants’ knowledge of translation equivalents underlies both group differences between monolingual and bilingual infants’ use of disambiguation, as well as individual differences amongst bilinguals.

Having ruled out several alternate explanations, we interpret these findings as support for the lexicon structure hypothesis (Byers-Heinlein & Werker, 2009), that the development of disambiguation is supported by infants’ learning of one-to-one mappings between words and concepts. For monolinguals, knowledge of one-to-one mappings increases as their vocabularies grow, explaining why disambiguation and vocabulary size are correlated in monolingual infants (Bion et al., 2013; Graham et al., 1998; Houston-Price et al., 2010; Mervis & Bertrand, 1994). Bilinguals regularly learn translation equivalents, which result in a vocabulary that is characterized by many-to-one mappings between words and concepts. Our finding that bilinguals do not show disambiguation from the same age as monolinguals unless they happen to have learned relatively few translation equivalents is consistent with this interpretation. Our results clarify previous findings with respect to bilinguals’ use of disambiguation. We have demonstrated that the development of disambiguation is not affected by the learning of two language systems per se, but by the structure of individual infants’ lexicons as they learn words and build their vocabularies (see also McMurray et al., 2012, for a computational model congruent with the current empirical results).

A possibility that cannot be ruled out by the current data is that experience instead influences whether infants use an already-available heuristic (e.g. Markman, 1992). The current study tested infants at ages 17–18 months as this is the youngest age at which disambiguation has been reliably shown by monolinguals. If disambiguation does indeed first develop at this age, then demonstrated differences between monolinguals and bilinguals support the position that the learning of one-to-one mappings is important during the initial development of disambigua-
tion. Yet, there are some hints of early precursors to disambiguation in children who are less than a year old (Dewar & Xu, 2007; Mather & Plunkett, 2010; Xu, Cote, & Baker, 2005). At the same time, these results have been difficult to reconcile with other studies in which 1- and 1.5-year-olds are unsuccessful in disambiguation tasks (Bion et al., 2013; Halberda, 2003). If it is eventually confirmed that infants younger than 17–18 months do show disambiguation, it would be essential to compare monolingual and bilingual infants younger than 17–18 months to investigate whether these groups also differ at earlier ages.

The lexicon structure hypothesis is not the first developmental account of disambiguation to emphasize the role of early word learning. In their novel-name–nameless-category (N3C) account, Mervis and Bertrand (1994) put forward the notion that disambiguation develops when children have the insight that each object has a name. Under N3C, children’s seek out a “novel name” when they see an object that exemplifies a “nameless category”. However, N3C is not consistent with the current results. The N3C account predicts that just like monolinguals, bilinguals should develop disambiguation as their vocabularies grow, because encountering multiple labels for each object does not negate the notion that each object should have a name. Thus, N3C would predict individual differences amongst bilinguals as a function of their vocabulary size, but not as a function of their knowledge of translation equivalents.

4.1. Disambiguation beyond infancy

While this and other studies (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010) have demonstrated that some bilinguals do not use disambiguation in infancy, there is considerable evidence that older bilinguals do use disambiguation and related word learning heuristics, although in some cases less consistently than same-aged monolinguals do (Byers-Heinlein et al., in preparation; Davidson & Tell, 2005; Davidson et al., 1997; Diesendruck, 2005; Frank & Poulin-Dubois, 2002; Merriman & Kutlesic, 1993). Further, there is some evidence that early differences between monolingual and bilingual children’s use of disambiguation attenuate across development (Bialystok et al., 2010). How might we reconcile these findings with older bilingual children with our claim that the structure of the lexicon influences the development of disambiguation?

Monolingual children’s disambiguation abilities become increasingly sophisticated over development. Thirty-month-old monolinguals, but not younger children, are able to retain mappings between words and objects established through disambiguation (Bion et al., 2013; Horst & Samuelson, 2008). By age 3 children can disambiguate not only the meaning of a novel noun, but also the meaning of a novel verb (Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996). Further, preschool children are able to integrate multiple sources of information in their use of disambiguation, for example taking into account information about the class of a word (i.e. count nouns versus proper nouns; Hall, Quantz, & Persoage, 2000). Preschoolers are also able to consider pragmatic information, such as a speaker’s knowledge or ignorance, in disambiguating a novel noun (Diesendruck, 2005; Diesendruck, Hall, & Graham, 2006; Diesendruck & Markson, 2001).

It may be that this emerging ability to integrate information from multiple sources eventually allows bilingual children to notice that words and concepts have a one-to-one structure in each language. And indeed, preschool and school-aged bilinguals tend to associate a novel noun with a novel object only when the distracter object has a label in that same language (within-language disambiguation), but not when two labels are in different languages (between-language disambiguation; Au & Glusman, 1990). As the lexicon continues to develop (see Curtin, Byers-Heinlein, & Werker, 2011, for a recent discussion of the organization of the bilingual lexicon), an explicit understanding that they are learning two languages emerges in bilingual children (see Genesee, Boivin, & Nicoladis, 1996; Genesee, Nicoladis, & Paradis, 1995, for evidence of pragmatic differentiation of their two languages by bilingual toddlers). At this point, the structure of their dual lexicon might support the use of disambiguation appropriately within a language, and may even help bilingual children actively seek translation equivalents for words they already know and to learn new words (see Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013, for work relating knowledge of translation equivalents to speed of lexical access in bilingual toddlers).

4.2. Conclusions and future directions

Our results provide convincing support for the hypothesis that the structure of the emerging lexicon in infancy influences the use of disambiguation. Left unanswered is whether experience learning words with one-to-one mappings is sufficient for the development of a disambiguation heuristic (as experience is claimed to be for the development of a shape bias in early vocabulary, e.g. Perry & Samuelson, 2011), or whether experience works in tandem with a more general cognitive bias that might be available to both humans and animals (e.g. MacWhinney, 1989; Markman, 1992). Our results have also raised questions about how metalinguistic knowledge influences children’s use of disambiguation, for example bilinguals’ eventual ability to use disambiguation appropriately in each language. Subsequent research investigating learners with different types of experience, and at younger and older ages, will be helpful in pursuing these exciting questions.

Acknowledgements

We thank Marisa Cruikshank, Vashti Garcia, Vivian Pan, Jessica Deglau, and Jasmine Cady for their assistance with data collection, Jacqueline Chong for her work on translation equivalents, Alexandra Polonia for her assistance with manuscript preparation, and Padmapriya Kandhadai for her comments on an earlier draft of this work. We are grateful to the parents and infants who participated in this study. This research was supported by doctoral fellowships from the National Sciences and Engineering Research Council of Canada (NSERC), the O’Brien Foundation, the Killam Trust, and the Canadian Federation of University Women to KBH, and by Grants from the Social Sciences and
Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2013.05.010.

References


