Developmental Psychology
2015, Vol. 51, No. 10, 1501–1515

0012-1649/15/$12.00 http://dx.doi.org/10.1037/a0039474

English- and Mandarin-Learning Infants' Discrimination of Actions and Objects in Dynamic Events

Jie Chen and Twila Tardif
University of Michigan

Rachel Pulverman
Delaware State University

Marianella Casasola
Cornell University

Liqi Zhu
Chinese Academy of Sciences

Xiaobei Zheng and Xiangzhi Meng
Peking University

The present studies examined the role of linguistic experience in directing English and Mandarin learners' attention to aspects of a visual scene. Specifically, they asked whether young language learners in these 2 cultures attend to differential aspects of a word-learning situation. Two groups of English and Mandarin learners, 6–8-month-olds (n = 65) and 17–19-month-olds (n = 91), participated in 2 studies, based on a habituation paradigm, designed to test infants' discrimination between actions and objects in dynamic events. In Study 1, these stimuli were presented in silence, whereas in Study 2, a verbal label accompanied videos. Results showed that 6–8-month-olds could discriminate action changes but not object changes, whereas 17–19-month-olds could discriminate both types of changes. However, there were only very subtle cross-linguistic differences in these patterns when the scenes were presented together with a verbal label. These findings show strong evidence for universal developmental trends in attention, with somewhat weaker evidence that the differences in the types of words Mandarin- versus English-learning children produce or are exposed to affect attention to different aspects of a scene in the first 2 years of life.

Keywords: attention, word learning, objects and actions, English, Mandarin

Word learning requires establishment of a mapping between words and concepts. Before this mapping can occur, children need to attend to and identify conceptual units from perceptual components in the complex scenes that surround them in their everyday lives (Waxman & Lidz, 2006; Yu, Ballard, & Aslin, 2005). Attentional processes have generally been acknowledged as key to the development of children's earliest concepts, involving of a focusing of attention, perceptual analysis, and the ultimate extraction of meaning (Mandler, 2006). Similarly, attention is foundational to early word learning, and which elements of a complex scene children direct their attention to is a key issue in language acquisition and development. However, past studies examining the relationship between early attention and language acquisition have focused primarily on children's attention to objects in object-word acquisition (Kannass & Oakes, 2008), with relatively little research focused on differential attention to the various referents involved in a complex scene (i.e., actions vs. objects; but see Göksun et al., 2011; Katerelos, Poulin-Dubois, & Oshima-Takane, 2011). The present studies extended this question by asking whether infants exposed to different ways of talking about and interacting with the world (i.e., different languages and cultures) might also show differences in the elements they attend to in a complex, dynamic scene. They also asked when such differences might emerge—before or after the onset of word production—and whether infants' attention might also be affected by the number of words they have already learned.

On the one hand, attentional patterns should undergo a universal developmental trajectory, and attention does not distribute equally across all referents in a dynamic scene (Colombo, 2002; Göksun et al., 2011; Kannass & Oakes, 2008; Lakusta, Wagner, O'Hearn, & Landau, 2007). On the other hand, there is also evidence to suggest

Editor's Note. Nirmala Rao served as the action editor for this article.—JSE

This article was published Online First August 24, 2015.

Jie Chen and Twila Tardif, Department of Psychology, University of Michigan; Rachel Pulverman, Department of Psychology, Delaware State University; Marianella Casasola, Department of Human Development, Cornell University; Liqi Zhu, Institute of Psychology, Chinese Academy of Sciences; Xiaobei Zheng and Xiangzhi Meng, Department of Psychology, Peking University.

We acknowledge the generous support of National Institute for Child Health and Human Development Grant 1R03 HD60139-01A1 to Twila Tardif and National Science Foundation of China Grant 31070917 to Liqi Zhu in facilitating this research. We also thank the many children, families, and research assistants in both locations for making this research possible.

Correspondence concerning this article should be addressed to Twila Tardif, Department of Psychology, University of Michigan, 1012 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043. E-mail: twila@umich.edu

that the specific language a child is exposed to affects the particular elements that the child attends to, particularly as he or she begins to acquire words for those elements. In general, researchers have tended to find these patterns later in development rather than earlier, such that the universals of perception precede any differentiations in attention based on the language that is heard in infants' immediate environments.

Presumed universals of attention include evidence that infants under 2 years of age have been found to attend to moving objects with distinct shapes more than to static, more amorphous objects. At 11 months, infants fail to discriminate the ground (e.g., railroad track vs. road) but discriminate moving figures (e.g., adult vs. child) in dynamic events (Göksun et al., 2011); at 12 months, infants devote more attention to the goal than to the source of an action (Lakusta et al., 2007). In addition, infants pay attention to or discriminate some elements in dynamic events earlier than others, and their sensitivity to other components actually decreases with age (Bahrick, Gogate, & Ruiz, 2002; Casasola & Cohen, 2002; Göksun et al., 2011; Lakusta et al., 2007). In explaining these findings, the concept-to-language hypothesis proposes that infants' event perception starts from a common base such that all infants are sensitive to differences in components even though the concepts are not lexicalized in their native language. As they acquire language, attention to some event components may be dampened as other components are highlighted (Göksun et al., 2011).

Several studies have shown that, once infants begin to acquire language, they also begin to discriminate events in ways that are specific to the language that they are learning. English-learning infants, for instance, distinguish containment events before support events, and they are able to recognize support relations as they begin to acquire spatial terms for the types of relationships depicted (Casasola & Cohen, 2002). Moreover, English learners decrease in sensitivity, whereas Korean learners, who are learning a language that highlights this distinction, maintain sensitivity to tight-fit versus loose-fit containment relations as their vocabularies grow (Choi, 2006).

For other aspects of events, such as spatial representations of manner and path, exposure to different languages has been shown to influence attentional patterns in ways that may further guide language learning (Pulverman, 2005; Pulverman, Golinkoff, Hirsh-Pasek, & Buresh, 2008). Moreover, Göksun et al. (2011) found that the developmental changes in discriminating ground versus figure are also influenced by language experience. Specifically, English-learning infants, who learn a language in which there are no linguistic distinctions for different categories of grounds, had reduced sensitivity to these distinctions with development. In contrast, Japanese-learning infants, who are immersed in a language that makes linguistic distinctions between the two types of grounds, maintain the distinctions. Together, these findings suggest that attentional biases toward elements that are highlighted in one's native culture may be wide-ranging and begin even at the earliest stages of learning a language.

Although studies show developmental changes and cross-cultural differences in attention to spatial representations (e.g., figure, ground, manner, path), no study has examined the influence of language or culture for attentional differences in the relative acquisition of object versus action words in a complex dynamic event. One study that laid the groundwork for such an examination is Katerelos et al. (2011), in which infants' relative attention to the actions versus the objects (agents) taking part in dynamic events in English- and French-learning infants was examined. In this study, 18-month-old English- and French-learning infants were first habituated with events in which an animated object (either animallike or vehiclelike) followed a particular motion path (best described as "jumping over" vs. "bouncing up and down"). Infants were then presented with a number of test trials, based on a habituation-switch design, that systematically tested discrimination of the objects and the actions. Katerelos et al. found that in the absence of a label, infants looked significantly longer at the object-switch and action-switch trials than at the baseline trial, indicating that infants paid attention to both the objects and the actions and were able to differentiate between the two actions as well as the two objects. When the scenes were presented together with a verbal label, however, English- and French- learning infants looked significantly longer at the object-switch and the word-switch trials than at the baseline trial, but their looking time to the action-switch trial did not differ from baseline. This finding suggests that both English- and French- speaking infants tend to pay more attention to the object taking part in the action than to the action itself when presented in the context of a verbal label. Importantly, this study did not examine developmental changes in infants' attentional preferences. Thus, it is not clear whether becoming a word learner changes the pattern of basic attentional processes to objects and actions. Nonetheless there is reason to suspect, given a number of previous studies in this area, that such changes may occur and that they may be related to the very process of having learned words (Casasola & Cohen, 2002; Gliga, Volein, & Csibra, 2010; Göksun et al., 2011; Lakusta et al., 2007). Both English- and French-learning infants have been shown to exhibit strong noun biases in their early vocabularies (Bates et al., 1994; Bornstein et al., 2004; Tardif et al., 2008; Tardif, Gelman, & Xu, 1999). Thus, it is not surprising that objects, especially objects that appear to be agentive and moving, may be a focus of attention for 18-month-old infants learning these languages, when given a verbal label.

The current studies aimed to extend these questions by examining the role of attention as it applies to cross-cultural differences in word learning between English- and Mandarin-speaking children. There are numerous differences between English and Mandarin that lead to differences in the frequencies with which noun and verb types and tokens are heard by native language learners of these two languages. Across several studies of monolingual and bilingual Mandarin speakers, native Mandarin-speaking adults have been found to produce more verbs and fewer object nouns (both types and tokens) relative to their English-speaking counterparts, even when controlling for identical contexts (W. H. Chan & Nicoladis, 2010; Chen, Tardif, Meng, & Setoh, 2009; Lee & Naigles, 2005; Sandhofer, Smith, & Luo, 2000; Tardif et al., 1999; Tardif, Shatz, & Naigles, 1997). In addition to the absolute frequencies of nouns and verbs, there are multiple ways in which verbs are made salient in Mandarin and are less salient in English, including differences in the syntactic properties that allow for the omission of subjects and objects in Mandarin sentences (Lee & Naigles, 2005; Tardif et al., 1997); in the semantic properties such that Mandarin, but not English, verbs tend to be highly imageable and specific in meaning (Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif, 2009; Pulverman, Tardif, Rohrbeck, & Chen, 2010; Tardif, 2006); and in the nonlinguistic cues caregivers provide to children (Snedeker, Li, & Yuan, 2003).

Moreover, there have been several reports of cross-linguistic differences in both the production and the comprehension of terms for these concepts, even in young children (C. C. Chan et al., 2011; Tardif et al., 2008; but see Imai & Gentner, 1997). Nonetheless, these data are somewhat more controversial and certainly more nuanced. First, in multiple studies comparing children's productive use of nouns and verbs in English and Mandarin, every study has found that Mandarin-learning monolinguals produce more verbs than English-learning monolinguals at similar stages of vocabulary development (Tardif et al., 2008; Waxman et al., 2013). Moreover, English–Mandarin bilinguals also produce more verbs, and produce them at earlier stages of vocabulary development, in Mandarin than they do in English (W. H. Chan & Nicoladis, 2010; Levey & Cruz, 2003). This does not mean, however, that Mandarin-learning children produce more verbs than object nouns or that English-learning children produce more nouns than verbs at all stages of vocabulary development or for all measures of vocabulary. Rather, the large majority of studies have reported that English-learning children produce more nouns, and object nouns in particular, than verbs (Bornstein et al., 2004; Gentner, 1982; Tardif et al., 1999; Waxman et al., 2013; but see Bloom, Tinker, & Margulis, 1993). The reports on Mandarin-learning children are more mixed, such that some studies, some measurement methods (e.g., parent report, vocabulary checklist), some stages of vocabulary production, and some contexts (e.g., book reading) have resulted in reports of more object nouns than verbs (Gentner, 1982; Levey & Cruz, 2003; Tardif et al., 1999), whereas other studies, using other measurement methods (e.g., direct observation, vocabulary checklists), sampling children at different stages of vocabulary development and in different contexts (e.g., toy play), have reported the production of more verbs than object nouns (W. H. Chan & Nicoladis, 2010; Tardif, 1996; Tardif et al., 1999, 2008).

The data for comprehension are no less complex. Specifically, with 3–5-year-olds, Imai, Haryu, and Okada (2005; Imai et al., 2008) found that English-, Japanese- and Mandarin-learning children, regardless of language, succeeded in mapping novel nouns to objects but not verbs to actions and that Mandarin-speaking children had the most difficulty in a novel verb condition. In contrast, C. C. Chan et al. (2011) examined novel word mapping with 14- and 18-month-old infants and found that Mandarin-learning infants learned word–action associations earlier than word–object associations, whereas English-learning infants in the same experimental paradigm learned word–object associations earlier than word–action associations. Waxman et al. (2013) found that both English- and Mandarin-learning 24-month-olds could learn to map verbs in a novel context.

Despite the differences in both input and acquisition of nouns and verbs across languages, the question of whether attention can be shifted from objects to actions is a fundamental one. The present set of studies aimed to examine whether developmental changes would be seen in discriminating the actions and objects involved in a dynamic event and whether these developmental patterns would be similar or different across languages. In particular, we were interested in how the process of learning a language may affect attentional behaviors among children who have not yet begun this process (i.e., prelinguistic 6-month-olds), and among children who are in the middle of an explosive phase of early vocabulary development (i.e., 18-month-olds). As discussed earlier, English learners produce mostly nouns and hear a disproportionate number of noun types in this early word-learning phase. Thus, we expected that they would show greater attention to objects at 18 months of age than they do at 6 months of age. An important question, however, was whether this would be equally true for their attention to actions. Given that English emphasizes nouns and that English-learning children produce many more nouns than verbs, particularly in their first few months of producing words (i.e., from 12 to 18 months), one might expect them to show diminished attention to actions, relative to objects, at this age. Alternatively, it could be that actions are highly salient (Bahrick et al., 2002), and because we chose to focus on static objects, which are more the focus of the cross-linguistic difference between English and Mandarin than agentive objects (see Tardif et al., 2008), English learners might not show a decrease in attention to actions over this time period.

Mandarin learners show a more balanced pattern of nouns and verbs in their early vocabularies (Tardif, 1996; Tardif et al., 2008), are exposed to a disproportionate number of verb types and verb tokens in the language that they hear, and map novel words to actions earlier than to objects (C. C. Chan et al., 2011). Thus, it was not clear whether they would transition from paying more attention to actions and then show a more balanced pattern of attending to both actions and objects, if they would show a pattern that parallels their very earliest words (i.e., attention to actions), or—because of universals in early perceptual development that may be independent of language—if they would show a pattern that is more like English (i.e., attention to actions and moving objects; Bahrick et al., 2002). In any case, it is also important to consider whether attentional patterns shift in the presence of a verbal label (e.g., Casasola & Bhagwat, 2007; Katerelos et al., 2011) or if similar patterns of attention to actions and objects prevail regardless of age, language being learned, or the presence of a linguistic label, thus suggesting strong universal patterns in the development of attention to the actions versus the objects of a dynamic scene.

Given these questions, we designed two parallel studies to examine English- and Mandarin-learning infants' attention to objects versus actions. In Study 1, we examined infants' attention to different components of novel scenes presented in the absence of linguistic labels. If the language children are learning drives their attentional mechanisms at a pervasive level, then we would expect English and Mandarin learners to show different patterns of attention to these two elements (i.e., actions and objects) of the scene, although this may not appear until they have actually begun to comprehend and produce distinct words for the specific types of concepts. Specifically, we would expect Mandarin-learning 17–19-month-olds to pay more attention to action changes than similar-age English learners or 6-month-olds learning either language. Conversely, we would expect English-learning 17–19-month-olds to pay more attention to object changes than would the other groups of children. If, however, perceptual development alone drives attentional patterns, we would expect to find minimal or no differences across the two language groups, although we might still find differences in the extent to which different elements are attended to by younger versus older infants (Bahrick et al., 2002; Göksun et al., 2011).

In Study 2, we explored whether presenting a linguistic stimulus—in this case, a single novel word—introduced any additional influences on children's attentional patterns to the objects versus

the actions in a dynamic scene. Previous studies have shown that words promote detection of category-based commonalities, which have consequences for conceptual organization (Waxman & Markow, 1995), and that words, but not other sounds, promote the formation of conceptual categories (Balaban & Waxman, 1997; Gogate, Walker-Andrews, & Bahrick, 2001; Sloutsky & Lo, 1999) and even abstract categorical representations of spatial relations (Casasola, 2005; Casasola, Bhagwat, & Burke, 2009). Study 2 was more specifically focused on whether the presence of a novel word—with no morphological or contextual information and, thus, of ambiguous syntactic class—would drive attention away from some components (e.g., from the action) and toward others (e.g., the object) within the event, detected by comparing the attention allocated to each component. We hypothesized that if language draws attention to the concepts labeled most frequently in that language, then the presence of a novel word should bias children's attention toward those elements that the language most frequently labels (i.e., actions for Mandarin, objects for English). Thus, we would expect to find the greatest differences in the extent to which different elements are attended to by English- versus Mandarin-learning infants in the presence of a verbal label (i.e., in Study 2 more than in Study 1). Conversely, if linguistic cuing does not play a significant role, we would expect minimal or no differences across the two studies.

Study 1: English- and Mandarin-Learning Infants' Discrimination of Actions and Objects in a Dynamic Event Presented in Silence

Method

Participants. A total of 136 Chinese and American infants participated in this study, 79 of whom were used for analysis. All participants were healthy, full-term monolingual infants with no history of visual or auditory problems. In both cultures, parents read and signed consent forms approved by the universities where the research was conducted. Specifically, 19 (eight male, 11 female) 6–8-month-old English learners (M_{age} = 6.61 months, SD = 0.50) and 21 (10 male, 11 female) 17–19-month-old English learners (M_{age} = 17.40 months, SD = 0.48) were recruited from an infant database. In parallel, 16 (eight male, eight female) 6–8-month-old (M_{age} = 6.62 months, SD = 0.55) and 23 (12 male, 11 female) 17–19-month-old (M_{age} = 17.76 months, SD = 0.66) Mandarin-learning infants were selected, who participated in the study in Beijing, China, after responding to flyers posted in neighborhoods and advertisements placed on a parenting website. Data provided by additional 29 English-learning infants (13 6–8-month-olds and 16 17–19-month-olds) and 28 Mandarin-learning infants (seven 6–8-month-olds and 21 17–19-month-olds) were not included for the following reasons: parental interference such as pointing, talking, looking at the screen, and so forth (English: n = 2; Mandarin: n = 6); child fussiness during testing (English: n = 9; Mandarin: n = 6); receiving the maximum number of habituation trials but failing to reach the habituation criteria (English: n = 8; Mandarin: n = 4); success in reaching the habituation criteria but showing dishabituation on the control trial (English: n = 2; Mandarin: n = 2); nonrecovery from the posttest novel stimulus (English: n = 2; Mandarin: n = 1); background noise (English: n = 0; Mandarin: n = 2); bilingual/wrong age (English: n = 3; Mandarin: n = 4); and experimenter error (English: n = 3; Mandarin: n = 3). These attrition rates are comparable to those in habituation studies with similar-age infants using similar types of video stimuli (Casasola & Cohen, 2000; C. C. Chan et al., 2011; Pulverman et al., 2008; Werker, Cohen, Lloyd, Casasola, & Stager, 1998).

Materials. A novel 3-s videotaped event without sound was presented to infants in a habituation paradigm in which a young woman (Caucasian for U.S. infants, ethnic Chinese for Chinese infants) performed a novel action on a novel object. The events looped 10 times, so infants viewed a specific event for a maximum of 30 s. The novel actions were transitive motions with clear paths toward the objects and with manners that involved no changes to the object. The objects were bright, bicolored figures with smooth edges and included no primary shapes. There were a total of eight different dynamic events, with each element (person, object, action) having two possible forms. At the end, all infants saw a 30-s video of a young woman juggling three colored balls as a posttest trial (see Appendix A).

Although we did not expect the 6–8-month-olds to be able to comprehend or produce any words, we were concerned that they appear to be developing normally in both groups and, thus, developed an early language questionnaire given to both groups (see Appendix B). This instrument was developed from the "early signs of comprehension" on the MacArthur Communicative Development Inventory (MCDI; Fenson et al., 1993) and included additional questions on the child's development of vocalizations and babbling. There were 10 common items in both English and Mandarin versions, including four items on producing sounds (e.g., making "aaa" or "ooo" sounds), three items on understanding first signs (e.g., responding when name is called), and three items on understanding short sentences (e.g., "Don't touch").

The MCDI; (Fenson et al., 1993) and the toddler form of the Chinese Communicative Development Inventory (CCDI; Tardif, Fletcher, Zhang, Liang, & Zuo, 2009) were used for the 17–19-month-old infants. Parents were asked to report on which words their children could spontaneously say. In the standardized short forms of these measures, English had 101 total words and Mandarin 113 words, which were ordered, but not separated into, semantically related sections (i.e., object names, actions, people terms, prepositions, descriptive terms, and so on). In addition, because the number of action terms on the short forms is relatively small, we added the entire action word section of the MCDI and CCDI long forms (toddler version). From this, and the object nouns on the short forms, the total number of object nouns and action verbs were calculated both in terms of their raw scores and their opportunity scores (e.g., total object nouns reported divided by the maximum number of object nouns), by taking consideration of different numbers of words of each type for each of the two languages.

Apparatus. The experiment took place in a dimly lit and sound-attenuated laboratory room. The labs in both locations were designed to be as similar as possible, with a standardized visual angle of 20°. In the U.S. lab, infants sat on their caregivers' laps facing a 13-in. color computer monitor located approximately 3 feet away; in Beijing, infants sat on their caregivers' laps facing a 25-in. TV screen located approximately 5.5 feet away. Auditory levels and lighting were also standardized to be approximately 65

dB (±5 dB), 26 lx (±3 lx) while the stimuli were on, and 20 lx (±3 lx) when the stimuli were off (because of laptop background).

Design and procedure. Infants were randomly assigned to a habituation scene from eight possible person–action–object combination scenes. The between-subjects variables were age (6–8 months vs. 17–19 months) and language group (English learning vs. Mandarin learning); test type (control, action change, object change, posttest) was a within-subject variable.

Prior to the experiment, infants played with a set of toys in a playroom until the experimenters and caregivers felt that they were acclimated to the environment. During this period, one of the experimenters explained the study in detail to the caregiver and instructed the caregiver to sign a consent form and to complete a language development questionnaire (an early word survey for 6–8-month-olds [see Appendix B] and the short form version of the Words and Sentences MCDI for 17–19-month-olds).

Prior to entering the testing room, caregivers were instructed to close their eyes and avoid talking, pointing, or otherwise influencing their infants' attention throughout the duration of the video. Infants and their caregivers were then taken to the testing area, and infants were seated on their caregivers' laps facing the monitor while the experimenter presented the stimuli via the Habit X program (Version 1.0; Cohen, Atkinson, & Chaput, 2004).

Habituation trials began with an attention getter, an animated geometric shape appearing at the center of the screen in synchrony with a siren sound, to draw infants' attention. As soon as the infants oriented toward the screen, they were introduced to the dynamic scene. The scenes for the habituation phase were chosen randomly from eight potential events, all of which were identical across habituation trials for a particular infant but counterbalanced across infants (e.g., Young Woman A directing her chin in a downward motion perpendicular to the table to touch the top of a triangular green object). The habituation phase ended when an infant's looking time across three consecutive trials decreased by 50% from his or her total looking time in the initial three trials or when a maximum of 20 trials had been presented. Individual habituation trials ended when an infant looked away for 2 s or more during the trial.

Once the habituation criterion was reached, infants were presented with three test trials in which none (control trial) or one of the elements changed (action change trial, object change trial). For instance, with the stimulus just described, the infant would be presented with a scene in which Young Woman A uses her chin to touch the top of a triangular green object (control); followed by a scene in which Young Woman A moves two fingers, extended like a pair of scissors, of her left hand parallel to the table toward the triangular green object (action change); and then a scene in which Young Woman A directs her chin downward to touch the top of an arc-shaped red object (object change). Note that all objects and actions were designed to be simple yet nonnameable in both languages (see Appendix A). The order of these trials was counterbalanced across infants.

The test trials were followed by a posttest novel stimulus—a young woman juggling three brightly colored balls. This trial was designed to provide a control to alternative explanations in case infants' attention remained uniformly low throughout the test phase. All habituation and test trials were conducted in total silence, without any labels or other sounds provided.

Coding and reliability. Infants' looking times were calculated online by a live experimenter who was blind to the stimuli appearing on the screen and who coded an infant's eye movements through a curtain hole. The Habit X program, designed to present audiovisual stimuli and record looking times in infant studies (Cohen et al., 2004), was used. The experimenter pressed a designated computer key when an infant looked at the screen and released the key when the infant looked away. If looking-away time reached 2 s, the event disappeared, and the attention getter resumed, starting the next trial.

To ensure consistency across both the Chinese and U.S. labs, the experimenters in both locations were trained by using an identical protocol and practice videos. Experimenters were not allowed to code real data in either lab until they had met the dual requirement of 95% reliability with five successive videos coded by previously trained experimenters and 95% reliability with three successive videos coded by themselves 1 week prior to the comparison. Interrater reliability for the primary data was calculated from 15 videos across different coders, and the online–offline coding correlation was 99.5%.

Results and Discussion

Infant language development. There were minimal differences in the early language questionnaire given to infants in both cultures, with both groups in the typical development range on this instrument. More specifically, out of 13 items for English-learning infants, 51.3% of questions (range: 30.8%–92.3%) were answered "Yes," and out of 11 items for Mandarin-learning infants, 65.6% of questions (range: 36.4%–90.9%) were answered "Yes." For the 10 items that were common to both the English and the Mandarin forms, no difference was found, with mean scores of 5.61 (SD = 1.69) for English-learning infants and 6.35 (SD = 1.82) for Mandarin-learning infants.

For the older infants, no difference was found between the two language groups either in the raw scores or the word opportunity scores for both the total short-form vocabulary and the total number of nouns that the infants were reported to be able to "say," as shown in Table 1. In contrast, and as expected from numerous previous studies on vocabulary differences in English- and Mandarin-learning toddlers, Mandarin-learning infants were reported to say more verbs than English-learning infants—again, both in terms of raw scores and word opportunity scores (ps < .05).

Developmental and cross-linguistic effects on attention. Prior to performing analyses specific to our hypotheses, the eight different videos were examined for initial salience with a one-way analysis of variance (ANOVA) comparing average looking time in the first three trials across the various videos and for the number of habituation trials required for each of the eight scenes. Overall, no differences were found in salience across the eight dynamic scenes shown to either the English or the Mandarin learners, nor were differences found across any of the 16 dynamic scenes shown across the two cultures. Specifically, the mean looking times during the first three trials of habituation were 17.24 s (SD = 7.26 s) and 18.54 s (SD = 6.02 s) for 6–8-month-old and 17–19-month-old English-learning infants, respectively, and 17.56 s (SD = 8.18 s) and 21.61 s (SD = 7.08 s) for 6–8-month-old and 17–19-month-old Mandarin-learning infants, respectively. Moreover,

Table 1
Means (With Standard Deviations in Parentheses) for Raw and Word Opportunity Scores of 17–19-Month-Olds by Language in Studies 1 (Silent) and 2 (Word Label)

Language and study	Age (months)	Total number of words	Word opportunity score: Total	Total number of nouns	Word opportunity score: Nouns	Total number of verbs	Word opportunity score: Verbs
English							
Study 1	17.40 (0.48)	18.00 (10.74)	.178 (.106)	9.04 (6.85)	.192 (.146)	5.70 (6.21)	.055 (.060)
Study 2	17.43 (0.67)	21.21 (15.37)	.210 (.152)	11.42 (9.70)	.243 (.206)	9.71 (14.16)	.094 (.137)
Mandarin							
Study 1	17.76 (0.66)	22.81 (22.10)	.202 (.196)	9.71 (9.61)	.237 (.234)	35.19 (45.97)	.157 (.205)
Study 2	17.52 (0.31)	25.25 (27.76)	.223 (.246)	11.46 (13.53)	.279 (.330)	35.54 (51.69)	.159 (.230)

Note. Word opportunity = number of words an infant can say divided by the number of words in the form.

across cultures, ages, and videos, there were no significant differences in the final three habituation trials, in which English learners dropped to an average of 39.46% (SD = 9%) and Mandarin learners dropped to an average of 40.78% (SD = 8%) of initial looking time. In addition, to exclude the possibility of fatigue, we conducted a 2 (recovery: control trial vs. posttest trial) × 2 (language: English vs. Mandarin) × 2 (age: 6–8 months vs. 17–19 months) repeated-measures ANOVA to examine whether infants demonstrated increased interest in the posttest trial. The main effect of recovery was highly significant, $F(1, 75) = 402.90$, $p < .001$, $\eta_p^2 = .843$, and no other main effects or interactions were found for this trial, eliminating the possibility of general fatigue by the end of the task, even for 6-month-olds.

The main purpose of this study was to explore infants' attentional patterns through their discrimination of the actions and objects involved in dynamic events across ages and cultures. No difference was found between male and female participants, so all of the following analyses were collapsed across gender. Looking times to each of the test trial types by age and language are presented in Figure 1. A repeated-measures ANOVA was carried out to examine infants' mean looking durations, with test type (control, action change, object change) as a within-subject variable and age (6–8 months vs. 17–19 months) and language (English vs. Mandarin) as between-subjects variables. As expected, there were main effects of test type, $F(2, 150) = 29.28$, $p < .001$, $\eta_p^2 = .281$, and age, $F(1, 75) = 8.50$, $p < .01$, $\eta_p^2 = .102$. The Age × Test Type interaction was also significant, $F(2, 150) = 4.83$, $p < .01$, $\eta_p^2 = .061$. A pairwise comparison with Bonferroni corrections further indicated that, overall, infants looked significantly longer at the change trials than at the control trial and that 17–19-month-olds looked more during the change trials, but not the control trial, than did 6–8-month-olds. In addition, collapsing across two languages, the 6–8-month-olds looked significantly longer at the action change trial than at the control trial ($p < .05$), whereas the 17–19-month-olds looked significantly longer at both the action change and object change trials ($ps < .001$) than at the control trial and also spent longer times looking at the action change than at the object change trial ($p < .05$).

Although no difference in any test type between languages was found, we conducted planned comparisons on the looking behavior of infants at each age and within each language group because of our interest in the effects of children's developing language abilities together with the language being spoken. At 17–19 months of age, both English and Mandarin learners looked longer at the action and object change trials than at the control trial ($ps < .05$). In contrast, at 6–8 months of age, only the Mandarin learners looked longer at the object and action change trials than at the control trial ($ps < .05$), whereas English learners showed only a marginal difference between the action change and control trials ($p = .10$).

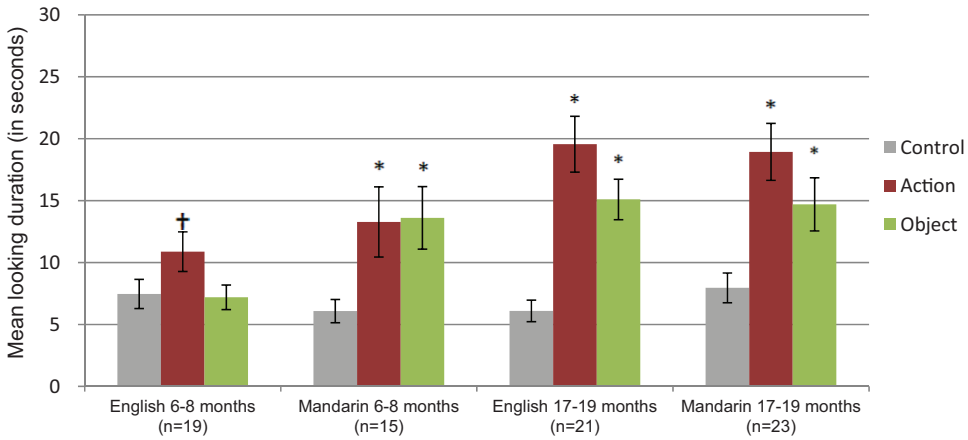

Figure 1. Mean looking durations on test trials as a function of age and language in silence (Study 1). Error bars represent standard errors. † $p < .10$. * $p < .05$. See the online article for the color version of this figure.

Effects of individual language development on attention. To more directly test the potential effect of language ability on attention, we performed a repeated-measures ANOVA with test type as a within-subject variable and language and vocabulary size as between-subjects variables. A median split within each language was used to divide participants into high- and low-vocabulary groups (English median: 19.0; Mandarin median: 16.5). Again, results revealed only a significant effect of test type, $F(2, 74) = 27.13$, $p < .001$, $\eta_p^2 = .423$, with no main effect of vocabulary or interactions among language, vocabulary, and test type.

In summary, we found evidence for developmental differences in attention such that 17–19-month-olds paid more attention to both the action and the object changes in a dynamic scene than did 6–8-month-olds. In addition, the 6–8-month-olds had particular difficulty discriminating the object changes. In many ways, these data replicate those found by Bahrick et al. (2002), in which 5.5-month-old infants discriminated and remembered actions, but not objects, in dynamic events. However, the Bahrick et al. events involved familiar actions and objects, whereas our events involved novel actions and objects. Nonetheless, it is significant that both studies resulted in similar findings for 6-month-old infants.

This study is novel in that we attempted to examine differences in these patterns as a function of the particular linguistic environments that the children experienced in their day-to-day lives. On the one hand, one might have expected to find significant differences in the older children because they were well on their way to learning language-specific patterns of vocabulary, with English-learning children's focus largely on object nouns, with a relative paucity of verbs, and Mandarin-learning children's focus spread across both verbs and nouns. On the other hand, there are several reasons why we might not have found these differences in this particular design even if attentional patterns do differ across languages.

First, it is possible that although attention during a language-learning situation may differ, attention in the absence of language may not differ across languages even though it may differ across age groups. From a simply adaptational perspective, infants in all cultures might be expected to pay attention first to things that move or other immediately relevant changes in their environments. Objects that neither move nor are ingested by other agents are not likely to receive high amounts of attention simply because they do not have high relevance for survival. As an infant gets older and more mobile, however, many more types of objects may become relevant. Objects that are manipulated by others, even though they are not ingested and do not move of their own accord, could still elicit attentional and discriminative mechanisms. Further, greater attention to changes in a dynamic event is also likely to enable finer discriminations and additional learning, with or without language. Thus, in the absence of a linguistic label, one might expect infants to show similar patterns of attention, regardless of their language environment. Providing a linguistic label, however, may begin to alter a universally relevant set of attentional mechanisms by providing a specific spotlight to direct attention to elements that the language most frequently labels (i.e., actions for Mandarin, objects for English). Thus, an important follow-up would be to examine these same patterns of attention in the presence of a verbal label. This we did in Study 2, which is described next.

Study 2: English- and Mandarin-Learning Infants' Discrimination of People, Actions, and Objects in a Dynamic Event in the Presence of a Linguistic Label

Method

Participants. A total of 125 infants participated in Study 2, with 78 infants included for analysis. Although they were recruited through the same baby databases, none of these infants participated in Study 1. The final sample included 16 (seven male, nine female) 6–8-month-old English learners ($M_{age} = 7.15$ months, $SD = 0.52$) and 23 (12 male, 11 female) 17–19-month-old English learners ($M_{age} = 17.43$ months, $SD = 0.69$) from monolingual families in a small, Midwestern university town and 15 (eight male, seven female) 6–8-month-old ($M_{age} = 6.70$ months, $SD = 0.65$) and 24 (12 male, 12 female) 17–19-month-old ($M_{age} = 17.52$ months, $SD = 0.31$) Mandarin-learning infants from monolingual families in the northwest university district of Beijing, China. Data provided by an additional 30 English-learning infants (six 6–8-month-olds and 24 17–19-month-olds) and 17 Mandarin-learning infants (three 6–8-month-olds and 14 17–19-month-olds) were not included for the following reasons: parental interference (English: $n = 1$; Mandarin: $n = 3$), fussiness during testing (English: $n = 5$; Mandarin: $n = 2$), failure to reach habituation criteria (English: $n = 2$; Mandarin: $n = 0$), success in reaching the habituation criteria but dishabituation on the control trial (English: $n = 1$; Mandarin: $n = 0$), nonrecovery from the posttest novel stimulus (English: $n = 10$; Mandarin: $n = 11$), background noise (English: $n = 1$; Mandarin: $n = 0$), bilingual/wrong age (English: $n = 6$; Mandarin: $n = 0$), and experimenter error (English: $n = 4$; Mandarin: $n = 1$).

Materials and procedures. The video stimuli used in Study 2 were identical to those used in Study 1. The only difference between Study 1 and Study 2 was the inclusion of linguistic labels. In the current study, novel monosyllabic words were presented simultaneously with the video during the habituation phase in the word condition. The words were single-syllable forms constructed from phonotactically regular patterns that appeared frequently in English- and Mandarin-learning infants' earliest acquired words on the basis of norming data from the MCDI and the CCDI (Fenson et al., 1993; Tardif et al., 2008). The novel word presented to the U.S. infants was *sug*, followed by simple expressions such as "wow!" or "ooh!" The novel word presented to Chinese infants was *fai1*, followed by similar simple expressions (e.g., "wa!" or "oh!"). Only the exclamations were presented in the test trials (see the audio stimuli in Appendix C). Female native speakers of English and Mandarin recorded exemplars of each label in infant-directed prosody inside a sound-attenuated booth. The audio stimuli were standardized at 70 dB (±5 dB) using a sound meter. All the procedures in Study 2 were the same as those in Study 1.

Results

As with Study 1, we expected to find both developmental and cross-linguistic differences in infants' discrimination of the elements. Moreover, if a novel label biases infants to attend to one element more than another, then infants in Study 2 (stimuli presented with verbal labels) should have demonstrated different patterns of discrimination than did the infants in Study 1 (who

viewed the events in silence). Specifically, for English-learning infants, hearing a novel word was expected to direct attention to the objects in the events, particularly at 18 months (Katerelos et al., 2011; Waxman, 1999), whereas Mandarin speakers were expected to focus more on actions, particularly when the events were presented together with a verbal label.

Infant language development. Similar to Study 1, no significant differences were found in the final three habituation trials across languages, ages, or videos. In addition, the main effect for the recovery trial was highly significant, $F(1, 74) = 318.05$, $p < .001$, $\eta_p^2 = .811$, and no other main effects or interactions were found, eliminating the possibility of general fatigue by the end of the task. Again, no differences were found between male and female participants, so all of the following analyses were collapsed across gender.

As with Study 1, the 17–19-month-old Mandarin-learning infants were reported to say more verbs than the English-learning infants ($p < .05$), but no difference was found in terms of total short-form vocabulary or in the number of nouns produced. In addition, results from the early language questionnaire showed that out of 13 items for English-learning infants, 55.3% of questions (range: 38.5%–84.6%) were answered "Yes," and out of 11 items for Mandarin-learning infants, 61.4% of questions (range: 36.4%–100.0%) were answered "Yes." Again, there were no differences for the 10 common items, with mean scores of 5.88 ($SD = 1.36$) for English-learning infants and 6.00 ($SD = 2.10$) for Mandarin-learning infants.

Developmental and cross-linguistic effects on attention. To explore infants' discrimination of the actions and objects involved in the dynamic events across ages and cultures, we performed a repeated-measures ANOVA to examine infants' mean looking durations, with test type (control, action change, object change) as a within-subject variable and age (6–8 months vs. 17–19 months) and language (English vs. Mandarin) as between-subjects variables (see Figure 2). Similar to Study 1, there was a main effect of test type, $F(2, 148) = 18.73$, $p < .001$, $\eta_p^2 = .202$; a main effect of age, $F(1, 74) = 3.95$, $p < .05$, $\eta_p^2 = .051$; and a significant Test Type × Age interaction, $F(2, 148) = 4.71$, $p < .05$, $\eta_p^2 = .060$. Overall, infants looked significantly longer at the change trials than at the control trial; 17–19-month-olds looked significantly longer than 6–8-month-olds; and, because a significant Test Type × Age interaction was found, pairwise comparisons further revealed that the 6–8-month-olds looked significantly longer at the action change than at the control scene ($p < .05$), whereas the 17–19-month-olds looked significantly longer at both the object change and action change trials relative to the control trial ($ps < .001$). However, unlike in Study 1, in which the scenes were presented in silence, we did not find a main effect of looking time for action versus object changes, which suggests that providing word labels may have neutralized infants' preference for looking at actions.

Interestingly, there was a main effect of language when the novel label was provided, $F(1, 74) = 7.47$, $p < .01$, $\eta_p^2 = .092$, such that Mandarin-learning infants looked significantly longer than English-learning infants. However, because we were specifically interested in attention to actions and objects rather than the control situation, we examined the patterns of looking behavior as planned comparisons across action and object changes separately with both older and younger infants. In this comparison, a univariate ANOVA showed that 17–19-month-old Mandarin-learning infants looked significantly longer at action changes than did 17–19-month-old English-learning infants, $F(1, 45) = 4.31$, $p < .05$, $\eta_p^2 = .087$. No language effect was found in either the action or the object changes with the younger infants.

Moreover, when the data from the two studies were combined to examine directly the labeling effect by using the difference score between looking time for the change versus control trials to account for any differences in baseline looking (see Figure 3), a univariate ANOVA with study (i.e., scenes presented in silence vs. with a label) as a between-subjects variable revealed that 17–19-month-old English-learning infants spent less time looking at the action changes, $F(1, 42) = 5.93$, $p < .05$, $\eta_p^2 = .124$, when presented with a verbal label for these dynamic scenes. In contrast, no significant differences were found across studies for the Mandarin learners. We believe that these findings are consistent with

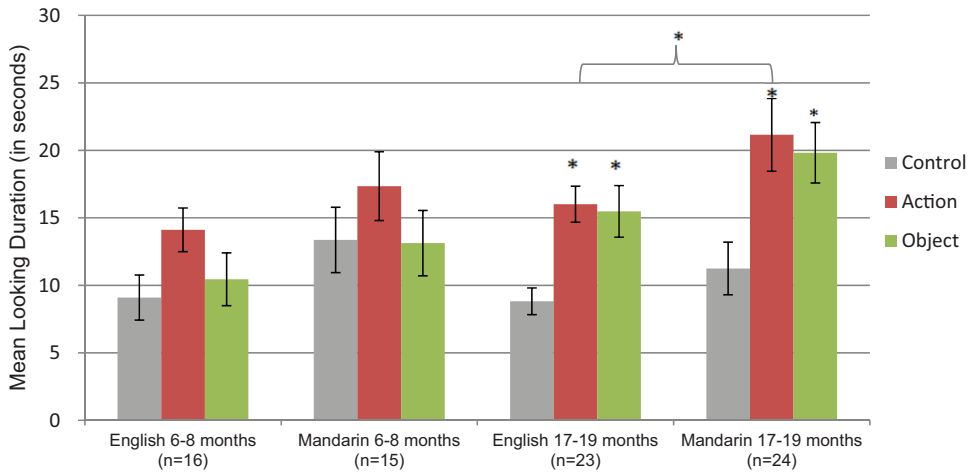

Figure 2. Mean looking durations on test trials as a function of age and language when scenes were presented with a novel word (Study 2). Error bars represent standard errors. * $p < .05$. See the online article for the color version of this figure.

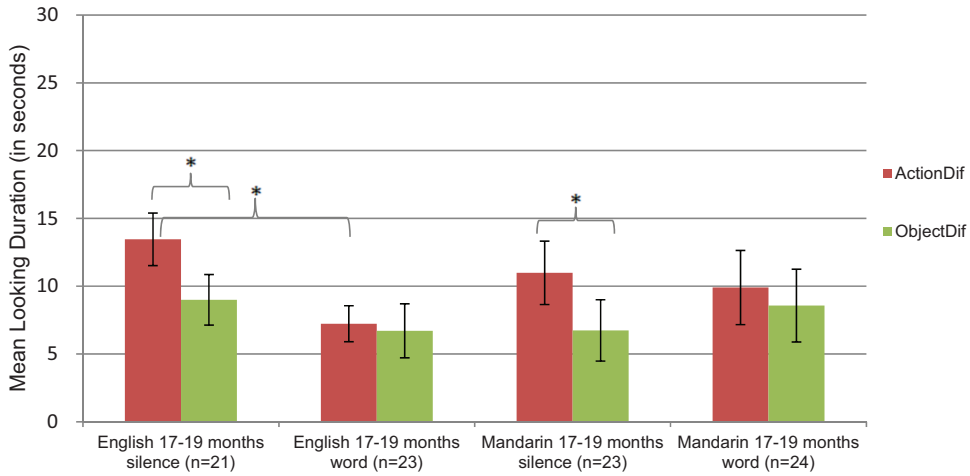

Figure 3. Mean looking duration difference (Dif) scores (change trials − control trials) as function of language and condition for 17–19-month-olds. Error bars represent standard errors. * $p < .05$. See the online article for the color version of this figure.

the result that providing a verbal label with a complex scene reduced sensitivity to action changes, and this was particularly true for infants learning English. Again, 6–8-month-old learners revealed no effect of label in either the action or the object change conditions, regardless of the language to which they were exposed.

Effects of individual language development on attention. As with Study 1, a repeated-measures ANOVA with test type as a within-subject variable and language and median-split vocabulary group (English median: 16 words; Mandarin median: 18 words) as between-subjects variables was conducted to test the potential effect of individual differences in vocabulary on attention to either the object or the action changes. Similar to Study 1, only a main effect of test type was found, $F(2, 86) = 23.49$, $p < .001$, $\eta_p^2 = .361$. There were no interactions between vocabulary size and looking behavior for either the English- or the Mandarin-learning infants.

In sum, the results of Study 2 were similar to those of Study 1 in showing a developmental effect of discriminating specific elements in a dynamic event. More interestingly, these results provide preliminary evidence for the role of linguistic labels in directing infants' attention to elements that the language most frequently labels, but in a number of subtle ways. First, results showed that 17–19-month-old Mandarin-learning infants looked significantly longer at action changes than did English-learning infants. In addition, it is clear from Study 2 that providing a verbal label appeared to change the pattern of attention for 17–19-month-old English learners such they no longer noticed the action changes to the same degree as did the Mandarin learners. Moreover, the direct comparison between Study 1 (silence) and Study 2 (word label) found that 17–19-month-old English-learning infants spent less time looking at the action changes when they were presented with a verbal label during habituation than they did when the scenes were presented in silence.

General Discussion

The present research was motivated by the fact that a fundamental prerequisite for learning words is children's ability to attend to specific aspects of dynamic events in a complex scene—the quintessential word-learning challenge for every child worldwide. These studies aimed to investigate the cognitive mechanisms underlying the mapping from aspects of a complex scene to the words that children hear. We aimed to determine whether differences in English and Mandarin learners' early exposure to language, as well as consistent differences in the extent to which they are able to comprehend and produce nouns and verbs, would be paralleled by similar differences in attentional patterns to different aspects of the complex scenes that children experience in their day-to-day lives. Consistent with differences in adult-to-child speech and, to a lesser extent, with cross-linguistic differences in early word production (Tardif, 1996; Tardif et al., 1997, 2008), there is some evidence to suggest that Mandarin-learning infants learn word–action associations earlier than word–object associations, whereas English-learning infants in the same experimental paradigm learn word–object associations earlier than word–action associations (C. C. Chan et al., 2011). However, the C. C. Chan et al. study did not separate infants' ability to associate words to actions or objects from their attentional patterns to the actions and objects in scenes. Thus, it was not clear whether attention alone might have contributed to this effect or if it was the specific association of words to concepts that resulted in the cross-linguistic differences.

The findings from the present studies not only clarify and complement those earlier findings, they contribute to a broader theoretical framework of word learning and fit well with the view that the developmental process of word learning happens across multiple levels, each of which may have its own developmental trajectory and pattern of cross-linguistic differences. Specifically, the predominantly universal attentional patterns between the English- and Mandarin-learning infants found in the current studies provide evidence that cultural differences in mapping words to actions and objects occur at a level that is beyond attention per se, at least with very young children. Nonetheless, even at 17–19 months of age, specific language experience does play some role, but only when children are given a linguistic context, even the

minimal linguistic context of a single nonsense word that has phonotactic properties consistent with words in the child's native language.

In general, the findings from present study provide strong evidence for developmental changes and much weaker evidence, but some interesting points for further investigation, regarding cross-cultural differences in infants' attentional patterns. Developmentally, we found significant changes in attentional patterns that cut across the two languages such that 6–8-month-olds could discriminate different actions by showing greater attention to scenes that showed changes in the action, relative to a previously presented scene (control), but showed no difference in attention to scenes in which different objects appeared. In contrast, 17–19-month-olds paid attention to both of the change trial conditions (action and object). These data echo other findings (Pulverman et al., 2008) that English-learning children certainly do not suffer from an inability to notice actions or changes in actions in dynamic events. Even English-learning 6-month-olds could discriminate—and, in fact, only discriminated—between contrasting actions in the current studies. Rather, in the present studies, as in previous work (e.g., Bahrick et al., 2002), young infants appeared to have trouble discriminating static objects in a dynamic scene—perhaps because these were simply less salient than the actions or because moving objects and agents are more primary and requiring of attention than are stationary ones. Older infants may have succeeded because they had more attention to allocate to the event components and were, therefore, able to discriminate actions as well as objects or because they had simply had greater exposure to a variety of objects in their environments that are worthy of exploration, and, thus, objects had become more salient.

The Role of Language Experience in Attention to Actions Versus Objects

Despite these commonalities in attention, we also found evidence supporting the role of specific language experience in attention when even minimal linguistic stimuli were presented at the same time as the complex scenes. First, a cross-cultural difference was found in attentional patterns to the aspects of a scene attended to in the presence of a verbal label for the older infants. Specifically, 17–19-month-old Mandarin-learning infants paid more attention to the action changes than did 17–19-month-old English learners, even though learners from both cultures could reliably discriminate both of these components by this age. Because no external or contextual cues were provided in this experiment, the cross-cultural differences in attentional patterns manifested here can only be attributed to implicit heuristics based on children's language-learning experiences. Thus, when linguistic labels were offered in Study 2, older (17–19-month old) English-learning infants, whose vocabularies are dominated by nouns, showed less sensitivity to the action changes than they did when these same scenes were presented in silence. In contrast, Mandarin-learning infants showed a nonsignificant but interesting increase in attending to both the object and the action changes when these data were compared with those of infants who were presented with the same scenes in the absence of a verbal label (Study 1). In both English and Mandarin, the strong attentional pull of actions was clearly moderated when the exact same scenes were presented together with a verbal label. Although these data are only suggestive, and Study 2 offered only a very minimal linguistic context (a single word), it is possible that linguistic labels may influence children's attention not only by directing it toward items that are frequently labeled in a language but also by directing it away from elements (e.g., actions) that are less frequently labeled.

Word Learning Is a Developmental Process at Multiple Levels: Attention Versus Mapping

The findings at the attentional level are consistent with and complement other findings in the literature. Specifically, in C. C. Chan et al.'s (2011) study, 18-month-old Mandarin-learning infants could establish an association between novel words and actions but not between novel words and objects. In the current study, 17–19-month-old Mandarin-learning infants succeeded in discriminating different actions in the dynamic event and looked significantly longer at action changes than did English-learning infants when under the spotlight of linguistic labels. Such an attentional pattern parallels more general cross-linguistic findings on early word learning and provides further evidence of the relationship between cognitive and language development (Casasola, Bhagwat, & Ferguson, 2006; Mandler, 2006).

However, despite this ability to discriminate and parse actions, English-learning infants are generally unable to associate words to actions in the same types of experimental word-learning studies in which they are able to associate words to objects (C. C. Chan et al., 2011; Childers & Tomasello, 2002). This suggests that the differential ease of acquiring different types of words across languages might in fact be attributed more to the mapping process than to attentional processes on their own (Gentner & Boroditsky, 2001; Waxman & Lidz, 2006). C. C. Chan et al. (2011) argued that the difficulty Mandarin learners demonstrated in their object condition was less likely to be about encoding the two objects and was more likely to reflect Mandarin learners' readiness to form mappings between objects and labels during the habituation phase (pp. 1466–1467). Importantly, the current studies provided further evidence for this interpretation. In particular, the findings that 17–19-month-old children in both language groups discriminated both the actions and the objects may indicate that, by this age, children have no difficulty in encoding these components.

Limitations and Conclusions

There are some limitations in the current study that lead one to wonder whether there might still be more subtle attentional differences across children exposed to these two very different types of linguistic environments. First, although it is clear that the 6–8-month-olds in our studies were not at ceiling on the task, the 17–19-month-olds may have been at ceiling for this task. This is important, because if language experience is relevant to attention, even in the absence of a linguistic stimulus, we would expect to see stronger effects for infants who have had more linguistic experience, and, thus, we would need to test the hypothesis in older, rather than younger, children. Thus, it is important to ensure that the older children were not performing at ceiling on the task. Bahrick et al.'s (2002) study examined 5.5-month-olds' discrimination and memory for events in which women were performing different activities but included both 1-min and 7-day delays for test trials. The authors found that infants could discriminate and

memorize the actions but not the faces. Our study included no delays and examined 6–8- and 17–19-month-olds, with both age groups older than those examined in the Bahrick et al. experiments. Thus, one way of looking at whether more subtle attentional differences will show up in a harder task would be to add a delay to the test trials, as was done by Bahrick et al., or to examine more subtle comparisons across conditions. For instance, instead of simply testing changes to a single element, one could examine changes of various types and have multiple test trials that specifically contrast whether participants attended to person, action, or object changes. To examine these more subtle changes, one could also examine additional trials that included, for instance, both person and action changes with the same object, person and object changes with the same action, and so on to see which specific changes (in one- vs. two-change events) resulted in the greatest increases in looking. Although this goes beyond the scope of the present research, if infants detect actions and objects as independent elements of an event, then they should notice changes in action regardless of whether the person or the object has changed, and they should notice changes in object regardless of whether the person or the action has changed.

In addition, it is not clear that 17–19-month-olds are the only appropriate older children to examine for this question. Although they clearly have much more linguistic experience than do 6–8-month-olds, it may be that English- and Mandarin-learning 17–19-month-olds experience more similarities in their language-learning environments than do younger children. Indications that this may be the case come from both case studies of the adult-to-child speech of individual English- and Mandarin-learning children from 6 to 18 months of age (Chen et al., 2009) and large-scale studies of English- and Mandarin-learning children's vocabularies on parent-reported vocabulary checklists (Bornstein et al., 2004; Tardif et al., 2008). In both types of data, the cross-linguistic differences are greatest when children are at the earliest stages of producing words (i.e., 10–14 months or even the first 10 words) compared with when they have 50 or more words in their vocabularies (usually around 16–18 months). This is an interesting hypothesis, although it is also true that even the children who were below the median on overall vocabulary at 17–19 months showed exactly the same patterns of attention as those who were above the median. Thus, if there were any differences in an intermediate age group (e.g., 12–14 months), it would be important to tease apart the effects of vocabulary per se from those of general cognitive development and amount of exposure to the ambient language. It could also be that even older children and adults are more sensitive to the effects that the ambient language have had on their attention to objects versus actions in a complex scene. To test this hypothesis, however, a different paradigm would need to be used such that the close-to-ceiling effects observed in the current 17–19-month-olds' performance would no longer be present.

Finally, the experimental stimuli and paradigm used in this study were ideal for presenting maximally identical stimuli to infants across a variety of ages and two very different cultures. As such, there would be minimal effect of experimenter variance as a function of personality, cultural, or age-related concerns. However, habituation methods with video stimuli for this age group have very high attrition rates and run the risk of being difficult to translate to more naturalistic contexts. Moreover, if the children who do not successfully complete such studies (e.g., through a failure to habituate, through excessive fussiness) are systematically different from those who do complete the studies, then there is a risk that the conclusions drawn from infants who complete the studies are not representative of a larger population. Although both of these issues are real limitations with the paradigm, the primary focus in the current studies was on cross-linguistic and cross-age comparisons. Given that there were no significant differences in the numbers of infants who did not complete the studies across language groups, it is unlikely that these issues had much impact on such comparisons.

Contrary to our initial hypotheses, the second of the present studies did not provide strong evidence that a bare word can direct attention to the specific elements involved in an event. And yet, there was clear evidence that even this minimal stimulus did change the way in which older infants attended to complex scenes. However, the changes were neither entirely consistent nor entirely inconsistent with the pattern of vocabulary learning that children from English- and Mandarin-speaking backgrounds have shown. One explanation for the difference between our findings and the overall word-learning patterns is that although words may help direct infants' attention to the commonalities of isolated objects, the effect may drop off in the face of dynamic events that involve more than one element—hence, we did not see strong increases in attention to objects that were most predominant in the English-learning children. Another possibility could be that although words may facilitate attending and mapping to actions for Chinese children versus to objects for English children, changing only the elements (and not the labels) did not surprise the infants as much as an incorrect mapping would have (per C. C. Chan et al., 2011). Future studies might examine this possibility by introducing a word, or perhaps a full sentence, in test trials as well as in habituation trials (although not the target word presented in habituation or they would become word-mapping experiments). Relevant to this, although the bare words used in the Study 2 failed to drive attention, it is possible that a richer linguistic context (e.g., full sentences, labels with morphological information) would exert a stronger influence on attention. Thus, yet another possibility for future studies would be to further investigate the role of minimal versus richer linguistic contexts on attention.

In conclusion, the present studies offer insight into how English- and Mandarin-learning infants' attentional patterns contribute to the process of word learning. Our findings revealed strong developmental changes, clear effects of adding a word on attentional preferences in older infants, and only very subtle cross-cultural differences (only in the presence of novel labels for 17–19-month-olds) in infants' attentional patterns. This evidence reinforces, and yet limits the range of hypotheses available to explain, previous findings demonstrating cross-cultural differences in children's word comprehension, mapping, and production as well as caregivers' emphases on nouns and verbs. Taken together, the current findings suggest the importance of taking an integrative approach to word learning and to considering the learning process as occurring at multiple levels (attentional patterns, mapping, comprehension, production, and—ultimately—much deeper comprehension as complex patterns of meaning are constructed for synonyms and related words) and the possibility of linguistic experience entering at any of these levels and at various points in development.

References

Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development, 73,* 1629–1643. http://dx.doi.org/10.1111/1467-8624.00495

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64,* 3–26. http://dx.doi.org/10.1006/jecp.1996.2332

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., . . . Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21,* 85–123. http://dx.doi.org/10.1017/S0305000900008680

Bloom, L., Tinker, E., & Margulis, C. (1993). The words children learn: Evidence against a noun bias in early vocabularies. *Cognitive Development, 8,* 431–450. http://dx.doi.org/10.1016/S0885-2014(05)80003-6

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., . . . Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75,* 1115–1139. http://dx.doi.org/10.1111/j.1467-8624.2004.00729.x

Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development, 76,* 279–290. http://dx.doi.org/10.1111/j.1467-8624.2005.00844.x

Casasola, M., & Bhagwat, J. (2007). Do novel words facilitate 18-month-olds' spatial categorization? *Child Development, 78,* 1818–1829. http://dx.doi.org/10.1111/j.1467-8624.2007.01100.x

Casasola, M., Bhagwat, J., & Burke, A. S. (2009). Learning to form a spatial category of tight-fit relations: How experience with a label can give a boost. *Developmental Psychology, 45,* 711–723. http://dx.doi.org/10.1037/a0015475

Casasola, M., Bhagwat, J., & Ferguson, K. T. (2006). Precursors to verb learning: Infants' understanding of motion events. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 160–190). New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195170009.003.0007

Casasola, M., & Cohen, L. B. (2000). Infants' association of linguistic labels with causal actions. *Developmental Psychology, 36,* 155–168.

Casasola, M., & Cohen, L. B. (2002). Infant categorization of containment, support and tight-fit spatial relationships. *Developmental Science, 5,* 247–264. http://dx.doi.org/10.1111/1467-7687.00226

Chan, C. C., Tardif, T., Chen, J., Pulverman, R. B., Zhu, L., & Meng, X. (2011). English- and Chinese-learning infants map novel labels to objects and actions differently. *Developmental Psychology, 47,* 1459–1471. http://dx.doi.org/10.1037/a0024049

Chan, W. H., & Nicoladis, E. (2010). Predicting two Mandarin–English bilingual children's first 50 words: Effects of frequency and relative exposure in the input. *International Journal of Bilingualism, 14,* 237–270. http://dx.doi.org/10.1177/1367006910363059

Chen, J., Tardif, T., Meng, X., & Setoh, P. (2009). The influence of adult input on children's early word learning: A case study of a Mandarin-speaking child. *Acta Psychologica Sinica, 41,* 715–725. http://dx.doi.org/10.3724/SP.J.1041.2009.00715

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38,* 967–978. http://dx.doi.org/10.1037/0012-1649.38.6.967

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language, 26,* 207–232. http://dx.doi.org/10.1177/0142723706060748

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). *Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0)* [Computer software]. Austin: University of Texas.

Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. *Current Directions in Psychological Science, 11,* 196–200. http://dx.doi.org/10.1111/1467-8721.00199

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., . . . Reilly, J. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, cognition and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511620669.010

Gliga, T., Volein, A., & Csibra, G. (2010). Verbal labels modulate perceptual object processing in 1-year-old children. *Journal of Cognitive Neuroscience, 22,* 2781–2789. http://dx.doi.org/10.1162/jocn.2010.21427

Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word comprehension: An ecological–dynamic systems view. *Developmental Science, 4,* 1–18. http://dx.doi.org/10.1111/1467-7687.00143

Göksun, T., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Konishi, H., & Okada, H. (2011). Who is crossing where? Infants' discrimination of figures and grounds in events. *Cognition, 121,* 176–195. http://dx.doi.org/10.1016/j.cognition.2011.07.002

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62,* 169–200. http://dx.doi.org/10.1016/S0010-0277(96)00784-6

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development, 76,* 340–355. http://dx.doi.org/10.1111/j.1467-8624.2005.00849_a.x

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development, 79,* 979–1000. http://dx.doi.org/10.1111/j.1467-8624.2008.01171.x

Kannass, K. N., & Oakes, L. M. (2008). The development of attention and its relations to language in infancy and toddlerhood. *Journal of Cognition and Development, 9,* 222–246. http://dx.doi.org/10.1080/15248370802022696

Katerelos, M., Poulin-Dubois, D., & Oshima-Takane, Y. (2011). A cross-linguistic study of word-mapping in 18- to 20-month-old infants. *Infancy, 16,* 508–534. http://dx.doi.org/10.1111/j.1532-7078.2010.00064.x

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development, 3,* 179–197. http://dx.doi.org/10.1080/15475440701360168

Lee, J. N., & Naigles, L. R. (2005). The input to verb learning in Mandarin Chinese: A role for syntactic bootstrapping. *Developmental Psychology, 41,* 529–540. http://dx.doi.org/10.1037/0012-1649.41.3.529

Levey, S., & Cruz, D. (2003). The first words produced by children in bilingual English/Mandarin Chinese environments. *Communication Disorders Quarterly, 24,* 129–136. http://dx.doi.org/10.1177/15257401030240030401

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language, 36,* 405–423. http://dx.doi.org/10.1017/S0305000908009008

Mandler, J. M. (2006). Actions organize the infant's world. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 111–113). New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195170009.003.0005

Pulverman, R. (2005). *The relationship between language development and event processing: Lexical acquisition and attention to manner and path*. Unpublished doctoral dissertation, University of Delaware.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108,* 825–830. http://dx.doi.org/10.1016/j.cognition.2008.04.009

Pulverman, R., Tardif, T., Rohrbeck, K. L., & Chen, P. (2010, March). *The effect of specificity on verb acquisition: Evidence from English and Mandarin Chinese*. Poster session presented at the 17th Biennial International Conference on Infant Studies, Baltimore, MD.

Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27,* 561–585. http://dx.doi.org/10.1017/S0305000900004256

Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology, 35,* 1478–1492. http://dx.doi.org/10.1037/0012-1649.35.6.1478

Snedeker, J., Li, P., & Yuan, S. (2003, July). *Cross-cultural differences in the input to early word learning*. Paper presented at the 15th Annual Conference of the Cognitive Science Society, Mahwah, NJ.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32,* 492–504. http://dx.doi.org/10.1037/0012-1649.32.3.492

Tardif, T. (2006). But are they really verbs? Chinese words for action. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 477–498). New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195170009.003.0019

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology, 44,* 929–938. http://dx.doi.org/10.1037/0012-1649.44.4.929

Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Zuo, Q. H. (2009). *The Chinese Communicative Development Inventory (Putonghua and Cantonese versions): Manual, forms, and norms*. Beijing, China: Peking University Medical Press.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development, 70,* 620–635. http://dx.doi.org/10.1111/1467-8624.00045

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24,* 535–565. http://dx.doi.org/10.1017/S030500099700319X

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition, 70*(3), B35–B50. http://dx.doi.org/10.1016/S0010-0277(99)00017-7

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. J. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives, 7,* 155–159. http://dx.doi.org/10.1111/cdep.12032

Waxman, S. R., & Lidz, J. (2006). Early word learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 299–335). Hoboken, NJ: Wiley.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29,* 257–302. http://dx.doi.org/10.1006/cogp.1995.1016

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34,* 1289–1309.

Yu, C., Ballard, D. H., & Aslin, R. N. (2005). The role of embodied intention in early lexical acquisition. *Cognitive Science: A Multidisciplinary Journal, 29,* 961–100. http://dx.doi.org/10.1207/s15516709cog0000_40

(*Appendices follow*)

Appendix A

Examples of the Contrasted Objects and Actions From the Eight Dynamic Events and the Posttest Stimuli

We received signed consent from the women depicted in Figure A1 for their likenesses to be published in this article.

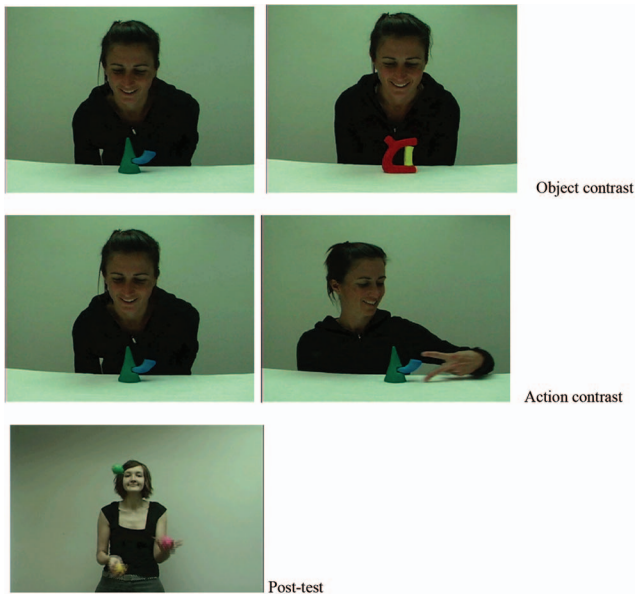

Figure A1. Example stimuli. Contrasted objects included a green inverted cone with blue side tail (top left) and a red *C* with a stub and a yellow brace (top right). Contrasted actions included using the chin in a vertically descending motion to lightly touch the top of an object (middle left) and horizontally approaching the side of an object with the left hand, using two fingers extended like a pair of scissors (middle right). The posttest (bottom) was a 30-s video of a young woman juggling three colored balls. See the online article for the color version of this figure.

Appendix B

Early Language Questionnaires for English- and Mandarin-Learning Infants

These questionnaires were used to determine that 6–8-month-old participants appeared to be developing normally. The English (see Figure B1, top) and Mandarin versions (see Figure B1, bottom) contained 10 items in common: four on producing sounds, three on understanding first signs, and three on understanding short sentences.

(*Appendices continue*)

EARLY LANGUAGE QUESTIONS
(6-month-olds)

Before children start to talk, they begin to produce "baby talk" sounds and react to some words and phrases. Below is a list of common examples. Please answer for each question whether or not your child has begun to produce the reaction listed below.

Figure B1.

Infant Sounds Produced		
1. Makes "aaa" or "ooo" sounds?	No	Yes
2. Makes "ba" or "da" sounds?	No	Yes
3. Makes "ma" sounds?	No	Yes
4. Makes "ba ba ba" or "da da da" sounds?	No	Yes
First Signs of Understanding		
5. Respond when name is called. (e.g., by turning and looking at source)	No	Yes
6. Respond to "no" or "don't" (by stopping what he/she is doing, at least for a moment).	No	Yes
7. React to "there's mommy/daddy" by looking around for them.	No	Yes
Infant Early Understandings of Sentences		
8. Daddy's/mommy's home.	No	Yes
9. Want up?/Up?/Want me to hold you.	No	Yes
10. Come here/come on.	No	Yes
11. Look/look here.	No	Yes
12. Don't touch.	No	Yes
13. Open your mouth.	No	Yes

初期语言发展问卷

小孩子开始讲话之前，会开始发一些咿咿语的音和对一些字句有反应。以下有一些普通的例子。请对每项问题回答如果您的孩子有没有像以下的反应？

婴儿的发音		
1. 发出"啊"或"呜"的声音？	No	Yes
2. 发出"爸"或"大"的声音？	No	Yes
3. 发出"妈"的声音？	No	Yes
4. 发出"爸爸爸"或"大大大"的声音？	No	Yes
婴儿的初期语言理解		
5. 叫他的名字时会有反应（例如：转向及看声音来源的方向）	No	Yes
6. 别人说"别"时会有反应（例如：暂停做某件事）	No	Yes
7. 当听到"妈妈/爸爸在哪里？"会向周围找。	No	Yes
婴儿初期听懂的句子（有适当的反应）		
8. 抱抱	No	Yes
9. 看看	No	Yes
10. 别动	No	Yes
11. 张嘴	No	Yes

Appendix C

Audio Stimuli in the Word Condition for Both Cultures in Study 2

Language	Habituation phase	Test phase
English	"Sug, wow, sug, ooh . . ."	"Wow, ooh . . ."
Mandarin	"Fai1, wa, fai1, ou . . ."	"Wa, ou . . ."

Note. All stimuli were looped five times.

Received May 21, 2013
Revision received August 18, 2014
Accepted November 6, 2014 ■