Applied Psycholinguistics **40** (2019), 555–583
doi:10.1017/S014271641800070X

How does language proficiency affect children's iconic gesture use?

MEGHAN ZVAIGZNE
McGill University

YURIKO OSHIMA-TAKANE
McGill University and University of Victoria

MAKIKO HIRAKAWA
Chuo University

Received: September 12, 2017 Revised: October 4, 2018 Accepted: October 5, 2018

ADDRESS FOR CORRESPONDENCE
Yuriko Oshima-Takane, Department of Psychology, University of Victoria, Victoria, British Columbia, P. O. Box 1700 STN CSC, V8W 2Y2, Canada. E-mail: yuriko.oshima-takane@mcgill.ca

ABSTRACT
Previous research investigating the relationship between language proficiency and iconic gesture use has produced inconsistent findings. This study investigated whether a linear relationship was assumed although it is a quadratic relationship. Iconic co-speech gesture use by 4- to 6-year-old French–Japanese bilinguals with two levels of French proficiency (intermediate and low) but similar levels of Japanese proficiency was compared with that of high-proficiency French monolinguals (Study 1) and Japanese monolinguals with similar proficiency to the bilinguals (Study 2). To control the information participants communicated, a dynamic referential communication task was used; a difference between two cartoons had to be communicated to an experimenter. Study 1 showed a significant quadratic relationship between proficiency and iconic gesture use in French; the intermediate-proficiency bilinguals gestured least among the three proficiency groups. The monolingual and bilingual groups with similar Japanese proficiency in Study 2 gestured at similar rates. It is suggested that children gestured for different reasons depending on their language proficiency and the cognitive resources available for the task.

Keywords: bilingual; children; French; iconic gestures; Japanese; language proficiency; monolingual

People gesture spontaneously when speaking by moving their hands, arms, and sometimes other body parts. This is true of children and adults who speak different languages and come from different cultures (McNeill, 1992). According to McNeill (1992), there are four types of co-speech gestures: iconics, metaphorics, deictics, and beats. Iconic gestures visually represent aspects of concrete events and objects. For example, when describing a Slinky® toy, a speaker may make a circular hand shape and move it along a path in semicircles to demonstrate it jumping. Metaphoric gestures look like iconics; however, they depict abstract concepts concretely (e.g., a speaker pushes down on his or her shoulders while describing how a big project feels like a weight on his or her shoulders). Deictics (points) indicate objects

or locations, and beats are simple rhythmic movements produced in time with speech or for emphasis. The present study focused on iconic gestures.

Researchers have examined factors that affect speakers' use of iconic co-speech gestures, including the frequently examined factor of language proficiency. Despite extensive examination, the relationship between language proficiency and iconic gesture use remains unclear due to inconsistent findings. The present study investigated how language proficiency influences children's use of iconic co-speech gestures in French and Japanese by controlling for factors that might have contributed to the inconsistent findings.

Given that iconic gestures can provide supplemental information about concrete objects and events through visual representation, one might expect that speakers' use of them relates to their language proficiency. However, two opposing views of the nature of this relationship exist. One view holds that speakers with low proficiency will use iconic gestures frequently when they have difficulty with verbal expression (Alibali, Kita, & Young, 2000; Krauss & Hadar, 1999; Nicoladis, 2007). For example, speakers might have problems accessing words, formulating utterances, or expressing the intended information clearly and unambiguously. If speakers use iconic gestures when they encounter difficulty expressing themselves, it can help overcome the problem. That is, gesturing can help retrieve words (Hadar, Dar, & Teitelman, 2001; Rauscher, Krauss, & Chen, 1996), compensate when specific words do not readily come to mind or are unavailable (Kita & Özyürek, 2003), organize information into well-structured utterances (Alibali et al., 2000; Kita, 2000), and/or clarify or disambiguate an unclear message (Holler & Beattie, 2003; Kidd & Holler, 2009; So, Demir, & Goldin-Meadow, 2010). Iconic gestures might serve these cognitive and compensatory roles because they are visual, concrete, and communicative (Hostetter, 2011; McNeill, 1992). Highly proficient speakers who presumably do not encounter such difficulties as often as less proficient speakers might not use iconic gestures to the same extent or for the same reasons.

The contrasting view is that speakers with high language proficiency use iconic gestures more often than those with low proficiency (Gullberg, 1999; Nicoladis, 2002; Nicoladis, Mayberry, & Genesee, 1999). One explanation for this is that iconic gestures can help communicate complex information, and highly proficient speakers are more likely than less proficient speakers to attempt to express complex ideas (Nicoladis, 2002; Nicoladis et al., 1999). In this scenario, gesturing can help speakers organize and structure their ideas for expression. While iconic gesture use in this case is like that described for the opposing view (i.e., to aid the expression of information), for high-proficiency speakers it is the level of complexity of their intended message that motivates the use of iconic gesture, not the lack of language proficiency. In a similar vein, proficient speakers may gesture with the goal of communicating important information via the visual channel (Hostetter & Alibali, 2007; Kendon, 1980; McNeill & Duncan, 2000). For instance, a dress's silhouette might be communicated more clearly via iconic gesture than via speech. In addition, because proficient speakers can speak fluently, they could devote some cognitive resources to gesturing with the aim of engaging or entertaining their addressee (Hostetter, 2011; Hostetter & Alibali, 2007; Kendon, 1994, 2001).

Researchers have studied iconic co-speech gesture use in bilinguals because they often have unequal exposure to and, thus, unequal proficiency in each language (e.g., Mayberry & Nicoladis, 2000). Bilinguals can provide unique insights about gesture use because they act as their own controls. Performance in each language can be contrasted while other factors (e.g., intelligence and personality) are controlled for. When bilinguals have unequal proficiency, iconic gesture use can be compared across their dominant and nondominant languages. When they are equally proficient in each language, similar gesture use in both languages is expected, provided that other factors are controlled for. To date, only Nicoladis, Pika, Yin, and Marentette (2007) found that bilingual adults gestured more in their nondominant than their dominant language; these were English–Chinese bilinguals. The authors suggested that iconic gestures were used to aid lexical access when language proficiency was low. Other studies of bilinguals have found the opposite, that child and adult bilinguals use more iconic gestures when using their dominant than their nondominant language (English–Spanish adult bilinguals in Gregersen, Olivares-Cuhat, & Storm, 2009; French–Swedish adult bilinguals in Gullberg, 1999; French–English child bilinguals in Nicoladis, 2002 and Nicoladis et al., 1999). Gregersen et al. (2009) speculated that the bilinguals in their study intended to provide redundant information in speech and iconic gesture in their dominant language to avoid ambiguity. Nicoladis and her colleagues (Nicoladis, 2002; Nicoladis et al., 1999) suggested that bilingual children's iconic gestures reflected the more complex information conveyed by more complex utterances (operationalized as mean length of utterance) produced in the dominant language compared to the less complex utterances in the nondominant language. Based on Gullberg's finding that bilingual adults gestured more in their dominant than their nondominant language during a story-retelling task, Kendon (2001) suggested that speakers often want to create a vivid and interesting experience for the listener. They can do this in their dominant language because they are linguistically fluent and do not have to focus on finding words and producing grammatically appropriate sentences. They may draw on other expressive resources to supplement and complement the meaning conveyed in speech.

Other bilingual studies failed to show a link between proficiency level and iconic gesture use in each language. That is, rates of iconic gesture use were similar across both languages (Arabic–English child bilinguals in Abdalla, 2015; English–Spanish adult bilinguals in Marcos, 1979; English–Hindi adult bilinguals in Nagpal, Nicoladis, & Marentette, 2011; French–English child bilinguals in Nicoladis, Pika, & Marentette, 2009; English–Spanish adult bilinguals in Sherman & Nicoladis, 2004). Some researchers have also correlated gesture rates in each language with measures of language proficiency (e.g., Peabody Picture Vocabulary Test scores, a measure of receptive vocabulary in Nicoladis, 2002; number of word types used in Sherman & Nicoladis, 2004). With one exception (a positive correlation for English–Spanish adult bilinguals using their native language of English in Sherman & Nicoladis, 2004), no studies have reported significant correlations (Nagpal et al., 2011; Nicoladis, 2002; Nicoladis et al., 2009; Sherman & Nicoladis, 2004).

A few studies investigated the relationship between language proficiency and iconic gesture use in monolinguals, also with inconsistent results. Colletta, Pellenq, and Guidetti (2010) found that as French-speaking children's narrative skills developed with age, their use of iconic gestures increased. However, Frick-Horbury and Guttentag (1998) found no difference in iconic gesture rate between high- and low-proficiency groups of English-speaking adults, while Hostetter and Alibali (2007) found that groups of high- and low-proficiency English-speaking adults gestured more than a group with intermediate proficiency.

Based on these conflicting findings from bilingual and monolingual research, iconic gesture use does not relate to language proficiency in a simple linear way. Rather, iconic gesture use may depend on several factors such that the relationship with language proficiency is complex. The extant studies differ with respect to several important variables related to language use that may themselves affect gesture use, and these differences could have contributed to the conflicting results. One factor that might account for the inconsistent relationship is the nature and difficulty of the tasks used to elicit speech and iconic gestures (Nicoladis, 2007; Nicoladis et al., 2007). A variety of tasks have been used, and several allowed participants considerable freedom in what to talk about (e.g., story-retelling in Gullberg, 1999; free play in Nicoladis, 2002). For example, if bilinguals do not feel comfortable talking about something in one language because they lack relevant words or expressions, they may omit the information altogether (Gregersen et al., 2009). As a result, the speech samples obtained in each language may differ in content, detail, and complexity. This could, in turn, have affected gesture rates. Consequently, the comparisons that have been made across languages may not have been based on comparable data.

Another potential confound in the bilingual studies is that comparisons were made across different languages (e.g., English, French, and Spanish). Several studies have found cross-linguistic differences in the frequency and way that iconic gestures are used (e.g., Efron, 1972; Kita & Özyürek, 2003; McNeill & Duncan, 2000). For instance, Kita and Özyürek (2003) reported that iconic gesture use was higher by Turkish and Japanese speakers than English speakers describing a cartoon cat rolling down a hill in a story-retelling task. Many Turkish and Japanese speakers expressed the cat's rolling movement and downward trajectory in two separate clauses (e.g., the cat is descending the hill while rolling) and produced two separate gestures conveying manner and trajectory of movement. In contrast, many English speakers described the same scene in one clause (e.g., the cat is rolling down a hill) and produced one iconic gesture combining manner and trajectory information. The authors suggested that this cross-linguistic difference was found because English is a language typologically different from Turkish and Japanese for expressing manner and trajectory of movement. Such findings on cross-linguistic differences in gesture use suggest that bilinguals may use gestures at different rates when speaking in each of their languages, independent of their proficiency level in each language. Thus, it is important to consider the languages spoken when conducting bilingual studies. Ideally, comparisons across groups with different levels of proficiency should be made within a language, such as the comparison of speakers at high, intermediate,

and low proficiency levels that Hostetter and Alibali (2007) did with English monolingual speakers.

Several issues regarding language proficiency may have also contributed to the inconsistent findings from bilingual studies. First, the participants varied in their relative levels of language proficiency. Studies with unbalanced bilinguals generally found a proficiency-related difference in gesture rate (e.g., Gregersen et al., 2009; Gullberg, 1999; Nicoladis, 2002), except for Marcos (1979), while studies with relatively balanced or advanced bilinguals have not (Nagpal et al., 2011; Nicoladis et al., 2009; Sherman & Nicoladis, 2004). These differences between unbalanced and balanced bilinguals would be expected, however. When bilinguals differ in their level of proficiency in each language, and if any potential language-specific differences in iconic gesture use are controlled for, they should gesture at different rates when speaking each language. In contrast, when bilinguals have a similar level of proficiency in each language and there are no language-specific differences in iconic gesture use, they should gesture at similar rates when using each language.

Second, exactly what constituted high or low language proficiency in these studies is unknown because researchers did not always provide clear definitions of proficiency (Gullberg, 2012), and proficiency measures varied extensively across studies (e.g., an assessment by a panel of native bilingual speakers in Gullberg, 1999; the Peabody Picture Vocabulary Test, a measure of receptive vocabulary in Nicoladis, 2002; and mean length of utterance in Nicoladis et al., 1999). Thus, in the extant bilingual research, the term *language proficiency* refers to different kinds of language abilities, and this might have contributed to the inconsistent findings. Furthermore, several language proficiency measures such as number of word types and mean length of utterance are problematic in that they were not independent of the task in which gesture rates were examined. That is to say, proficiency was assessed by how the participant did on the task, but this was also related to how much he or she gestured while doing the task. An independent measure of proficiency should be used because task difficulty itself has been shown to influence gesture rates (Gullberg, 2012; Nicoladis et al., 2007). The Peabody Picture Vocabulary Test used by Nicoladis (2002) is an independent measure of vocabulary; however, it assesses vocabulary comprehension and does not necessarily relate to iconic gesture use, as shown in her study. It seems that the most appropriate measure of language proficiency should assess productive language rather than language comprehension because spoken language and gesture production are of particular interest.

Given the variation across studies in task demands, languages spoken, and type and level of language proficiency, it is not surprising that a clear link between language proficiency and iconic gesture use has yet to emerge. The present study was designed to address these issues and provide a more stringent evaluation of how proficiency and iconic gesture use are related in children compared to previous studies. More specifically, we conducted two studies with French–Japanese bilingual and French or Japanese monolingual children. This language combination has not been investigated in previous bilingual studies of iconic gestures.

In Study 1, we investigated how language proficiency affects children's iconic gesture use by comparing three groups of children who differed in their level of French proficiency. The groups were (a) French monolinguals who had a high level of French proficiency, (b) French–Japanese bilinguals who were dominant in French and had an intermediate level of French proficiency, and (c) French–Japanese bilinguals who were dominant in Japanese and had a low level of French proficiency. All groups were similar in age to control for level of cognitive development. We made within-language comparisons across three groups of children (rather than just two) to avoid a language-specific confound should there be differences in gesture use that are specific to French or Japanese. All previous studies, with the exception of Hostetter and Alibali (2007), only compared gesture use across two different or two similar levels of proficiency. Hostetter and Alibali compared English monolingual adults at three different levels of proficiency and found a quadratic relationship between proficiency and gesture use. The high- and low-proficiency groups gestured more than the intermediate-proficiency group. Because language proficiency increases with age, it is difficult to find three different groups of monolingual children who are similar in age but differ in language proficiency. Thus, to obtain three different language proficiency groups of similar age, we included monolingual and bilingual children. This allowed us to examine whether language proficiency and iconic gesture use have a linear relationship, a quadratic relationship, or no relationship.

In Study 2, the Japanese data from the bilingual groups in Study 1, as well as a group of Japanese monolinguals, were analyzed to compare iconic gestures produced by three groups of children who did not differ in their level of Japanese proficiency: (a) Japanese monolinguals, (b) French–Japanese bilinguals who were dominant in Japanese, and (c) French–Japanese bilinguals who were dominant in French. It was expected that these three groups with similar levels of proficiency in Japanese would not differ in gesture use if language proficiency influences gesture use.

To examine how much children use iconic gestures when describing information they need to communicate, we used a referential communication task in which children were asked to describe one of two animated cartoons to the experimenter. Unlike tasks such as story-retelling or free play that typically do not require that specific information be communicated, the referential communication task we used allowed us to control for the amount and type of verbal descriptions. At the same time, it makes fewer demands on verbal abilities than a task like story-retelling, which also requires narration skills. Thus, we could include bilingual children with relatively low levels of proficiency in one language.

As a measure of language proficiency, we used a productive vocabulary test independent of the referential communication task. We operationalized language proficiency as lexical proficiency, referring to the level of vocabulary knowledge that the children had, because vocabulary knowledge includes vocabulary size, accessibility of known words, and the amount of semantic knowledge associated with known words (Crossley, Salsbury, McNamara, & Jarvis, 2010). It seems

reasonable to assume that the children we studied would use iconic gestures to aid lexical access and retrieval when speaking because they are in the midst of acquiring their lexical skills. The bilinguals in particular may have had some difficulty with this owing to reduced exposure to each of their languages (Gollan, Montoya, Cera, & Sandoval, 2008). There is some evidence that adult English monolinguals use iconic gestures to aid lexical retrieval; Rauscher et al. (1996) found that when describing animated cartoons, speech with spatial information, but not with nonspatial information, became less fluent when speakers were prevented from gesturing. Dysfluencies were within clauses, so they suggested that the participants had difficulties with lexical access. Hostetter and Alibali (2007) used phonemic fluency and semantic fluency as proficiency measures and found that only phonemic fluency was related to iconic gesture use. Phonemic fluency indicates how efficiently one can organize and navigate his or her lexicon, skills used online when speaking to plan what to say next, whereas semantic fluency indicates one's efficiency at lexical access and retrieval. In the case of their adult participants, the high- and low-proficiency speakers used iconic gestures to help organize their ideas for expression rather than to help recall specific words to use and may not have gestured to aid lexical retrieval. However, semantic fluency is likely a more appropriate index of language proficiency for our referential communication task and child participants. At a minimum, they had to communicate one feature of an animal that could be expressed with one word, and grammatically correct sentences were not necessary. Thus, phonemic fluency might not be an appropriate measure of language proficiency. In addition, explicit links between iconic gesture use and semantic fluency have yet to be demonstrated in children.

STUDY 1

In this study we hypothesized a quadratic relationship between language proficiency and iconic gesture use. We predicted that the groups with high and low proficiency in French would gesture more than the group with intermediate proficiency, similar to what Hostetter and Alibali (2007) found with English-speaking monolingual adults.

Method

Participants. Three groups of children, aged 4 years, 2 months to 6 years, 7 months, participated in this study. They were: (a) 15 French monolinguals, (b) 9 French–Japanese bilinguals who were dominant in French, and (c) 8 French–Japanese bilinguals who were dominant in Japanese. The focus in this study was on the children's use of French, and thus, the bilinguals who were dominant in French are referred to as French-dominant bilinguals and the bilinguals who were dominant in Japanese are referred to as French-nondominant bilinguals. Two subgroups of French monolinguals were included, one matched on age and the other matched on vocabulary score, described further below.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

All bilingual children were recruited from a Japanese language center in Montreal, Canada. To ensure that they had limited exposure to and knowledge of languages other than French and Japanese, we requested information about their language experiences from their parents in a questionnaire (see Materials in the Methods section). Children who were reported to have exposure to French and Japanese combined for at least 90% of the time were included; there were two exceptions, one child in each group, who had French and Japanese exposure 70% to 80% of the time. Despite having reportedly lower exposure to French and Japanese, the number of verbal descriptions produced in their nondominant language in the referential communication task were within the range of other children. Furthermore, their vocabulary scores were within the same age level on the vocabulary test as other bilingual children who had low scores. These children were included because of the difficulties recruiting bilingual children of the appropriate age with the unique French–Japanese language combination, and the results did not change without these children. All bilingual children were exposed to Japanese from birth because their mothers were native Japanese speakers. Sixteen bilinguals were exposed to French from birth and one was exposed to French from the age of 1 year when starting French daycare as his father spoke English to him. The third language that most children were reportedly exposed to was English, which is common in Montreal. English exposure was typically from music and television, which is hard to avoid in Montreal.[1] They may also have had some English exposure at daycare, from nannies or if the parents spoke English to each other.

The bilinguals' performance on French and Japanese versions of the Expressive One-Word Picture Vocabulary Test (described below; Academic Therapy Publications, 2000) was used to determine their language dominance. More specifically, children were classified as French-dominant if their score on the French version was 6 or more points higher than their score on the Japanese version, and vice versa for classifying them as French-nondominant. We chose a difference of 6 or more points as the cutoff for determining language dominance based on the idea that the majority of items per age level should be correct, and there are 10 items for each age level in this vocabulary test. In addition, the children's performance on the referential communication task and information provided by their parents on a questionnaire (both described in Materials below) were used to confirm the children's dominance classification. Four children (two in each group) appeared relatively balanced in each language in that their vocabulary scores differed by only 1 or 2 points, but they were classified as French-dominant or French-nondominant based on their parents' report about the children's language exposure and their performance on the referential communication task. In the referential communication task, these children appeared to have more difficulty providing appropriate words for the animal features or making themselves understood by the experimenter while speaking in their nondominant language (e.g., "*like a star*" for the square shaped bug or "*the thing that hurts when you touch it*" for the spiky fish). When we removed these children from the analyses, the results did not change; therefore, we have included them in all analyses to keep a similar number of children in each group and avoid statistical complications and compromises.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

Table 1 summarizes the mean ages and French vocabulary scores by group. The French-dominant group ($M = 6;0$) was significantly older than the French-nondominant group ($M = 5;1$), $t\ (15) = -2.64$, $p = .019$, and the French-dominant group scored significantly higher on the French vocabulary test ($M = 34.89$) than the French-nondominant group ($M = 13.75$), $t\ (15) = -4.69$, $p < .001$. All children were proficient enough in French to complete the referential communication task.

The French monolinguals were from Montreal and were contacted through a database of French-speaking families interested in participating in developmental studies. As with the bilinguals, the monolinguals may have been exposed to languages other than French. Thus, to minimize the influence of other languages, all monolingual children had to have been exposed to French at least 90% of the time, as reported by their parents on a questionnaire (see below).

We formed two subgroups of French monolinguals: (a) monolinguals who were matched with the French-dominant bilinguals on age (i.e., age-matched French monolinguals) and (b) monolinguals who were matched with the French-dominant bilinguals on vocabulary score (i.e., vocabulary-matched French monolinguals). This allowed us to examine possible effects of age differences and a difference between monolinguals and bilinguals. Of the monolinguals, three participants in the age-matched group were included in the vocabulary-matched group. As shown in Table 1, the age-matched French monolinguals did not differ significantly from the French-dominant bilinguals in age ($M = 5;7$ vs. $6;0$), $t\ (16) = 1.57$, $p = .137$. However, they scored significantly higher on the vocabulary test ($M = 51.67$ vs. 34.89), $t\ (16) = 3.26$, $p = .005$. Thus, as anticipated, the French-dominant bilinguals were less lexically proficient than the age-matched French monolinguals in French. The age-matched French monolinguals and French-nondominant bilinguals did not differ significantly in age ($M = 5;7$ vs. $5;1$), $t\ (15) = 1.48$, $p = .160$, but the age-matched monolinguals scored significantly higher than the French-nondominant bilinguals on the French

Table 1. *Mean ages and French vocabulary scores of the monolingual and bilingual children in Study 1*

			Age (y;m)		French vocabulary score (raw)	
	Proficiency group	*N* (male)	*M*	*SD*	*M*	*SD*
Monolinguals	Age-matched (High)	9 (5)	5;7	0;6	51.67	12.05
	Vocabulary-matched[a]	9 (5)	5;1	0;7	35.56	9.50
Bilinguals	French-dominant (Intermediate)	9 (5)	6;0	0;6	34.89	9.66
	French-nondominant (Low)	8 (4)	5;1	0;10	13.75	8.81

Note: Three children were in both the age-matched and the vocabulary-matched French monolingual groups.
[a]The vocabulary-matched French monolinguals had an intermediate level of proficiency.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

vocabulary test ($M = 51.67$ vs. 13.75), t (15) = 7.32, $p < .001$. In contrast, the vocabulary-matched French monolinguals did not differ from the French-dominant bilinguals in their vocabulary scores ($M = 35.56$ vs. 34.89), t (16) = 0.148, $p = .885$, but differed significantly in age ($M = 5;1$ vs. 6;0), t (16) = –3.210, $p = .005$. The vocabulary-matched monolinguals did not differ significantly from the French-nondominant group in age ($M = 5;1$ vs. 5;1), t (15) = 0.755, $p = .462$, but their vocabulary scores differed significantly ($M = 35.56$ vs.13.75), t (15) = 4.615, $p < .001$.

Materials

REFERENTIAL COMMUNICATION TASK. A referential communication task was used to elicit verbal descriptions and gestures from the children. In this task the child and experimenter sat facing each other, each with his or her own computer screen; they could not see one another's screens. Two cartoon scenes with animated animals were displayed side by side on the screens. The scenes were identical except for one difference in the animal's manner of motion, their shape, or their size. For example, in one scene, a round fish swam across the screen, and in the other scene, a spiky fish swam across the screen. See Table 2 for descriptions of the scenes, and Figure 1 for sample still image pairs. One scene was the target, indicated by a star above it. The child had to communicate the difference between the animals' features in each scene so that the experimenter could correctly guess the target scene.[2] When the experimenter guessed correctly, the star on the child's screen smiled and flashed. If the experimenter incorrectly guessed the nontarget scene, the star on the child's screen showed a confused face to indicate that the experimenter's choice was wrong.

Three pairs of scenes were used as practice items and eight pairs as test items, half of which differed in manner of motion and half differed in physical features. The test items were separated into two blocks in which the order of the four items (two manner of motion and two physical feature) remained the same. The order of the blocks was counterbalanced across participants.

Table 2. *Feature differences for the animals in the pairs of cartoon scenes*

Scene feature	Feature difference[a]	Animal for French	Animal for Japanese
Visual feature	Fat, thin	Bird	Mouse
	Fluffy, smooth	Cat	Dog
	Square, round	Bug	Turtle
	Spiky, smooth	Fish	Lizard
Motion feature	Jumping, running	Frog	Rabbit
	Rolling, sliding	Dog	Pig
	Swing, jump	Monkey	Squirrel
	Flapping wings, still wings	Bird	Butterfly

[a]Feature differences listed first were those of the target cartoon scenes.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

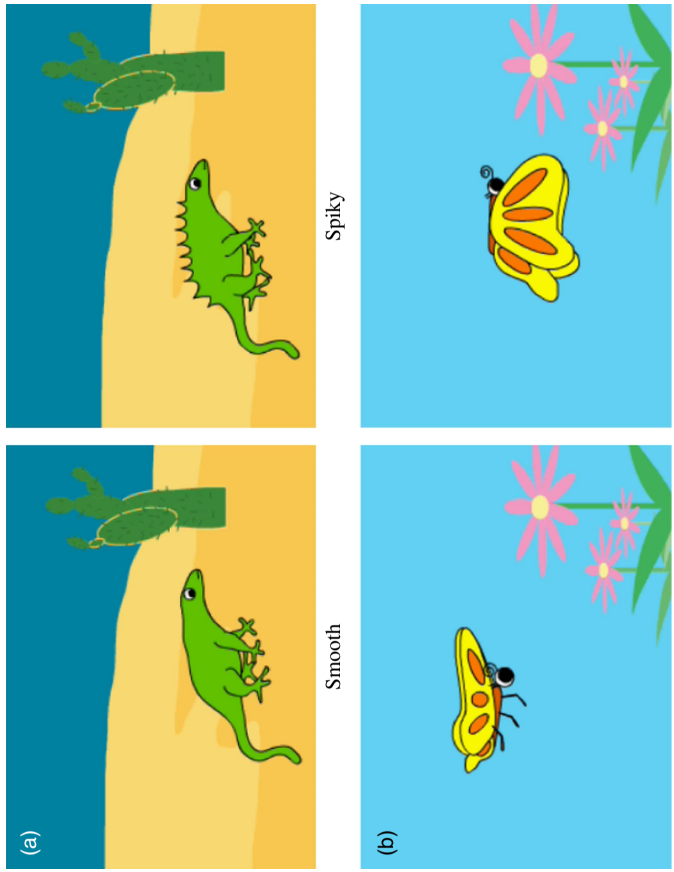

Figure 1. Still image examples of cartoon scene pairs that depict the animals' (a) visual features and (b) motion features.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

VOCABULARY TEST. The Expressive One-Word Picture Vocabulary Test (Academic Therapy Publications, 2000) was administered in French to the monolinguals and in French and Japanese to the bilinguals to assess the children's expressive language proficiency, and to determine the bilinguals' dominant language. In this test, children were shown pictures of objects or actions, and they had to name the object, a part of an object, the action, or a category that encompassed the objects.

This test was developed and normed for English-speaking children in the United States. We created French and Japanese answer keys and modified the administration and scoring procedures for all groups to make it more appropriate for our participants.[3] The French answer key was based on the French version translated by the Groupe coopératif en orthophonie région Laval-Laurentides-Lanaudière (1995). For new items in the version that we administered, native French speakers were consulted to determine the most common and appropriate name(s) for the items. To create the Japanese answer key, native Japanese speakers were consulted to decide the most common and appropriate name(s) for the items. They also identified which items were culturally inappropriate and thus problematic for Japanese children. Pilot testing was done with French and Japanese speakers to ensure the appropriateness of both answer keys. There are no norms for French- or Japanese-speaking children; thus, we compared our participant groups on raw vocabulary scores. That is, the total number of correct items was not adjusted in any way to standardize the scores. Appendix A provides further details of the administration and scoring modifications that were applied for all groups to ensure consistency in assessing the children's level of vocabulary.

PARENTAL QUESTIONNAIRE. The parents completed questionnaires requesting demographic and language experience information. Questions asked about the children's exposure to different languages from birth at home, daycare, and preschool, and from various sources such as television and books. There were also questions asking parents to estimate the percentage of language exposure to French, Japanese, and other languages based on a typical week at the time of testing (see an English version of the parental questionnaire provided in the online-only Supplementary Material).

Procedure. The parents were informed that this study had been approved by the university research ethics committee and were asked to read and sign the consent form prior to testing. They moved to the control room with a one-sided mirror, and the children were tested individually in a large university playroom. The children completed the referential communication task first, followed by the vocabulary test in a single session. All sessions were video-recorded. Bilingual children had a French and a Japanese session between 2 and 3 weeks apart. The order of these sessions was counterbalanced across participants.

For the referential communication task, the experimenters[4] told the children that they would play a guessing game. The experimenters showed the children

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

what would happen with the practice trials. It was explained that there were two animated cartoon scenes that were the same except for one difference, and one cartoon had a star above it. The child had to find the difference between the cartoon scenes and give the experimenter clues so she could guess which had the star. The scenes played repeatedly until the experimenter selected one scene based on the children's verbal and gestural descriptions by pressing a computer key. To minimize any negative experiences, which could happen if the experimenter chose the incorrect item, the experimenter asked prompting questions to encourage the children to describe how the target scene differed from the nontarget. The questions were not directly related to the scene difference (e.g., *how is the one with the star different from the other one?*; *I think they look different. Look at them again and tell me more*.). A maximum of four prompts were given for each item if needed.[5] Because this task was designed to elicit children's speech and iconic gestures for communicating specific information and not to assess children's verbal and nonverbal skills, the experimenters interacted with the children naturally and tried to keep them engaged and focused on the task as if they were playing an interesting game together. The children received a sticker for their help after each item, regardless of whether their descriptions were sufficient or not, to motivate them to continue.

Transcription and coding. The speech produced by the children and experimenters during the referential communication task was transcribed verbatim from the video by native or near native speakers of French using the CHAT transcription system (MacWhinney, 2000). All initial coding was done by a pair of coders[6] who were not familiar with gesture research and were trained by the graduate student on the project team. They were native or near native speakers of French. While watching the videos, the coders identified verbal descriptions and coded them as such in the transcripts, and then they coded them as being accompanied by an iconic gesture or not using the CHILDES coder mode (MacWhinney, 2000). A verbal description was defined as a phrase in which the child described or attempted to describe one animal's motion or physical feature. The length of the children's responses to describe the animals varied. Some children provided responses with two phrases or clauses, each describing one animal in the pair (e.g., *the one with the star looks like a square and the other one looks like a circle*). Such a response was considered to be two verbal descriptions, and each was coded separately. Some children produced very brief utterances (e.g., *the spiky one*) or one-word utterances (e.g., *rolling* or *spiky*), which were considered single verbal descriptions.

In addition, to examine whether there were differences in the quality of verbal descriptions across proficiency groups, the verbal descriptions were coded for how specific and precise they were by classifying them into one of three categories: precise, imprecise, and other. Precise descriptions included appropriate words that clearly described the manner of movement, size, or shape of the animals that distinguished the target from the nontarget scene (e.g., *rolling* or *spiky*). The imprecise descriptions included descriptions where the child did not provide enough information for the experimenter to determine the scene referred

to (e.g., "*the bird is flying*" while watching the animations in which both birds were flying but one was flapping its wings; "*he moves like this, it looks like this*" while showing the movements with hands). All other descriptions were classified as other.

Any iconic gesture produced by the children that expressed the manner of motion, size, or shape of the animal was coded as such. Children produced iconic gestures with their hands or arms and sometimes their whole body, like acting out what the animal in the cartoon was doing (e.g., jumping around the room to show how the frog jumped). We did not examine precisely which word(s) the iconic gestures accompanied because children often gestured over an entire description (McNeill, 1992).

The second coder verified the first coder's coding against the video and inserted a second coding line right below the first coding lines when she did not agree. Then the graduate student on the project team who was familiar with gesture research but did not know the hypothesis of the present study verified all the coding lines against the videos. As in some previous gesture research (e.g., Pika, Nicoladis, & Marentette, 2006), the intercoder reliability rates were not calculated, and any discrepancies and coding issues were instead extensively discussed among the coders and the graduate student verifier by watching the videos until they reached 100% agreement on the codes to ensure that all the coding lines included in the analysis were consistent and reliable. The few coding lines on which no consensus was reached were recoded as undecided and excluded from the analysis. The FREQ program in CHILDES was run on the coded transcripts (MacWhinney, 2000). For each child, this counted the total number of verbal descriptions coded (tokens of coded clauses), the total number of words used in the verbal descriptions (word tokens), and the total number of verbal descriptions accompanied by iconic gestures (tokens of coded clauses with iconic gesture codes).

Analyses. Although the bilinguals had a French and a Japanese session, only the French sessions were analyzed in Study 1. There was variation across the children with respect to the number of verbal descriptions that they produced to describe the animals' features. To control for this variation, rates of overall iconic gesture use were calculated for each child by dividing the number of verbal descriptions they produced with an accompanying gesture by the total number of verbal descriptions they produced. This method for calculating gesture rate is similar to that done by Colletta et al. (2010) and Kita and Özyürek (2003), who used clauses as the speech unit for calculating gesture rate. The mean numbers, standard deviations, and ranges of iconic gesture use are provided as additional information in Appendix B.

Results and discussion

Table 3 shows the means for the number of overall verbal descriptions and word tokens, and Figure 2 displays the proportions of iconic gesture use (i.e., gesture rates) for the high- (age-matched French monolingual), intermediate- (French-

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

Table 3. *Mean numbers, standard deviations, and ranges of French overall verbal descriptions and word tokens in the verbal descriptions by group in Study 1*

Proficiency group		Overall descriptions			Word tokens		
		M	*SD*	Range	*M*	*SD*	Range
Monolinguals	Age-matched (High)	18.56	7.20	9–27	133.22	61.81	63–215
	Vocabulary-matched[a]	21.33	6.91	10–30	152.33	51.17	76–217
Bilinguals	French-dominant (Intermediate)	22.67	7.11	9–31	168.78	46.87	87–266
	French-nondominant (Low)	21.25	5.42	13–29	133.63	62.99	60–274

[a]The vocabulary-matched French monolinguals had an intermediate level of proficiency.

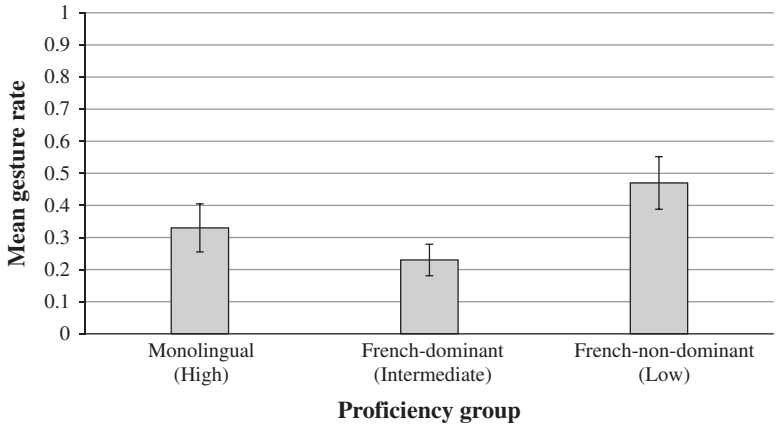

Figure 2. Mean gesture rates with French verbal descriptions by the children in Study 1. Error bars represent standard errors of the means. The ranges of the gesture rates were 0.04–0.76 for the age-matched French monolinguals, 0.03–0.44 for the French-dominant bilinguals, and 0.26–0.93 for the French-nondominant bilinguals.

dominant bilingual), and low- (French-nondominant bilingual) language proficiency groups. There was no main effect of proficiency group on the number of overall verbal descriptions, F (2, 23) = 0.880, p = .428, as well as word tokens, F (2, 23) = 0.112, p = .345, indicating that the groups did not differ significantly in the amount of verbal descriptions produced. To test our hypothesis that the relationship between language proficiency and gesture use is quadratic, planned comparisons for trend analysis were conducted on the gesture rates with proficiency group as the between-subjects factor (Ferguson & Takane, 1989). As predicted, the quadratic trend was significant, t (23) = 2.04, p = .0265,[7] one-tailed, and the linear trend was not significant, t (23) = 1.421 p = .0845, one-tailed.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

The results show a quadratic trend among the three proficiency groups wherein the high-proficiency group ($M = .33$, $SD = .23$) and the low-proficiency group ($M = .47$, $SD = .23$) tended to gesture more than the intermediate-proficiency group ($M = .23$, $SD = .15$). This result provides support for the view that lower levels of proficiency would be associated with higher levels of gesture use. The low-proficiency group might have had difficulty accessing appropriate words and formulating precise, specific verbal descriptions due to their relatively low French proficiency (Gollan et al., 2008). The following example from our data illustrates this possibility: *c'est le c'est le c'est le poisson qui a qui a des piques là*, "it's the it's the it's the fish who has who has the spikes there." Producing iconic gestures may have helped them find suitable specific words and arrange them in an utterance that clearly conveyed the required information to the experimenter (Alibali et al., 2000; Kita, 2000).

The result that the high-proficiency group showed slightly higher gesture rates than the intermediate-proficiency group is consistent with Hostetter and Alibali's (2007) previous finding with English-speaking monolingual adults. These highly proficient children presumably did not have as much difficulty describing the pertinent information as the less proficient children did, and thus, they likely gestured for different reasons as suggested by Kendon (2001).

To provide support for our proposition that the high- and low-proficiency groups gestured for different reasons, we analyzed the quality of the children's verbal descriptions. The mean proportions of precise and imprecise descriptions by group are summarized in Table 4. Separate one-way analyses of variance (ANOVAs) were conducted on the proportions of precise and imprecise descriptions with proficiency group as the between-subjects factor. The results showed a main effect of proficiency for the precise descriptions, F (2, 23) = 11.305, $p < .001$, and for the imprecise descriptions, F (2, 23) = 6.863, $p = .005$. Least significant difference post hoc tests showed that the low-proficiency group produced significantly fewer precise descriptions than the high-proficiency group, t (15) = –3.897, $p = .001$, and the intermediate-proficiency group, t (15) = –3.515, $p = .007$. Furthermore, they produced significantly more

Table 4. *Mean proportions of precise, imprecise, and other verbal descriptions per group in Study 1, with standard deviations in parentheses*

Proficiency group		Precise	Imprecise	Other
Monolinguals	Age-matched (High)	.82 (.11)	.03 (.05)	.15 (.09)
	Vocabulary-matched[a]	.74 (.13)	.04 (.07)	.22 (.10)
Bilinguals	French-dominant (Intermediate)	.75 (.06)	.05 (.10)	.19 (.05)
	French-nondominant (Low)	.52 (.18)	.21 (.16)	.27 (.14)

[a]The vocabulary-matched French monolinguals had an intermediate level of proficiency.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

imprecise descriptions than the other two groups: high-proficiency group, t (15) = 3.485, p = .003; intermediate-proficiency group, t (15) = 2.807, p = .013. However, the high- and intermediate-proficiency groups did not differ significantly in the use of precise (p = .113) or imprecise descriptions (p = .753). In addition, imprecise descriptions were produced by seven of eight children in the low-proficiency group, but only three children in each of the more proficient groups produced them.

These results suggest that the referential communication task was difficult for the low-proficiency group but not the more proficient groups. Nonetheless, the fact that the high-proficiency group gestured slightly more often than the intermediate-proficiency group suggests that the referential communication task was relatively easy for the high-proficiency group.

Some may argue that the differences in gesture use were due to a difference between monolinguals and bilinguals, or the age difference between the two bilingual groups, rather than the differences in language proficiency (Gollan et al., 2008). To examine these possibilities, we compared the gesture rates and preciseness of verbal descriptions between the vocabulary-matched French monolinguals and French-dominant bilinguals, both of intermediate proficiency. We reasoned that if proficiency influences gesture use and the quality of verbal descriptions, monolingual and bilingual children with similar levels of proficiency would have similar gesture rates and quality of verbal descriptions, even if they differed in age. They did not differ significantly in their overall verbal descriptions (p = .692), word tokens (p = .487), preciseness of verbal descriptions (ps > .835), or gesture rates (M = .26 vs. .23, SD = .21 vs. .15, p = .667). The vocabulary-matched monolinguals (intermediate proficiency) were similar in age to the French-nondominant bilinguals (low proficiency), and they produced significantly more precise, t (15) = 2.883, p = .001, and fewer imprecise descriptions, t (15) = –3.224 p = .006, than the French-nondominant bilinguals, although the difference in gesture rates was marginally significant (M = .26 vs. .47, SD = .21 vs. .23), t (15) = 2.014, p = .062. These results provide some support for our proposition that the differences in gesture use were due to differences in language proficiency. Thus, it would seem that the referential communication task in the present study was not sufficiently difficult to distinguish the high- and intermediate-proficiency groups in their verbal descriptions, although they differed significantly in their level of French proficiency. A more difficult referential communication task might reveal differences in the preciseness of verbal descriptions between these groups.

STUDY 2

In Study 2 we analyzed the Japanese data from the two bilingual groups described in Study 1 and compared them to a group of Japanese monolinguals, all of whom had a similar level of proficiency in Japanese. We expected that the children's gesture rates when speaking Japanese would be similar across participant groups.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

Methods

Participants. The same 17 French–Japanese bilingual children described in Study 1 were included in these analyses. The focus in this study was on their Japanese descriptions; therefore, we refer to the bilinguals who were dominant in Japanese as Japanese-dominant bilinguals ($n = 8$) and the bilinguals who were dominant in French as Japanese-nondominant bilinguals ($n = 9$). To clarify, the French-dominant bilinguals from Study 1 are referred to here as Japanese-nondominant bilinguals, and the French-nondominant bilinguals from Study 1 are referred to here as Japanese-dominant bilinguals. A group of Japanese monolingual children ($n = 8$) was also recruited to match the Japanese-dominant bilinguals on age. They were recruited from a daycare in Tokyo, Japan, and were exposed to Japanese at least 90% of the time as reported on the language questionnaire (see Materials in Methods below).

Table 5 summarizes the mean ages and Japanese vocabulary scores for the bilingual and monolingual groups. As intended, the Japanese monolinguals and Japanese-dominant bilinguals did not differ significantly in age, $t\ (14) = 0.24$, $p = .815$. However, the Japanese-nondominant bilinguals were significantly older than both the Japanese-dominant bilinguals, $t\ (15) = 2.64$, $p = .019$, and Japanese monolinguals, $t\ (15) = 3.35$, $p = .004$. A one-way ANOVA showed that the three groups did not differ significantly in their Japanese vocabulary scores, $F\ (2, 22) = 1.13$, $p = .340$; thus, they had similar levels of proficiency in Japanese. It should be noted that the Japanese-dominant bilingual group was, on average, a year younger than the Japanese-nondominant bilingual group. It is likely that these children were dominant in Japanese because, due to their younger age, they spent most of their time at home with their Japanese-speaking mother and had relatively less exposure to French speakers in daycare, preschool, or kindergarten compared to the other bilingual group. The year-older Japanese-nondominant bilingual group spent relatively more time in French daycare, preschool, or kindergarten, which contributed to their dominance in French. Furthermore, most bilingual children were taking Japanese classes. Thus, it is not likely the case that the Japanese-nondominant bilinguals had an overall general proclivity to both languages compared to the Japanese-dominant bilinguals. Rather, it is more likely the amount of exposure to French that the children received that contributed to

Table 5. *Mean ages and Japanese vocabulary scores of the monolingual and bilingual children in Study 2*

		Age (y;m)		Japanese vocabulary score (raw)	
Participant group	*N* (male)	*M*	*SD*	*M*	*SD*
Japanese monolinguals	8 (3)	5;2	0;5	26.75	4.03
Japanese-dominant bilinguals	8 (4)	5;1	0;10	25.00	5.04
Japanese-nondominant bilinguals	9 (5)	6;0	0;6	20.00	14.82

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

their proficiency profiles. The Japanese monolinguals were selected to match the Japanese-dominant bilinguals on age, and their vocabulary size level was similar to that of both bilingual groups, although they were about a year younger than the Japanese-nondominant bilinguals.

Materials. Japanese versions of the referential communication task, vocabulary test, and language questionnaire were used. A different set of cartoons in the referential communication task was created for the Japanese sessions (same shape, size, or manner but different animals and backgrounds, see Table 2) along with a Japanese answer key for the vocabulary test that was comparable to the French versions used in Study 1.[3] The language questionnaire from Study 1 was translated into Japanese as well. These Japanese versions were used for the Japanese monolinguals and the bilinguals during their Japanese session.

Procedure. The procedure for the bilinguals' Japanese sessions was the same as in Study 1, except that they saw the Japanese version of cartoons in the referential communication task. The Japanese monolinguals did the vocabulary test prior to the referential communication task and were tested in a room at their daycare. During the referential communication task, they received a set of stickers after the first four items and then after the last four items.

Transcription and coding. The transcription and coding procedures were the same as in Study 1, except that the JCHAT transcription system was used to transcribe the speech samples (Oshima-Takane, MacWhinney, Sirai, Miyata, & Naka, 1998). A pair of native Japanese speakers transcribed and coded the referential communication task sessions, and the graduate student verified all the coding with the help of the professor on the project team in Montreal who was a native Japanese speaker. Any coding issues and discrepancies were discussed extensively until 100% agreement on the codes was reached following the procedure used in Study 1.

Analyses. The analyses were similar to those in Study 1.

Results and discussion

Table 6 presents the mean number of verbal descriptions and word tokens and Table 7 presents the mean proportions of precise and imprecise descriptions. The ANOVAs performed on the number of overall descriptions, word tokens, and proportions of precise and imprecise descriptions did not show main effects of participant group: overall, F (2, 22) = 0.815, $p = .456$; word tokens, F (2, 22) = 1.563, $p = .232$; precise, F (2, 22) = 1.968, $p = .164$; imprecise, F (22) = 0.210, $p = .812$. Thus, the three groups did not differ significantly in any of the verbal description measures.

Figure 3 displays the mean gesture rates for each group. To verify that these groups who did not differ in proficiency or verbal descriptions also did not differ in gesture use, a one-way ANOVA was conducted on the mean gesture rates with

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

Table 6. *Mean numbers, standard deviations, and ranges of Japanese overall verbal descriptions and word tokens in the verbal descriptions by group in Study 2*

	Overall descriptions			Word tokens		
Participant group	*M*	*SD*	Range	*M*	*SD*	Range
Japanese monolinguals	22.13	5.99	13–32	140.75	68.93	69–249
Japanese-dominant bilinguals	20.50	10.66	8–37	102.25	65.70	47–228
Japanese-nondominant bilinguals	17.78	2.91	14–23	94.56	31.01	62–152

Table 7. *Mean proportions of precise, imprecise, and other verbal descriptions per group in Study 2, with standard deviations in parentheses*

Participant group	Precise	Imprecise	Other
Japanese monolinguals	.63 (.23)	.13 (.13)	.24 (.12)
Japanese-dominant bilinguals	.46 (.24)	.20 (.20)	.34 (.11)
Japanese-nondominant bilinguals	.42 (.25)	.22 (.21)	.37 (.15)

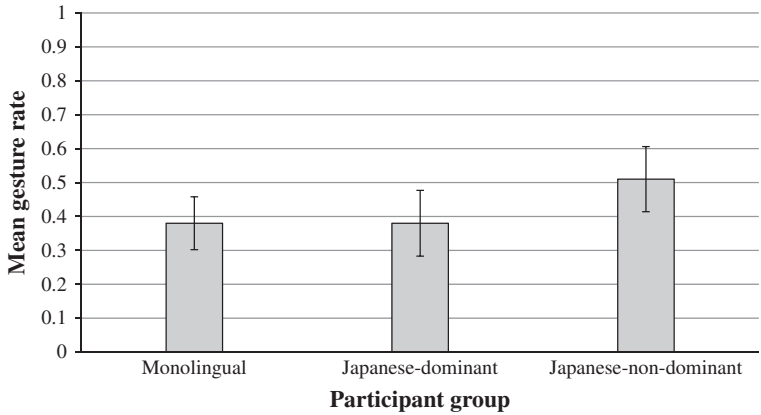

Figure 3. Mean gesture rates with Japanese verbal descriptions by the children in Study 2. Error bars represent standard errors of the means. The ranges of the gesture rates were 0.04–0.70 for the Japanese monolinguals, 0.00–0.75 for the Japanese-dominant bilinguals, and 0.00–0.68 for the Japanese-nondominant bilinguals.

participant group as a between-subjects factor. As predicted, there was no main effect of participant group (monolinguals: $M = .39$, $SD = .22$; Japanese-dominant: $M = .38$, $SD = .28$; Japanese-nondominant: $M = .52$, $SD = .29$), $F(2, 22) = 0.431$, $p = .655$. Trend analyses confirmed that the groups did not differ in their gesture

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

use as both the linear trend, F (1, 22) = 0.41, p = .842, and the quadratic trend, F (1, 22) = 0.821, p = .375, were not significant.

These results support our hypothesis that children with a similar level of language proficiency produce verbal descriptions of similar quality and have similar rates of iconic gesture use, whether they are monolinguals or bilinguals. In addition, the fact that the Japanese monolinguals and Japanese-dominant bilinguals were significantly younger than the Japanese-nondominant bilinguals suggests that the quadratic trend of gesture rates in French found in Study 1 cannot be attributed to the age difference between the bilingual groups.

GENERAL DISCUSSION

Previous studies investigating the effect of language proficiency on iconic gesture use have found contradictory results. We argued that variation across studies in task demand, language spoken, and type and level of language proficiency are possible reasons for these contradictory findings; we designed the present study to address these issues. To address the task demand issue, we used a dynamic referential communication task appropriate for 5-year-olds because it makes fewer demands on verbal abilities compared to story-retelling tasks used in past studies. In addition, it controls for amount and type of information the speaker needs to provide to complete the task. To examine if proficiency and gesture use have a linear or quadratic relationship, we compared children with three different or similar levels of language proficiency within the same language by including bilingual and monolingual children, instead of comparing bilinguals' gesture use in their dominant and nondominant languages. Finally, unlike past studies, we used a measure of expressive lexical proficiency independent from the task that examined co-speech gesture use. As a result, the present study showed a clear link between language proficiency and iconic gesture use.

The most important finding was that the relationship between language proficiency and iconic co-speech gesture use was quadratic, as shown in Study 1. Among three groups, the intermediate-proficiency group had the lowest gesture rate and the low-proficiency group had the highest gesture rate. The high-proficiency group showed slightly higher gesture rates than the intermediate group. This U-shaped pattern is similar to the quadratic trend found in Hostetter and Alibali's (2007) study with English monolingual adults in three proficiency groups.

As discussed earlier, the low-proficiency group likely had some difficulty describing the animals' motions/features precisely, as they produced fewer precise and more imprecise verbal descriptions than the high- and intermediate-proficiency groups. Iconic gestures may have facilitated lexical access (Hadar et al., 2001; Rauscher et al., 1996) or supplemented their verbal message (Holler & Beattie, 2003; Kita & Özyürek, 2003; So et al., 2010).

The high- and intermediate-proficiency groups showed no significant differences in their use of precise and imprecise descriptions, and likely had no trouble describing the animals' motions/features. Yet, the high-proficiency group

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

gestured slightly more than the intermediate group. This suggests that the referential communication task was relatively easy for the most proficient group and their cognitive resources were not taxed; as a result, they could put their remaining resources toward producing gestures that enhanced their verbal descriptions and/or to ensure that the experimenter understood their message (Hostetter, 2011; Hostetter & Alibali, 2007; Kendon, 1994, 2001). For instance, many monolingual children described the bird flapping its wings precisely and moved their arms up and down like the bird's wings. They maybe considered this large gesture effective at conveying the bird's movement. Thus, while both the high- and the low-proficiency groups gestured often, it is likely that they gestured for different reasons.

The intermediate-proficiency group gestured least despite producing descriptions of similar preciseness to the high-proficiency group. This suggests that their level of proficiency was just adequate for the referential communication task. Thus, they did not have extra cognitive resources to produce more gestures like the most proficient group and also did not need to gesture often to help overcome communication difficulties like the least proficient group. This explanation was further supported by the finding that they did not differ from the vocabulary-matched French monolinguals in their use of gestures or preciseness of their verbal descriptions, even though their age differed by 1 year. In addition, the finding in Study 2 that similar gesture rates were produced by Japanese monolinguals and Japanese-dominant and Japanese-nondominant bilinguals with a similar level of language proficiency provides further support for our interpretation that the quadratic trend found in Study 1 was not due to the age difference between the French-dominant (intermediate proficiency) and French-nondominant (low proficiency) groups, or to a difference between monolinguals and bilinguals. It appears that a 1- to 2-year age difference does not affect gesture rates when language proficiency is similar for monolinguals and bilinguals between 4 and 6 years of age. Future research should investigate whether language proficiency is a stronger predictor of iconic gesture use than age by testing other age groups.

Previous bilingual studies compared gesture use in unbalanced bilinguals' dominant and nondominant languages or balanced bilingual's two languages; these comparisons could not test for a nonlinear relationship between language proficiency and gesture use because comparisons were only made between two proficiency levels. As discussed previously, some studies found that bilinguals gestured more in their dominant than their nondominant language (Gregersen et al., 2009; Gullberg, 1998; Nicoladis, 2002; Nicoladis et al., 1999), one found the opposite (Nicoladis et al., 2007), and others have found no relationship (Abdalla, 2015; Nagpal et al., 2011; Nicoladis et al., 2009; Sherman & Nicoladis, 2004). These conflicting findings can be reconciled if we think of each study as offering a partial glimpse into the proficiency–gesture relationship. If we consider the findings of all previous studies together, the quadratic relationship between language proficiency and iconic gesture rate comes into view. These findings underscore the complexity of co-speech gestural behavior and show that co-speech gestures are multifunctional. That is, speakers use gestures both when they

have communication difficulties and when they are capable of communicating their message clearly but wish to enhance it or the interaction experience for their addressee(s), as suggested by Kendon (2001) and later by Hostetter and Alibali (2007).

While the quadratic relationship between speech and gesture use mirrors Hostetter and Alibali's (2007) findings with adults, it is also reminiscent of the u-shaped curves found in various areas of child development (e.g., language acquisition and problem solving; Plunkett & Marchman, 1993; Siegler, 2004). In these studies, children's behavior is correct and similar during early and later stages of development. However, there is a middle stage characterized by incorrect behavior, and this produces an overall u-shaped curve. For example, children learning English first produce the correct past-tense forms of irregular verbs (e.g., *went* and *came*); they then go through a period where they over-regularize these verbs and produce incorrect forms such as *goed* or *comed*; and finally they produce the correct forms again (Plunkett & Marchman, 1993). When such a pattern is found, developmental researchers aim to understand the underlying knowledge or processes that are reflected in the changing behaviors.

The present study did not examine correct or incorrect behavior, but rather the co-expression of speech and iconic gestures. What knowledge or processes might have given rise to the u-shaped co-production of speech and iconic gestures that we observed for the children? Likely candidates are the children's vocabulary knowledge (i.e., whether it was sufficient in relation to that required for the referential communication task; Gershkoff-Stowe & Thelen, 2004), the extent to which their cognitive resources were tapped by the processes involved in dyadic communication (Kendon, 1980), and their understanding of how iconic gestures may be used in communicative contexts (Namy, Campbell, & Tomasello, 2004). Initially, children may use gestures only to help overcome difficulties they encounter when expressing their verbal messages, and once they are able to express themselves without difficulty, they use gestures less often. When children learn that people also gesture with the intent to communicate and supplement their verbal message, and they have the cognitive resources to do this, they may also gesture with this purpose. Thereafter, whether or not a child or adult uses iconic gestures, and the reason(s) for doing so, may be dependent on the speaker's cognitive capacity, verbal skills, and the communicative context. That is, an individual speaker's iconic gesture use is expected to vary in a u-shaped manner across different communication contexts and across development.

Our hypotheses about how task difficulty and cognitive load contribute to the use of co-speech iconic gestures need to be tested in future studies. The gesture rates of monolingual children with a similar proficiency level can be compared across several referential communication tasks that vary in complexity, requiring different levels of proficiency for the children to easily complete. Children would be expected to use iconic gestures at different rates and for different reasons depending on how the difficulty level of the task compares to the child's level of proficiency (Gershkoff-Stowe & Thelen, 2004), and how much the task taxes his or her cognitive resources (Siegler, 2004). To verify that children's use of iconic gestures shifts with increasing language proficiency, monolingual children need

to be followed longitudinally from approximately 3 to 6 years of age, without changing the difficulty level of the task. As children's language proficiency increases with age, a u-shaped curve for gesture use should emerge that reflects how and why the children are using iconic gestures. Such additional studies would provide a more comprehensive understanding of the effects of language proficiency on gesture use and would also demonstrate that gesture use can change within the same individuals depending on their cognitive capacity and the difficulty level of the communication task undertaken.

It is also important to directly test Kendon's (2001) communication enhancing hypothesis concerning proficient speakers' iconic gesture use. This may be done by designing an experimental study that analyzes whether a listener's engagement in the interaction with a proficient speaker is affected by his or her use of iconic gestures in general, as well as different types of iconic gestures such as those conveying manner of motion versus physical attributes.

A limitation in the present study is the small sample size. We studied iconic gesture use by French–Japanese bilingual children and respective monolinguals because this unique language combination had not been investigated. However, it was difficult to recruit appropriately aged bilingual children who had sufficient proficiency in each language to complete the referential communication task. We spent 3 years collecting data from bilingual children. Because gesture rates may differ across languages due to different linguistic structures (e.g., Efron, 1972; Kita & Özyürek, 2003; McNeill & Duncan, 2000), we analyzed the bilingual subgroups' performance separately and compared verbal descriptions and gesture use within the same language. As a result, the present study yielded clearer results despite the small sample size. Future research should examine whether the present findings can be replicated with bilingual and monolingual children who speak a language other than French or Japanese. To fully understand the relationship between language proficiency and iconic gesture use in children, more bilingual studies with different language combinations in different age groups are needed, and the present study is one attempt toward this.

A potential limitation in cross-linguistic research is the influence of culture. It has been widely believed that French-speaking adults gesture more than Japanese-speaking adults (Sekine, Stam, Yoshioka, Tellier, & Capirci, 2015), though more clear evidence is needed. Previous research with children has found similarities and differences across cultural groups. For example, Pettenati, Sekine, Congestrì, and Volterra (2012) found that 2-year-old Japanese and Italian children produced iconic gestures at similar rates and gestured actions more often than objects in a picture-naming task. However, Italians gestured size and shape information more often, and the Japanese's action gestures were more precise. They suggested that children in different cultures develop co-speech iconic gestures similarly to support speech production, only minimally influenced by culture. There is the possibility that cultural differences influenced the 4- to 6-year-old children's gesture rates in the present study. The Japanese monolinguals were growing up in Tokyo, Japan, while the French monolinguals and French–Japanese bilinguals were growing up in Montreal, Canada. It is difficult to assess how much Japanese culture the bilingual children were exposed to; however, we

know that they had DVDs, books, and music from Japan, and attended Japanese classes at the Japanese language center. To tease apart influences of language proficiency and culture on children's gesture use, future research should compare groups of bilingual children with similar levels of language proficiency who are growing up surrounded by the culture of each language, and then compare them with monolinguals in each cultural environment.

APPENDIX A

Modifications to the administration and scoring of the Expressive One-Word Picture Vocabulary Test

1. All of the participants began the test at Item 10, the level appropriate for 3-year-olds, because the bilingual children might have had less vocabulary than their same-age peers.
2. There were a maximum of 18 items that the children were tested on but were omitted when scoring because they were deemed to be culturally specific (e.g., windmill).
3. The test ended when the children failed six consecutive items (as per normal administration procedures), or five consecutive items were failed twice. This was because the bilingual children would often fail multiple items in a row and then get one correct.
4. Due to the modification of the end of the test, the children's raw vocabulary scores were calculated beginning from Item 10 until five consecutive items were failed.

APPENDIX B

The mean number, standard deviation, and rage of iconic gestures by group and by language

Table A1.

		French[a]			Japanese[b]		
Participant group		*M*	*SD*	Range	*M*	*SD*	Range
Bilinguals	French-dominant (Japanese-nondominant)	4.56	2.70	1–8	9.33	5.45	0–15
	French-nondominant (Japanese-dominant)	10.00	7.03	6–27	9.88	8.84	0–21
Monolinguals	French age-matched	6.44	5.77	1–19	—	—	—
	French vocabulary-matched	5.56	5.39	1–19	—	—	—
	Japanese	—	—	—	8.88	5.79	1–17

[a]Planned comparisons for trend analysis on the number of iconic gestures in the two bilingual groups and age-matched monolingual group in French: quadratic t (23) = 1.646, $p = .057$ one-tailed; linear t (23) = 1.355, $p = .094$ one-tailed.
[b]Two bilingual children (one Japanese-dominant and one Japanese-nondominant) produced no gestures in the Japanese session.

https://doi.org/10.1017/S014271641800070X Published online by Cambridge University Press

ACKNOWLEDGMENTS

This research was supported by grants from the Social Sciences, Humanities Research Council, Fonds de Recherché du Quebec Société et Culture, and a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science. We thank the Montreal Japanese Language Centre, Japanese Canadian Cultural Center of Montreal, Momiji daycare, and the children and parents who participated. We are grateful to Alexis Haikalis, Alisa Koyama, Azusa Yamaguchi, Andrew Pinchefsky, Carrie Cheung, Chigusa Kurumada, Kayo Nakamura, Kristin Fasciano, Kyoko Sato, Galia Pogossova, Gill Fruchter, Hisako Noguchi, Irina Pivneva, Leila Salem, Mari Umeda, Michi Tomoda, Patricia Groleau, Serina Nishioka,Yasuko Senoo, Yuri Fujisaki, and Kimiko Matushima, who helped with participant recruitment and testing, transcription, and coding. We also thank Fred Genesee, Marie Lippeveld, Rachel Mayberry, and Tom Shultz for helpful discussion and comments.

NOTES

1. Although all children were exposed to more than two languages to some extent, including language through music, books, and television, we do not know to what extent they knew any language(s) other than French and Japanese. This is because the number of languages children were exposed to is not the same as the number of languages they knew. Thus, we used the term "bilinguals" although they were not bilinguals in a strict sense.
2. A cardboard wall was used for half the trials and placed between the child and experimenter to eliminate reciprocal visibility. The order of the visibility conditions (visible vs. nonvisible) was counterbalanced across participants. This manipulation was done to compare the children's use of gesture for themselves versus for their addressee, and we present this data elsewhere.
3. A team of two professors (native Japanese speakers), one graduate student (a near-native French speaker), and three assistants (one French native speaker and two native Japanese speakers) worked on developing the French and Japanese versions of the vocabulary test.
4. A total of four French-speaking and five Japanese-speaking assistants administered the French and Japanese sessions of the bilingual experiments, respectively, over a period of 3 years. A total of three French-speaking and three Japanese-speaking assistants administered the French and Japanese monolingual experiments, respectively, over a period of 2 years. For each experimental session, one assistant administered the referential communication task and vocabulary test while the other assistant operated the video recorder in the control room (Study 1) or in the daycare room (Study 2). More than two assistants were needed to collect the present data for the bilingual and monolingual sessions because some assistants left after 1 or 2 years. All the experimenters who conducted the bilingual experimental sessions in Montreal were trained by one of the professors on the project team. Those who conducted the French monolingual experimental sessions were trained by the graduate student on the project team, and those who conducted the Japanese monolingual experimental sessions in Japan were trained by the other professor on the team. All training was done following the same instructions for the experimental session.

5. There were no significant differences in the total number of prompts used by the experimenter during the referential communication task among groups.
6. A total of three French speakers and four Japanese speakers coded French and Japanese transcripts, respectively, over a period of 3 years. They were trained by the graduate student and/or one of the professors on the project team.
7. Statistical analyses were done on proportions that were arc sine transformed, even though the assumption of homogeneity of variance was not violated, because the distribution of proportions tends to violate the normal assumption. However, for ease of interpretation, the original mean proportions are presented in the tables and figures. A significance level of .05 was used for all statistical tests. One-tailed tests were used for the planned comparisons of trend analysis and two-tailed tests were used for the remaining statistical tests.

REFERENCES

Abdalla, F. (2015). *Language dominance and gesture production in Arabic-English bilingual children*. Paper presented at the International Conference on Bilingualism, Malta.

Academic Therapy Publications. (2000). *Expressive One-Word Picture Vocabulary Test* (2nd ed.). Novato: Academic Therapy Publications.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, *15*, 593–613.

Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication*, *52*, 565–576.

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, *28*, 561–580.

Efron, D. (1972). *Gesture, race and culture; a tentative study of the spatio-temporal and "linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions*. The Hague: Mouton.

Ferguson, G. A., & Takane, Y. (1989). Multiple-comparison procedures. In G. A. Ferguson & Y. Takane (Eds.), *Statistical analysis in psychology and education* (6th ed., pp. 321–329). New York: McGraw-Hill.

Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology*, *111*, 43–62.

Gershkoff-Stowe, L., & Thelen, E. (2004). U-shaped changes in behavior: A dynamic systems perspective. *Journal of Cognition and Development*, *5*, 11–36.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*, 787–814.

Gregersen, T., Olivares-Cuhat, G., & Storm, J. (2009). An examination of L1 and L2 gesture use: What role does proficiency play? *Modern Language Journal*, *93*, 195–208.

Groupe Coopératif en orthophonie région Laval-Laurentides-Lanaudière. (1995). *Épreuve de compréhension de Carrow-Woolfolk et Épreuve de dénomination EO-WPVT-R: Adaptation et normalisation.* Unpublished manuscript, Montréal.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse. A study of Learners of French and Swedish*. Lund: Lund University Press.

Gullberg, M. (1999). Communication strategies, gestures, and grammar. *Acquisition et Interaction en Langue Étrangère*, *2*, 61–71.

Gullberg, M. (2012). Bilingualism and gesture. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (2nd ed., pp. 417–437). New York: Wiley-Blackwell.

Hadar, U., Dar, R., & Teitelman, A. (2001). Gesture during speech in first and second language. Implications for lexical retrieval. *Gesture*, *1*, 151–165.

Holler, J., & Beattie, G. (2003). Pragmatic aspects of representational gestures. Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, *3*, 127–154.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*, 297–315.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, *7*, 73–95.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The Relationship of verbal and nonverbal communication* (pp. 207–227). The Hague: Mouton.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, *27*, 175–200.

Kendon, A. (2001). Gesture as communication strategy. *Semiotica*, *135*, 191–209.

Kidd, E., & Holler, J. (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science*, *12*, 903–913.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.

Kita, S., & Özyurek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.

Krauss, R. M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. In L. S. Messing & R. Campbell (Eds.), *Gesture, speech, and sign* (pp. 93–116). Oxford: Oxford University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analysis talk*. Mahwah, NJ: Erlbaum.

Marcos, L. R. (1979). Nonverbal behavior and thought processing. *Archives of General Psychiatry*, *36*, 940–943.

Mayberry, R., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, *9*, 192–196.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D., & Duncan, S. D. (2000). Growthpoints in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.

Nagpal, J., Nicoladis, E., & Marentette, P. (2011). Predicting individual differences in L2 speakers' gestures. *International Journal of Bilingualism*, *15*, 205–214.

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A u-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, *5*, 37–57.

Nicoladis, E. (2002). Some gestures develop in conjunction with spoken language development and others don't: Evidence from bilingual preschoolers. *Journal of Nonverbal Behavior*, *26*, 241–266.

Nicoladis, E. (2007). The effect of bilingualism on the use of manual gestures. *Applied Psycholinguistics*, *28*, 441–454.

Nicoladis, E., Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology*, *35*, 514–526.

Nicoladis, E., Pika, S., & Marentette, P. (2009). Do French-English bilingual children gesture more than monolingual children? *Journal of Psycholinguistic Research*, *38*, 573–585.

Nicoladis, E., Pika, S., Yin, H., & Marentette, P. (2007). Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, *28*, 721–735.

Oshima-Takane, Y., MacWhinney, B., Sirai, H., Miyata, S., & Naka, N. (1998). *CHILDES manual for Japanese* (2nd ed.). Nagoya: JCHAT Project.

Pettenati, P., Sekine, K., Congestrì, E., & Volterra, V. (2012). A comparative study on representational gestures in Italian and Japanese children. *Journal of Nonverbal Behavior*, *36*, 149–164.

Pika, S., Nicoladis, E., & Marentette, P. F. (2006). A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer? *Bilingualism: Language and Cognition*, *9*, 319–327.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21–69.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, *7*, 226–231.

Sekine, K., Stam, G., Yoshioka, K., Tellier, M., & Capirci, O. (2015). Cross-linguistic views of gesture usage. *Vigo International Journal of Applied Linguistics*, *12*, 91–105.

Sherman, J., & Nicoladis, E. (2004). Gestures by advanced Spanish-English second-language learners. *Gesture*, *4*, 143–156.

Siegler, R. S. (2004). U-shaped interest in u-shaped development—And what it means. *Journal of Cognition and Development*, *5*, 1–10.

So, W. C., Demir, O. E., & Goldin-Meadow, S. (2010). When speech is ambiguous, gesture steps in: Sensitivity to discourse-pragmatic principles in early childhood. *Applied Psycholinguistics*, *31*, 209–224.