Let me show you what I know: Image, word and gesture in two-year-old children

by

Sarah Brown

A thesis submitted in partial fulfillment of the requirements for the degree of

Bachelor of Arts

in Psychology

Department of Social Science, Augustana Faculty
University of Alberta

Abstract

The iconic gesture used by two-year-old children is indicative of their knowledge and communicative abilities about objects and events. In this study, this knowledge was tested by studying the gesture produced by children with multiple language exposure and children with exposure to only one language, with the hypothesis that there would be a difference between the two groups. The Picture Naming Game (PiNG task) was used to elicit verbal responses as well as spontaneous iconic gesture from two-year-old children. There were no significant differences between the two groups of children; however, children used gesture to communicate information about how an object is to be used or how an event takes place, rather than simpler size/shape information. These results suggest that children are capable of creating complex and holistic communicative events.

This thesis is an original work done by Sarah Brown. This research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name “Picture Naming Game Multilingual Study”, No. PRO00049315, June 11, 2014.

Acknowledgements

I would like to thank my family and friends for the incredible support that they have provided throughout my university career. I would also like to extend thanks to the children who participated in the study, as well as their parents for helping with the data collection process. Finally, I would like to extend a special thank you to Dr. Paula Marentette. Without the hours of hard work you put in, this research would not have been possible.

Table of Contents

Page

List of Tables and Figures

Page

Let me show you what I know: Image, word and gesture in two-year-old children

Speech and gesture are closely linked. This can be seen most clearly when studying iconic gestures in young children. Children do not begin to use iconic gestures until they have attained a certain, though modest, level of fluency in that language. The iconic gestures that are used are a function of the lexical development that children have obtained.

Noun and verb word learning occurs differently for children during language acquisition, and certain aspects of lexical development occur differently for monolingual and bilingual children. Nouns are learned first and often make reference to concrete objects while predicates make reference to more dynamic and often abstract events or states of being. Due to these differences in lexical development, once could expect that there would be differences in the frequency with which monolingual and bilingual children gesture and the form they use when shown and image of an object versus and image of an event.

In this study I used a task designed to elicit verbal responses from two-year-old children to study the co-speech gesture that was produced when they made reference to an image of either an object or an event. The Picture Naming Game (PiNG task) asked children to name sets of pictures. Spontaneous co-speech/bimodal gesture was often produced as children named the images they were shown. Using this task, I was able to analyze whether or not the form and frequency of children's gestures differed when the lexical referent was an object versus an event, with the intention of drawing a conclusion about the function of

young children's gesture. In addition, I examined the overall difference between the gesture produced by monolingual children and children with multiple language exposure, with the hypothesis that the frequency of gesture would be greater among the children with multiple language exposure and that the frequency with which they used unimodal gesture would be greater than the frequency with which monolingual children did.

In this literature review, I examined how co-speech iconic gesture aids speech by allowing a child to use two modes of representation to communicate. In order to do this, I examined the differences between noun and verb word learning during early language acquisition in two- to three-year-olds as well as lexical development in bilingual children of the same age, with the assumption that differences in lexical development might have lead to differences in gesture. The research has shown that there is a clear link between speech and gesture in both adults and children, making it important to consider lexical development when studying children's gesture.

Children's Gesture

Speech and gesture go hand-in-hand during the development of language in children. Four types of gesture have been discussed throughout the literature: beats, conventional, deictic and iconic (McNeill, 1992). Beats typically display meaningless, emphatic motions. They have primarily been studied in adults because it appears that young children do not use them (McNeill, 1992). Conventional gestures have an agreed-upon cultural significance (e.g., waving bye-bye) and are typically not accompanied by speech (Iverson, Capirci, Volterra, & Goldin-Meadow, 2008). Deictic

gestures involve the intent to pick out a specific referent, typically by pointing. Of the four types of gesture, children use deictic gestures more frequently than others. Lastly, iconic gestures typically accompany speech and serve to represent something in the world. Iconic gestures are distinct from other forms of gesture. They are almost always accompanied by speech, distinguishing them from conventional and deictic gestures. Furthermore, beats and deictic gestures do not have a representational action associated with them at the climax of the gesture, making iconic gestures unique compared to the other forms (McNeill, 1992). Out of the four types of gesture that have been studied, iconic and deictic gestures are of particular interest to researchers who study children because these two forms of gesture lead to complex, co-speech utterances.

Deictic gestures have been shown to precede spoken language by many months. Much research has been done studying the deictic gestures that children produce at an early age and how they relate to the onset and early development of spoken language (Goldin-Meadow, 2014; Iverson et al., 2008; Özçalişkan, & Goldin-Meadow 2005). Children are able to point with intent within the first year of their life, before they are able to create the same intention using words (Goldin-Meadow, 2014). Supplementary deictic gesture plus word combinations facilitate the development of two-word utterances during language learning by eliciting them from caregivers (e.g. a child points to a pair of glasses and says “mommy”. Following this the caregiver says, “Yes, those are mommy’s glasses.”). In this case, the deictic gesture points out the second object of intent, allowing the child to create a combination of gesture and word before they are able to put two words together

using spoken language (Iverson et al., 2008). Likewise, Goldin-Meadow, 2014; Özçalişkan, as well as Goldin-Meadow, 2005 suggested that deictic gestures + speech combinations allow children to expand their communication by giving them two different modes of representation to use when interacting with others, also removing some of the cognitive strain of communicating with only one mode.

Spontaneous deictic gestures can be seen before a child is two years of age, however, children are far less likely to create spontaneous iconic gestures before about the age of two (Behne, Carpenter, & Tomasello, 2014; McNeill, 1992). Iconic gestures are arguably the most difficult form of gesture for children to produce and comprehend. Iconic gestures are used to enrich a person's explanation of an object or event (McNeill, 1992). Iverson, Capirci and Caselli (1994) suggest that because iconic gesture and speech work together to create one meaning, that the gesture supplements the speech rather than reproduces it. For example, a child describing the activity 'swimming' may gesture using their arms outstretched in front of them, stroking towards their body to help explain the act of swimming. In this circumstance, the gesture serves to supplement the speech, allowing the child to enhance their description by physically showing how the act occurs.

In this study the Picture Naming Game (PiNG task) was used to elicit spoken responses from children (Bello, Caselli, Pettenati, & Stefanini, 2010). Much of the research that has been done studying the iconic gesture produced by young children has been conducted using this task. The PiNG task was created as a measure of lexical development, but has been used as a means to collect spontaneous co-speech gesture from children in a controlled setting with a fixed set of images. Given this

brief description of the PiNG task, I will now examine research that has utilized this task as a means of collecting and analyzing children's gesture.

Stefanini et al. (2009) found that correct verbal naming responses increased with age but gesture use decreased among the older Italian children in the study. The number of representational gestures produced by children of all ages was far less than the number of deictic gestures produced. In addition, young children between the age of 2;0 and 3;0 used the greatest number of representational gestures. Almost all of the children who participated in the Stefanini et al. (2009) study used some sort of representational gesture at some point during the task and in all cases when a representational gesture was produced; it served to supplement the speech that it accompanied. They suggest that the gesture creates an association to the image that is not only used to help the child communicate, but is also used as a tool for the child to understand the image.

Pettenati, Stefanini and Volterra (2010) looked at the representational gestures produced by 45 Italian children between the ages of 2;0-3;0 as they were asked to complete the PiNG task. The researchers found that the most common way for a child to label an image was using unimodal speech. However, in certain cases, the child was able to correctly label an image but still used a gesture in conjunction with the speech, bimodally describing the image. These findings suggest that speech and gesture as two modalities are linked and that they serve to complement or supplement one another in certain situations.

Similarly, Hall et al. (2013) studied gesture in Australian children between the ages of 18-31 months suggesting that gesture is used along with speech to allow

children to further label objects or actions that they may not yet have a spoken label for. They suggest that the use of gesture creates an "action-based link" between the image and the word that is being spoken, allowing the child a second modality to help them understand and label the images that they are shown.

Children are more likely to create an action-based gesture rather than a size/shape gesture (Behne et al., 2014; Pettenati et al., 2010; Stefanini et al., 2009). Pettenati et al. (2010) found that the majority of gestures used by the children while they completed the PiNG task were action-based gestures that depicted what was being shown in the image, or an aspect of an action being shown in an image. Children were far less likely to create a gesture that described the size and/or shape of an object. This lack of size/shape gestures is significant because it showed that the actions depicted by children through their gesture were not always linked explicitly to actions being shown in the images. In other words, children used action-based gestures even if the image that they were looking at depicted something that was stationary.

Behne et al. (2014) showed that children understand action-based gestures at an earlier age than size/shape gestures, providing explanation for why greater numbers of action-based gestures are used by young children. Although this study did not make use of the PiNG task, their findings regarding the gesture produced by the 20-21 month and 27-month-old children showed the same pattern as other studies. They found that the younger children (20-21 months) were less likely to use iconic gesture than were the 27-month-old children, but that all children of this age were able to create spontaneous iconic gestures to aid in the completion of a task.

Behne et al. (2014) suggest that young children use iconic gesture to fill in gaps in their lexical repertoire as they attempt to complete the task.

Gestures serve one of two functions for young children as they complete the PiNG task. First, they serve as a substitute for a lexical gap, allowing the child to describe an image for which they have no label (Behne et al., 2014; Hall et al., 2013). Second, they add supplementary or complementary information to the speech that they accompany (Hall et al., 2013; Pettenati et al., 2010; Stefanini et al., 2009). Given that there is an assumed difference between the lexical development of monolingual children and children with multiple language exposure, as will be discussed in more detail shortly, this leads to the question; will bilingual children's gesture differ from monolingual children's due to this difference in lexical ability?

Very little research has been done looking specifically at young, bilingual children's gesture. This area of research is significant because, as will be discussed in more detail below, there are clear differences in the lexical development of young monolingual and bilingual children. There is a strong link between speech and gesture, suggesting that a difference in lexical development may be reflected in the gesture produced.

Nicoladis, Mayberry and Genesee (1999) conducted a study in which they looked at 5 bilingual boys longitudinally from the ages of 2;0- 3;6. Similar to research already discussed, the children in this study were more likely to use deictic gestures; iconic gestures were more likely to be used as the child's lexicon grew. All of the children were being raised in a simultaneous dual language environment, so

they were not primed to use more gesture in one language over the other to compensate for lack of vocabulary skills.

Although there is a clear lack of research in the area of bilingual children's gesture use, a number of cross-cultural studies have been done with young children, with particular interest invested in cultures that are assumed to use lots or very little gesture. Cross-cultural research is relevant to bilingual individuals because these studies have shown that there are differences in the frequency with which different cultures use gesture (Iverson et al., 2008; Marentette et al., under review; Pettenati et al., 2012). This is pertinent to the current study because bilingual children often come from cross-cultural families, which could influence the amount of gesture that they use.

Iverson et al. (2008) looked at the differences in gesture between Italian and American children longitudinally, between 0;10-1;9 years-of-age. They found that Italian children gestured more frequently than American children. Furthermore, they found that Italian children were more likely to use iconic gestures, while American children used far more deictic gestures. This study clearly showed the relationship that exists between lexical development and gesture as the Italian children used more gesture but had smaller lexicons than American children who used fewer gestures but had larger lexicons. The differences in lexical development suggest that the iconic gestures that young children use serve as a way for children to fill gaps in their lexicons with a second modality when spoken words fail them. While the previously mentioned Behne et al. (2014) study found similar results, this

study went beyond to show a cross-cultural difference, rather than simply an age difference.

The results found by Iverson et al. (2008) also suggest that the environment in which a child is raised influences their lexical development and, by association, the amount of gesture that they will use. Italian children are constantly exposed to iconic gesture in their everyday lives, as the culture that they live in is very gesture-rich. American children, by contrast, live in a culture that is far less reliant on iconic gesture. While the results support the hypothesis that culture makes a difference to gesture use in children, it is important to note that the number of participants in this study was quite small (n = 6, Iverson et al., 2008). This study looked at frequency differences between different groups of participants, but did not examine gesture form.

A cross-cultural study by Pettenati, Sekine, Congestrì and Volterra (2012) looked at the gesture produced by Japanese and Italian children in comparison to one another using the PiNG task. They found that children were far more likely to use representational gestures when asked to label an image of an action than when asked to label an object. Pettenati et al. (2012) also found that Italian children produced more gestures without any accompanying speech, while Japanese children produced gestures with an incorrect verbal label. These results suggest that in these circumstances, children in both cultures use gesture to help them label the images when their lexicons fail them. This study adds to the previous research done by Behne et al. (2014) and Iverson et al. (2008) by examining the differences in gesture

frequency, across two very different cultures, one that is assumed to be gesture rich and one that is not.

Marentette et al. (under review), looked at the differences in the form and frequency of gestures across a group of Canadian and a group of Italian children as they performed the PiNG task. Replicating results found by Iverson et al. (2008), they found that Italian children gestured significantly more than Canadian children. What makes this study unique was the focus on gesture form as well frequency. Similar to Pettenati et al. (2010), they found that children were least likely to use size/shape gestures, but that Italian children were more likely to use them than Canadian children. In this study, gesture form was coded into four different categories, hand-as-object, hand-as-hand, size-and-shape, and own-body. This unique coding scheme allowed the researchers to examine what gesture form was most common among two-year-old children. They found that children were able to provide a wide range of information about the images that they were presented with, using iconic gesture to represent complex ideas.

The following sections will discuss lexical development in both monolingual and bilingual children and why there is a need to study the gesture produced by both in comparison to one another. It appears, based on the research that has been reviewed, that two-year-old children, both monolingual and those with multiple language exposure use iconic gesture to help fill in gaps in their lexicons when completing activities like the PiNG task.

Lexical Development in Monolingual Children

Children learn nouns and verbs at different points during the process of lexical development. Research shows a clear noun bias in early lexical development for monolingual children (Golinkoff & Hirsh-Pasek, 2008; Kauschke & von Frankenberg, 2008; Klassert, Gagarnia, & Kauschke, 2014; Longobardi, Rossi-Arnaud, Spataro, Putnick, & Bornstein, 2014). This noun bias, first introduced by Gentner (1982) suggests that there is a disposition among young children, across a variety of languages, to learn nouns before verbs. Furthermore, Hoff (2005) showed that, not only do children learn nouns earlier than verbs, but they also learn more nouns than verbs in early lexical development.

Nouns represent objects or beings in the world while verbs are relational and represent something more complex, often involving the interaction of objects in an environment or with other objects (Golinkoff & Hirsh-Pasek, 2008; Longobardi et al., 2014). If a child did not have a basic understanding of nouns, it could be predicted that they would be unable to use verbs because they would lack the understanding necessary to put the two word categories together. For example, in order for a child to use the verb 'put', in the instance of "put the book on the shelf", they must have an understanding of the object they want to put somewhere (the book) as well as an understanding of the location where they want that object to be put (on the shelf). In other words, they must have an understanding of the two nouns, book and shelf, before they can understand and be able to use the verb 'put' in this context. Golinkoff and Hirsh Pasek (2008) suggested that because verbs include these complicated interactions, they require a higher level of cognitive processing. (i.e.,

verbs may require a child to pick out smaller components of a complex event in order to make sense of it). Therefore, this noun bias is logical, as nouns are less complex than verbs and do not require interactions with other objects to be meaningful.

The noun bias has been studied in monolingual children across many different languages, including Italian, Russian, German, and English, and found to be robust in children ranging in age from one-and-a-half-years to eight-years (Golinkoff & Hirsh-Pasek, 2008; Kauschke & von Frankenberg, 2008; Klassert et al., 2014; Longobardi et al., 2014). However, most research has been done with children around the age of one-and-a-half- to three-years, as this is the time when children experience the greatest lexical growth. According to Hoff (2005), between the ages of 18-30 months children's lexicons expand from 50 to between 200 and 500 words, making this an ideal age to study lexical development. In addition, this is the age at which children expand their lexicons to include verbs (Longobardi et al., 2014; Gentner, 1982; Golinkoff & Hirsch-Pasek, 2008). Given the research that was discussed earlier about children's gesture, Stefanini, Bello, Caselli, Iverson and Volterra (2009) found that when children were asked to name pictures, they were more likely to use nouns than verbs, supporting the idea of a noun bias in young children.

The context of a given situation has also been shown to have an impact on the type of language used by children (Longobardi et al., 2014). In this study, the language that children used in two different situations was studied. When children were reading, they used far more nouns than verbs. This makes sense, as the majority of the content in children's picture books are nouns, making them the

majority of words produced by both child and caregiver. Conversely, when children were at play, the number of nouns and verbs used was shown to be more equal, (Longobardi et al., 2014).

Lexical Development in Bilingual Children

While the above research studied primarily monolingual children, this study also considered the lexical development of children with MLE, as it has been shown to differ slightly from monolingual lexical development. There are two different ways that a child can become bilingual. First, and most widely studied, is simultaneous dual language acquisition. This is where the child learns multiple languages at the same time (Arnberg, 1987; Sebastián-Gallés, 2008; MacLeod, Fabiano-Smith, Boegner-Pagé & Fontolliet, 2012). The second way a child can become bilingual is through successive dual language acquisition, where they learn one language first and a second at a later time (Arnberg, 1987; Sebastián-Gallés, 2008).

Simultaneous dual language acquisition is of great interest to those studying language development in children. Successive dual language acquisition cannot be studied with young children because one language must be learned first and they are too young to have mastered one language. However, researchers have also studied simultaneous duel language acquisition to uncover whether or not learning two languages is damaging to linguistic development (Arnberg, 1987; Sebastián-Gallés, 2008; MacLeod et al., 2012). The overwhelming consensus is that bilingual children follow the same developmental trends as their monolingual peers but may lag very slightly and briefly behind. This lag stems from the children's need to split

their cognitive attention between two or more languages (Arnberg, 1987; Sebastián-Gallés, 2008; Hoff et al., 2012; Pearson, 2009). Moreover, because bilingual children are afforded less exposure to each of their given languages than monolingual children are with one language, they require more time to gain the same level of lexical development (Arnberg, 1987).

The vocabularies of each of a bilingual child's languages are small compared to those of their monolingual peers although bilingual children also show a disposition to learn nouns earlier and in greater numbers than verbs (Hoff et al., 2012; Scheele, Leseman, & Mayo, 2010). When bilingual children complete tests designed to study their vocabulary in only one language, they perform more poorly than matched monolingual children. However, when both of their lexicons are considered, their full vocabulary is equal to or greater than that of monolingual children. Thus, when allowed to use both languages, they perform at the same level as the monolingual children (Arnberg, 1987; Sebastián-Gallés, 2008; Hoff et al., 2012; Romaine, 1989; Scheele et al., 2010). Around the age of three or four, this linguistic lag decreases as the child begins to understand that they speak two different languages and can learn to distinguish between the two. Once this distinction has been made, the language of each lexicon will 'catch up' with those of their monolingual peers (Romaine, 1989).

David & Wei (2005) found that the two languages that a bilingual child learns between the ages of one-to-three, follow the same developmental trends that one language does for a monolingual child; in other words, they found that the two languages developed in a parallel way and in much the same way as monolingual

children. A bilingual child's lexicon is comprised of the same linguistic categories as a monolingual child, namely nouns and verbs. These categories develop at the same speed and in the same proportion in each language, no matter the number of words in each individual lexicon (David & Wei, 2005). This study shows that differences and similarities have been documented in the way that monolingual and bilingual children learn language(s).

Although bilingual children under the age of three display a noun bias similar to that of monolingual children, Klassert et al. (2014) show that the tendency to use more nouns than verbs when completing a picture-naming task decreases earlier than in monolingual children. In their study, Klassert et al. found that bilingual children from the ages of four-to-seven showed fairly balanced abilities when it came to naming both nouns and verbs in a picture-naming task, while the Russian and German monolinguals in the study still showed the typical noun bias. This study showed that there are differences in the way that monolingual and bilingual children's lexicons develop. However, the general developmental trends of a bilingual child's two lexicons have been shown to be relatively similar.

Clearly there is a difference in the way that monolingual and bilingual children's lexical development occurs and because there is a strong link between language and gesture, it make sense that there could be a difference in the frequency and form of gesture that is produced by these young children. No study that I am aware of has looked at both the form and frequency of gestures produced by both monolingual and bilingual children.

This Study

There is a difference in the frequency with which children learn nouns and verbs in their speech and the time in their lexical development when each category is learned. In light of this and the close relationship that exists between speech and gesture, particularly during language acquisition, this study focused on children's spoken responses and gesture when they were shown an image of an object versus an image of an event using the PiNG task. Because nouns and verbs are learned differently during lexical development, it stands to reason that the co-speech, iconic gesture that was created when a child named an object or event image would differ in accordance with what word type was produced. Two-year-old children were ideal to study in this circumstance because of the increase in lexical development that takes place between the ages of two and three and because this lexical development co-occurs with the emergence of iconic gesture.

I further proposed that examining the form of children's co-speech gesture would allow me to consider the form and frequency of two-year-old children's gesture. This was significant because it allowed for insight into how much young children actually know about the images that they were shown as they completed the picture-naming task, as well as whether or not multiple language exposure made a difference to the gesture produced. Only one study examined the specific representational strategies used in young children's iconic gesture (Marentette et al., under review).

I looked at both gesture production between two matched groups of children and then differences in the information conveyed through their gesture. Hypotheses

one, two and three made comparisons between children with multiple language exposure (MLE) and monolingual children, looking at gesture production by child. Hypotheses four and five looked at the gesture produced across all three groups of participants.

First, I predicted that children with MLE will use greater numbers of co-speech gestures overall. The research discussed earlier suggests that a child who is actively learning more than one language typically has two lexicons that, each on their own, are smaller than the lexicon of a monolingual child. It was predicted that they would gesture more to help compensate for a lack of vocabulary. Children around the age of two were particularly appropriate to examine this hypothesis, because although the two lexicons that a child with MLE may have developed are both smaller on their own than the single lexicon of their monolingual peers, they experience a similar vocabulary spurt around this age.

Second, I predicted that children with MLE would differ in their pattern of modality use in response to the images. Specifically, that they would have more instances of unimodal gesture. Because this task was administered in English, they may have had smaller English lexicons than the monolingual children, leading them to use more gesture to compensate and possibly resulting in the use of gesture without speech. Furthermore, because of this supposed difference in lexicon size between the two groups (Hoff et al., 2005), I predicted that children with MLE would be less accurate in their verbal responses than the matched monolingual group.

Third, replicating and expanding on previous findings, it was hypothesized that both monolingual children and children with MLE would use more gesture when asked to name an event rather than an object. An object is straightforward to name and as such, it was predicted that any co-speech gesture that accompanied the naming would be simple as well. Furthermore, children were expected to know more noun words than verbs, suggesting that they would be able to name more object images. On the contrary, events are dynamic and relational. Supported by results found by Stefanini et al. (2009), it was predicted that more gesture would be utilized to help the child understand and describe the dynamic relationships that were pictured.

For the final two hypotheses, the analysis of gesture changed from one that looked at gesture production by child, between the matched groups, to one that looked at the gestures themselves, across all of the participants.

Fourth, it was predicted that if the child's spoken response does not correctly label the image they are shown, any gesture that is used would match the word type that was spoken and not the image type shown. For example, if a child were shown an image of a fork (an object) and instead of labeling it as a fork, they responded with, "eating", any gesture that accompanied this utterance would represent the word spoken rather than the object in the image. The definition of an iconic gesture, as described by McNeill (1992), suggests that they are most often co-speech and that the speech and gesture work together to provide information. Based on this definition, the speech and gesture should always match, in order for the communication to make sense.

Fifth, when both monolingual children and children with MLE were shown an image of an object or an event and asked to name what they were being shown, it was predicted that the any co-speech gesture that may have accompanied the spoken response would provide information about what the object was used for or how an event took place. More specifically, when a child named an image of an object the gesture that I expected to accompany this type of utterance would represent **what** the object was used for. Conversely, when a child named an image of an event I expected any gesture that accompanied the spoken response would represent **how** an event was completed. This was expected because nouns (objects) are simpler and do not require as much elaboration whereas verbs (events) are more complicated and might require more elaboration using gesture.

Methods

Participants

In this study, two groups of children were studied. First, eight children with multiple language exposure (MLE), and eight matched monolingual children were studied to address hypotheses 1, 2 and 3. Second, a group of 11 monolingual comparison children were added to the sample to allow me to consider hypotheses 4 and 5 in light of a more expanded dataset, as children tend not to gesture in large numbers (Marentette et al., under review). The demographic information about the children is provided in Table 1.

-- insert Table 1 here --

MLE was determined by the parent who was asked to fill out a basic information form. This form asked questions such as, “Is your child regularly

exposed to a second language?" and "If yes, by whom and how often?" This additional language exposure information is shown in Table 2.

-- Insert Table 2 here --

Participants were recruited largely through word of mouth. Posters were put up in local public buildings such as the public library and personal connections to families with young children were utilized. Camrose is a small community so recruiting a larger group of children was difficult. Due to this, all children with multiple language exposure were included in the MLE dataset. All of the children were exposed to English. Only two children in my MLE sample came from homes that used a different language other than English consistently.

Materials

PiNG. The Picture Naming Game (PiNG) has been widely used in recent research to elicit gestures from young children (Bello et al., 2010). This task utilizes two sets of 22 triplet picture sets including two training items for each, with the first set of 22 triplet image sets that are shown to the children depicting nouns and the second set depicting predicates. The participants are shown the triplet sets in a fixed order, while the three images within each set are presented in a random order. One of the images is a comprehension target, one is a production target, and the final image is a distractor as shown in Figure 1.

The PiNG task is used to collect both comprehension and production data. For the comprehension target, the participant is asked to identify one image (e.g., "Can you show me the cat?"). Following this, the comprehension target and the

distractor are removed and the participant is asked to name the remaining picture (e.g., “What is this a picture of?”).

-- insert Figure 1 here --

The task was administered individually and the children were videotaped so that the data could be coded for any gestures produced. The children provided behavioural assent as the task was performed, while the caregiver, present in the room at the time of the task, filled out the official consent form.

Other than pointing to the images to focus a participant’s attention, the experimenter avoided gesture use of any kind. The two sets of images were tested separately with a break in between as required. The task took approximately 20 minutes per child.

MB-CDI. The MacArthur-Bates Communicative Development Inventory – Words and Sentences is a parental report that was filled out by the parents of children less than 30 months of age (Fenson et al., 2007). The MacArthur-Bates Communicative Development Inventory – III is a second parental report that was filled out for children over 30 months of age (Fenson et al., 2007). These inventories allowed me to compare monolingual and bilingual children’s English lexicon size. The parent was instructed only to fill out the words that they have actually heard their child produce before (not the ones that they believe that they understand).

Coding

Verbal Accuracy. I coded whether or not children would label the picture that they were being shown correctly. Children were given two chances to correctly label the production image. If they answered correctly by the second try, that

answer was recorded. If neither answer given was correct, the first response was recorded. If a child's answer was unclear or if they did not speak, it was recorded in a separate category. In certain cases, multiple answers were considered as possible responses (e.g., a glass can also be called a cup in Canadian English).

Gesture Frequency. All manual representational gestures produced after the first image was placed before the children were coded. Only gestures that were produced with the hands were coded as iconic. Posture, body movements and facial expressions were not considered. Deictic gestures were not coded. I coded iconic gestures that occurred with and without speech. Any gestures that were produced while naming any image, performed with the card in hand or with an empty hand, and not performed as an imitation of an adult's gesture, were coded (Pettenati et al., 2010).

In all cases, I recorded repetitions of gestures. In the majority of instances, the speech that accompanied a repeated gesture differed (e.g., the gesture was rubbing hands together for the target word "washing" and the child repeated the gesture three different times, saying the three different words, "washing," "soap," and "scrubbing"). I hypothesized that the image, the label that the child gave it and any gesture that accompanied the image matched up, so it was necessary to record repeated gestures that occurred with differing speech, as there was differences in the information conveyed among the repetitions.

Representational Strategy. In order to code representational strategy, I looked at what the gesture was conveying about the object or event in the image, ignoring the words that were spoken by the child.

Modality. There were four different categories that could be coded for modality. The first was unimodal speech. In this instance, the child provided relevant information about the image using only spoken words. The second category was unimodal gesture. In this instance, the child provided relevant information about the image using only gesture. The third category was bimodal speech and gesture. In this instance, the child provided relevant information about the image using both spoken words and accompanying gesture. The fourth category was no response/unintelligible response. This category was added for those instances where the child produced no response or if the response that they produced was not codable.

Representation. In the cases where gesture was produced, whether or not it was accompanied by speech, I was interested in the information that gesture provided about the image the child was being shown. Gesture was coded into one of five categories.

What. This first representational category included instances when the gesture represented what an object looked like. For example, the child was shown an image of an umbrella and responded by holding their arm above their head with their fingers spread out. The gesture shows the experimenter what was depicted in the image in a complementary way to the image.

Where. This second category included gesture that represented where an object was used, typically in relation to the body. For example, the child was shown an image of a hat and responded by touching their head. In this case, the gesture showed where the hat is used on the body, providing supplementary information.

How. This category represented how an object was to be used or how an event took place. For example, in the case of an object, the child was shown an image of a fork and responded by gesturing with their fist moving towards their mouth (mimicking moving a fork towards the mouth). An example in the case of an event was when the child was shown an image of a person in a swimming pool and responded with by gesturing by moving their arms in a front crawl motion. In both instances, the gesture provided more information about how the object was used or how the task took place.

Combinations. Some gesture represented more than one of the previously described categories simultaneously. For example, the child was shown an image of a comb, responded with the word "comb" and touched their head with spread fingers, pulling them through their hair. In this case, the gesture represents all three of the above categories by showing **what** (fingers spread to represent the comb), **where** (touching their head) and **how** (pulling fingers through hair, showing the act of combing).

Uncodable. This category included instances where the gesture was uncodable. In these cases, the gesture that was produced could not be coded as representing what, where, how or any combination of the three. For example, the child was shown an image of someone driving a car and gestured by clasping their palms together and rotating their wrists and clasped hands. In this instance, the gesture was unrecognizable as a representation of any of the above three categories. This only occurred 2/54 times in my data.

Reliability

Of the videos of eight children who had second language exposure, four were randomly chosen to be coded for reliability on whether or not a gesture occurred as well as what the gesture represented. Using Cohen's Kappa, which accommodates for chance matching, we found a very good measure of reliability, $\kappa = 0.80$, $\kappa_{max} = 0.80$. The percent agreement was 98.75%.

Results

The demographics between the matched monolingual and MLE children were carefully controlled. An independent sample t-test showed that there was no significant difference between the ages of the matched monolingual and MLE groups, $t(4) = 0.94$, $p > 0.05$ (see Table 1). There was no difference between the genders of the matched monolingual and MLE groups. There was no significant difference between the MB – CDI scores of the matched monolingual and MLE groups, $t(4) = 0.98$, $p > 0.05$.

The first three hypotheses focus on the comparison between the matched monolingual and MLE groups, using the gesture production by child as the unit of analysis.

The first hypothesis predicted that children with MLE would gesture more frequently than monolingual children. Looking only at the matched MLE and monolingual children by child, a t-test showed no significant difference between the two groups. Children with MLE produced 22 gestures ($M = 2.75$, $SD = 4.41$) while monolingual children produced 17 ($M = 2.13$, $SD = 2.52$), $t(14) = 0.75$, $p > 0.05$.

The second hypothesis predicted that children in the MLE group would have different patterns of modality use as they completed the task. Specifically, it was predicted that children with MLE would have more instances where they used unimodal gesture to describe the images they were being shown. A chi square showed no significant differences between the two matched groups in terms of the patterns of responses they gave, X^2 (3) = 2.00, $p > 0.05$. Among the matched monolingual children, there was one instance of unimodal gesture and among the MLE children there were 4. The different types of modalities are shown in Figure 2 across the two groups. The numbers of the different response types are shown in Table 3.

Following this a second analysis was done looking specifically at instances of unimodal speech to determine how accurate the children were when naming the images with the prediction that monolingual children would be more accurate than children with MLE because of the presumed lexical differences. A chi square showed that there was no significant difference between the matched monolingual and MLE groups in terms of verbal accuracy, X^2 (3) = 3.41, $p > 0.05$. The average numbers of responses by accuracy are shown in Figure 3.

-- insert Table 3, Figure 2 and Figure 3 here --

The third hypothesis predicted that children across all of the groups would gesture more when shown an image of an event rather than an object. A 2 (group: monolingual, MLE) × 2 (picture type: object, event) ANOVA showed that there was no significant main effect for the picture type ($F (1, 14) = 1.13$, $p = 0.31$) or for the

group that the children belonged to ($F(1, 14) = 0.07$, $p = 0.80$). There was no significant interaction. A box plot of these main effects is shown in Figure 4.

-- insert Figure 4 here --

The final two hypotheses, analyze representational strategy across the gesture produced by all three groups of children, using the gesture itself as the unit of analysis rather than the child.

The fourth hypothesis predicted that if there were an instance where a child did not correctly label the image that they were being shown, any gesture that accompanied the speech would match the speech, not the image. In 100% of cases where gesture was produced along with recognizable speech (40/54 gestures), the gesture matched the speech produced, not the image. For the 14 other gestures, the speech (or lack thereof) was unrecognizable and therefore, unable to be coded. The number of image/speech matches and mismatches (verbal accuracy) can be seen in Figure 3. As can be seen in this Figure, there were many instances where the children mislabeled the images they were shown. In all of these instances, if gesture was produced, the gesture matched the speech that it accompanied and not the image that was mislabeled.

The fifth hypothesis examined children's gestural representation when they were shown an image of an object vs. an event. I predicted that when shown images of objects, children would provide information in their gesture about **what** the object looked like. Conversely, when shown images of events, their gesture would provide information about **how** an event took place. Because of the small sample size, I will show the raw representation data in two tables.

I considered this hypothesis in two ways. As can be seen in Figure 5, there are three different elements to be considered when analyzing the responses provided when children label images: the image, the spoken response given and any gesture that accompanies the response. First, I considered my data with regards to the relationship between the image and the gesture, regardless of the speech produced by the child. Following this, a second analysis considered my data with regards to the relationship between the speech and the gesture, regardless of the image shown.

Shown in Table 4 are the 54 gestures produced, separated by representational strategy considering the target picture. The gestures produced when children named object images tended to give information about **how** or **where** an object was used (13/23). The gestures produced when children named images of events tended to give information about **how** to complete a task (19/31). It was rare for the children to gesture only about **what** the object looked like when shown an image of an object as I had predicted they would (3/23). When the children named images of objects, there were 7 instances of combinations and when they named events, there were 5. These will be considered in more detail in the discussion section of the paper.

Shown in Table 5 are the 54 gestures that were produced, organized considering the child's speech. If the child's speech was considered, the representational pattern turned out to be quite similar. The most common information found in the children's gestures when naming objects was **where** the object was used (8/17). The representation most common in their gestures when

naming events was **how** the event took place (20/23). A third category was necessary for instances where the child's speech did not fit into either category. This third category included instances where the child did not respond with either a noun or a verb (i.e., one instance of adjective response), where their speech was unrecognizable (n = 10), or where they did not provide a spoken response at all (n = 3). In the instances where the child did not respond with either a noun or a verb, the tendency in the gesture was to provide information about **how** the object was to be used or the event took place (5/14). When children provided a noun as a response, there were 5 instances of combinations, when they provided a verb there were 2 instances and when they did not provide either a noun or a verb there were 5. These will be considered further in the discussion section.

-- insert Table 4 and 5 here --

Discussion

Multiple Language Exposure

The analysis of these data showed that there was no difference between the children with multiple language exposure and the monolingual children.

The first three hypotheses considered children's gesture by comparing a matched group of monolingual children and children with MLE. The gesture in these hypotheses was considered by child, making a direct comparison between the two groups. Based on the research about children's lexical development and the differences in lexical development between bilingual and monolingual children, these hypotheses predicted that there would be a difference in the verbal accuracy and gesture produced by these two groups.

The analysis that resulted from these first three hypotheses made one thing very clear: there was no difference between the two matched groups of children. It appears that simply being exposed to intermittent multiple languages had no impact on lexical development on the children in this study, as the MLE group and the monolingual group showed similar verbal accuracy and patterns of gesture use.

There is clearly a difference between the children with MLE in this sample and children actively learning more than one language as discussed earlier in the literature review. The research shows that there is a difference in lexical development between children who are actively learning more than one language and children who are only receiving input from one language (Hoff et al., 2012). However, from the data that were analyzed in this study, it appears that multiple language exposure did not affect children's development in the same way as active bilingual input.

The lack of significant results surrounding which images children gestured more for (object or event) is particularly interesting as the children in the monolingual group were a subset of a pre-existing dataset that showed this tendency for children to gesture with greater frequency when shown an image of an event rather than an image of an object. This suggests that the children in this sample may have been an odd subset of children, who happened to not display this bias. Furthermore, because these results have been shown in previous studies (Marentette et al., under review; Stefanini et al., 2009), it is possible that the lack of significant results, for this and for the first two hypotheses, may have come from the fact that my sample size was quite small. In the previous two studies, the sample

sizes were considerably larger than in this study. In the case of these first three hypotheses, I analyzed the data from 16 children across two groups, collecting a total of 39 gestures, making the sample size too small to compare my results to the studies that came before this one.

Representational Strategy

Now I will consider gesture with regards to the representational information that was conveyed in each gesture that the children produced. For hypotheses four and five, I shifted my analysis from one that looked at gesture by child between two matched groups to one that looked at the representational strategy demonstrated in the gesture itself, across all of the groups.

The data that I reported about children's gesture representation show that children provided complex information in their iconic gesture when naming images for the PiNG task. This suggests that children understand what is required for them to use multiple modalities to communicate in a way that provided meaningful information to the experimenter.

In the fourth hypothesis, it was predicted that children would always match their gesture to the speech, rather than the image they were shown. McNeill (1992) discussed that the link between speech and gesture was meaningful and indeed, this was found to be true in all circumstances where children produced gesture with intelligible accompanying speech. The iconic gesture used by adults is always co-speech, as adults typically have extensive vocabularies; however, in certain cases children did not always use accompanying speech or the speech that they used was incoherent. In the cases where the speech was recognizable, the children's speech

and gesture did not provide differing information in any circumstance. Instead, their gestures most often served to supplement the speech produced by providing related information. This suggests that in cases where they use gesture, speech and gesture work together in order for the communication to make sense.

When children were asked to name images of objects, they were expected to respond with a noun label. However, they did not always label the image in this expected way. There were cases the child was shown an image of an object and provided different information about the image. For example, one child was shown an image of a fork and responded with the verb “eating”. Although the expected response was the noun “fork” the response given provided relevant information about the use of the object.

The analysis also showed that children were unlikely to provide simple information about what an object looked like in their gesture. Instead, they were likely to provide more complex, supplementary information about where an object was to be used, relative to the body, or how an object was to be used. In the example above, the child did not gesture about the object in the image, instead they provided information about eating, linked to their spoken response. Because the child was describing something to the experimenter that was not shown in the image, they may have considered it necessary to include a supplementary gesture to help elaborate on their speech.

When children were asked to name images of events, they were expected to respond using more verbs because of the complex nature of the event images. Also, their gesture was expected to depict supplementary information about how an

event would take place, allowing the child to enrich their labels and explanations. The analysis showed that again, that children did not always label the image as was expected. Because of the complex nature of events, there were many different ways that a child could discuss these images. For instance, there are many ways that a child could respond to an image of someone riding on a merry-go-round. The expected response was "spinning" but the child could respond with other verbs like "playing" or "hanging (on)" instead. In another instance, when the children responded to the image of a person swimming, most responded with the word "swimming", which was the expected response. However, there was an instance where one child responded with the noun, "glasses" (referring to swimming goggles) and gestured to their eyes. In these two different cases, the alternative responses provide information about the image, although none were the expected response.

Second, this analysis showed, again, that children were most likely to provide supplementary information in their gesture to enrich the information given in the speech. For example, if the child was shown an image of a person swimming, their gesture served to add action more information to their "swimming" response. Responding with the word "swimming" tells the experimenter exactly what is in the image, but provides no more information than that.

A subset of children's gestures provided a combination of the different categories of representation. These gesture provided more holistic information than in any of the cases where the gesture represented only one of the representational categories. In the case of 'combing', when the child pulled their spread fingers

through their hair, they provided all the information the experimenter could ever ask about the action of combing. Gesture like this was described in the current study as a combination of representation information. In a previous study, they were described as 'hand-as-object' (Marentette et al., under review). They showed what the comb looked like, where on the body it was used and how it was used. A gesture like this, that accompanied the speech "comb" or "combing", suggests that children know more than just a simple name and are very interested in communicating all of the knowledge that they have about the object to the experimenter.

I argue that children's communication as they complete the PiNG task is comprised of three different aspects; the image, the speech and any gesture (see Figure 5). After considering the fifth hypothesis in two ways, with regards to the image-gesture match and then with regards to the speech-gesture match, it became clear that it was the second analysis that provided important information about children's gesture. As was found in hypothesis four, children prioritize connection to their speech (not the image shown) when they use iconic gesture. Based on the analysis, I argue that children understand that their speech and gesture must match in order for their communication to make sense and to allow them to provide complex or holistic information about the images that they are shown.

Representation Information

Previous studies considered the representation information that is contained in children's iconic gesture (Marentette et al., under review; Pettenati et al., 2010). In both of these studies it was found that children were more likely to create action based gestures (e.g. pulling fingers through hair to represent 'combing') than

size/shape gestures (e.g. holding palms up to show the shape of a pair of gloves). Pettenati et al. (2010) found, similar to the current study, that the creation of action based gestures was not necessarily linked to action shown in the images.

The current study looked at the representational information within the gestures by considering what the information in the gestures was communicating along with the accompanying speech. Rather than considering the gesture in terms of size/shape or action-based, it was considered in terms of what, where or how. However, these categories appear to be fairly similar, with **what** comparable to **size/shape**. In this case, simple information is expected. **How** is comparable to **action-based** gestures. In this case, the information is intended to represent the action displayed in an image.

The data in this study clearly show a preference for gestures that are action-based and give complex information about how objects are used or events take place, as has been shown in previous studies. However, the coding scheme that was used in previous research was either too broad (Pettenati et al., 2010) or too conflated (Marentette et al., under review) to allow for an examination of the aforementioned hypotheses. Based on comparisons with previous symbolic processing literature, the coding scheme in this study focused on more specific representation information, across all iconic gesture produced. The analysis of my data showed, as also found in the Marentette et al. (under review) study, that children were most likely to create gestures that included how (action) information. In any instances where the children produced what (shape) gestures, it was likely to include how/action information as well.

Taken together, this study has expanded on previous research about children's gesture use by showing that children are capable of communicating very complex information using both their speech and gesture. These data show that children are most likely to use supplementary speech plus gesture communication and that the gesture used is most likely to provide information about how an object is to be used or how an event takes place. This suggests that children are capable of complex interactions and that they may be aware of what information might be useful to the person with whom they are communicating.

Communication Abilities

McNeill (1992) proposed that adults use the combination of speech and iconic gesture to create holistic communicative events. Adults are capable of providing rich information within speech alone; however, the addition of gesture provides another layer of complexity, making the communicative event bigger than in the case of only speech. Previous research has shown that children use supplementary iconic gesture plus word combinations to communicate complex information as they complete cognitive tasks (Özçalişkan and Goldin-Meadow, 2005; Iverson et al., 2008). In both of these studies, they showed that the use of two modalities allowed children to communicate more information than they would be able to using only one.

In the current study, I found results that suggest that children use two modalities in much the same way as McNeill (1992) discussed that adults do. The finding that they were unlikely to provide simple information about the size or shape of an object without other action information as well, suggests that children

used gesture as a way to supplement their speech, providing a more holistic communicative event. More specifically, the current research goes further, showing that children are able to communicate a great deal; however, due to lack of vocabulary at the young age of two, that complex communication stems from a gesture plus speech combination.

Conclusion

The speech that children use when naming images in the PiNG task did not provide information about how the objects are used or how the events take place, gesture was used as a way to supply this information to the researcher. In order to help elaborate on the their spoken response to the image, they include how information about the object or event itself within their gesture. Based on my data, I propose that children, like adults, are capable of providing complex information about these images, using multimodal communication.

Limitations

In this study, two significant limitations were noted. First, only two of the children in the sample were actively learning two languages in a simultaneous way. It is possible that the majority of the children in my study actually fell in the category of simultaneous dual language learning and that may learn to use more than one language later in their lives. However, up until the time of their testing, they simply hadn't had enough time or exposure to multiple languages at their young age. The lack of consistent and simultaneous exposure to multiple languages made it impossible to consider these children "bilingual". I had thought it possible

that a small amount of MLE might make a difference to the gesture used, but this turned out to be untrue.

The second limitation that was noted was that the sample size was extremely small. In the two comparison groups, used for the first three hypotheses, data analyzed from only 16 children was considered. In the last two hypotheses, data from 27 children was considered, with a total of only 54 iconic gestures provided. This is a very small number of children and gestures. The small numbers made it difficult to find significant results in the first three hypotheses and impossible to conduct statistical analysis in the case of the fifth hypothesis.

Implications for Future Research and Conclusions

Given these limitations, future research conducted with children who were actively learning more than one language would be interesting as significant results might be noted if their lexicons were impacted by the simultaneous language development. Furthermore, a larger sample size would allow for more gestures to be collected, possibly allowing my fifth hypothesis to be statistically analyzed.

This study provided original data about representational strategy in two-year-old children's gesture use. This study contributed to the already existing literature on children's gesture use by showing that children are likely to use speech and gesture together to create complex and holistic communicative events.

References

Arnberg, L. (1987). *Raising children bilingually: The pre-school years.* Bank House, England: Multilingual Matters Ltd.

Behne, T., Carpenter, M., & Tomasello, M. (2014). Young children create iconic gestures to inform others. *Developmental Psychology, 50*(8), 2049-2060. doi: http://dx.doi.org/10.1037/a0037224

Bello, A., Caselli, M. C., Pettenati, P., & Stefanini, S. (2010). *Parole in Gioco*. Firenze, Italia: Giunti.

David, A., & Wei, L. (2005). The composition of the bilingual lexicon. In Cohen, J., McAlister, K. T., Rolstad, K., & MacSwan, J. (Eds.), *Proceedings of the 4th International Symposium of Bilingualism* (pp. 594-607). Somerville, MA: Cascadilla Press.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories*. Baltimore, MD: Brookes Publishing.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kucazaj, II (Eds.), *Language Development Volume 2: Language, Thought, and Culture* (pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Goldin-Meadow, S. (2014). Widening the lens: What the manual modality reveals about language, learning and cognition. *Philosophical Transactions of the Royal Society B, 369*, 1-11. doi: http://dx.doi.org/10.1098/rstb.2013.0295

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12*(10), 397-403. doi: 10.1016/j.tics.2008.07.003

Hall, S., Rumney, L., Holler, J., & Kidd, E. (2013). Associations among play, gesture and early spoken language acquisition. *First Language, 33*(3), 294-312. doi: 0.1177/0142723713487618

Hoff, E. (2005). Language experience and language milestones during early childhood. In K. McCartney & D. Phillips (Eds.), *Blackwell Handbook of Early Childhood Development.* Blackwell Publishing (Blackwell Reference Online). doi: 10.1111/b.9781405120739.2005.x

Hoff, E., Core, C., Place, S., Rosario, R., Señor, S., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1-27. doi: 10.1017/S0305000910000759

Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development, 9*, 23-43.

Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language, 28*(2), 164-181. doi: 10.1177/0142723707087736

Kauschke, C., & von Frankenberg, J. (2008). The differential influence of lexical parameters on naming latencies in German. A study on noun and verb picture naming. *J Psycholinguist Res, 37*, 243-257. doi: DOI 10.1007/s10936-007-9068-5

Klassert, A., Gagarina, N., & Kauschke, C. (2014). Object and action naming in Russian- and German-speaking monolingual and bilingual children.

Bilingualism: Language and Cognition, 17(1), 73-88. doi: 10.1017/S1366728913000096

Longobardi, E., Rossi-Arnaud, C., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2014). Children's acquisition of nouns and verbs in Italian: Contrasting the roles of frequency and positional salience in maternal language. *Journal of Child Language,* FirstView, 1-27. doi: 10.1017/S0305000913000597

MacLeod, A., Fabiano-Smith, L., Boegner-Pagé, S., & Fontolliet, S. (2012). Simultaneous bilingual language acquisition: The role of parental input on receptive vocabulary development. *Child Language Teaching and Therapy, 29*(1), 131-142. doi: 10.1177/0265659012466862

Marentette, P., Pettenati, P., Bello, A., & Volterra, V. (under review). Strategies for symbolic representation: A cross-cultural examination of two-year-olds' gestures.

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

Nicoladis, E, Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology, 35*(2), 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. doi: 10.1007/s10936-009-9121-7

Özçalişkan, Ş, & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition, 96,* B101-B113. doi: 10.1016/j.cognition.2005.01.001

Pearson, B. Z. (2009). Children with two languages. In E. L. Bavin (Eds.), *The Cambridge handbook of child language* (pp. 379- 397). New York: Cambridge University Press.

Pettenati, P., Sekine, K., Congestrì, E., & Volterra, V. (2012). A comparative study on representational gestures in Italian and Japanese children. *Journal of Nonverbal Behaviour, 36,* 149-164. doi: 10.1007/s10919-011-0127-0

Pettenati, P., Stefanini, S., & Volterra, V. (2010) Motoric characteristics of representational gestures produced by young children in a naming task. *Journal of Child Language, 37*(4), 887-911. doi: 10.1017/S0305000909990092

Romaine, S. (1989). *Bilingualism.* Cambridge, MA: Blackwell Publishers.

Scheele, A. F., Leseman, P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*, 117-140. doi: 10.1017/S0142716409990191

Sebastián-Gallés, N. (2008). Bilingualism. In *Encyclopedia of infant and early childhood development* (Vol. 1, pp. 172-182). San Diego, CA: Elsevier Inc.

Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M., & Volterra, V. (2009). Co-speech gestures in a naming task: Developmental data. *Language and Cognitive Processes, 24*(2), 168-189. doi: 10.1080/01690960802187755

Table 1

Participant's Demographic Information.

	Multiple Language Exposure	Matched Monolingual	Comparison Monolingual
Age	Range: 26-34 mos	Range: 26-34 mos	Range: 24-33 mos
	M = 29.75	M = 29.63	M = 27.91
	SD = 2.99	SD = 3.16	SD = 3.18
Gender	5 Female	5 Female	4 Female
	3 Male	3 Male	7 Male
MB-CDI Percentile	Range: 5-95%	Range: 10-99%	Range: 5-99%
	M = 49.57	M = 49.00	M = 52.1
	SD = 23.95	SD = 31.94	SD = 29.3
	Missing: 1	Missing: 2	Missing: 1

Table 2

Multiple Language Exposure.

Child	Additional Language Exposure
MLE 1	French input
MLE 2	German input
MLE 3	French input*
MLE 4	French input*
MLE 5	French input
MLE 6	French input
MLE 7	German input
MLE 8	Twi input

Note. * = Simultaneous and Consistent Dual Language Exposure

Table 3

Number of Different Response Types by Child.

	Unimodal Speech	Unimodal Gesture	Bimodal	No Response
MLE 1	7	3	0	30
MLE 2	27	0	3	10
MLE 3	20	0	14	6
MLE 4	13	1	0	26
MLE 5	36	0	0	4
MLE 6	32	0	1	7
MLE 7	35	0	0	5
MLE 8	28	0	0	12
Total MLE	198	4	18	100
Monolingual 1	32	0	5	3
Monolingual 2	34	0	5	1
Monolingual 3	36	0	1	3
Monolingual 4	10	0	0	30
Monolingual 5	12	0	0	28
Monolingual 6	18	1	5	16
Monolingual 7	26	0	0	14
Monolingual 8	31	0	0	9
Total Monolingual	199	1	16	104

Table 4

Gesture Representation Information Organized by Target Image

	Target Image: Object	Target Image: Event
What	3	1
Where	5	4
How	8	19
What/How	3	0
Where/How	1	5
What/Where/How	3	0
Unknown	0	2
Total	**23**	**31**

Table 5

Gesture Representation Information Organized by the Child's Speech.

	Speech: Noun	Speech: Verb	Adjective or No response or Uncodable speech
What	2	0	2
Where	8	1	0
How	2	20	5
What/How	2	0	1
Where/How	1	2	3
Where/How/What	2	0	1
Unknown	0	0	2
Total	**17**	**23**	**14**

.

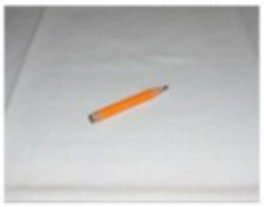
	Comprehension	Production	Distractor
Noun	 "Where is the cat?"	 "What is this?	
Predicate - event	 "Who is building?"	 "What is she doing?"	
Predicate -descriptor	 "Which one is short"	 "How is this one?"	

Figure 1. Picture Naming Game (PiNG task).

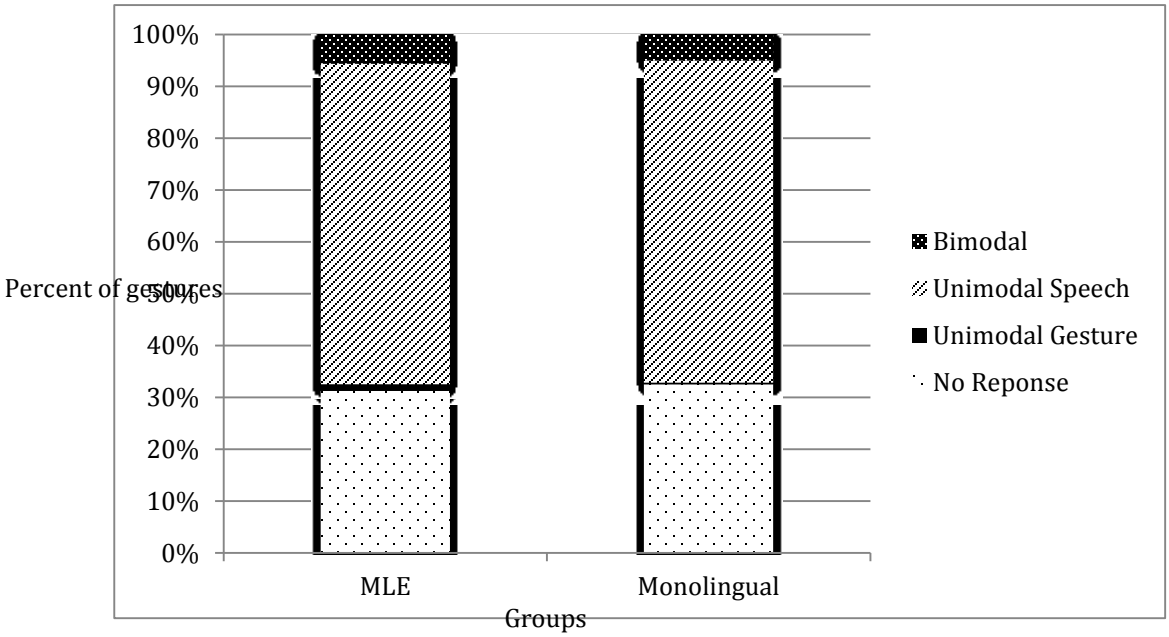


Figure 2. Gesture types by group.

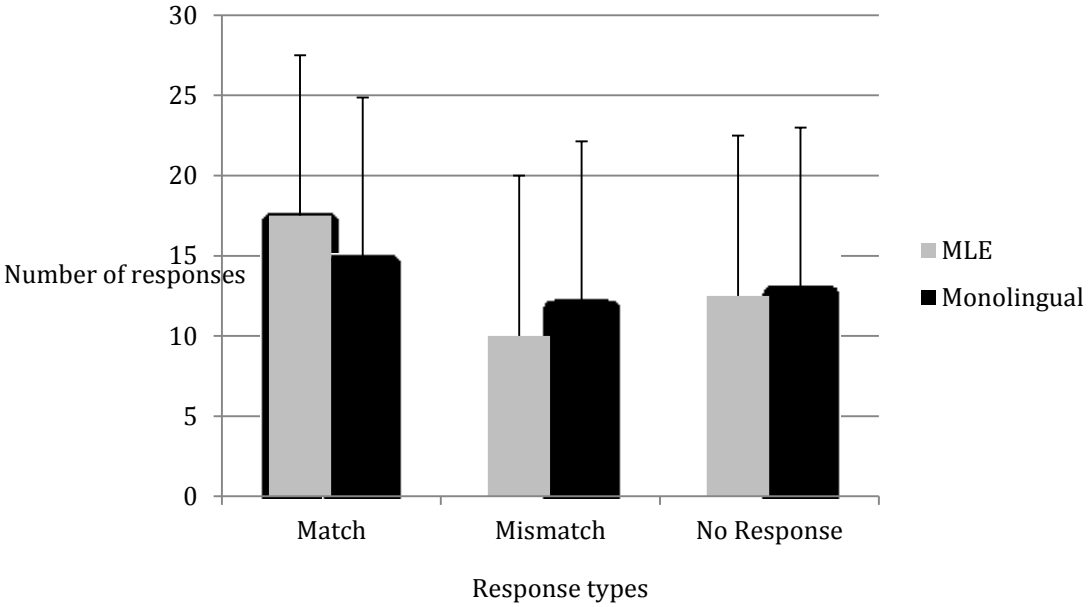


Figure 3. Average number of response types across the matched groups.

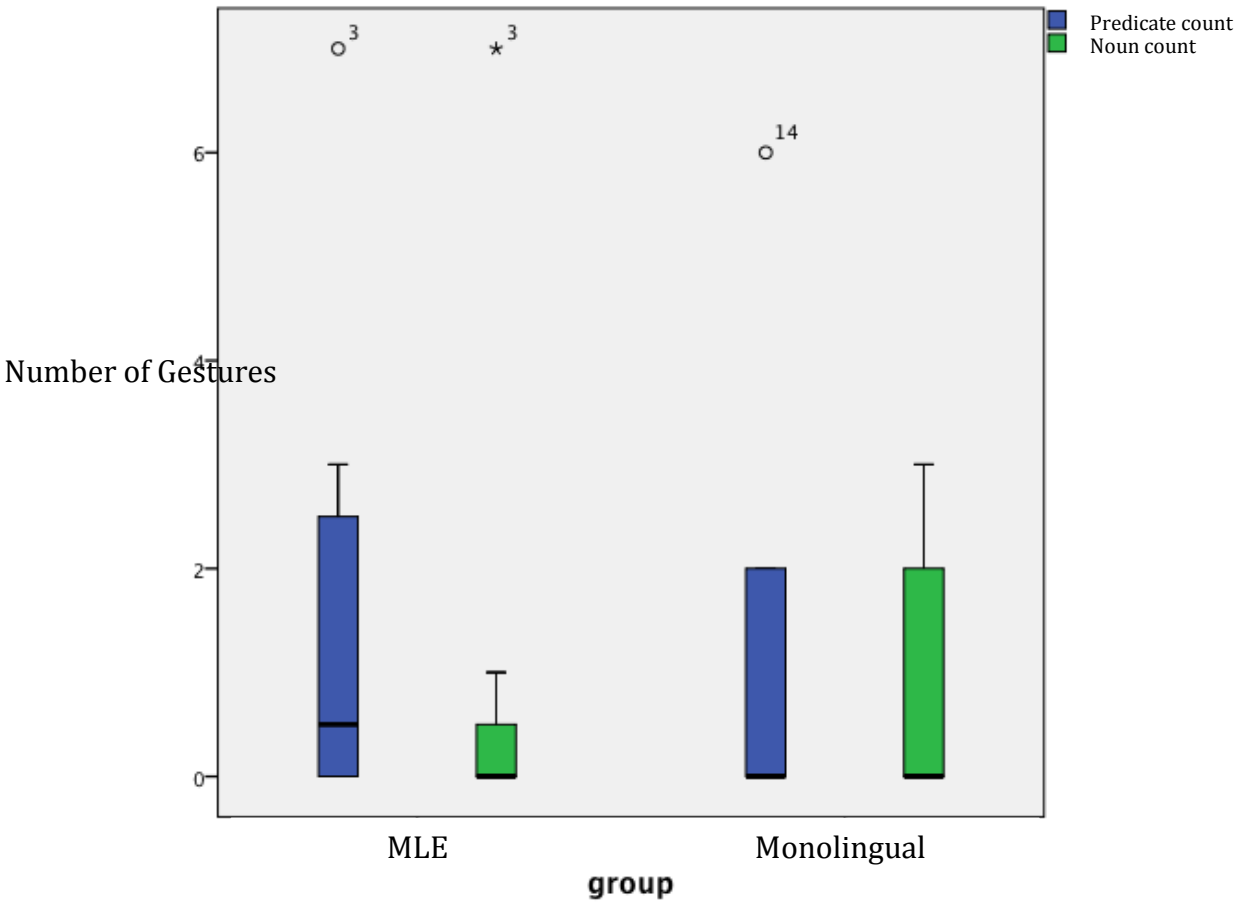


Figure 4. Box plot of 2 × 2 ANOVA by group and picture type. The bold line represents the median of each group while the upper side of each box, above the median up to the top of the error bar represents 50% of the data. The asterisk and two circles near the top represent the outliers (the children who produced far above average numbers of gestures).

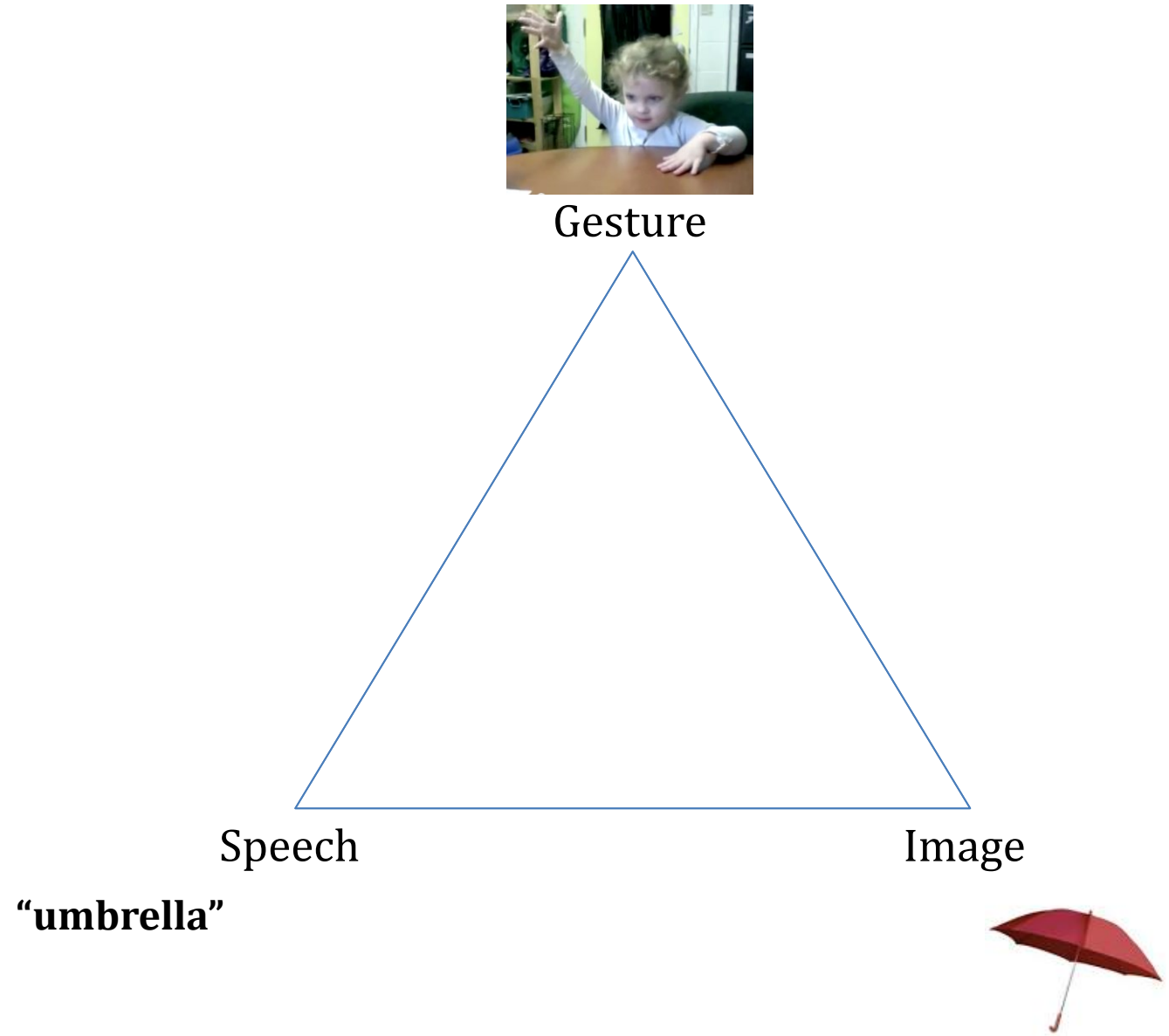


Figure 5. The three modes of representation.