

# When speech is ambiguous, gesture steps in: Sensitivity to discourse-pragmatic principles in early childhood

WING CHEE SO  
*National University of Singapore*

ÖZLEM ECE DEMIR and SUSAN GOLDIN-MEADOW  
*University of Chicago*

Received: July 16, 2008      Accepted for publication: March 14, 2009

## ADDRESS FOR CORRESPONDENCE

Wing Chee So, Department of Psychology, National University of Singapore, BLK AS4 9 Arts Link,  
Singapore. E-mail: psyswc@nus.edu.sg

## ABSTRACT

Young children produce gestures to disambiguate arguments. This study explores whether the gestures they produce are constrained by discourse-pragmatic principles: person and information status. We ask whether children use gesture more often to indicate the referents that have to be specified (i.e., third person and new referents) than the referents that do not have to be specified (i.e., first or second person and given referents). Chinese- and English-speaking children were videotaped while interacting spontaneously with adults, and their speech and gestures were coded for referential expressions. We found that both groups of children tended to use nouns when indicating third person and new referents but pronouns or null arguments when indicating first or second person and given referents. They also produced gestures more often when indicating third person and new referents, particularly when those referents were ambiguously conveyed by less explicit referring expressions (pronouns, null arguments). Thus Chinese- and English-speaking children show sensitivity to discourse-pragmatic principles not only in speech but also in gesture.

Young children often underspecify their intended referents starting in the two-word stage and continuing until 4 to 5 years of age, whether or not their ambient language permits underspecification (Allen, 2000; Serratrice, 2005; Valian, 1991). For example, a child might say “ $\phi$  eat cookies” ( $\phi$  refers to the omitted eater) or “I like this one” (“this one” refers to a particular puzzle), even when it is not clear from the context who is doing the eating or what the child likes. However, children routinely gesture when they talk (Goldin-Meadow, 2003; McNeill, 1992,

2005), and might be using gesture to compensate for their underspecification in speech. The child could, for example, point to the girl munching on a cookie while saying “eat cookies,” or point to his favorite puzzle while saying “I like this one.”

Gesture and speech develop together during the early language learning period (e.g., Volterra, Caselli, Capirci, & Pizzuto, 2005). After age 2, children construct an integrated speech–gesture system as they acquire their language (Butcher & Goldin-Meadow, 2000; Mayberry & Nicoladis, 2000) and distribute information across speech and gesture modalities. They convey information in gesture that is not conveyed in speech; for example, when engaging in spontaneous conversations (Goldin-Meadow & Butcher, 2003; Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005), when telling a story (Demir & So, 2006), or when solving a problem (Church & Goldin-Meadow, 1986; Goldin-Meadow, 2005; Goldin-Meadow, Alibali, & Church, 1993).

The question we address here is whether children use their gestures to clarify a referent that is ambiguous in speech but should be (on discourse grounds) specified. Specifically, we ask whether the way young children gesture is constrained by the discourse-pragmatic rules underlying the language they are learning. We look, in particular, at two discourse-pragmatic features that determine whether a referent needs to be specified: person (first or second vs. third) and information status (given vs. new; Clancy, 1993; Greenfield & Smith, 1976). A referent needs to be specified by overt arguments, like nouns, when it is a third person (as opposed to first or second person) or when it was not previously mentioned (i.e., new as opposed to given information). Imagine that a child has been talking to his mother about a puzzle and says, “I like this one.” It is perfectly clear which object the child likes in this context. Because the puzzle is given information, it does not need to be fully specified. Now consider a child who has not been talking about the puzzle and wants to tell his mother that he likes the puzzle. Because the puzzle is new information, it needs to be specified for the child to be fully informative. He could say, “I like this puzzle.” However, he could also use the less specified sentence along with a gesture; for example, “I like this,” said while pointing at the puzzle. If children’s gestures are discourse appropriate (in this case, sensitive to the new/given status of the referent), the child should be more likely to point at the puzzle in the second scenario than in the first.

Do children use gestures more often to indicate referents that have to be specified than referents that can be underspecified or omitted? To address this question, we videotaped Chinese- and English-speaking children spontaneously interacting with an adult, and examined their sensitivity to first or second versus third person and to given versus new information in their expression of referents in speech and gesture. English is a subject-prominent language (Bloom, 1990; Hyams & Wexler, 1993) and, as such, does not generally permit argument omission.<sup>1</sup> In contrast, Chinese is a null argument language that allows argument omission governed by discourse-pragmatic factors (Li & Thompson, 1979; Tsao, 1990). This grammatical difference between English and Chinese allows us to investigate whether children’s use of gesture is sensitive to discourse-pragmatic principles. The following discourse example demonstrates the grammatical difference between English and Chinese.

Speaker A:	$f_1$	chi1fan1	Le		mei2you3?
	$f_1$	eat lunch	ASP		not
		Have you eaten lunch?			
Speaker B:	$f_1$	he2	Li3 xian1sheng4		chi1fan1。
	$f_2$	with	Mr. Lee		eat lunch
		I have eaten lunch with Mr. Lee.			
Speaker B:	$f_3$	ji4de	bu4	ji4de	ta1 le?
	$f_3$	remember	Not	remember	him INT
		Do you remember him?			
Speaker A:	$f_4$	dang1ran2	ji4de	$f_5$	le。
	$f_4$	of course	remember	$f_5$	INT
		Of course I remember him.			

In Chinese, underspecified or omitted arguments are used to refer to entities that are retrievable from the discourse; nouns are used to refer to entities that are not retrievable (e.g., Huang, 1984, 1994; Lee & Naigles, 2005; Li & Thompson, 1979; Tsao, 1990). First and second person are always active in discourse and are thus retrievable and can be omitted (e.g.,  $\phi_1$ ,  $\phi_2$ ,  $\phi_3$ ,  $\phi_4$ ; Chafe, 1994, 1996; Dimitriadis, 1995). In contrast, third person is not always active in discourse; whether it is explicitly mentioned depends on whether it is new or given information. If the third-person referent was not previously mentioned (i.e., it is new information), it is not retrievable from discourse and should be expressed in an overt argument (e.g., Li3 xian1sheng4). However, if the third-person referent was previously mentioned (i.e., it is given information), it can be omitted (e.g.,  $\phi_5$ ) or presented in a less explicit form such as a pronoun (e.g., ta1<sup>2</sup>; Chafe, 1994, 1996; Dimitriadis, 1995).

Unlike Chinese, English is relatively strict in terms of representing arguments and generally does not allow argument omission (see Chomsky, 1981). As a result, explicit forms must be used to indicate even the given information in the preceding example (i.e., first and second person). However, as in Chinese, discourse factors *do* influence the choice of referential expression in English (e.g., Chafe, 1976, 1994, 1996; Huang, 1994; Levinson, 1987, 1991). For example, less explicit forms are typically used instead of more explicit forms for previously mentioned referents (*I* and *you* are used, rather than the speakers' names, in the preceding example). Thus, personal pronouns tend to be used in English in situations where omission is possible in Chinese.

In the experiment reported here, we expected to replicate previous findings in the literature and find that both Chinese- and English-speaking children would show sensitivity in their speech to the discourse-pragmatic principles *person* and *information status* that underlie their respective languages. The question of interest is whether children also show sensitivity to these two discourse features in the gestures that they produce along with speech.

## METHOD

### *Participants*

The participants were six English-speaking and six Mandarin Chinese-speaking children, living in Chicago, Illinois, and Nanjing, China, respectively, in

middle-class homes. The English-speaking children were an average of 4 years, 1 month (4;1) old (range = 3;7–5;2). The Chinese-speaking children were an average of 3;11 (range = 2;10–4;11).<sup>3</sup> None of the children had any major sensory or hearing problems, and none of the children or caregivers knew a conventional sign language. Families were recruited from postings and were paid for their participation.

### *Procedure*

The children participated in free-play activities and spontaneous conversations with their caregivers (mothers, fathers, or grandparents) and an experimenter. Both caregivers and experimenter were instructed to interact naturally with the children. A bag of toys, books, pictures, and puzzles was brought to each taping session to facilitate communication. The session lasted for approximately 45 min for each child (ranging from 30 min to an hour, depending on the attention span of the child) and was videotaped.

### *Speech coding*

All conversations between children and caregivers or experimenter were transcribed by research assistants who were native speakers of English or Mandarin Chinese. All transcripts were then checked by a second coder who was also a native speaker. Breaths, pauses, and speech dysfluencies such as self-interruptions, self-corrections, and repetitions were included in the transcriptions. The stream of speech was segmented into utterances. Utterances that contained syntactic questions, imitations, unintelligible sounds, songs, or poems were excluded from the analyses.

Our unit of speech analysis was the clause. A clause is a grammatical unit that expresses propositions,<sup>4</sup> and includes a predicate<sup>5</sup> (Crystal, 1980; Hartmann & Stork, 1972; Pei & Gaynor, 1954). We analyzed clauses containing predicates that described actions involving a subject *and* either a direct object (e.g., *I eat an apple*, *wo3 chi1 ping2guo3*) or indirect object (*I go to school*, *wo3 qu1 xue3xiao1*). Utterances containing more than one clause connected by a conjunction, for example, *and* (*hai3you2*), or *but* (*bu2guo4*), were separated into two clauses.

The following types of clauses were excluded from the database: (a) clauses that did not contain either a direct or indirect object, for example, *I go* (*wo3 qu4*); (b) clauses containing copula verbs, for example, *is* (*shi1*), because copulas are optional in Chinese; (c) ditransitive clauses containing both direct and indirect objects, for example, *I give a pen to you* (*wo3 ge2 ni3 bi3*) because our goal was to compare clauses of equal syntactic complexity; (d) clauses with serial verb structure (applicable only in Chinese), for example, *wo3 na2 zhe4ge4 jiao1 ge2 ni3* (*I hold this give to you*), again to insure equality in syntactic complexity; and (e) clauses containing grammatical omission of subjects (applicable only in English), including imperatives (e.g., *Open this door!*), *wanna* questions (e.g., *Wanna eat this?*); implied first person declaratives in past tense (e.g., *Got it!*), and progressive participles in response to questions (e.g., *brushing teeth*).

Thus, each clause contained a subject and an either a direct or indirect object that could *potentially* be expressed in a complete description of the action. We

identified the subject and object within each clause and assigned each to the following categories according to the form of the expression used to refer to it: null argument, personal pronoun, for example, *he*, *she*, *it* (*ta1*),<sup>6</sup> demonstrative pronoun, for example, *this*, *that* (*zhe4ge4*, *ne4ge4*), or noun, for example, *cat*, *dog* (*mao1*, *gou2*).<sup>7</sup> Subjects and objects were also classified according to whether the referent was a first or second versus third person. Referents were further classified according to information status: first and second persons were assumed to be given information (Chafe, 1994, 1996); a third-person referent was considered *given* if it was mentioned somewhere in the preceding 20 utterances and *new* if it had not been mentioned (Chafe, 1987; Du Bois, 1987).

### *Gesture coding*

We analyzed the gestures that co-occurred with the relevant clauses and determined whether the subject and object within each clause were identified in gesture. We followed Goldin-Meadow and Mylander (1984; see also Iverson & Goldin-Meadow, 2005; and Özçalışkan & Goldin-Meadow, 2005) in excluding hand movements that involved direct manipulation of an object (e.g., placing a toy on a floor) or were part of a ritualized game (e.g., putting a puzzle in a puzzle slot) from the database. There were three types of gestures: (a) *object-referring iconic gestures*<sup>8</sup> that bear a resemblance to the referents they represent (e.g., two hands flapped at shoulders, classified as a reference to *bird*); (b) *pointing gestures* that refer to objects, people, or places by singling out the referent (e.g., index finger point to a bottle, classified as a reference to *bottle*); and (c) *hold-up gestures* that refer to objects by raising them in the air (e.g., hold-up bottle, classified as a reference to *bottle*). The purpose of a hold-up gesture is not to manipulate the object, but to draw the interlocutor's attention to the object (Gullberg, de Bot, & Volterra, 2008).

Each gesture was then assigned a semantic meaning. The semantic meaning of a gesture was determined by its form in conjunction with the speech in the clause with which it occurred. For example, two hands flapping at shoulders produced in conjunction with the clause, "The bird eats a worm," was assumed to refer to the bird. If the gesture was not accompanied by a clause containing a word that expressed its referent, context or form was used to determine the gesture's meaning. The meaning of a point or a hold-up gesture depended on the context of interpretation; for example, a point at a puzzle was assumed to refer to the puzzle. The meaning of an iconic gesture depended on its form; for example, a curved palm moving toward the mouth was assumed to refer either to a glass or to the action of drinking.<sup>9</sup>

The proportion of referents conveyed in gesture was calculated as the total number of referents conveyed in gesture, divided by the total number of referents conveyed in speech and/or gesture. All proportions were subjected to an arcsine transformation before statistical analysis.

### *Reliability*

A subset of each transcript (20%) in English and Chinese was independently coded by a second research assistant, who was a bilingual speaker in English

and Chinese and was trained to code speech and gesture. Reliability was 98% for the English-speaking children ( $N = 120$ ) and 97% for the Chinese-speaking children ( $N = 140$ ) for identifying target clauses; 100% for the English-speaking children ( $N = 236$ ) and 100% for the Chinese-speaking children ( $N = 262$ ) for classifying references to subjects and objects according to speech form (noun, pronoun, etc.); 100% for the English-speaking children ( $N = 236$ ) and 100% for the Chinese-speaking children ( $N = 262$ ) for determining first-, second-, and third-person status of the referents; 90% for the English-speaking children ( $N = 236$ ) and 93% for the Chinese-speaking children ( $N = 262$ ) for determining information status of the referents; 85% for the English-speaking children ( $N = 236$ ) and 84% for the Chinese-speaking children ( $N = 262$ ) for identifying gestures; 95% for the English-speaking children ( $N = 201$ ); 90% for the Chinese-speaking children ( $N = 220$ ) for determining types of gestures; and 92% for the English-speaking children ( $N = 201$ ) and 88% for the Chinese-speaking children ( $N = 220$ ) for identifying the semantic meaning of gestures.

## RESULTS

We analyzed all clauses containing predicates that described actions involving both subjects *and* either direct or indirect objects.<sup>10</sup> There were no significant differences in the number of clauses produced by the English-speaking children ( $M = 80.67$ ,  $SD = 41.01$ ) and the Chinese-speaking children ( $M = 120$ ,  $SD = 43.16$ ),  $t(10) = 1.62$ , *ns*. However, the Chinese-speaking children produced gestures for a greater proportion of the referents they indicated ( $M = 0.18$ ,  $SD = 0.05$ , range = 0.13–0.24) than the English-speaking children ( $M = 0.08$ ,  $SD = 0.04$ , range = 0.04–0.15),  $t(10) = 3.26$ ,  $p < .009$ .

The goal of this study was to explore the role of two discourse-pragmatic features, person and information status, in referential expressions in speech and gesture. We look first at how discourse-pragmatic features influenced lexical choices in speech and we then turn to gesture.

### *Speech*

We first examined how subjects and objects were expressed in speech. Figure 1 shows the distribution of lexical choices in speech in the English- and Chinese-speaking children. We conducted a repeated measures analysis of variance (ANOVA) with proportion of arguments expressed as the dependent variable, and with type of referential expression (null, personal pronoun, demonstrative pronoun, noun) as a within-subject independent variable, and language (English, Chinese) as a between-subject independent variable. We found a significant effect of referential expression,  $F(3, 30) = 12.18$ ,  $p < .001$ , a marginal effect of language,  $F(1, 10) = 3.62$ ,  $p < .08$ , and a significant interaction,  $F(3, 30) = 9.82$ ,  $p < .001$ . As expected, the Chinese-speaking children produced null arguments (i.e., they omitted arguments) more often than the English-speaking children,  $t(10) = 8.94$ ,  $p < .001$ . The English-speaking children produced personal pronouns more often than the Chinese-speaking children,  $t(10) = 4.46$ ,  $p < .001$ . In fact, the proportion of null arguments that the Chinese-speaking children produced was similar to the

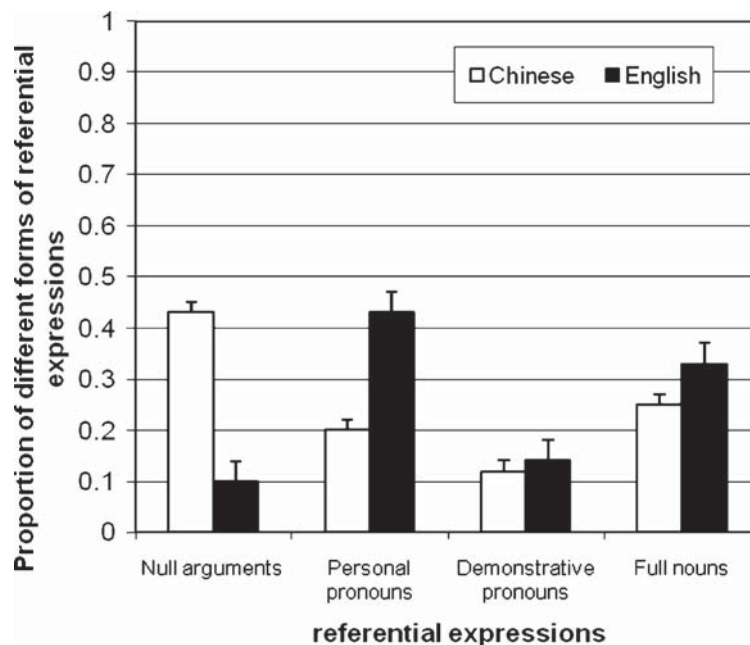


Figure 1. The distribution of different forms of referential expressions in speech produced by Chinese-speaking and English-speaking children.

proportion of personal pronouns that the English-speaking children produced. No significant differences were found between the groups in demonstrative pronouns,  $t(10) = .64, ns$ , or full nouns,  $t(10) = 1.58, ns$ .

Thus, children in the two language groups did not differ in how often they produced explicitly specified referents: they used nouns equally often. They differed only in the type of less specified expressions they used: English-speaking children used personal pronouns as their preferred form, and Chinese-speaking children used null arguments. The two groups of children had learned to use the less specified term appropriate to the language each was acquiring. Because null arguments, personal pronouns, and demonstrative pronouns are all less explicit than nouns, we grouped them together into a *non-noun* category in the following analyses.

We next ask whether person and information status affect the explicitness of the children's referential expressions. We classified referents into three types: (a) first or second person (which were assumed to be given), (b) third-person given, and (c) third-person new. We found that 0.37 ( $SD = 0.07$ ) of the referents that the Chinese-speaking children produced were first or second person, 0.31 ( $SD = 0.09$ ) were third-person given, and 0.32 ( $SD = 0.09$ ) were third-person new; comparable numbers for the English-speaking children were 0.36 ( $SD = 0.07$ ), 0.31 ( $SD = 0.03$ ), and 0.33 ( $SD = 0.06$ ).

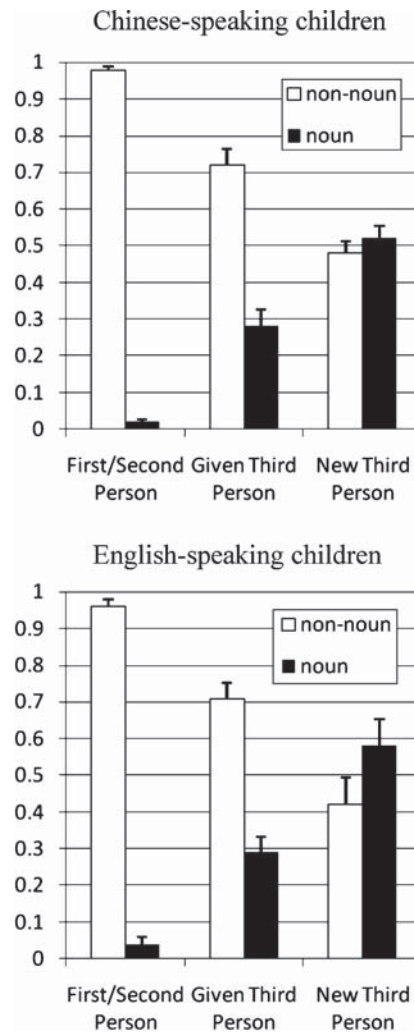


Figure 2. The distribution of nouns and nonnouns for first or second person, third-person given, and third-person new referents produced by Chinese-speaking children and English-speaking children.

We expected that the children would be sensitive to discourse-pragmatic factors. We hypothesized that, of the three categories, third-person referents that were new to the context would be the least known to a listener and thus should be explicitly specified more often than third-person referents that were given, followed by first- or second-person referents. Figure 2 presents the proportion of nouns and nonnouns that the children used in each referential category.



We conducted a repeated measures ANOVA with the proportion of nonnouns that the children produced as the dependent variable and referential category (first or second person, third-person given, and third-person new) as a within-subject independent variable, and language (Chinese, English) as a between-subject independent variable. We found a significant effect of referential category,  $F(2, 20) = 121.50$ ,  $p < .0001$ , no effect of language,  $F(1, 10) = .54$ , *ns*, and no interaction,  $F(2, 20) = .10$ , *ns*. Bonferroni-adjusted pairwise comparisons showed that children in both language groups produced nonnouns more often when referring to first or second person than when referring to third-person given ( $p < .0001$ ) and third-person new ( $p < .0001$ ). They also produced nonnouns more often when referring to third-person given than when referring to third-person new ( $p = .001$ ).

Thus, children in both language groups tended to use less specified forms (pronouns, null arguments) for referents that did not need to be specified (first or second person, third-person given). Of importance, null arguments were used in the same way in Chinese-speaking children as personal pronouns in English-speaking children: 61% ( $SD = 0.10$ ) of the null arguments that the Chinese-speaking children used referred to first or second person and third-person given referents, referents that did not need to be specified. Similarly, 67% ( $SD = 0.15$ ) of the personal pronouns that the English-speaking children produced referred to first or second person and third-person given referents.

Children in both languages tended to use nouns when they needed to; that is, when discourse required that the referents be specified (third-person new referents). Still, 40% of the third-person new referents that they produced were conveyed by nonnouns and thus were underspecified. Our next question was whether the children used gesture to help disambiguate these underspecified forms.

### *Gesture*

Children in both groups produced all three types of gestures and in roughly the same proportions: pointing ( $M = 0.65$ ,  $SD = 0.11$ ), hold-up ( $M = 0.30$ ,  $SD = 0.12$ ), iconic ( $M = 0.05$ ,  $SD = 0.06$ ) in the Chinese-speaking children; and pointing ( $M = 0.53$ ,  $SD = 0.29$ ), hold-up ( $M = 0.37$ ,  $SD = 0.32$ ), iconic ( $M = 0.10$ ,  $SD = 0.12$ ) in the English-speaking children.

Figure 3 displays the proportion of expressions indicating first or second person, third-person given, third-person new referents that were accompanied by gesture in the two groups of children. We conducted a repeated measures ANOVA with proportion of expressions accompanied by gesture as the dependent variable, and referential category (first or second person, third-person given, third-person new) as a within-subject independent variable, and language (English, Chinese) as a between-subject independent variable. We found a significant effect of referential category,  $F(2, 20) = 34.68$ ,  $p < .0001$ , a significant effect of language,  $F(1, 10) = 60.50$ ,  $p < .0001$ , and no interaction,  $F(2, 20) = 2.79$ , *ns*. Overall, Chinese-speaking children produced gestures more often than English-speaking children ( $p < .0001$ ; perhaps because Chinese caregivers produce more gestures when interacting with their children than American caregivers; Goldin-Meadow & Saltzman, 2000). However, both groups of children produced gestures more often

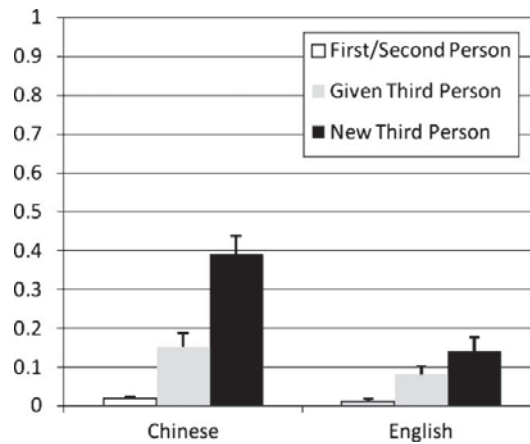


Figure 3. The proportion of first or second person, third-person given, and third-person new referents accompanied by gesture in (left) Chinese-speaking children and (right) English-speaking children.

when indicating third-person new referents than when indicating third-person given referents ( $p = .024$ ) and first- or second-person referents ( $p < .0001$ ). They also produced gestures more often when indicating third-person given referents than when referring to first- or second-person referents ( $p = .002$ ).<sup>11</sup>

We are now able to address our final question: are gesture and speech working as an integrated system to specify referents? We focused on referents that need to be specified (i.e., new third persons), which were frequently accompanied by gesture (see Figure 3). We asked whether nonnouns were accompanied by gesture more often than nouns, the pattern we would expect if the children were using their gestures to adjust for the fact that nonnouns are underspecified relative to nouns.

Figure 4 presents the proportion of third-person new referents conveyed by nouns or nonnouns (null arguments, pronouns) that were accompanied by gestures. We conducted a repeated ANOVA with proportion of referential expressions accompanied by gesture as the dependent variable, and type of referential expression (nonnoun, noun) as a within-subject independent variable, and language (Chinese, English) as a between-subject independent variable. We found a significant effect of type of referential expression,  $F(1, 10) = 20.28, p = .001$ , a significant effect of language,  $F(1, 10) = 23.33, p = .001$ , and no interaction,  $F(1, 10) = 1.22, p = ns$ . Given the earlier analyses, it was not surprising that Chinese-speaking children produced gestures proportionally more often than English-speaking children. More importantly, when indicating third-person new referents, both groups of children produced gestures more often with nonnouns than with nouns, as we would expect if gesture is being used to further specify underspecified referents.

To summarize, both Chinese- and English-speaking children produced gestures more often to indicate referents that should be specified in discourse, particularly when those referents were conveyed by potentially ambiguous words.

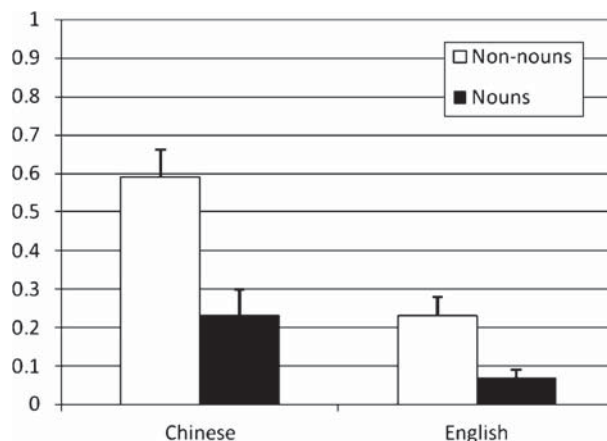


Figure 4. The proportion of third-person new referents conveyed by nouns (black bars) and nonnouns (null arguments, pronouns, white bars) that were accompanied by gesture in (left) Chinese-speaking children and (right) English-speaking children.

## DISCUSSION

Our study explored whether English- and Chinese-speaking children display sensitivity to two discourse-pragmatic features (person and information status) in speech and in gesture. In terms of speech, we found that, despite the fact that Chinese is a null argument language and English is a subject-prominent language, Chinese-speaking and English-speaking children produced fully specified referring expressions (i.e., nouns) required by the discourse equally often. However, the children displayed sensitivity to the discourse requirements of their respective languages in the underspecified expressions they produced: Chinese-speaking children omitted arguments for their underspecified forms, whereas English-speaking children produced pronouns. Of importance, both groups of children produced their underspecified forms in appropriate discourse contexts, more often for referents that did not need to be fully specified (first or second person and third-person given referents) than for referents that did need to be specified (third-person new referents). Thus, irrespective of the language they were learning, children produced fully specified forms (nouns) more often when expressing referents that needed to be specified than when expressing referents that did not need to be specified (third-person new referents > third-person given referents > first- or second-person referents).

Previous research has found that children learning null argument languages are particularly sensitive to discourse-pragmatic features (e.g., Allen, 2000; Allen & Schroder, 2003; Serratrice, 2005). Italian- and Inuktitut-speaking children tend to use overt arguments to indicate third person and new referents but null arguments to indicate first or second person and given referents. Similar findings have been reported in other languages (Korean: Clancy, 1993; Hindi: Narasimhan, Budwig, & Murty, 2005; Romance: Paradis & Navarro, 2003). Our study builds on this

research by extending the phenomenon to Chinese, another null argument language. Of interest, children who are learning a subject-prominent language, like English, appear to be as sensitive to discourse-pragmatic principles as children learning Chinese, a null argument language, suggesting that sensitivity to discourse is universal across language learners.

It was important that the children also displayed sensitivity to discourse in their gestures, producing precisely the same pattern as they displayed in speech. They produced gestures for referents that needed to be specified more often than they produced gestures for referents that did not need to be specified (third-person new referents > third-person given referents > first- or second-person referents). Thus, the children were paying attention to discourse-pragmatic information when deciding when to use gesture.

Moreover, when children used underspecified forms in their speech to refer to referents that needed to be specified (i.e., when they used null arguments and pronouns, as opposed to nouns, to refer to third-person new referents), the children produced gestures along with these nonnouns. In other words, gesture stepped in to clarify potentially ambiguous speech and did so equally often in children learning English and Chinese (Allen, 2008; but see Guerriero, Oshima-Takane & Kuriyama, 2006, for a comparison of children learning English and Japanese).

How do young children develop sensitivity to discourse-pragmatic principles? Previous work has found that caregivers use discourse-pragmatic strategies when talking to their children (Clancy, 1993; Paradis & Navarro, 2003). However, few studies have investigated the relation between parental input and the development of children's sensitivity to discourse-pragmatic features. One exception is a longitudinal study by Guerriero et al. (2006). They followed English- and Japanese-speaking children for more than a year, observing conversations between the children and their parents. The English-speaking parents showed consistent language-specific discourse patterns in their referential expressions in speech, but the Japanese-speaking parents did not. In turn, the English-speaking children developed discourse-pragmatic strategies earlier than the Japanese-speaking children. These findings suggest that children may learn about discourse-pragmatic features in their language from their parents' speech.

However, unlike the children in our study, adults do not appear to routinely use gesture to clarify potentially ambiguous referring expressions. So, Kita, and Goldin-Meadow (2009) showed English-speaking adults vignettes of two stories and asked them to retell the stories to an experimenter. Because none of the protagonists or objects in the story was present, the adults could not use points at real-world objects to indicate referents in the story. However, they did use points at space to indicate particular referents. A gesture was considered to identify a referent if it was produced in the same location as the previous gesture for that referent. The adults frequently used gesture location to identify referents. However, they used gesture to identify referents that were already specified in speech, and not to clarify referents that were ambiguous in speech (even though they produced a number of expressions that did not fully specify the referent). In other words, the adults did not use gesture to disambiguate speech, as the children in our study did.

However, there are many differences between the So et al. (2009) study and ours. First, the participants in So et al. were telling stories; our participants were engaged in spontaneous conversation. Second, the participants in So et al. were pointing at empty spaces, which were used to stand for particular referents; our participants were pointing at real objects and people in the room. Third, the participants in So et al. were adults; our participants were children. The differences between adults and children may stem from the different types of discourse examined in the two studies (displaced story telling vs. here and now conversation). Thus, it is possible that adults do use gesture to disambiguate their underspecified speech when they engage in spontaneous conversations where points can be directed at real-world objects.

Alternatively, using gesture to disambiguate underspecified speech may be a characteristic of early childhood, one that disappears as children become more proficient speakers. According to this view, children use gesture differently from adults simply because they have not yet fully mastered lexical specification in speech. This phenomenon would then be another instance of gesture preceding, and perhaps propelling, advances in speech (Iverson & Goldin-Meadow, 2005) and in other cognitive tasks (Goldin-Meadow, 2003; Goldin-Meadow et al., 1993). Comparable data from adults speaking null argument and subject-prominent languages engaged in a here and now conversation is needed to fully test this view.

Our findings also have clinical implications. Children with language impairments often have difficulty producing sentences with complex argument structure (Grela, 2003). In line with our findings, these children might be able to use gesture to specify referents that they are not able to specify in speech. Children whose language development is impaired for a variety of reasons (focal brain injury; Sauer, Levine, & Goldin-Meadow, 2009; Down syndrome: Caselli et al., 1998; Stefanini, Caselli, & Volterra, 2007; Stefanini, Recchia, & Caselli, 2008; specific language impairment: Evans, Alibali, & McNeill, 2001; Fex & Mansson, 1998) have been shown to use gestures to compensate for their communicative deficiencies.

To summarize, we found that children in the early stages of language learning use speech and gesture to identify referents and do so in accordance with discourse-pragmatic principles. They use nouns and gestures more often when indicating referents that need to be specified than when indicating referents that can be inferred from context. Moreover, when speech is less specific than it needs to be, gesture is often used to fill the breach, whether the child is learning English or Chinese.

#### ACKNOWLEDGMENTS

This research was supported by Provost Research Funding Grant R-581-000-074-133 (to W.C.S.) at the National University of Singapore and NIH Grant RO1 00491 (to S.G.-M.) at the University of Chicago. We thank Wenping Xue, Stacy Steine, Zachary Mitchell, and Carolyn Mylander for help in collecting data in Nanjing and Chicago; and Lim Jia Yi, Tee Can Shou Joseph, Lee Yingqi, Tan Wenlin, Elizabeth Sarah Ragen, Chew Xin Ying Ivane, and Kirrthana Krishnamoorthy for help in coding the data. The coding system for parts of speech and gesture was established in the first author's dissertation.

## NOTES

1. Under some circumstances, argument omission is permitted in English; see the examples in the Method Section.
2. In this example, ta1 is more appropriate than a null argument to identify Mr. Lee. A null argument might be mistakenly understood as the event of Speaker B having lunch with Mr. Lee. Note, however, we did not aim to study the differences between pronouns and null arguments in Chinese in the present study. Both pronouns and null arguments were considered less explicit forms of referential expressions.
3. The English-speaking children were somewhat older than the Chinese-speaking children. To be certain that age was not responsible for any differences found between the groups, we redid all of the analyses on three Chinese- and three English-speakers matched for age. We found that the patterns in this matched sample were identical to those reported below.
4. A proposition is the meaning content of units within the clause.
5. A predicate is the portion of a clause, excluding the subject, that expresses something about the subject.
6. The pronunciation of the pronouns referring to animate and inanimate entities is the same, that is, ta1.
7. Any combinations of demonstrative pronoun and nouns were assigned to the noun category.
8. Iconic gestures are also known as characterizing (Goldin-Meadow & Mylander, 1984) or representational (Gullberg, de Bot, & Volterra, 2008) gestures.
9. The children produced very few iconic gestures overall: 5% of the English-speaking children's gestures and 10% of the Chinese-speaking children's gestures were iconic.
10. Following the criteria described in the speech coding section, we excluded 461 clauses in the English-speaking children and 684 clauses in the Chinese-speaking children.
11. As in our previous analyses, Chinese-speaking children used null arguments in the same way as English-speaking children used personal pronouns. The Chinese-speaking children produced more gestures with their null arguments for referents that needed to be specified (third-person new,  $M = 0.45$ ,  $SD = 0.17$ ) than for referents that did not need to be specified (third-person given,  $M = 0.08$ ,  $SD = 0.07$ ; first or second person,  $M = 0.01$ ,  $SD = 0.03$ ). English-speaking children showed precisely the same pattern for personal pronouns: third-person new ( $M = 0.26$ ,  $SD = 0.27$ ) versus third-person given ( $M = 0.04$ ,  $SD = 0.07$ ), and first or second person ( $M = 0.01$ ,  $SD = 0.01$ ).

## REFERENCES

- Allen, S., & Schroder, H. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In J. W. Du Bois, L. Kumpf, & W. Ashby (Eds.), *Preferred argument structure: Grammar and architecture for function*. Amsterdam: John Benjamins.
- Allen, S. E. M. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, 38, 483–521.
- Allen, S. E. M. (2007). Interacting pragmatic influences on children's argument realization. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 191–210). Mahwah, NJ: Erlbaum.
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 24, 721–734.

- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–257). Cambridge: Cambridge University Press.
- Caselli, M. C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. *Journal of Speech, Language, and Hearing Research, 41*, 1125–1135.
- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic Press.
- Chafe, W. L. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 21–53). Amsterdam: John Benjamins.
- Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chafe, W. L. (1996). Inferring identifiability and assessability. In T. Fretheim & J. Gundel (Eds.), *Reference and referent accessibility* (pp. 37–46). Amsterdam: John Benjamins.
- Clancy, P. M. (1993). Preferred argument structure in Korean acquisition. In E. V. Clark (Ed.), *Proceedings of the 25th Child Language Research Forum*. Stanford, CA: CSLI Publications.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43–71.
- Crystal, D. (1980). *A first dictionary of linguistics and phonetics*. Boulder, CO: Westview.
- Demir, E., & So, W. C. (2006). What's hidden in the hands? How children use gesture to convey arguments in a motion event. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 31st Boston University Conference on Language Development* (pp. 172–183). Somerville, MA: Cascadilla Press.
- Dimitriadis, A. (1995). When pro-drop languages don't: On overt pronominal subjects in Greek. *Penn Working Papers in Linguistics, 2*, 45–60.
- Du Bois, J. W. (1987). The discourse basis of ergativity. *Language, 63*, 805–855.
- Evans, J., Alibali, M., & McNeill, N. M. (2001). Divergence of verbal expression and embodied knowledge: Evidence from speech and gesture in children with specific language impairment. *Language and Cognitive Processes, 16*, 309–331.
- Fex, B., & Mansson, A.-C. (1998). The use of gestures as a compensatory strategy in adults with acquired aphasia compared to children with specific language impairment (SLI). *Journal of Neurolinguistics, 11*, 191–206.
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S. (2005). The two faces of gesture: Language and thought. *Gesture, 5*, 239–255.
- Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*, 279–297.
- Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85–107). Mahwah, NJ: Erlbaum.
- Goldin-Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: The effects and non-effects of parental input on early language development. *Monographs of the Society for Research in Child Development, 49*(No. 207).
- Goldin-Meadow, S., & Saltzman, J. (2000). The cultural bounds of maternal accommodation: How Chinese and American mothers communicate with deaf and hearing children. *Psychological Science, 11*, 311–318.
- Greenfield, P. M., & Smith, J. H. (1976). *The structure of communication in early language development*. New York: Academic Press.
- Grela, B. G. (2003). Production based theories may account for subject omission in normal children and children with SLI. *Journal of Speech-Language Pathology and Audiology, 27*, 221–228.
- Guerriero, A. M. S., Oshima-Takane, Y., & Kuriyama, Y. (2006). The development of referential choice in English and Japanese: A discourse-pragmatic perspective. *Journal of Child Language, 33*, 823–857.
- Gullberg, M., de Bot, K., & Volterra, V. (2008). Gestures and some key issues in the study of language development. *Gesture, 8*, 149–179.

- Hartmann, R. R. K., & Stork, F. C. (1972). *Dictionary of language and linguistics*. London: Applied Science.
- Huang, C. T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531–574.
- Huang, C. T. J. (1994). On null subjects and null objects in generative grammar. *Linguistics*, 33, 1081–1123.
- Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 24, 421–459.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 367–371.
- Lee, J., & Naigles, L. R. (2005). Input to verb learning in Mandarin Chinese: A role for syntactic bootstrapping. *Developmental Psychology*, 41, 529–540.
- Levinson, S. (1991). Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics*, 27, 107–161.
- Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics*, 23, 379–434.
- Li, C. N., & Thompson, S. (1979). Third-person pronouns and zero-anaphora in Chinese discourse. In T. Givon (Ed.), *Syntax and semantics: Discourse and syntax* (Vol. 12, pp. 311–335). New York: Academic Press.
- 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 gesture reveals about thought*. Chicago: University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- Narasimhan, B., Budwig, N., & Murty, L. (2005). Argument realization in Hindi caregiver–child discourse. *Journal of Pragmatics*, 37, 461–495.
- Özçalışkan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, 96, B101–B113.
- Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of input? *Journal of Child Language*, 30, 371–393.
- Pei, M. A., & Gaynor, F. (1954). *A dictionary of linguistics*. New York: Philosophical Library.
- Sauer, E., Levine, S. C., & Goldin-Meadow, S. (2009). *Early gesture predicts language delay in children with pre- and perinatal brain lesions*. Manuscript submitted for publication.
- Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics*, 26, 437–462.
- So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, 33, 115–125.
- Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language*, 101, 208–221.
- Stefanini, S., Recchia, M., & Caselli, M. C. (2008). The relationship between spontaneous gesture production and spoken lexical ability in children with Down syndrome in a naming task. *Gesture*, 8, 197–218.
- Tsao, F. (1990). *Sentence and clause structure in Chinese: A function perspective*. Taipei: Student Book.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40, 21–81.
- Volterra, V., Caselli, M. C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature–nurture: Essays in honor of Elizabeth Bates* (pp. 3–40). Mahwah, NJ: Erlbaum.