

When Do Speakers Use Gestures to Specify Who Does What to Whom? The Role of Language Proficiency and Type of Gestures in Narratives

Wing Chee So · Sotaro Kita · Susan Goldin-Meadow

Published online: 22 January 2013
© Springer Science+Business Media New York 2013

Abstract Previous research has found that iconic gestures (i.e., gestures that depict the actions, motions or shapes of entities) identify referents that are also lexically specified in the co-occurring speech produced by proficient speakers. This study examines whether concrete deictic gestures (i.e., gestures that point to physical entities) bear a different kind of relation to speech, and whether this relation is influenced by the language proficiency of the speakers. Two groups of speakers who had different levels of English proficiency were asked to retell a story in English. Their speech and gestures were transcribed and coded. Our findings showed that proficient speakers produced concrete deictic gestures for referents that were not specified in speech, and iconic gestures for referents that were specified in speech, suggesting that these two types of gestures bear different kinds of semantic relations with speech. In contrast, less proficient speakers produced concrete deictic gestures and iconic gestures whether or not referents were lexically specified in speech. Thus, both type of gesture and proficiency of speaker need to be considered when accounting for how gesture and speech are used in a narrative context.

Keywords Gesture · Concrete deictic gesture · Iconic gesture · Discourse · Referential identification · Speech-gesture coordination

W. C. So (✉)

Department of Educational Psychology, Faculty of Education, The Chinese University of Hong Kong,
Room 530, Sino Building, Shatin, New Territories, Hong Kong
e-mail: wingchee@cuhk.edu.hk

S. Kita

School of Psychology, University of Birmingham, Birmingham, UK

S. Goldin-Meadow

Department of Psychology, University of Chicago, Chicago, IL, USA

Introduction

To produce a comprehensible and cohesive narrative, speakers have to refer to the characters in the story in the ways that allow the listener to unambiguously identify who is doing what to whom (Lyons 1977; Bosch 1983; Garrod 2001). Speakers can use nouns, pronouns, or zero anaphora to make these references (e.g., Fox 1987; Givon 1984). The ways these devices are used vary as a function of discourse context (e.g., Arnold 2000; Francik 1985). For example, speakers tend to use pronouns (less attenuated forms of referring expressions) to specify given referents but nouns (specific forms of referring expressions) to specify new referents (e.g., Ariel 1990; Chafe 1994; Givon 1985; Lambrecht 1994).

In addition to using nouns and pronouns to keep track of the characters in a narrative, speakers can use gestures (Gullberg 1998, 2003, 2006; So et al. 2005; Yoshioka 2008), which are tightly integrated with the speech they accompany (McNeill 1992, 2005). Recently, So and colleagues (So et al. 2009) examined the ways speakers gesture in relation to speech. Specifically, they explored the circumstances under which native English-speakers use iconic gestures (gestures that depict the actions, motions, or shapes of entities) and abstract deictic gestures (gestures that point to the abstract locations associated with entities) to identify referents. They found that speakers were more likely to use iconic or abstract deictic gestures to identify referents when those referents were also specified in speech than when they were not specified. In other words, when speech fails to lexically specify referents, iconic or abstract deictic gestures also tend not to identify those referents. Lexical specificity in speech thus appears to go hand-in-hand with referential identification in iconic or abstract deictic gestures.

However, what remains to be explored is whether the parallel relationship between gesture and speech applies not only to iconic and abstract deictic gestures but also to concrete deictic gestures. The first aim of the present study is to address this question. Like iconic and abstract deictic gestures, concrete deictic gestures identify referents and do so very directly (e.g., index finger points to an entity). Previous work has shown that concrete deictic gestures often identify referents that are under-specified in the accompanying speech, thus supplementing co-occurring speech. For example, in a study by Bangerter (2004), proficient English-speakers identified photos of faces to a listener who could also see the photos. When the speakers *reduced* the specificity in their speech, they used concrete deictic gestures to point at the target photos. Conversely, when they elaborated their speech, they tended not to produce concrete deictic gestures. As another example, So and Lim (2012) found that, when talking to young children, caregivers who were proficient in Mandarin pointed to the objects more often when they were not specifying those objects in speech than when they were specifying them in speech. Based on these findings, we expected that, unlike iconic and abstract deictic gestures whose content goes hand-in-hand with the content of speech, concrete deictic gestures would trade off with speech and would be produced more often when the co-occurring speech failed to specify a referent than when it specified a referent.

However, the way speakers produce concrete deictic gestures in relation to speech might be influenced by their language proficiency. The second aim of the present study is to investigate the circumstances under which less proficient speakers and proficient speakers produce concrete deictic gestures in relation to speech. Related to this line of research, previous studies on reference tracking have shown that early-staged second language learners speak and gesture differently in their stronger language (i.e., first language, L1) and in their weaker language (i.e., second language, L2). In a series of studies by Gullberg (1998, 2003, 2006), early-staged second language learners were found to use specific referential expressions (i.e., nouns) when referring to the newly introduced characters *and* previously mentioned

characters in L2, resulting in over-explicit reference. For example, a second language learner might say in L2, “*The gentleman gave the lady a basket. Then the gentleman kissed the lady.*” In this sentence, the speaker uses nouns rather than pronouns (i.e., she does not say, *Then he kissed her*) when referring back to the male and female. Not only did these speakers over-specify referents in speech when using L2, but they also marked those references in their iconic and abstract deictic gestures whereby each referent was associated with different spatial locations (Gullberg 2003, 2006). In the example just given, the second language learner would point to her left when referring to the male character and to her right when referring to the female character, locations previously established as referring to the man vs. the woman. Interestingly, however, the early-stage second language learners spoke and gestured differently in their L1. They used pronouns instead of nouns to identify referents in their L1, and they did not over-mark the referents with gestures (see also Yoshioka 2008).

We know from these findings that language proficiency can influence the way speakers gesture in relation to speech. However, Gullberg and Yoshioka examined only iconic and abstract deictic gestures. The present study will investigate whether proficient and less proficient speakers differ in how they use concrete deictic gestures in relation to speech.

The present study thus aims to investigate (1) whether concrete deictic gestures bear a different semantic relation to speech than iconic and abstract deictic gestures, and (2) whether this gesture-speech relationship is influenced by language proficiency. To address these two questions, we adopted So et al.’s (2009) paradigm and observed two groups of bilingual for whom English was a second language—one group that was proficient in English, and one that was less proficient in English.

Method

Participants

Fifty English–Mandarin bilingual undergraduate students (22 males), naïve to sign language,¹ were recruited through postings throughout the campus. The participants were 18–23 years old and were born and grew up in Singapore. All were Singaporean Chinese² and started learning English after age three in school. They used English and Mandarin both at home and in school. All participants received research credit for their participation.

Singapore is a multicultural and multilingual country. It has four official languages (English, Mandarin-Chinese, Tamil, and Malay). A bilingual policy was adopted in 1966 and is considered to be fundamental to Singapore’s well-being (Pakir 1998). According to this policy, every child should learn English and one of the three mother tongues.³ English is recognized as the “working language” in education and work; the other three official languages are the “mother tongues” of the major ethnic groups (Dixon 2005; Pakir 1999).

Since 1987, English has been the sole medium of instruction for all content subjects in school. The mother tongues are taught as mandatory classroom subjects but are not themselves used to teach content subjects (Shepherd 2005). British Standard English is considered

¹ Conventional sign languages such as American Sign Language use space grammatically as part of their pronominal systems (Padden 1988). We therefore included in our study only those participants who had no knowledge of sign language to avoid the possibility that the way they used space in their co-speech gestures might have been influenced by their knowledge of sign.

² According to the Singapore Department of Statistics, the population of Singapore in 2000 contained 77 % Chinese, 14 % Malays, 8 % Indians and the rest other races.

³ Ministry of Education Parliamentary Speech by Tony Tan Keng Yam., March 1986, Singapore.

more prestigious than Singapore Colloquial English (SCE), which evolved in school children of different language backgrounds in English schools in the early part of the twentieth century (Shepherd 2005). The bilingual policy has greatly improved the literacy rates of Singaporeans in English (Pakir 1998). However, among the Singaporean Chinese, there has been a decline in Mandarin-Chinese⁴ because many of these families speak English at home.

Procedure

We began the experiment by administering a self-report language proficiency questionnaire developed by Lim et al. (2008) for use in Singapore. This tool measures Asian bilinguals' language proficiency in four different modalities (speaking, listening, reading, and writing) and the frequency of language use in three domains (home, school, and other social arenas). In this study, each participant was required to answer seven questions that included: (1) an evaluation of his/her English speaking and understanding proficiency on a 5-point scale (e.g., "How proficient are you in speaking English?" 5 indicates natively proficient; 1 indicates very few words); (2) an estimate of how often he/she used English at home, in school, and in other social settings on a 5-point scale (e.g., "How often do you use English at home?" 5 indicates always; 1 indicates never); and (3) an estimate of the number of hours he/she spoke and listened to English on a particular day (e.g., "How often do you speak in English on a given day?" 5 refers to 10 h or above; 4 refers to 8 to less than 10 h; 3 refers to 5 to less than 8 h; 2 refers to 2 to less than 5 h, 1 refers to less than 2 h).

Participants were then tested individually. They were shown a video of a story involving two men in variety of motion events (e.g., the first man drops a rock on the foot of the second man; the second man removes a noose from the neck of the first man; the first man throws the second man into the water). The video lasted for approximately 30 s and was presented twice. After participants had watched the video, a naïve listener entered the room and was seated face-to-face with the participant across a table. The participant was then asked to describe the story as comprehensively as possible in English to the listener. Participants were told that the listener had not watched the video before and that he/she was a proficient English speaker. The participant was given pictures of the protagonists in the story, which they were free to refer to during their retelling; the pictures were placed on the table and visible to both speaker and listener. The entire session was videotaped.

The procedures used in this study differed from those in So et al. (2009) in three respects. First, in So et al. (2009), participants were shown segments of the cartoon and, after each segment, were asked to describe the segment. In our study, participants saw the entire videotape before retelling the story. We made this change so that participants had an opportunity to produce coherent and cohesive stories. Second, in So et al. (2009), the listener was the experimenter who had watched the videotape along with the participant. In our study, the listeners had not watched the stories. We made this change so that participants had to be specific in their references to the characters in the story. Third, in So et al. (2009), participants were not provided with the pictures of the protagonists. In our study, participants were given the pictures to refer to as needed during their retellings. We made this change in order to provide participants with the opportunity to produce concrete deictic gestures.

Language Dominance Assessment

A bilingual's language proficiency is often related to the amount of the time he/she is exposed to or engaged in a conversation in a particular language (Mayberry and Nicoladis 2000). Thus,

⁴ "Missing the bilingual boat." The Straits Times, November 27th, 2009.

for each participant, we summed the rating scores for language proficiency and for frequency of use in English. The average rating score for all participants was 28.32 (SD = 7.74; range from 15 to 35). Our goal was to compare the speech and gestures produced by participants whose English was more vs. less proficient. To maximize the difference in language proficiency between groups, we included the twenty participants whose rating scores were at the extremes of the distribution, 10 at the high end and 10 at the low end. The 10 (4 males) who got the highest rating scores (range from 32 to 35) and were considered proficient in English. These participants rated themselves as natively proficient in speaking and understanding English, and also used English very often in school, home and social settings. In addition, they spoke in and listened to English for at least 8 h on a given day. The 10 participants (5 males) who got the lowest rating scores (range from 15 to 19) were considered less proficient in English. These participants rated themselves as having a fair level of proficiency in English. They always used English in school, but occasionally or even rarely used it at home and other social settings. The majority of participants in this group only spoke in and listened to English for less than 5 h per day.

To verify the level of language proficiency of both groups of bilinguals, a native English speaker, who was blind to the hypothesis of the study, was asked to listen to their speech and judge their *oral* language proficiency. The judge reported that participants who had the highest rating scores sounded proficient in English whereas those who had the lowest rating scores did not.

Speech Coding

As in [So et al. \(2009\)](#), we restricted our analyses to references to the human characters in the story (i.e., the two male protagonists). We classified all references to the protagonists as containing either pronouns (e.g., *he*, *him*) or nouns (e.g., *man*, *goofy-looking man*). Since we were interested in the participants' ability to lexically specify characters over a span of discourse, we analyzed only those nouns and pronouns that referred back to previously mentioned characters (i.e., maintained references). A referent was considered *given* if it was mentioned somewhere in the *preceding* 20 utterances and *new* if it had not been mentioned in the *preceding* 20 utterances ([Chafe 1987](#); [Du Bois 1980](#)). For each participant, we calculated the proportion of nouns (or pronouns) produced when referring back to the two male protagonists, that is, the number of nouns (pronouns) divided by the total number of maintained references to the characters.

We also determined whether each noun and pronoun *lexically* specified its referent in the discourse.⁵ Pronouns were considered to lexically identify referents when they referred to referents that had been mentioned earlier. For example, in the sentence, “*This guy with a walking stick walked down the stairs and he saw this guy with the loop around his neck*”, the speaker used the pronoun, “*he*”, to lexically refer back to the guy with a walking stick. In contrast, in next sentence, “*Then he pushed him down into the sea*”, the speaker used “*he*” to refer to one of the guys and “*him*” to refer to another one but did not make it clear who pushed whom into the sea. Thus, these two pronouns did not lexically specify their referents. Note that referential under-specification could, and did, occur with nouns as well as pronouns, e.g., “*This guy was trying to help the other guy up*”. The nouns here, “*this guy*” and “*the other guy*”, were ambiguous and thus did not lexically identify their referents. In contrast,

⁵ Besides lexical specificity of referential expressions, reference tracking also depends on other factors such as presence of contrastive stress, shared knowledge of speakers and listeners, locations of spoken references in the sentences, processing preferences for proforms, and shared knowledge. However, this study focused on lexical specificity of spoken references.

nouns such as “the drunken man”, “Charlie”, and “the tall guy” did lexically identify their referents.⁶

For each participant, we calculated the proportion of nouns (pronouns) that lexically specified their referents, that is, the number of nouns (pronouns) that lexically specified their referents, divided by the total number of nouns (pronouns) referring to maintained referents.

Gesture Coding

We coded the gestures that co-occurred with spoken references (i.e., with nouns or pronouns) (Levy and Carol 2000). Co-occurrence was coded when the stroke phase or the post-stroke hold of the gesture was produced along with the spoken reference. The stroke phase is the part of the gesture that conveys meaning (McNeill 2005). After identifying the gestures that co-occurred with spoken references, we described their forms (hand shape, location, and movement) and meanings.

We classified the gestures that accompanied spoken references into the following categories: (1) concrete deictic gestures referred to the characters by indicating their pictures (e.g., index finger points to a picture of one of the male protagonists); (2) iconic gestures referred to the characters by evoking qualities (either attributes or actions) of the characters (e.g., right hand draws mustache, referring to one of the male characters; curved palm moves from left to right, referring to an action of pushing performed by the drunken guy); (3) abstract deictic gestures referred to the characters by pointing to the abstract locations that had been previously associated with the characters (e.g., index finger points to the speaker’s left hand side, referring to one of the male protagonists); and (4) beats involved simple hand movements that followed the rhythm of speech (e.g., index finger flips outward).

Our analyses focused on the concrete deictic gestures, iconic gestures, and abstract deictic gestures as they could identify referents and/or associated actions of the referents (McNeill 1992; So et al. 2009). A concrete deictic gesture was considered to identify a previously mentioned referent if it was directed towards the picture of the protagonist. We used criteria developed by Senghas and Coppola (2001; see also Padden 1988) to determine whether an abstract deictic or iconic gesture was used to identify a referent. We assessed the spatial location of each abstract deictic or iconic gesture produced for a particular character in relation to the spatial location of the previously produced gestures for that character. A gesture was considered to identify the referent if it was produced in the same location (left, right, center, top or bottom relative to the location of a speaker) as the previous gestures for that referent.

We used the form of a gesture, in conjunction with the speech in the clause with which it occurred, to determine the character that the gesture identified. Take, for example, an index finger pointed to the left hand side (a location previously identified with the drunken man), which was produced in conjunction with the clause, “The drunken man (point to the left) is walking down the stairs.” Both the gesture and the speech referred to the drunken man; in this case, the gesture added little information to the information conveyed in speech. As another example, consider a right hand curved palm moved across space, which was produced along with the clause, “The drunken guy (iconic gesture moves across space) pushed him into the water.” The gesture and the speech again referred to the drunken man; however, in this case, the gesture added information (the fact that the man moved) to the information conveyed in speech.

⁶ This coding system was validated in So et al. (2009). Twenty naïve listeners were played the audio (and not the video) portion of the tapes and asked to identify the protagonists in the two stories. Listeners identified significantly fewer protagonists in a story containing two male protagonist (the M-M story) than in a story containing a male and female protagonists (the M-W story), suggesting that the speakers did, indeed, use speech to lexically specify referents less often in the M-M story than in the M-W story.

The proportion of spoken references accompanied by different types of gestures (iconic gestures, abstract deictic gestures, and concrete deictic gestures) was calculated for each participant as follows: The total number of the spoken references that were accompanied by a specific type of gesture, divided by the total number of spoken references.

Reliability was assessed by having a second coder transcribe a subset of the story. Agreement between coders was 93 % ($N = 205$) for determining whether spoken referential expressions lexically specified referents; 85 % ($N = 191$) for identifying gestures and describing their forms; 90 % ($N = 162$) for assigning referents to gestures; and 85 % ($N = 146$) for determining whether gestures were produced in the same location as the previously produced gestures for the same referent. All the disagreement was resolved after discussion with the first author.

Results

All proportional data were subjected to an arcsine transformation before statistical analysis. We first explored how often proficient and less proficient English speakers lexically specified the story protagonists in speech. We then examined the gestures that they used, asking whether they were used for referents that were or were not lexically specified in speech. Recall that our analyses focused exclusively on the protagonists that had been previously mentioned (i.e., references back to the previously mentioned protagonists). Altogether, speakers who were proficient in English referred back to the previously mentioned protagonists in speech 23.30 ($SD = 7.01$) times, whereas speakers who were less proficient in English referred back 24.50 ($SD = 5.52$) times.

Lexical Specificity in Referential Expressions in Speech

Of all the spoken references produced by speakers who were proficient in English, 69.56 % ($SD = 9.82$ %) were pronouns and 30.44 % ($SD = 5.14$ %) were nouns. For spoken references produced by speakers who were less proficient in English, 42.24 % ($SD = 7.49$ %) were pronouns and 57.76 % ($SD = 8.29$ %) were nouns. Proficient speakers produced pronouns more often (nouns less often) than less proficient speakers, $t(18) = 4.01$, $p < .001$.

The proportion of nouns that lexically specified their referents was higher in less proficient speakers, $M = 87.49$ % ($SD = 12.93$ %), than in proficient speakers, $M = 47.84$ % ($SD = 8.93$ %), $t(18) = 5.82$, $p < .001$. However, the proportion of pronouns that lexically specified their referents was comparable in both groups, proficient speakers: 37.38 % ($SD = 6.38$ %); less proficient speakers: 38.52 % ($SD = 6.38$ %), $t(18) = 1.12$, $p = ns$. Thus, compared to the proficient speakers, less proficient speakers tended to produce nouns as opposed to pronouns, and the majority of those nouns lexically specified their referents.

The question we next ask is how both groups of speakers gestured in relation to speech. Specifically, we were interested in examining whether concrete deictic gestures, iconic gestures, and abstract deictic gestures bore different semantic relationships to speech, and whether less proficient speakers gestured in relation to speech differently from proficient speakers.

Identifying Referents in Gestures

Speakers in both groups gestured. On average, proficient speakers produced 18.80 ($SD = 8.80$) gestures (7.98 concrete deictic gestures, 9.23 iconic gestures, and 1.59 abstract deictic

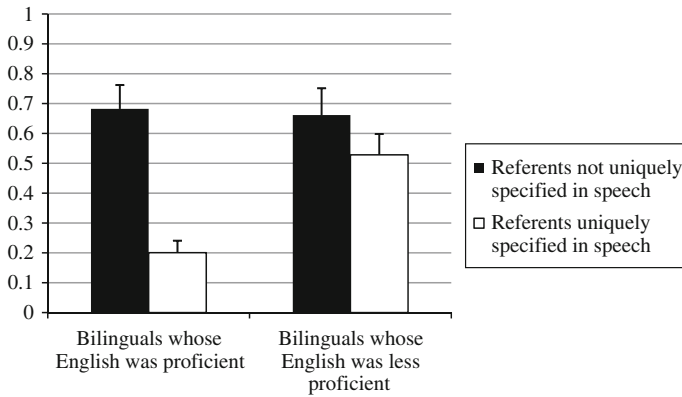


Fig. 1 The proportion of spoken references that were accompanied by concrete deictic gestures as a function of the specificity of the accompanying speech in speakers whose English was proficient (*left bars*) and speakers whose English was less proficient (*right bars*). Spoken references that lexically specified the characters are displayed in the *white bars*; spoken references that did not lexically specify the characters are displayed in the *black bars*. The numbers on this figure represent the original untransformed values

gestures). Less proficient speakers produced 19.40 ($SD = 11.80$) gestures (10.02 concrete deictic gestures, 7.56 iconic gestures, and 1.82 abstract deictic gestures) when referring back to the protagonists. Both groups of speakers thus produced very few abstract deictic gestures: we therefore focused the remainder of our analyses on concrete deictic gestures and iconic gestures.

Concrete Deictic Gestures

We begin with the spoken references that were accompanied by concrete deictic gestures. All concrete deictic gestures were directed towards the pictures of the protagonists, and thus were considered to identify their referents. Proficient speakers produced concrete deictic gestures along with 42.38 % ($SD = 8.19\%$) of their spoken references, whereas less proficient speakers produced concrete deictic gestures along with 54.12 % ($SD = 10.83\%$) of their spoken references.

Figure 1 shows the proportion of spoken references accompanied by concrete deictic gestures as a function of lexical specificity in both groups of speakers. A mixed ANOVA, with proportion of spoken references accompanied by concrete deictic gestures as the dependent variable, referents (lexically specified in speech, not lexically specified in speech) as the within-subject independent variable, and language proficiency (proficient in English, less proficient in English) as the between-subject independent variable, found a significant main effect of referents, $F(1, 18) = 16.99$, $p < .001$, a significant main effect of language proficiency, $F(1, 18) = 8.30$, $p < .01$, and a significant interaction, $F(1, 18) = 9.98$, $p < .005$. Overall, less proficient speakers produced concrete deictic gestures more often than proficient speakers. Proficient speakers were more likely to use concrete deictic gestures when referents were *not* lexically specified in speech than when they were, $t(9) = 5.47$, $p < .001$. However, this difference was not found in less proficient speakers, $t(9) = .65$, $p = ns$. Interestingly, when referents were lexically specified, less proficient speakers were more likely to produce concrete deictic gestures than proficient speakers, $t(18) = 5.29$, $p < .001$.

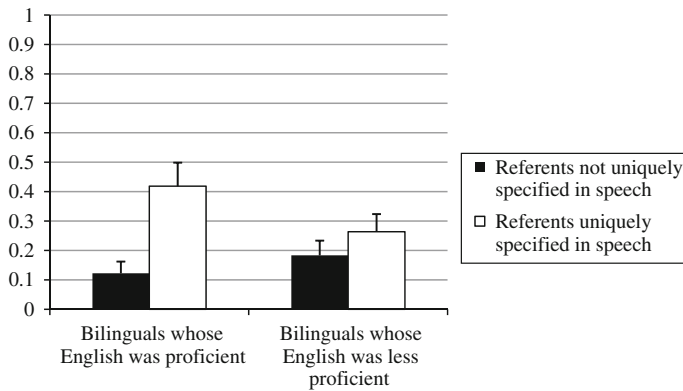


Fig. 2 The proportion of spoken references that were accompanied by iconic gestures as a function of the specificity of the accompanying speech in speakers whose English was proficient (*left bars*) and speakers whose English was less proficient (*right bars*). Spoken references that lexically specified the characters are displayed in the *white bars*; spoken references that did not lexically specify the characters are displayed in the *black bars*. The numbers on this figure represent the original untransformed values

Iconic Gestures

An iconic gesture was considered to specify a previously mentioned referent if it was produced in the same location as the previous gestures for that referent; 58.12% ($SD = 5.45\%$) of iconic gestures identified referents in proficient speakers, as did 52.39% ($SD = 4.82\%$) in less proficient speakers, $t(9) = 2.01$, $p = ns$. In approximately half of the referent-identifying gestures that both groups produced, the gesture added to the information conveyed in the accompanying speech ($M = .54$, $SD = .08$ for proficient speakers; $M = .51$, $SD = .07$, for less proficient speakers, $t(9) = .92$, $p = ns$); in the remaining referent-identifying gestures, the gesture conveyed essentially the same information as was conveyed in the accompanying speech. Proficient speakers were thus no more likely to use iconic gestures to elaborate on the information conveyed in speech than less proficient speakers.

Proficient speakers produced referent-identifying iconic gestures along with 24.52% ($SD = 5.75\%$) of their spoken references, whereas less proficient speakers produced referent-identifying iconic gestures along with 21.64% ($SD = 4.37\%$) of their spoken references. We next investigated the circumstances under which the two groups of speakers produced referent-identifying iconic gestures.

Figure 2 shows the proportion of spoken references accompanied by referent-identifying iconic gestures in both groups of speakers. A mixed ANOVA with the proportion of references to protagonists that were accompanied by referent-identifying iconic gestures as the dependent variable, referents (lexically specified in speech, not lexically specified in speech) as the within-subject independent variable, and language proficiency (proficient in English, less proficient in English) as the between-subject independent variable, found a significant effect of referents, $F(1, 18) = 13.43$, $p < .002$, no effect of language proficiency, $F(1, 18) = 2.96$, $p = ns$, and a significant interaction, $F(1, 18) = 6.52$, $p < .02$. Replicating So et al. (2009), we found that proficient speakers were more likely to produce iconic gestures when the referents were lexically specified in speech than when they were not specified, $t(9) = 4.39$, $p < .002$. In other words, they did not use their gestures to disambiguate unclear spoken references but rather to identify references that were lexically specified in speech. However, this pattern was not found in less proficient speakers, $t(9) = .78$, $p = ns$.

In addition, when referents were lexically specified, less proficient speakers produced referent-identifying iconic gestures less often than proficient speakers, $t(18) = 2.78$, $p < .01$.

Discussion

Our goal was to examine whether concrete deictic gestures and iconic gestures bear the same semantic relation to speech, and whether this gesture-speech relationship is influenced by language proficiency. Two groups of adult bilingual speakers, who varied in their proficiency of English (their second language in all cases) were asked to retell a story that involved two male characters. We found that less proficient speakers produced nouns more often (pronouns less often) than proficient speakers when referring to previously mentioned protagonists. In terms of gesture, we found that proficient speakers used concrete deictic gestures more often when referents were *not* lexically specified in speech than when they were, but showed the reverse pattern for iconic gestures; that is, they used them more often when referents *were* lexically specified in speech than when they were not. Less proficient speakers showed a similar pattern, but the differences were attenuated compared to the same differences in the proficient speakers, and did not reach statistical significance. In addition, less proficient speakers were more likely to produce concrete deictic gestures to indicate referents that were lexically specified than proficient speakers, suggesting that they over-specified referents in both speech and concrete deictic gestures.

Focusing first on the speech devices, we found that less proficient speakers were more likely to produce nouns than proficient speakers when referring to previously mentioned protagonists. This finding is in line with previous research on reference tracking in early-staged second language learners (Gullberg 2003, 2006; Yoshioka 2008). Specifically, second language learners often use specific referential expressions (i.e., nouns) in their weaker language not only for identifying newly introduced characters, but also for identifying previously mentioned characters. Interestingly, as we show here, this over-specification in the weaker language is also found in bilingual speakers who have been using their second language since early childhood. Thus, language proficiency affects the way both second language learners and early bilinguals identify referents in speech—both groups over-specify previously mentioned referents when speaking in their weaker or less dominant language.

Over-using nouns for maintained referents seems to violate an important discourse principle—that less attenuated forms of referring expressions (e.g., pronouns) are used to specify given referents, whereas specific forms (e.g., nouns) are used to specify new referents (e.g., Ariel 1990; Chafe 1994; Givon 1985; Lambrecht 1994). There are various explanations for the over-explicit reference tracking in the less proficient bilinguals in our study. One is that bilinguals who were less proficient in English might not have yet mastered reference-tracking (including the pronominal system) in English, and thus used nouns instead of pronouns to track referents by default (Gullberg 2003; Hendriks 2003). Another possibility is that less proficient bilinguals have a preference to be “hyper-clear” in identifying referents in their second language in order to reduce the ambiguity of those referents (Williams 1989). Finally, it is possible that proper use of English pronouns requires complex planning, which might be too demanding for bilinguals who are less proficient in their second language (Carroll et al. 2000)

Alternatively, the over-use of nouns might be attributed to the bilinguals’ two languages, i.e., Mandarin and English. It is very likely that the bilinguals who were less proficient in English (their second language) were more proficient in Mandarin (their first language). Unlike English where arguments must always be realized (Bloom 1990; Hyams and Wexler 1993),

Mandarin allows both subject and object ellipsis (Huang 1984; Tsao 1990). Moreover, Mandarin has a simple pronominal system compared to English—Mandarin has one form for the third person singular pronouns (i.e., *ta* for male and female), whereas English has two forms (i.e., he, she). In addition, Mandarin does not mark pronouns overtly for case, but English does. The bilingual speakers who were less proficient in English may have been reluctant to use pronouns to identify the protagonists when speaking in English simply because they had not yet mastered this relatively complex system in their second language, and thus fell back on using nouns.

We also found that speakers who were less proficient in English were more likely to lexically specify characters than speakers who were proficient. This result seems surprising in that the proficient speakers appeared to be producing ambiguous speech *more* often than the less proficient speakers. However, the effect may be due to the over-explicitness of the less proficient speakers, rather than to the under-explicitness of the proficient speakers. The less proficient speakers were, in fact, more likely to use elaborated nouns than the proficient speakers, even when those nouns were not needed. Six of the ten less proficient speakers routinely used labels such as *the gentleman* and *the small man* (instead of simply *the man*) for the two male protagonists. In contrast, all of the proficient speakers tended to use simple, yet potentially ambiguous, labels such as *this fellow*, *another man*, and *this guy*.

Both groups of speakers gestured when they talked. They produced both concrete deictic and iconic gestures in locations that had been previously used to identify particular characters. These two types of gestures bore different kinds of relationships to speech, but primarily in the speakers who were proficient in English. Speakers whose English was proficient produced concrete deictic gestures when the referents were *not* lexically specified in speech; in other words, they used these gestures to fill in the gaps left by their speech.

We have thus found that concrete deictic and iconic gestures bear tight, but *different*, semantic relations to the speech with which they occur in proficient speakers. We suggest that these different gesture-speech patterns may stem from the fact that concrete deictic and iconic gestures are generated by different underlying processes and have different discourse functions. Kita and his colleagues (Kita and Özyürek 2003; Kita et al. 2007; Özyürek et al. 2008) have suggested that iconic gestures are generated by a process that packages spatio-motoric imagery into units suitable for speaking (the Interface Hypothesis). In other words, iconic gestures reflect how speakers organize their thinking for the purpose of speaking (Slobin 1987, 1996). Under this view, iconic gestures ought to form parallel patterns found in speech, and they do—when speech lexically specifies a referent, an iconic gesture that co-occurs with the speech will also specify the referent (see also So et al. 2009).

Concrete deictic gestures, in contrast, may be generated by a process that plans multimodal communication as a whole and insures that the intended message is effectively conveyed by speech and gesture (Kita and Özyürek 2003). Under this view, speakers produce concrete deictic gestures to fulfill the communicative goal of encoding as much information in the intended message as possible (the Cross-modal Compensation Hypothesis, see de Ruiter 2006). Concrete deictic gestures, which are driven by physically present referents that are perceptually accessible to both the speaker and listener, provide an ideal vehicle to meet this communicative goal. Thus, when speech does not lexically specify a referent, co-occurring concrete deictic gesture can step in and perform the task of specifying (see So et al. 2010).

Unlike proficient speakers who clearly adjusted their use of concrete deictic gestures to the potential ambiguity in speech, less proficient speakers did not. Instead, less proficient speakers tended to produce concrete deictic gestures (and iconic gestures) regardless of whether the referents were lexically specified in speech. Why did less proficient speakers not use their gestures to distinguish between lexically specified and lexically unspecified referents?

The less proficient speaker's over-explicitness in speech might account for their over-use of concrete deictic gestures for referents that were already specified. By using nouns for maintained references, the less proficient speakers violate several pragmatic rules, such as the Gricean quantity principle (Grice 1975), conversational implicatures (Levinson 2000), relevance (Sperber and Wilson 1986), and the principle of recipient design (Sacks and Schegloff 1979), which is likely to increase ambiguity in who is doing what to whom for the listener. Less proficient speakers might then produce concrete deictic gestures to resolve this ambiguity.

However, the less proficient speakers in our study did not use their iconic gestures to over-specify referents, contrary to findings reported by Gullberg (2003, 2006). The less proficient speakers in our study even produced iconic gestures for specified referents less often than proficient speakers. One possible reason is that, when entities (e.g., pictures of protagonists) are physically present, it is easier to point at those entities than to produce iconic gestures for the entities. Less proficient speakers may find it easier to take this less demanding path (in our study, pointing at the pictures of the protagonists rather than producing iconic gestures for them). However, entities were not physically present in Gullberg's studies. Thus, participants might produce iconic gestures instead in order to identify the referents. As a result, it is difficult to compare the number of iconic gestures produced by the second language learners in Gullberg's studies to the number produced by the bilinguals in our study. Future work is needed to determine whether early-staged second language learners will reduce their production of iconic gestures when they have the opportunity to produce concrete deictic gestures.

To summarize, our findings showed that concrete deictic and iconic gestures bear a different semantic relation to the speech they accompany, particularly in proficient speakers. Proficient speakers tend to produce iconic gestures to further specify referents that are already specified in speech (i.e., they convey information that is redundant with the information conveyed in speech), and concrete deictic gestures to specify referents that are *not* already specified in speech (i.e., they add information to the information conveyed in speech). However, these patterns were attenuated in less proficient speakers, who were just as likely to produce gestures (both concrete deictic gestures and iconic gestures) for referents that were already specified in speech as for referents that were not specified in speech. The findings make it clear that both type of gesture and proficiency of speaker need to be taken into account to explain how speakers use gesture and speech in narrative discourse.

Acknowledgments This work was supported by National University of Singapore Provost Research Funding Grant No. R-581-000-074-133 to W.C. So. We thank Ho Sz Wha and Chan Yi Tsun for help in data collection and Lim Jia Yi and Chew Xin Ying Ivane for coding

References

- Ariel, M. (1990). *Accessing noun phrase antecedents*. London: Routledge.
- Arnold, J. E., Eisenband, J., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13–B26.
- Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, 15(6), 415–419.
- Bloom, P. (1990). Subjectless sentences in child's language. *Linguistic Inquiry*, 24, 721–734.
- Bosch, P. (1983). *Agreement and anaphora: A study of the roles of pronouns in discourse and syntax*. London: Academic Press.
- Carroll, M., Jorge, M.-S., Marzena, W., & Alessandra, B. (2000). The relevance of information organization to second language acquisition studies: The descriptive discourse of advanced adult learners of German. *Studies in Second Language Acquisition*, 22(3), 441–466.

- Chafe, W. L. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse*. Amsterdam: John Benjamins.
- Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.
- de Ruiter, J. P. (2006). Can gesticulation help aphasic people speak, or rather, communicate?. *Advances in Speech-Language Pathology*, 8(2), 124–127.
- Dixon, L. Q. (2005). Bilingual education policy in Singapore: An analysis of its sociohistorical roots and current academic outcomes. *International Journal of Bilingual Education and Bilingualism*, 8(1), 25–47.
- Du Bois, J. (1980). Beyond definiteness: The trace of identity in discourse. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural and linguistic aspects of narrative production* (pp. 203–274). Norwood: Ablex.
- Francik, E. P. (1985). Referential choice and focus of attention in narratives (discourse anaphora, topic continuity, language production). *Dissertation Abstracts International*, 46(4), 1363.
- Fox, B. (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.
- Garrod, S. (2001). Anaphora resolution. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavior sciences* (pp. 490–494). Amsterdam: Elsevier.
- Givón, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109–136). Amsterdam: Benjamins.
- Givón, T. (1985). Iconicity, isomorphism and non-arbitrary coding in syntax. In J. Haiman, *Iconicity in syntax* (pp. 187–219). Amsterdam: Benjamins.
- 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: Erlbaum Associates.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, Vol. 3* (pp. 41–58). New York: Academic Press.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund: Lund University Press.
- Gullberg, M. (1999). *Communication strategies, gestures, and grammar. Paper presented at the 8th EUROSLA Conference Paris*.
- Gullberg, M. (2003). Gestures, referents, and anaphoric linkage in learner varieties. In C. Dimroth & M. Starren (Eds.), *Information structure, linguistic structure and dynamics of language acquisition* (pp. 311–328). Amsterdam: Benjamins.
- Gullberg, M. (2006). Handling discourse: Gestures, referent tracking, and communication strategies in early L2. *Language Learning*, 56(1), 156–196.
- Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 291–326). Berlin: Mouton de Gruyter.
- Huang, C. T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15, 531–574.
- Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 24(3), 421–459.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 367–371.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for and interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.
- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22(8), 1212–1236.
- Lambrecht, K. (1994). *Information structure and sentence form. Topic focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Lim, V., Rickard Liow, S., Lincoln, M., Chan, Y. H., & Onslow, M. (2008). Determining language dominance in English–Mandarin bilinguals: Development of a self-report classification tool for clinical use. *Applied Psycholinguistics*, 29, 389–412.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT press.
- Levy, E. T., & Carol, A. F. (2000). The role of gestures and other graded language forms in the grounding of reference in perception. In D. McNeill (Ed.), *Language and gesture* (pp. 215–234). Cambridge: Cambridge University Press.
- Lyons, J. (1977). *Semantics, Vol. 2*. Cambridge: Cambridge University Press.
- Mayberry, R., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, 9(6), 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: The University of Chicago Press.
- Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, 44(4), 1040–1054.
- Padden, C. (1988). *Interaction of morphology and syntax in American Sign Language*. NY: Garland.
- Pakir, A. (1998). English in Singapore: The codification of competing norms. In S. Gopinathan, A. Pakir, H. W. Kam, & V. Saravanan (Eds.), *Language, society and education in Singapore: Issues and trends* (2nd ed., pp. 65–84). Singapore: Times Academic Press.
- Pakir, A. (1999). *Bilingual education with English as an official language: Sociocultural implications*. Georgetown University Roundtable on Language and Linguistics.
- Sacks, H., & Schegloff, E. A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 15–21). New York: Irvington Press.
- Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, 12, 323–328.
- Shepherd, J. (2005). *Striking a balance: The management of languages in Singapore*. Frankfurt am Main Germany: Peter Lang.
- Slobin, D. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the thirteenth annual meeting of the Berkeley Linguistic Society* (pp. 435–445). Berkeley: Berkeley Linguistic Society.
- Slobin, D. I. (1996). From “thought and language” to “thinking to speaking”. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.
- So, W. C., Coppola, M., Liccigarello, V., & Goldin-Meadow, S. (2005). The seeds of spatial grammar in the manual modality. *Cognitive Science*, 29, 23–37.
- So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Speech and gesture go hand-in-hand. *Cognitive Science*, 33, 115–125.
- So, W. C., Demir, E., & Goldin-Meadow, S. (2010). When speech is ambiguous gesture steps in: Sensitivity to discourse-pragmatic principles in early childhood. *Applied Psycholinguistics*, 31(1), 209–224.
- So, W. C., & Lim, J. Y. (2012). Point to a referent, and say, “What is this?” Gesture as a potential cue to identify referents in discourse. *Applied Psycholinguistics*, 33(2), 329–342.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Stokoe, W. C. (1960). Sign language structure: An outline of the visual communication systems of the American deaf. *Studies in Linguistics, Occasional Papers* 8.
- Tsao, F.-F. (1990). *Sentence and clause structure in Chinese: A function perspective*. Taipei: Student Book.
- Williams, J. (1989). Pronoun copies, pronominal anaphora and zero anaphora in second language acquisition. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition: Discourse and pragmatics* (pp. 153–189). Philadelphia: Multilingual Matters.
- Yoshioka, K. (2008). Gesture and information structure in first and second language. *Gesture*, 8(2), 236–255.