Cognitive Science 29 (2005) 1029–1043
Copyright © 2005 Cognitive Science Society, Inc. All rights reserved.

The Seeds of Spatial Grammar in the Manual Modality

Wing Chee So, Marie Coppola,
Vincent Licciardello, Susan Goldin-Meadow

Department of Psychology, University of Chicago

Received 22 December 2004; received in revised form 26 March 2005; accepted 4 April 2005

Abstract

Sign languages modulate the production of signs in space and use this spatial modulation to refer back to entities—to maintain coreference. We ask here whether spatial modulation is so fundamental to language in the manual modality that it will be invented by individuals asked to create gestures on the spot. English speakers were asked to describe vignettes under 2 conditions: using gesture without speech, and using speech with spontaneous gestures. When using gesture alone, adults placed gestures for particular entities in non-neutral locations and then used those locations to refer back to the entities. When using gesture plus speech, adults also produced gestures in non-neutral locations but used the locations coreferentially far less often. When gesture is forced to take on the full burden of communication, it exploits space for coreference. Coreference thus appears to be a resilient property of language, likely to emerge in communication systems no matter how simple.

Keywords: Sign languages; Homesign; Spatial grammar; Gestures; Language development; Resilient properties of language; Coreference

1. Introduction

Sign languages have evolved in a different biological medium from spoken languages. But despite striking differences in modality, they are structured like spoken languages at phonological (Brentari, 1998; Perlmutter, 1992; Sandler, 1989), morphological (Klima & Bellugi, 1979; Padden, 1988; Supalla, 1982), and syntactic (Liddell, 1980; Lillo-Martin, 1991; Neidle, Kegl, McLaughlin, Bahan, & Lee, 2000; Padden, 1981) levels. Indeed, many grammatical devices found in spoken languages are also found in sign languages (e.g., signaling who does what to whom by the order in which a word or sign is produced; Liddell, 1984). But sign languages employ at least one grammatical device that is not possible in spoken language—signaling differences in meaning by modulating where signs are produced in space.

Requests for reprints should be sent to Wing Chee So, Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637. E-mail: cwcso@uchicago.edu

All sign languages studied thus far use space to indicate referents and the relations among them (Aronoff, Meir, Padden, & Sandler, 2003; Supalla, 1995). These uses of space lay the foundation for maintaining coherence in a discourse. In American Sign Language (ASL), a signer can associate a spatial location with an entity and later articulate a sign with respect to that location to refer back to the entity. For example, after associating a location in space with Juliet, a signer can later produce a verb with respect to that space to refer to Juliet, without repeating the sign for Juliet (Padden, 1988). By using the same space for an entity throughout a discourse, signers maintain coreference. Coreference is an important function in all languages (Bosch, 1983) and considered a "core" property of grammar (Jackendoff, 2002). In spoken languages, coreference must be accomplished through nonspatial devices. For example, in the sentence, "Seymour always wins when he enters," the pronoun *he* refers back to, and is thus coreferential with, *Seymour.*[1]

Using space for coreference is found not only in well-established sign languages such as ASL, but also in newly emerging sign languages such as Nicaraguan Sign Language. In 1977 the first school for the deaf was established in Nicaragua. Deaf children had the opportunity to interact freely with one another for the first time, and they began to form a sign language (Kegl & Iwata, 1989).[2] Although less than 30 years old, Nicaraguan Sign Language uses spatial modulation for coreference (Senghas & Coppola, 2001).[3] This property thus appears to be basic to manual languages, young or old.

We ask here whether using space for coreference is so fundamental to language in the manual modality that it will be invented by individuals asked to create gestures on the spot. Exploiting a paradigm used in previous work (Casey, 2003; Dufour, 1993; Goldin-Meadow, McNeill, & Singleton, 1996), we instructed adults, all naive to sign language, to describe scenes using gesture and no speech. Previous work using this procedure found that English-speaking adults invent discrete gestures to represent objects and actions and produce these gestures in strings characterized by consistent order. Moreover, the order is one *not* typically found in English sentences (Gershkoff-Stowe & Goldin-Meadow, 2002). English speakers are not merely "translating" their spoken language into gesture—they are inventing a new system. These findings leave open the possibility that, when asked to communicate using gesture without speech, nonsigners might exploit the spatial properties of the manual modality for coreference.

It is, of course, impossible not to use space when using the hands to communicate. Moreover, English speakers have been found to produce gestures in non-neutral locations (i.e., away from the chest) when describing events (Casey, 2003) or retelling stories (Dufour, 1993), using gesture alone. Our question is whether speakers-turned-signers use space for coreference. To explore this question, we manipulated the events described. Some adults saw events presented in an order that told a story (connected events). Others saw the same events in random order interspersed with events from other stories (unconnected events). If adults are able to use space coreferentially, they should be able to establish a location for a character and use that location to refer to the character throughout the discourse. Adults describing connected events can use the same spatial framework throughout the story. But adults describing unconnected events must establish a new framework for each event and are thus less likely to use space coreferentially.

Finally, we ask whether space is used for coreference in the gestures speakers produce when they talk. Co-speech gestures are global in form with little combinatorial structure and, as such, are structured differently from sign languages (Goldin-Meadow, 2003a; McNeill, 1992). Nevertheless, speakers do use space in meaningful ways in gestures produced along with spatial descriptions (Emmorey, Tversky, & Taylor, 2000; Iverson, 1999) and will, at times, point toward a particular space when referring to the same character (Kendon, 2004; McNeill, 1992). To determine whether gesture is used coreferentially when it shares the burden of communication with speech versus when it is used on its own in place of speech, we asked English speakers to describe the same scenes with speech (gesture + speech) and without it (gesture alone).

2. Method

2.1. Participants and procedure

Eighteen English-speaking undergraduate students, naive to sign language, were recruited through postings and paid for their participation. Participants were individually shown 81 videotaped vignettes, each lasting 1 to 6 sec, culled from eight silent stories. Participants were randomly assigned to the connected or unconnected event conditions. Nine participants saw the 81 vignettes presented as eight stories (connected events). Participants were first shown all of the vignettes in a particular story without pauses so that they could get a sense of the plot. The vignettes in the story were then repeated one at a time and, after each vignette, participants were asked to describe the scene; they thus saw each vignette twice. Nine participants saw the 81 vignettes in random order, with vignettes from all eight stories scrambled throughout the presentation (unconnected events). Participants saw each vignette twice; after the second presentation, they were asked to describe the vignette before watching the next (Appendix A found at http://www.cognitivesciencesociety.org/supplements/ presents vignette orders for both conditions).

Participants were first asked to describe the vignettes to an experimenter using gesture without speech (gesture alone). Participants were then shown the vignettes again (either connected or unconnected events depending on their condition) and were asked to describe the scenes with speech; we assumed participants would gesture while speaking but did not mention gesture in the instructions (gesture + speech). We did not vary order of conditions because we wanted the events to be completely novel for the gesture-alone descriptions, the primary focus of our study.

We concentrated on one story containing 11 vignettes. The vignettes were presented in order in Slots 12 to 22 (out of 81) in connected events; these same 11 vignettes were presented in the following slots in unconnected events: 29, 56, 46, 44, 48, 42, 21, 36, 14, 58, and 51 (see Appendix A). The story involved two characters participating in a variety of motion events (e.g., man gives woman basket; man kisses woman; woman walks upstairs; man falls). We chose this story because the events were relatively easy to convey in gesture and often involved both characters. Participants thus had to distinguish between the characters when they appeared together and could use that distinction in future references.

Vignette-sequence (connected vs. unconnected) was analyzed as a between-subject factor and modality (gesture alone vs. gesture + speech) as a within-subjects factor. The session was videotaped.

2.2. Coding

Gestures were coded with the sound off. Gesture form was described in terms of parameters used to describe sign languages (Stokoe, 1960)—shape and placement of hand, trajectory of motion. Change in any one of these parameters marked the end of one gesture and the beginning of another. Gesture meaning was assessed in relation to the scene described (Gershkoff-Stowe & Goldin-Meadow, 2002). *Action* gestures described the action portrayed in the vignette (e.g., OPEN used to describe a door-opening action). *Object* gestures identified entities in the vignette (e.g., MOUSTACHE used to identify the man).[4] *Deictic* gestures were points at empty spaces (e.g., POINT to gesturer's right).

We used criteria developed by Senghas and Coppola (2001) to code use of space. Gestures were considered *spatially modulated* if they were produced in a non-neutral location (i.e., away from the chest area) or a location associated with an object or action, or if they incorporated either type of location into their movements. For example, a deictic was spatially modulated if it pointed to a space previously associated with an object.

To determine whether a spatially modulated gesture was used coreferentially, we compared that gesture to spatially modulated gestures preceding it within the vignette and referring to the same entity. If two gestures referring to the same entity shared a location, the second gesture was classified as *coreferential.* For example, the participant in Fig. 1 was describing the vignette "man gives woman basket." He first set up one person (man) on his body [G1] and a second person (woman) on his right [G2]. He then produced a GIVE gesture moving from a location in front of him (later identified as basket) to the location on his right (woman) [G3, which was coreferential with G2]. After producing a gesture for basket in the location in front of him [G4, which was coreferential with G3], he again produced a GIVE gesture moving from the basket location to the woman location [G5, which was coreferential with G2, G3, and G4].

We also examined whether space was used coreferentially *across* vignettes. For example, to describe the next vignette "man kisses woman," the participant in Fig. 2 did not set up woman anew but relied on the location established in the previous vignette to refer to her. To calculate coreference across vignettes, we focused on the man and woman characters who appeared together in six vignettes. We calculated the number of times participants placed their gestures for a character in the same location across the six vignettes (e.g., if a participant located the woman on his left in five of six vignettes, his coreference score was .83).

Reliability was assessed by having a second coder transcribe a subset of responses in each condition. Agreement between coders was 86% ($N = 144$) for identifying gestures and describing their form, 84% ($N = 123$) for assigning meaning to gestures, 92% ($N = 103$) for determining whether a gesture was spatially modulated, 86% ($N = 95$) for assessing coreference within vignettes, and 90% ($N = 32$) for assessing coreference across vignettes.

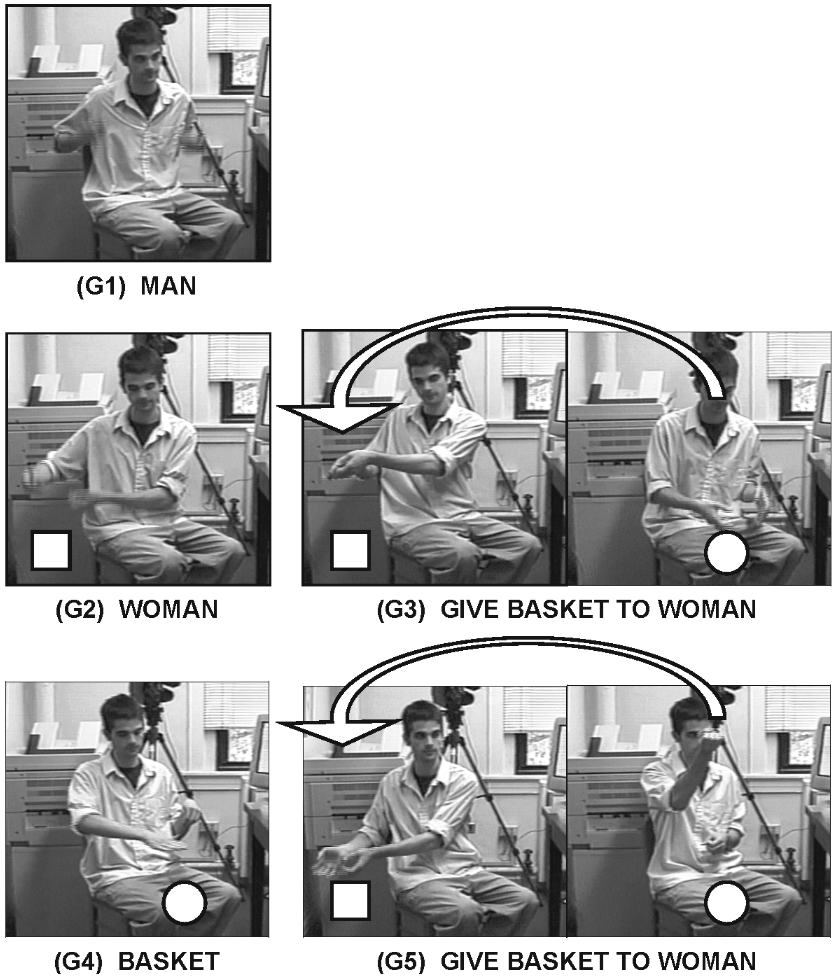

Fig. 1. An example of coreferential gestures within a vignette. The square represents the location established for the woman; the circle represents the location established for the basket. G3 (an Action gesture) is coreferential with G2 (an Object gesture). G4 (an Object gesture) is coreferential with G3. G5 (an Action gesture) is coreferential with G 2, 3, and 4.

3. Results

Table 1 presents the mean number of gestures produced with and without speech. Participants produced more gestures when describing vignettes in gesture alone than in gesture + speech, $F(1, 16) = 44.94$, $p < .00001$; there were no significant differences in number of gestures in connected versus unconnected events, $F(1, 16) = 1.26$, *ns,* and no interaction between modality and vignette sequence, $F(1, 16) = 2.86$, *ns.* Four participants produced no gestures at all in gesture + speech (1 connected, 3 unconnected).[5]

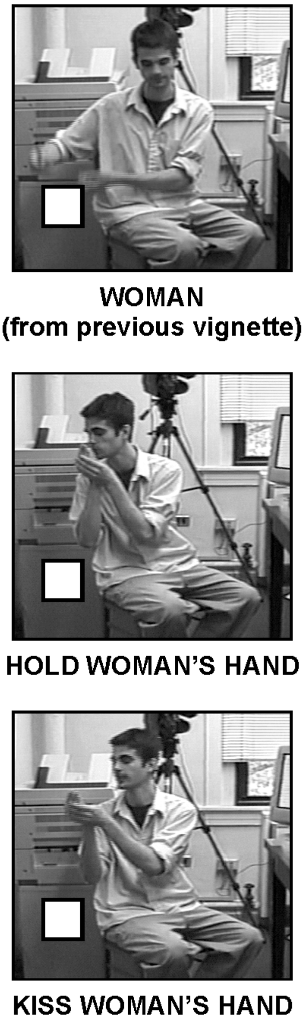

Fig. 2. Example of coreferential gestures across vignettes. The square represents the location that this participant established for the woman in his description of a previous vignette (top gesture). When describing the kissing vignette, he continued to use the location to his right as a placeholder for the woman (middle and bottom gestures). The gestures referring to the woman in the kissing vignette (middle and bottom gestures) are thus coreferential with the gestures referring to the woman in the previous vignette (top gesture and Fig. 1).

3.1. Gestures produced on their own

All 18 participants produced object and action gestures, but only 7 produced deictic gestures (2 connected, 5 unconnected) in gesture alone. Participants produced different numbers of object, action, and deictic gestures, $F(2, 32) = 24.16$, $p < .0001$—more objects than actions ($p = .05$) and deictics ($p < .0001$), and more actions than deictics ($p < .001$). There were no significant differences between connected versus unconnected events, $F(1, 16) = 1.96$, *ns,* and no interaction between gesture type and vignette sequence, $F(2, 32) = 1.67$, *ns.*

Table 1
Mean number and types of gestures produced in each of the four conditions

Modality	Events Described	Total Number of Gestures		Types of Gestures[a] Objects		Actions		Deictics	
		M	SD	M	SD	M	SD	M	SD
Gesture alone	Connected	48.22	20.22	26.44	16.13	21.00	4.30	0.89	1.83
	Unconnected	73.11	49.06	45.00	36.65	25.22	12.65	2.89	5.06
Gesture + speech	Connected	12.78	10.87	3.75	4.33	9.25	6.96	1.38	1.06
	Unconnected	13.78	18.97	7.83	8.38	12.50	12.15	0.33	0.82

[a]All 18 of the adults produced gestures in the gesture-alone condition, but only 14 produced gestures in the gesture + speech condition (8 who viewed the connected events and 6 who viewed the unconnected events); numbers of object, action, and deictic gestures are calculated only for those adults who produced gestures.

Fig. 3 (left panel) presents the proportion of gestures that were spatially modulated in gesture alone. Participants produced the same proportion of spatially modulated gestures whether describing connected or unconnected events, $F(1, 16) = 0.04$, *ns.*[6,7]

The question, however, is whether those gestures were used coreferentially. Fig. 3 (right panel) presents the proportion of spatially modulated gestures that were coreferential within vignettes in gesture alone. Participants used gestures coreferentially significantly more often when describing connected than unconnected events, $F(1, 15) = 19.32$, $p = .0005$.[8,9]

Turning next to spatially modulated gestures that were coreferential *across* vignettes, we found that participants were more likely to produce gestures for a given character in the same location when describing connected than unconnected events: .98 ($SD = .06$) versus .85 ($SD = .13$) for man, $F(1, 16) = 13.14$, $p = .002$; .93 ($SD = .22$) versus .54 ($SD = .20$) for woman, $F(1, 16) = 20.01$, $p < .0004$.

The lack of cohesive plot and the presence of multiple characters in unconnected events made it difficult to use a previously established location to identify a character. Participants could, however, use other means to identify characters. For example, in describing "man gives woman basket," one participant produced a MOUSTACHE gesture identifying the man immediately before producing a GIVE gesture that was *not* anchored in an identified space (i.e., his hands moved from one unidentified space to another). The juxtaposition of the two gestures effectively identified the man as the giver despite the fact that the GIVE gesture was not coreferential. To examine how often participants used this strategy, we extracted all of the spatially modulated action gestures and determined whether these gestures were immediately preceded or followed by an object gesture identifying one of the characters in the action. We expected that characters would be lexicalized (i.e., identified via an object gesture) when they were *first* introduced in the vignettes for both connected and unconnected events. However, we expected characters to be lexicalized in *subsequent* mentions more often in descriptions of unconnected events simply because these characters were less likely to be identified by a spatially coreferential gesture.

Fig. 4 presents the proportion of spatially modulated gestures that were accompanied by a lexicalizing gesture for a character (i.e., an object gesture that identified the character) in con-

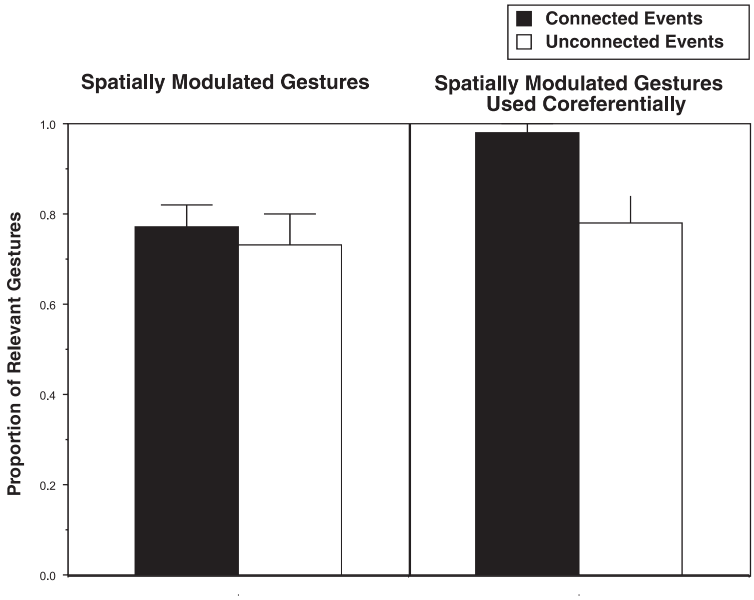

Fig. 3. The proportion of gestures that the adults produced in non-neutral locations (left panel) and the proportion of these spatially modulated gestures that the adults used to maintain coreference within a vignette (right panel) when describing Connected (black bars) and Unconnected (white bars) events in gesture without speech (Gesture-Alone).

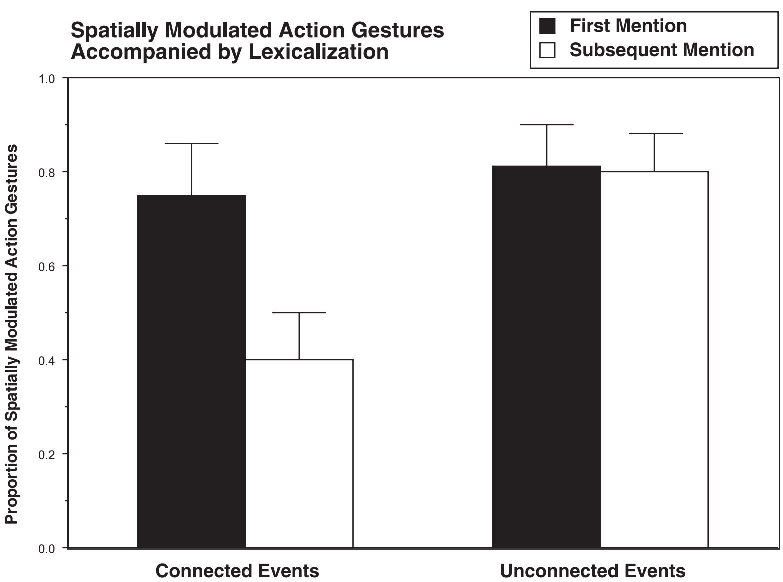

Fig. 4. The proportion of spatially modulated action gestures accompanied by lexicalization that the adults produced in their first mention of an entity (black bars) or in subsequent mentions of the entity (white bars) when describing Connected and Unconnected events in gesture without speech (Gesture-Alone).

nected and unconnected events. Black bars represent gestures used when a character was first introduced; white bars represent gestures used during subsequent mentions of the character. There was a significant effect of first versus subsequent mention, $F(1, 14) = 6.49$, $p = .02$, but as expected, this effect was mediated by condition.[10] Planned comparisons revealed that characters were identified via lexicalization less often in subsequent mentions than in first mentions in connected ($p = .007$) but not unconnected ($p = .53$) events. Indeed, lexicalization did not differ in descriptions of connected and unconnected events for first mentions ($p = .58$) but did for subsequent mentions ($p = .01$). Thus, when asked to describe unconnected events, participants found it relatively difficult to use space consistently to identify characters; however, they did manage to identify the characters by producing separate lexical gestures for those characters.

The findings thus far suggest that, when asked to describe events using gesture and no speech, adults spontaneously use space to refer back to characters. Moreover, they do so more often when they can exploit a coherent plot line. The final question we address is whether this same strategy appears in gestures that accompany speech.

3.2. Gestures spontaneously produced with speech

As mentioned earlier, 14 of 18 participants produced gestures when describing vignettes in speech. Thirteen produced action gestures (8 connected, 5 unconnected); 12 produced object gestures (6 in each condition); 7 produced deictic gestures (6 connected, 1 unconnected). Participants produced different numbers of object, action, and deictic gestures, $F(2, 24) = 12.88$, $p = .0002$—more actions than objects ($p = .05$), thus displaying a different pattern from gesture alone; but more actions ($p = .0001$) and objects ($p = .059$) than deictics, as in gesture alone. There were no significant differences between connected versus unconnected events, $F(1, 12) = .58$, *ns,* and no interaction between gesture type and vignette sequence, $F(2, 24) = .97$, *ns.*

To compare how space was used in gesture + speech versus gesture alone, we focused on the 14 participants who produced gestures in both conditions. Fig. 5 (left panel) presents the proportion of spatially modulated gestures that these 14 participants produced. Participants produced the same proportion of spatially modulated gestures in gesture alone and gesture + speech, $F(1, 12) = 2.25$, *ns;* there were no differences between connected and unconnected events, $F(1, 12) = .03$, *ns,* and no interaction between modality and vignette sequence, $F(1, 12) = .04$, *ns.*

Fig. 5 (right panel) presents the proportion of spatially modulated gestures that were coreferential in gesture alone versus gesture + speech.[11] There was an effect of modality, $F(1, 8) = 21.65$, $p < .002$, which interacted with vignette sequence, $F(1, 8) = 9.86$, $p = .01$. Participants used gestures coreferentially significantly more often in gesture alone than gesture + speech when describing connected ($p < .003$, Newman–Keuls) but not unconnected ($p = .32$) events. In this subset of the data, participants again used gestures coreferentially more often for connected than unconnected events in gesture alone ($p = .02$)—but not in gesture + speech ($p = .14$).

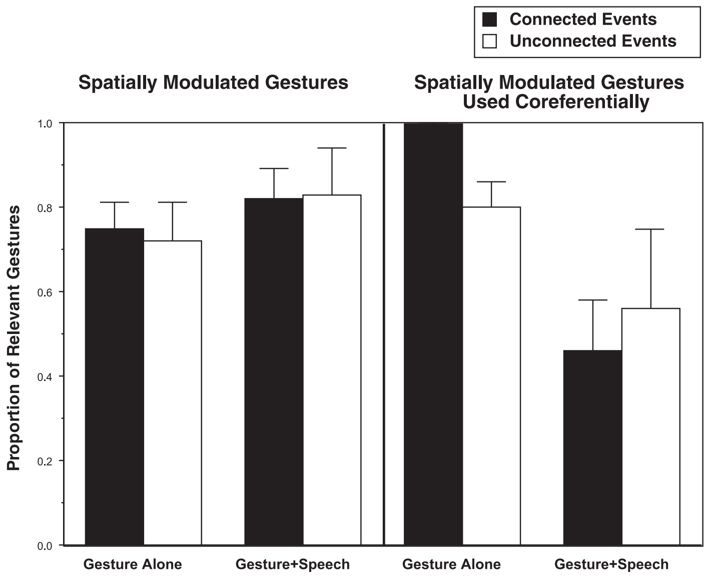

Fig. 5. The proportion of gestures that the adults produced in non-neutral locations (left panel) and the proportion of these spatially modulated gestures that the adults used to maintain coreference within a vignette (right panel) when describing Connected (black bars) and Unconnected (white bars) events in gesture with speech (Gesture+Speech) and without it (Gesture-Alone).

4. Discussion

Conventional sign languages use spatial modulations for coreference. We have shown here that spatial modulation is so fundamental to language in the manual modality that adults will introduce it into their gestures when asked to communicate using only their hands. They place gestures for particular entities in non-neutral locations and then use those locations later in the discourse to refer back to the entities. Moreover, they use this device more often when describing connected than unconnected events. The emergence of this particular spatial device in speakers-turned-signers is striking because, as we also demonstrate, speakers do not routinely use space coreferentially in the gestures they produce when they talk.

4.1. Coreference as a resilient property of language

Previous work has shown that, when asked to communicate using only gesture, English speakers concatenate gestures into strings characterized by consistent (and non-English) orders that signal who does what to whom (Gershkoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow et al., 1996). All languages, signed or spoken, display canonical word orders. Moreover, consistent word order is among the first properties to appear in linguistic systems emerging under less-than-ideal circumstances—for example, in signed and spoken linguistic systems developed by children and adults learning language beyond the critical period (Curtiss, 1977; Newport, 1991); in young sign languages transmitted across relatively few generations (Sandler, Meir, Padden, & Aronoff, 2005; Senghas, Coppola, Newport, & Supalla, 1997); and in home-sign gesture systems invented by deaf children whose hearing losses prevent them from acquiring speech and whose hearing parents have not exposed them to sign, that is, children who lack conventional linguistic input (Feldman, Goldin-Meadow, & Gleitman, 1978; Goldin-Meadow & Feldman, 1977). It is in this sense that word order is a resilient property of language (Goldin-Meadow, 2003b).

Our findings suggest that coreference may also be a resilient property of language. Coreference is a property of all conventional languages, although implemented differently in speech and sign. Sign languages use space to establish coreference, whether those sign languages are mature and established (Aronoff et al., 2003; Padden, 1988) or young and emerging (Senghas, 2003; Senghas & Coppola, 2001). Moreover, as shown here, speakers spontaneously use space for coreference when asked to communicate using gesture without speech.[12]

More important, speakers-turned-signers do not introduce all properties of language, or even all properties of spatial verb agreement systems, into their gestures. For example, they do not create gesture systems in which discrete hand shape and motion forms map consistently onto categories of meanings (i.e., morphological structure; Singleton, Morford, & Goldin-Meadow, 1993), a property that is developed by deaf children inventing home signs without input from a conventional language (Goldin-Meadow, Mylander, & Butcher, 1995). Morphology thus appears to be less resilient than coreference.

It may not be surprising that speakers-turned-signers exploit space as a representational device to maintain coreference. Spatial knowledge is among the earliest and most firmly established bodies of knowledge (e.g., Newcombe & Huttenlocher, 2000) and often serves as the source domain for abstract metaphors and analogies (e.g., Nunez & Lakoff, 1998). Indeed, some have argued that spatiomotor processing underlies all cognition, including abstract thought (Barsalou, 1999; Glenberg, 1997). Note, however, that the mappings used by our speakers-turned-signers were not particularly abstract. The adults established a gestural stage and used that stage to reenact the motion events they saw. The next question is whether nonsigning adults can take the step taken by all sign languages and use their gestural stage to convey abstract relations (e.g., can they use a location established for Juliet to indicate that Romeo is thinking about her, as opposed to physically interacting with her?).

4.2. Gesture with speech versus without it

If using space coreferentially comes so naturally to the manual modality, why then is not space used for this purpose more often in gestures that accompany speech? We suggest that co-speech gestures work differently from gesture used on its own simply because the gestures that accompany speech do not form a system unto themselves but are instead integrated into a single system with speech (Goldin-Meadow et al., 1996). This fact is nicely illustrated in our data. Adults produced more object than action gestures in gesture alone but more action than object gestures in gesture + speech. The gestures that speakers produce when they talk typically do not segment a proposition into discrete elements but rather represent the entire proposition with a single, global gesture (McNeill, 1992). An action gesture is often the best way to represent a proposition in its entirety because the action sets the stage for its participants. In contrast, when gesture is used in place of speech, it must take on all the roles that speech assumes and must refer not only to the proposition, but also to the participants involved in that proposition (a function it achieves with gestures for objects as well as actions).

Gestures produced along with speech thus form an integrated system with that speech. As part of this integrated system, co-speech gestures are frequently called on to serve multiple functions—for example, they not only convey propositional information (Goldin-Meadow, 2003a), but they also coordinate social interaction (Bavelas, Chovil, Lawrie, & Wade, 1992; Haviland, 2000) and break discourse into chunks (Kendon, 1972; McNeill, 2000). As a result, the placement of a series of co-speech gestures in a narrative reflects a variety of pressures, pressures that may compete with using those gestures coreferentially.

When asked to use gesture on its own, the adults in our study did not use gesture as they typically do but rather transformed it into a more language-like structure (cf. Goldin-Meadow et al., 1996). This transformation is reminiscent of the transformation that home-signing deaf children perform when they take the gestures that they see in the hearing world and turn them into a system with language-like structure (Goldin-Meadow, 2003b). Moreover, it may be the same sort of transformation that takes place when sign languages recruit gestures from hearing cultures (Senghas, Özyürek, & Kita, 2004). Thus, when a signer (either a life-long signer or one creating a system over a period of years or on the spot) borrows a gestural form from speakers, that form is not likely to be able to be used "off the rack" but will require alterations.

To summarize, when asked to communicate using gesture without speech, adults are able to exploit spatial properties of the manual modality for coreference. These findings suggest that coreference is itself a resilient property of language, one that is likely to be introduced into newly emerging communication systems no matter how simple.

Notes

1. Note that the *he* in this sentence can refer either to *Seymour* or to another contestant. In contrast, in "He always wins when Seymour enters," *he* cannot refer to *Seymour* and must refer to some other contestant. The regularities that govern when pronouns can and cannot refer to the same entity within a sentence are part of the rules of coreference, referred to as *government and binding* (Chomsky, 1981).
2. Prior to 1977 there were a small number of private clinics in Nicaragua for educating deaf children (Polich, 2001). The clinics did not foster interaction among children, and no conventionalized sign language emerged in these settings.
3. The first cohort of Nicaraguan signers (i.e., the children who created the sign language when they came together in 1977) and the second cohort (the children who entered the community later and learned the language from their peers) both used their signs coreferentially. However, the second cohort produced more spatial modulations per verb than the first cohort (Senghas & Coppola, 2001).
4. Participants at times used actions not displayed in the vignette to identify an object (e.g., PUT-ON-HAT used to identify the man who was wearing a hat but did not put it on during the vignette). These gestures were considered object gestures.
5. Interestingly, participants who produced many gestures without speech also produced many with speech: The correlation between mean number of gestures in gesture alone

and gesture + speech was .62 ($p < .006$, $N = 18$). However, the fact that the gesture + speech condition immediately followed the gesture-alone condition for all participants may have inflated the correlation. But note that conducting the two manipulations on the same day and in the same order works against the finding that gesture + speech differs from gesture alone.

6. All proportions were subjected to an arcsine transform before statistical analysis.
7. Participants were more likely to spatially modulate action (.97 connected, .94 unconnected) and deictic (.93 connected, .84 unconnected) gestures than object gestures (.54 connected, .51 unconnected). This pattern is reminiscent of conventional sign languages, although less extreme. For example, in ASL it is unusual for a noun to be produced in a non-neutral location (such a sign would be marked). In contrast, verbs are routinely signed in non-neutral locations.
8. The right panel of Fig. 3 excludes the one participant in unconnected events who produced only one spatially modified gesture per vignette and thus had no opportunity to use coreference within vignettes.
9. Participants produced more coreferential gestures when describing connected than unconnected events for all three types of gestures: objects (.98 vs. .75), actions (.98 vs. .80), and deictics (1.00 vs. .90).
10. Two participants in the unconnected events condition were excluded from this analysis because each produced fewer than three object gestures; if these participants are included, the patterns seen in Fig. 4 remain unchanged.
11. The right panel of Fig. 5 excludes the 1 participant in unconnected events (leaving $N = 5$) and the 3 in connected events (leaving $N = 5$) who produced only one spatially modified gesture per vignette and thus had no opportunity to use coreference within vignettes.
12. If coreference is a resilient property of language, home signers ought to display it. We know that home signers use space to indicate arguments in a proposition (Goldin-Meadow, Butcher, Mylander, & Dodge, 1994) and are exploring whether they also use space for coreference.

Acknowledgments

The research described in this article was supported by Grant R01 DC00491 from NIDCD to SGM, a Spencer/National Academy of Education Postdoctoral Fellowship to MC, an NIH Institutional Training Grant to the Committee on Human Development at the University of Chicago, and ROLE Grant REC 0337360 from NSF.

We thank Nelly Li for assistance in coding and Stephani Foraker for helpful comments on the manuscript.

References

Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2003). Classifier constructions and morphology in two sign languages. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 53–84). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes, 15,* 469–489.

Bosch, P. (1983). *Agreement and anaphora: A study of the roles of pronouns in discourse and syntax.* London: Academic.

Brentari, D. (1998). *A prosodic model of sign language phonology.* Cambridge, MA: MIT Press.

Casey, S. (2003). *"Agreement" in gestures and signed languages: The use of directionality to indicate referents involved in actions.* Unpublished doctoral dissertation, University of California, San Diego.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht, The Netherlands: Foris.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day "wild-child".* New York: Academic.

Dufour, R. (1993). *The use of gestures for communicative purpose: Can gestures become grammatical?* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Emmorey, K., Tversky, B., & Taylor, H. A. (2000). Using space to describe space: Perspectives in speech, sign, and gesture. *Spatial Cognition and Computation, 2,* 157–180.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language.* New York: Academic.

Gershkoff-Stowe, L., & Goldin-Meadow, S. (2002). Is there a natural order for expression semantic relations? *Cognitive Psychology, 45,* 375–412.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20,* 1–19.

Goldin-Meadow, S. (2003a). *Hearing gesture: How our hands help us think.* Cambridge, MA: Harvard University Press.

Goldin-Meadow, S. (2003b). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language.* New York: Psychology Press.

Goldin-Meadow, S., Butcher, C., Mylander, C., & Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: What's in a name? *Cognitive Psychology, 27,* 259–319.

Goldin-Meadow, S., & Feldman, H. (1977, July 22). The development of language-like communication without a language model. *Science, 197,* 401–403.

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review, 103,* 34–55.

Goldin-Meadow, S., Mylander, C., & Butcher, C. (1995). The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition, 56,* 195–262.

Haviland, J. (2000). Pointing, gesture spaces, and mental maps. In D. McNeill (Ed.), *Language and gesture* (pp. 13–46). New York: Cambridge University Press.

Iverson, J. M. (1999). How to get to the cafeteria: Gesture and speech in blind and sighted children's spatial descriptions. *Developmental Psychology, 35,* 1132–1142.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford, England: Oxford University Press.

Kegl, J., & Iwata, G. (1989). Lenguaje de Signos Nicaragüense: A pidgin sheds light on the "Creole?" ASL. In R. Carlson, S. Delancey, S. Gilden, D. Payne, & A. Saxena (Eds.), *Proceedings of the Fourth Annual Meeting of the Pacific Linguistics Conference* (pp. 266–294). Eugene: University of Oregon, Department of Linguistics.

Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177–210). Elmsford, NY: Pergamon.

Kendon, A. (2004). *Gesture: Visible action as utterance.* Cambridge, England: Cambridge University Press.

Klima, E., & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard University Press.

Liddell, S. (1980). *American Sign Language syntax.* The Hague, Netherlands: Mouton.

Liddell, S. (1984). THINK and BELIEVE: Sequentiality in American Sign Language. *Language, 60,* 372–392.

Lillo-Martin, D. (1991). *Universal grammar and American Sign Language: Setting the null argument parameters.* Dordrecht, The Netherlands: Kluwer Academic.

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

McNeill, D. (2000). Catchments and contexts: Non-modular factors in speech and gesture production. In D. McNeill (Ed.), *Language and gesture* (pp. 312–328). New York: Cambridge University Press.

Neidle, C., Kegl, J., McLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American Sign Language: Functional categories and hierarchical structure.* Cambridge, MA: MIT Press.

Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning.* Cambridge, MA: MIT Press.

Nunez, R., & Lakoff, G. (1998). What did Weierstrass really define? The cognitive structure of natural and ε–δ continuity. *Mathematics Cognition, 4,* 85–101.

Newport, E. L. (1991). Contrasting conceptions of the critical period for language. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 11–130). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Padden, C. (1981). Some arguments for syntactic patterning in American Sign Language. *Sign Language Studies, 32,* 237–259.

Padden, C. (1988). *Interaction of morphology and syntax in American Sign Language.* NY: Garland.

Perlmutter, D. (1992). Sonority and syllable structure in American Sign Language. *Linguistic Inquiry, 23,* 407–442.

Polich, L. (2001). Education of the deaf in Nicaragua. *Journal of Deaf Studies and Deaf Education, 6,* 315–326.

Sandler, W. (1989). *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language.* Dordrecht, The Netherlands: Foris.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of systematic grammatical structure in a new language. *Proceedings of the National Academy of Science, 102,* 2661–2665.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science, 12,* 323–328.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development, 18,* 511–531.

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997) Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. *Proceedings of Boston University Child Language Development, 21,* 550–561.

Senghas, A., Özyürek, A., & Kita, S. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science, 305,* 1779–1782.

Singleton, J. L., Morford, J. P., & Goldin-Meadow, S. (1993). Once is not enough: Standards of well-formedness in manual communication created over three different timespans. *Language, 69,* 683–715.

Stokoe, W. C. (1960). Sign language structure: An outline of the visual communications systems. *Studies in Linguistics, Occasional Papers, 8.*

Supalla, T. (1982). Structure and acquisition of verbs of motion and location in American Sign Language. Unpublished doctoral dissertation, University of California, San Diego.

Supalla, T. (1995). An implicational hierarchy in verb agreement in American Sign Language. Unpublished manuscript, University of Rochester, Rochester, NY.