The Seeds of Spatial Grammar: Spatial Modulation and Coreference in Homesigning and Hearing Adults

Marie Coppola and Wing Chee So
University of Chicago

1. Introduction

Signed and spoken languages, while articulated in very different modalities, are both structured at multiple levels of linguistic analysis (Emmorey, 2002). While many of the same grammatical devices are found in both modalities (e.g., word order), the manual modality of sign languages allows its users to modulate signs in space, a grammatical device not possible in the oral modality. Indeed, sign languages universally exploit space to express grammatical relations (Supalla, 1995). Sign languages use space not only at the level of individual sentences, but also across sentences to maintain discourse coherence. Space is used to refer back to previously mentioned referents (Liddell, 1980) – to establish co-reference, a 'core' property of human language (Jackendoff, 2002) typically accomplished through pronouns in spoken languages (e.g., in the sentence, "Sally always wins when she enters," the pronoun "she" refers back to, and is thus coreferential with, Sally).

The example from American Sign Language (ASL) in Figure 1 illustrates how space can be used to maintain coreference. The top panel depicts the verb 'give' moving from the signer to a location on his right. The bottom panel depicts the same verb 'give' but moving from the signer's right to his chest. In both sentences, the verb's movement conveys grammatical information – the entity associated with the sign's starting point is the agent/subject of the 'give' event, and the entity associated with the sign's endpoint is the recipient/object. Thus, the top sentence means "I give to him;" the bottom sentence means "He gives to me." Once the signer has associated an entity with a particular location, that location can be re-used in subsequent sentences to refer back to the entity. Thus, after having produced the top sentence in Figure 1, the signer's re-use of the location to his right in the bottom sentence serves to refer back to the entity previously associated with that location ('him'). However, because this location is the starting point of the verb 'give,' the entity ('he') is now the agent rather than the recipient of the sentence.

I give to him.

He gives to me.

Figure 1. Examples of spatial verb agreement and coreference in ASL.

Coreference seems to come so easily to the manual modality that we might expect it to be inevitable in the manual modality. One way to explore this hypothesis is to force speakers who do not know sign language to use their hands, rather than their mouths, to communicate. When asked to describe scenes without speaking, hearing adults who have had no experience with a sign language nevertheless produce a variety of language-like patterns in their gestures. They produce discrete gestures to refer to the elements in the scene, and they combine these gestures to form gesture strings (Goldin-Meadow, McNeill & Singleton, 1996). These gesture strings display regular ordering patterns that are not those of English (Gershkoff-Stowe & Goldin-Meadow, 2000). Do speakers-turned-signers also use space to maintain coreference?

Hearing adults asked to gesture without speech have been found to produce gestures in non-neutral locations (i.e., locations not directly in front of the gesturer's chest, as in Figure 1) when describing individual scenes (Casey, 2003) and re-telling stories (Dufour, 1993). The question, however, is whether these gesturers were able to re-use a spatial location to refer back to an entity across the discourse. In previous work, we asked hearing adults to describe a series of vignettes using speech in one condition and using only their hands in the other condition (So, Coppola, Licciardello & Goldin-Meadow, 2005). We found that the hearing gesturers did indeed use space to maintain coreference when asked to create gestures on the spot, more often than when producing gestures along with speech.

The hearing adults in these studies did not have a language model in which space is used to maintain coreference. They did, however, have a model for coreference in their spoken language (English). The question we address in this paper is whether coreference is so central to language that it will be incorporated into a communication system even if the communicator does not have a model for the property. To address this question, we need to examine individuals who have not been exposed to a conventional language model, signed or spoken.

The participants in our study were deaf, with hearing losses so severe that they could not acquire the spoken language that surrounded them. In addition, they had been born to hearing parents who had not exposed them to a

conventional sign language. Individuals in these circumstances use gestures, often called "homesigns," to communicate and those gestures display many of the properties of natural language. Extensive work by Goldin-Meadow and her colleagues has found that young homesigners, ages 2 to 5 years, can invent gesture systems that are language-like in many respects – the gestures have word-level structure, sentence-level structure, narrative structure, noun and verb categories, and can be used to describe the non-here-and-now as to make generic statements (Goldin-Meadow, 2003b). The homesigning children have been observed to use space systematically to mark an entity's semantic role (Goldin-Meadow, Mylander, Butcher & Dodge, 1994), but we do not yet know whether they use space to maintain coreference.

Adults who have been using their homesign systems for their entire lives, despite their continued lack of a conventional language model, might well have discovered that space can be used coreferentially. Adult homesigners are able to use spatial devices to mark grammatical contrasts (Coppola 2002; Coppola, Newport, Senghas & Supalla, 1997) and have been observed to modulate their gestures in space and make use of a variety of deictic devices (Coppola & So, 2005). Thus, we might reasonably expect homesigning adults to use space for coreference. To find out, we asked four deaf adults in Nicaragua who had not been exposed to Nicaraguan Sign Language (NSL) to describe a set of vignettes, and compared the homesigns they produced to the gestures created by four hearing adults asked to describe these same vignettes in our previous study (So et al., 2005).

2. Method
2.1. Participants

Four Nicaraguan homesigners (ages 18, 22, 27, and 27 years) participated in the study. At the time of data collection, none of them knew each other. They are all congenitally and profoundly deaf and had not acquired either a spoken language (due to their deafness) or a conventional community sign language (due to their lack of exposure to one)[1]. All four displayed extremely limited production and comprehension of spoken Spanish. They had had little to no formal education and had not received either hearing aids or oral instruction. None knew NSL, the language of the Deaf community in and around Managua, the capital. They communicated using a gesture system developed within the family. Their hearing family members gestured with them to varying degrees. Each homesigner had at least one person (a parent, sibling, or friend) who was fairly fluent in his or her gesture system, and with whom he or she gestured

---

1. Homesigners 1 and 4 have never met any users of Nicaraguan Sign Language (NSL). Homesigner 2, as an adult, had brief contact with two users of NSL, but has no NSL communication partners in his daily life. In adulthood, Homesigner 3 has occasionally visited the Deaf association in Managua but has not acquired even common NSL signs.

regularly. They had each been using their homesign system as their primary means of communication for their entire lives.

Data from four hearing adults who participated in the So et al. (2005) study, all undergraduates at the University of Chicago, were used as a comparative base. All four were native English speakers and naïve to sign languages. We selected the four participants who saw the same stimuli as our deaf participants and who had produced a total number of gestures closest to the mean.

2.2. Stimuli

The stimuli were 11 videotaped vignettes, each lasting 1 to 3 sec, from a single scene of a silent film featuring Charlie Chaplin (Figure 2 and Table 1). The vignettes primarily featured two characters participating in a variety of motion events that varied in number and type of elements (e.g., 1 argument: man falls; 2 arguments: man kisses woman; 3 arguments: man gives woman a basket).

Figure 2. An example of one vignette, "Man gives basket to woman."

Table 1. Vignettes making up the story.

Vignette number	Description of event	Vignette number	Description of event
1	Cat sits next to flowerpot on windowsill	6	Man grabs woman's hand
2	Man exits car	7	Man kisses woman's hand
3	Man leads woman under arch	8	Woman walks up stairs
4	Man doffs hat to woman	9	Woman enters room
5	Man gives basket to woman	10	Cat knocks flowerpot off windowsill
		11	Flowerpot hits man's head; man falls down

2.3. Procedure

The vignettes were presented to participants on a computer screen. Each participant first viewed all 11 vignettes sequentially to get a sense of the story. Participants then watched the vignettes one at a time; at the end of each vignette they were asked to describe it. Hearing Adults were instructed to use only their hands, and not their voices, to describe the vignette to an experimenter. Homesigning Adults gestured to a communicative partner who was one of his or

her main communication partners in everyday life (a sibling or a friend with whom they gestured frequently). All gestures were videotaped and transcribed.

2.4.   Coding
2.4.1. Gesture types

Data for participants in both groups were analyzed in the same manner. We analyzed the first response produced by the participant, and assigned each gesture a meaning with reference to the objects and events that appeared in the vignette. We then classified each gesture into one of the following Gesture Types[2]: *Act* gestures referred to actions (e.g., the hands move away from the body as if giving); *Entity* gestures referred to people or objects (e.g., the hands trace a moustache to refer to Charlie); and *Point* gestures indicated real objects or spatial locations (e.g., a point to a nearby flowerpot or to a location above one's head).

Participants sometimes used actions typically performed by an entity, actions typically performed on an entity, or attributes of an entity to refer to that entity. In these cases, we relied on the vignette itself to assign meaning. Consider the vignette in which a cat sits on a windowsill, with no human present, and no petting action taking place. If a participant described this vignette using a petting gesture, we classified the "petting" gesture as an Entity gesture for the cat simply because no petting action occurred. If a participant produced a gesture referring to a characteristic of an entity (such as "moustache" to refer to Charlie) and produced no other gestures that could plausibly refer to Charlie, we classified the "moustache" gesture as an Entity gesture for Charlie.

2.4.2. Spatial modulation and coreference

To determine whether participants were able to establish a spatial framework and use their gestures coreferentially, we first identified gestures that were spatially modulated. Gestures were considered *spatially modulated* if they: (1) were produced in an area not directly in front of the participant's chest; (2) moved away from or toward the participant; or (3) were used to associate a spatial location with a gesture. For example, the gesture for woman (G2) in Figure 3 is spatially modulated because it is produced away from the chest area.

A spatially modulated gesture was considered *coreferential* if the location used in the gesture was re-used to refer to the same entity in a later utterance

---

2. The participants produced some gestures that did not fit into one of these 3 categories (15% for both groups): gestures referring to an *Attribute* of an entity (e.g., size or shape), an *Aspect* of an event, or the *Number* of objects. Each type was produced less than 4% of the time, with the exception of *Aspect* gestures which were produced only by Homesigners (8%). In addition, the Hearing adults produced 20 (10%) *Don't-know* gestures (shrug, palm-up flip) which were excluded from further analyses.

describing that vignette[3]. Consecutive identical repetitions of a form were not counted. The response in Figure 3, produced by a Hearing Adult to describe the vignette in which Charlie gives the woman a basket, illustrates spatial modulation and coreference:

Figure 3. Spatially modulated and coreferential gestures. The square represents the location established for the woman; the circle represents the location established for the basket. G3 (an Act gesture) is coreferential with G2 (an Entity gesture). G4 (an Entity gesture) is coreferential with G3. G5 (an Act gesture) is coreferential with G2, 3, and 4.

2.5.   Results
2.5.1. Gesture Type

---

3. Young sign languages, such as NSL and Al-Sayyid Bedouin Sign Language (ABSL) also use spatial modulations. ABSL signers spatially modulate signs to contrast semantic roles (Padden, Meir, Sandler & Aronoff, 2005), but no studies of coreference over a discourse have been reported. Senghas & Coppola (2001) showed that early learners from Cohort 2 of NSL, but not those of Cohort 1, who provided their language input, used spatial modulations specifically for coreference. Their coding allowed multiple spatial modulations serving different functions (e.g., person, location) on a single verb, and calculated coreference based on the number of spatial modulations per verb. The present study did not consider the various functions of spatial modulations, and treated all modulations the same way. A more detailed analysis may reveal other differences between Homesigners vs. Hearing Adults, and between these two groups and Nicaraguan signers.

Homesigning Adults produced 289 gestures compared with 185 for Hearing Adults (t(6)=1.92, *ns*). A one-way ANOVA showed that the proportions of Act (F(1, 6))=1.32, *ns*), Entity (F(1, 6))=0.13, *ns*), and Point (F(1, 6))=0.67, *ns*) gestures produced by the two groups also did not differ.
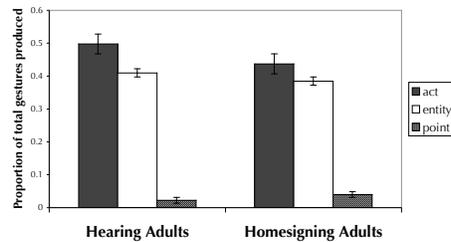

Figure 4. Gesture types.

2.5.2. Spatial modulation and coreference in Act gestures

Figure 5A (left graphs) presents the proportion of Act gestures that were spatially modulated and coreferential for both groups. Hearing Adults produced a larger proportion of spatially modulated Act gestures (.98) than did Homesigning Adults (.80) (t(6)=2.58, *p*<.05). Of these spatially modulated gestures, almost all were used coreferentially by both groups (Hearing Adults, 1.00 and Homesigning Adults, .95, (t(6)=1.73, *ns*).

2.5.3. Spatial modulation and coreference in Entity gestures

Figure 5B (right graphs) presents the proportion of Entity gestures that were spatially modulated and coreferential. Here the difference between groups was more pronounced: The Homesigning Adults produced significantly fewer spatially modulated Entity gestures (.13) than did the Hearing Adults (.66), (t(6)=2.79), *p*<.05. However, when participants in both groups used a spatially modulated Entity gesture, they used that gesture coreferentially all the time.
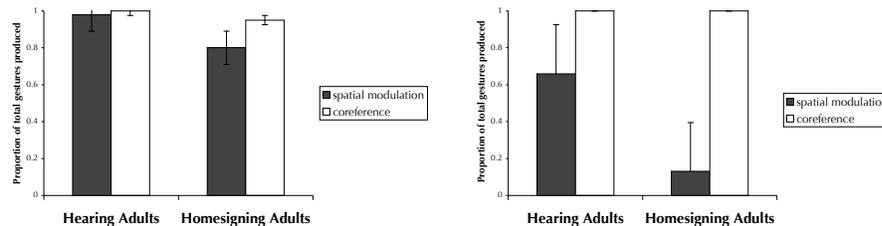

Figure 5. Spatial modulation and coreference in Act gestures (A, left graphs) and Entity gestures (B, right graphs).

3.   Discussion

The Hearing Adults, when asked to describe a scene using only gesture, as well as Homesigners, both spatially modulated the Act gestures they produced, and used those gesture coreferentially; that is, they routinely placed their act gestures in non-neutral locations, and then re-used those locations to indicate a previously mentioned referent. These findings suggest that it is not necessary to be exposed to a language model containing coreference in order to construct a communication system that makes use of coreference.

3.1.1. Does more space reflect more sophisticated structure?

Whereas the Hearing Adults spatially modulated Act gestures as often as Entity gestures, the Homesigners did not. However, when they spatially modulated a gesture, both groups used it coreferentially, whether it was an Act or Entity gesture. Why did the Hearing Adults use spatial modulation for both Act and Entity gestures while the Homesigners restricted their use to Act gestures? Figure 6 displays a response that was typical of the descriptions produced by the Hearing Adults in response to the vignette containing a cat sitting next to a flowerpot on a windowsill[4].

WINDOW	PETTING	WALKING
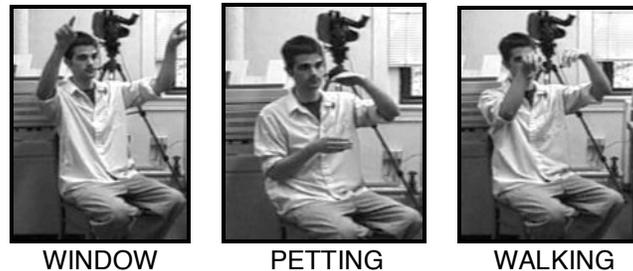
Figure 6. A Hearing Adult describing "A cat sits on a windowsill."

Two related aspects of Hearing Adults' responses distinguish them from those of Homesigning Adults. First, Hearing Adults tend to conflate gestures referring to objects (Entity gestures) and events (Act gestures). The communication of events, especially the simple physical events of the stimulus materials used in the current study, can relatively easily be conveyed by Act gestures that are very similar in form to the actual action in the vignette. However, because they do not have experience in communicating via gesture alone, without their voices, Hearing Adults do not readily produce gestures referring to objects. Rather, they tend to refer to objects (especially those in the role of patient) by invoking typical actions on them, such as the '*petting*' gesture

---

4. Because participants saw the entire story before responding, and knew that that the cat would knock over the flowerpot, they sometimes described the cat walking on the windowsill at the beginning of the story.

in Figure 6 produced to refer to the cat. Note that this petting gesture is ambiguous, unless one already knows the content of the vignette that the hearing participant intends to convey. Thus, the Hearing Adults' responses are characterized by ambiguous forms with multiple possible meanings: *"petting the cat"* or *"cat on windowsill"?*

The lack of lexical items in the gestural repertoire of the Hearing Adults, and their lack of experience in using gesture alone to communicate, leads them to produce pictorial, imagistic representations of events that are undifferentiated with respect to the discrete word classes typically found in more developed languages [5]. Hearing Adults also fail to distinguish gestures referring to Entities and Acts in terms of their spatial marking. In general, they do not observe restrictions on where either type of gesture can be produced in space. Counterintuitively, then, Hearing Adults' increased use of space conveys less, not more, information, and shows less, not more, language-like structure.

3.1.2. How far can Homesigners go in grammaticizing space?

Unlike the spontaneous gestures created by hearing adults, sign languages like ASL constrain form in a variety of domains: verb classes that vary in their patterns of obligatory spatial agreement, and in which semantic and grammatical roles do not always match (e.g., "backwards verbs" (Padden, 1988)); classifier constructions expressing verbs of motion and location (Supalla, 1986); and morphology that distinguishes nouns from verbs (Supalla & Newport, 1978). In ASL, some nouns and verbs share the same underlying form, which may clearly be derived from the typical path or movement of an event, or from a typical act on it, as in the Hearing Adults' productions. However, what distinguishes ASL from the gestures produced on the spot by Hearing Adults is the obligatory derivational morphology on nouns that have their origins in verbs.

While we don't expect Adult Homesigners, who are developing their gesture systems in the course of a single lifetime, without a language model, to show evidence of *all* the constraints shown by sign languages, these data indicate the early stages of a few. From the perspective of a developed sign language like ASL, Homesigners move beyond the pictorial representations of the Hearing Adults in the following ways: They distinguish gestures referring to Acts and Entities in terms of where they can be produced (Act gestures are relatively unconstrained, while Entity gestures are highly constrained).

What forms do these Entity gestures take? Do they display characteristics of the classifier constructions of sign languages? Classifier constructions are characterized by a reduced set of handshapes that correspond to semantic classes,

---

5. Note, however, that Hearing Adults' responses are not generally mimetic; they do produce sequential, discrete gestures, as has been shown in previous work (e.g., Goldin-Meadow, McNeill & Singleton, 1996; Gershkoff-Stowe & Goldin-Meadow, 2000). Here we highlight the fact that the gestures themselves are undifferentiated in terms of the relative roles being expressed.

restricted movements, and constraints on the combination of handshapes and movements across the two hands (Supalla, 1986). The example in Figure 7, describing the cat sitting on the windowsill, suggests that Adult Homesigners show inklings of the classifier systems found in established sign languages:

WHISKERS   SMALL-ANIMAL+ON-SURFACE

Figure 7. A Homesigning Adult (left) describing "A cat sits on a windowsill."

Homesigners' development of conventionalized lexical items and forms that refer to semantic classes removes the ambiguity prevalent in the hearing adults' gestures. WHISKERS is a lexicalized form that this participant consistently uses to refer to cats. In the two-handed form used to express "small-animal-on-surface," the left hand shows the surface of the windowsill, while the right hand, positioned above it, uses a handshape that this homesigner reliably uses in other contexts to refer to small animals like cats.

Hudson and Eigsti (2003) proposed the Lexical Competence Hypothesis based on a well-controlled laboratory study of pidgin language creation. They asked native English speakers to communicate with each other in a variety of tasks using only Modern Persian lexical items. The participants who knew more lexical items produced more complex grammatical structure. In accord with this hypothesis, having developed a repertoire of lexical items allows homesigners to focus on developing grammatical structure at other levels.

3.1.3. Conclusion

This study has a few limitations. First, we did not analyze the order that gestures were produced. Hearing Adults (Goldin-Meadow, 2003a) and Homesigning Adults (Coppola, Senghas, Newport & Supalla, 1997) use the order of their gestures to convey the relations of participants to the event, but little is known about the relationship between spatial devices and word order in gesture-alone discourse. Second, we compared native English-speaking American Hearing Adults to Nicaraguan Homesigning Adults; perhaps cultural effects are influencing our results. Homesigning Adults gestured to a friend or sibling, who often responded, whereas the Hearing Adults gestured to an

experimenter who did not respond to them. Finally, the current analyses are not directly comparable to other studies of coreference in emerging sign languages.

Future work will address these issues. Data have been collected from native Spanish-speaking Nicaraguan participants, providing a within-culture comparison that better controls for potential cultural effects. In addition, we will more closely examine the spatially modulated Entity gestures produced by participants in both groups, and expand our coding scheme to better account for potentially complex, two-handed productions that may display classifier-like characteristics. We can supplement analyses of these more narrative-like data with systematic data on the expressions of such complex motion events involving different types of elements. With these data, we can also address the issues of specificity in spatial modulation (e.g., verb classes that differentiate the grammatical functions of agreement; classifier constructions) and the relationship of spatial modulation to word order.

To conclude, these data show that all learners have "seeds" of spatial structure within them. They have the capacity to develop some degree of systematicity in the manual modality, even when such systematicity is not obligatory or present in their input, and regardless of the time spent "on task." Within a very short time, i.e., an hour-long session in the laboratory, Hearing Adults can develop limited spatial structure evidenced by their systematic use of the same spatial location to refer to the same entity. However, we observed strong effects of the amount of time using gesture alone as a means of communication. Homesigning Adults, who habitually use gestures over time, develop more language-like structure than Hearing Adults. We argue that Homesigning Adults' constraints on the spatial marking of Entity gestures represents a step toward the systems of spatial grammar found in established sign languages like ASL.

References

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

Casey, Shannon (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.

Coppola, Marie (2002). The emergence of the grammatical category of Subject in home sign: Evidence from family-based gesture systems in Nicaragua. Unpublished doctoral dissertation, University of Rochester, Rochester, NY.

Coppola, Marie, Senghas, Ann, Newport, Elissa L. & Supalla, Ted (1997). The emergence of grammar: Evidence from family-based sign systems in Nicaragua. Presented at the Boston University Conference on Language Development.

Coppola, Marie, & So, Wing Chee (2005). Abstract and Object-Anchored Deixis: Pointing and spatial layout in adult homesign systems in Nicaragua. *Proceedings of the Boston University Conference on Language Development 29*, 144-155. Boston: Cascadilla Press.

Coppola, Marie & Newport, Elissa L. (2005). Grammatical Subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences 102*, 19249-19253.

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

Emmorey, Karen 2002. *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Erlbaum.

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

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

Goldin-Meadow, Susan (2003b). *The resilience of language.* NY: Psychology Press.

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

Goldin-Meadow, Susan, Mylander, Carolyn, Butcher, Cynthia & Dodge, Mark (1985). Nouns and verbs in a self-styled gesture system: What's in a name? *Cognitive Psychology 27,* 259-319.

Hudson, Carla L. & Eigsti, Inge M. (2003). Lexical competency, the overt expression of grammatical structure, and contrasts between pidgins and creoles. *Journal of Pidgin and Creole Languages.*

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

Liddell, Scott K. 1980. *American Sign Language Syntax.* The Hague: Mouton.

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

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

Padden, Carol, Meir, Irit, Sandler, Wendy & Aronoff, Mark (2005). Against all expectations: Encoding subjects and objects in a new language. Presented at the International Association for the Study of Child Language, Berlin, Germany, July.

Senghas, Ann, Coppola, Marie, Newport, Elissa L. & Supalla, Ted (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. *Proceedings of the Boston University Conference on Language Development 21*, 550-561. Boston: Cascadilla Press.

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

So, Wing Chee, Coppola, Marie, Licciardello, Vincent & Goldin-Meadow, Susan (2005). The seeds of spatial grammar in the manual modality. *Cog. Sci. 29*, 1029-1043.

Supalla, Ted (1986). The classifier system in American Sign Language. In C. Craig (ed.), *Noun Classes and Categorization: Typological Studies in Language, 7*. Philadelphia: John Benjamins, 181-214.

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

Supalla, Ted & Newport, Elissa L. (1978). How many seats in a chair? The derivation of nouns and verbs in American Sign Language. In Patricia Siple (ed.), *Understanding language through sign language research,* 91-132. NY: Academic Press.