The Role of a Communication Partner in the Creation of a Gestural Language System

Lisa Gershkoff-Stowe & Susan Goldin-Meadow
University of Chicago

1. Introduction

Language genesis, like language change, is a social phenomenon arising, by most psycholinguistic accounts, from the persistent need to communicate with other human beings. For people to communicate, they must establish agreement as to the basic form and meaning of their shared language (Pinker and Bloom, 1990). The question of whether linguistic convergence results from innate grammatical knowledge (Pinker, 1994), functional processing constraints (Bates & MacWhinney, 1987; Studdert-Kennedy, 1991), self-organizing processes (Harris, 1991), or the pragmatics of social relations (Givón, 1979) remains intensely debated in the field.

In the present study we look to extrinsic factors as one potential source of linguistic structure, namely, the social influence of others. More specifically, we ask how the availability of an active language partner might shape the kinds of grammatical properties that emerge in a developing language system. We address this question in a controlled laboratory setting by presenting people with the task of having to communicate with each other but having no available language in common. This situation is created experimentally by instructing participants that they can use their hands and body to communicate, but not their mouths. In other words, they can gesture but not speak.

Previous research by Goldin-Meadow, McNeill, and Singleton (1996) employed a similar paradigm to explore aspects of formal language structure that might arise "online" in the spontaneous gestures of adults when asked to convey information without speech. However, in their study, the recipient of the information given by the gesturing subject was the experimenter, who provided neither feedback nor reciprocating linguistic input. Within this experimental context, Goldin-Meadow et al. found that subjects readily produced gestures that were both compositional and segmental in manner. However, their gestures did not form an integrated system of internally consistent and contrastive symbols. Rather, they relied on gestures that were most likely to capture some salient or distinctive property of the thing they wished to represent. Such gestures, therefore, were highly iconic or transparent in form, independently created to represent a particular referent, with little regard to how they fit into a single, coherent system of language.

The primary advantage of using icons to convey information is that they can be interpreted by others without previous agreement; general knowledge of the world is sufficient for shared understanding (Barber & Peters, 1992).

However, within a communicative context, there is likely to be continual pressure to find ways to stylize and condense one's gestures so that they become fast and easy to produce. At the same time, however, there remains competing pressure to use only the most transparent gestures -- gestures that are easy to understand and easy to learn (Slobin, 1977).

This study considers how both competing forces, together, might influence the development of a formal linguistic system. Accordingly, we test the hypothesis that the creation of a gesture language system is importantly shaped by the availability of an active communication partner -- one who both receives *and provides* input to another person. More specifically, we ask whether the participants in our study will (1) establish a systematic way of combining gestures and thus create an elementary syntax, and (2) develop a consistent and stable lexicon of gestures that reveals internal standards of form.

2. Method

We assigned participants to one of two experimental conditions: a Feedback condition and a No-Feedback condition. In the Feedback condition, two adults were randomly paired (holding gender constant) and each person alternated in the role of Gesturer and Listener. In the No-Feedback condition, only one person assumed the role of Gesturer and there was no Listener other than the experimenter who simply recorded the subject's responses. There were eight adults in the No-Feedback condition and eight pairs, or 16 adults, in the Feedback condition. Males and females were represented equally across conditions. All participants were students recruited from the University of Chicago and had no prior experience with conventional sign language.

Adults in both conditions were asked to describe a series of very brief scenes on videotape, using only gesture and no speech. The forty scenes were selected from the Verbs of Motion Production test (Supalla, 1982) which consists of animated objects moving in varying paths and manner across space. Half the scenes contained only one moving object (e.g., a porcupine wandering across the screen) while the other half contained one moving object and one stationary object (e.g., a girl jumping through a hoop).

In the Feedback condition, each person took turns watching a block of 10 vignettes, one at a time, and immediately gesturing the scene to the Listener, who had no knowledge of its content. After the first set of 10 vignettes, the Listener switched roles and gestured the next set of 10. Each participant thus had two turns as the Gesturer and two as the Listener.

Also in the Feedback condition, Listeners were encouraged to provide the Gesturer with feedback indicating how well they understood what was being conveyed. That is, they could "ask" the Gesturer questions or acknowledge their understanding through gestures (but no speech) of their own. After the scene was described to the satisfaction of both participants, they were instructed to write down what they thought the scene was about. This allowed us to determine how well the Listener actually understood the information conveyed by the Gesturer.

In the No-Feedback condition, participants saw the same 40 vignettes but gestured them, one at a time, to an experimenter who offered no response other than to say "okay" before going on to the next vignette.

3. Results

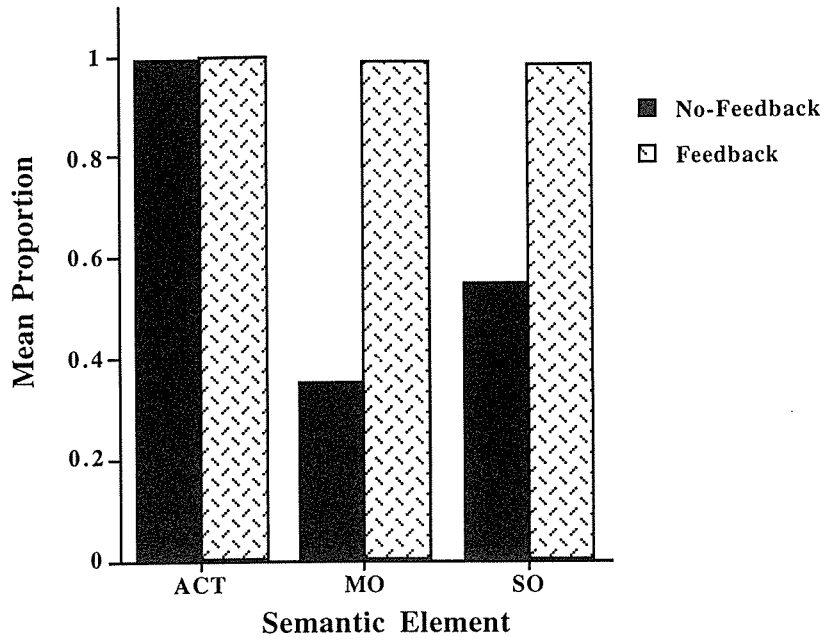
We found substantial differences in the gesture productions of the two groups. Table 1 presents the mean number of gesture tokens produced by subjects per vignette in both the Feedback and No-Feedback conditions and the number of different kinds of gestures they produced. Subjects with an active communication partner showed a tenfold increase in gesture frequency (mean=20) compared to those without a partner (mean=2). Moreover, as indicated by the number of different kinds of gestures produced, subjects in the Feedback condition did not simply repeat the same gesture over and over again but made several novel attempts to bring their meaning across; the mean number of gesture kinds was 8 in the Feedback condition and 2 in the No-Feedback condition. Also shown in Table 1 is the amount of time (in seconds) subjects spent gesturing (per vignette). Not surprisingly, subjects in the Feedback condition spent considerably more time describing each scene (mean=72) than subjects in the No-Feedback condition (mean=4).

TABLE 1. Number of Gesture Tokens and Kinds, and Time Spent Gesturing Per Vignette in the Two Conditions

	Feedback	No-Feedback
Gesture Tokens	20	2
Gesture Kinds	8	2
Time (sec)	72	4

We next compared how often subjects included each of the three semantic elements from the vignettes into their gestural descriptions, that is, the Action event (ACT), the Moving Object (MO), and the Stationary Object (SO). As shown in Figure 1, subjects in both conditions nearly always represented the action with a specific gesture: the mean proportion for both groups was 99%. Less consistent, however, was the proportion of vignettes in which subjects represented an object with a specific gesture. Only 35% of the Moving Objects were encoded by a separate gesture in the No-Feedback condition compared to almost 99% in the Feedback condition. Similarly, 55% of the Stationary Objects were lexicalized without feedback compared to 98% with feedback. Note, however, that while pragmatic factors increased gesture production in the Feedback condition relative to the No-Feedback condition, they were not key to motivating gestures for these thypes of semantic elements in the first place -- all three types of gestures were found in both conditions.

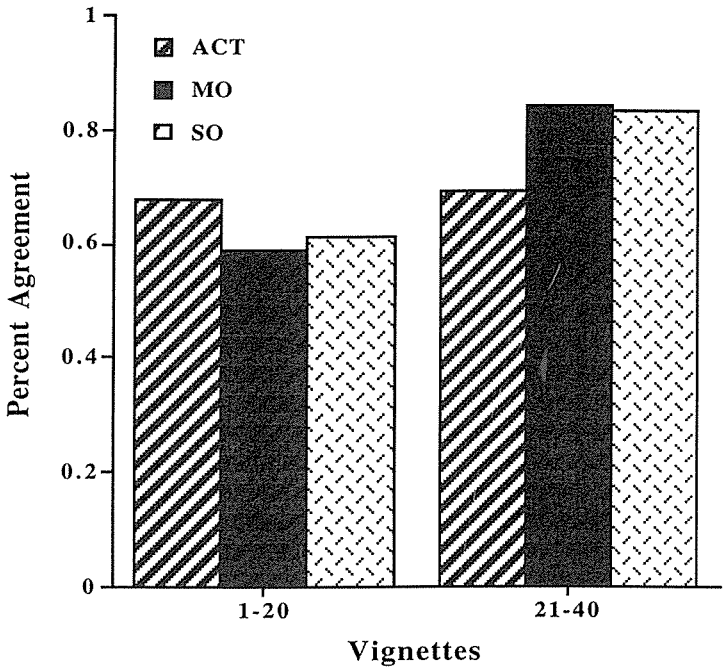
FIGURE 1. Type of Semantic Element Conveyed By Condition



Given these basic differences, then, it is important to ask how well Listeners understood the meaning of their partner's gestures. We can ask this question only for subjects in the Feedback condition, however, since, by definition, subjects in the No-Feedback condition did not have a partner receiving their gestures. Accordingly, Figure 2 shows the mean percent agreement between the Gesturer and the Listener in the Feedback condition on the basis of their written descriptions of each vignette. The data are presented separately for the Moving Object (MO), Stationary Object (SO), and Action (ACT) and are compared across time for the first and second set of 20 vignettes.

As this figure shows, overall, the Gesturers were quite successful in getting their message across in the manual modality. Moreover, Listener comprehension, at least for objects, improved over time: from 60% combined agreement for the Moving and Stationary Objects in the first set of vignettes, to 84% agreement in the second set. Surprisingly, however, comprehension of gestures that encode actions remained essentially unchanged at 68% agreement.

FIGURE 2. Listener Comprehension in Feedback Condition Based on Written Descriptions



So far we have considered how successful subjects were at conveying information about the identity of referents in the experimental vignettes. Yet for gesture to take on the form and function of language, it also must have some formal way of specifying the relations between those referents. In this regard, the Listeners in the Feedback condition had virtually no problem understanding their partners. Of the entire corpus of 320 vignettes, we found only three instances of confusion on the part of the Listener concerning "who did what to whom." The fact that errors in the interpretation of relations were extremely rare, at least in the Feedback condition, suggests that subjects had some systematic means for expressing the relation between elements within a sentence. As a first step, toward discovering that means, we asked whether subjects relied on a regular gesture order and, if so, whether it corresponded to canonical English word order.

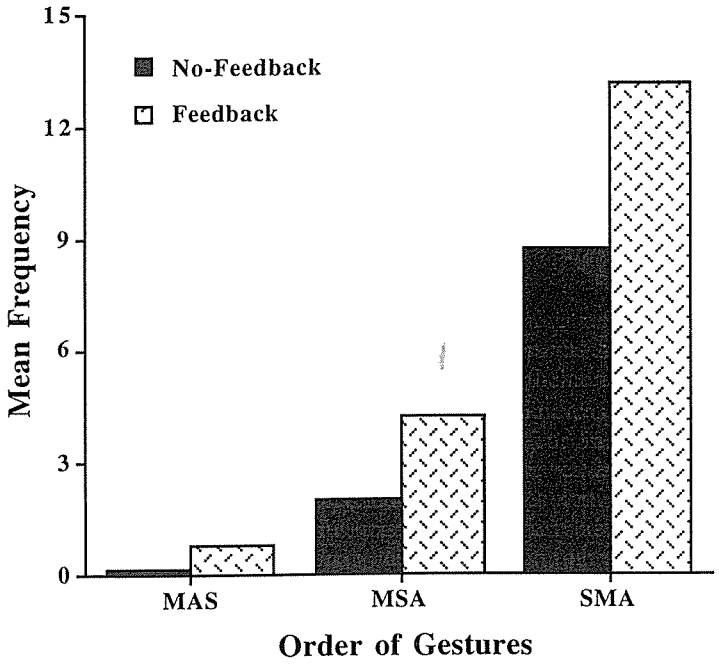
To address this question, we examined how often subjects placed each of the three semantic elements from the vignettes in a given position within a string of gestures. We did the analyses separately for vignettes containing one vs. two objects. Turning first to one-object vignettes -- that is, the 20 vignettes containing only a Moving Object and an Action -- we found that when subjects

used a separate gesture to name the Moving Object, it always came before the gesture for theAction. This was true for all subjects, regardless of condition. With respect to the set of 20 two-object vignettes -- those containing a Moving Object, an Action, and a Stationary Object -- a number of additional gestural orders were possible. Consequently, we found considerably more variability among subjects. Yet despite this increase in potnetial for variability, clear and consistent patterns were established by subjects in both groups.

We found that the subjects used onlythree of the six possible orderings of elements in both conditions (See Figure 3): (1) Moving Object-Action-Stationary Object (MAS) -- an arrangement corresponding to standard English word order, (2) Moving Object-Stationary Object-Action (MSA) and (3) Stationary Object-Moving Object-Action (SMA). As Figure 3 shows, there is a similar pattern of gestural production in the ordering of semantic elements regardless of condition. In particular, subjects in both groups demonstrated a strong preference for the SMA order (Feedback=13.13; No-Feedback=8.75) over the MSA order (Feedback=4.25; No-Feedback=2.0). Moreover, they rarely used the MAS order in their gesture strings (Feedback=.75; No-Feedback=.13).

These results suggest that subjects were not simply translating their gestures into standard English word order. Rather, they were creating their own distinct ordering of elements within a sentence. This is true for all subjects, regardless of whether they shared their gesture system with another person. It appears, then, that the particular structural regularities we find at the sentence level do not depend on either prior linguistic knowledge *or* a shared language system. Understanding the precise source of this structure, however, cannot be determined from the present findings. Rather, additional studies are needed to resolve this important issue.

FIGURE 3. Ordering of Semantic Elements for Two-Object Vignettes



As the vocabulary of individual subjects and subject pairs continued to grow, we next looked for possible structural regularities appearing at the *word* level. Specifically, we predicted that subjects in the Feedback condition would show evidence of a beginning morphology that would allow them to generate an unlimited number of signs using only a small set of gestures and some basic rules for combining them. We chose consistency of handshape as our dependent measure, determining to what extent subjects used the same handshape to represent a single class of objects. Five categories of objects were examined: people, animals, vehicles, airplanes, and trees. These five categories were represented by a variety of objects shown multiple times across the 40 vignettes. For example, within the category of airplanes, subjects saw a biplane, a military plane, a paper plane, and a plane made out of Leggo parts.

We determined handshape consistency by calculating the number of opportunities subjects had to view exemplars from each of the five categories and the percent of those vignettes represented by a single handshape. For example, if a person used the same handshape to represent all of the different airplanes mentioned above, he or she received a score of 1 for that category. In

contrast, a person who was less consistent, who, for example, used three different handshapes to encode the four airplane exemplars would earn a mean consistency score of .33. This hypothetical calculation is detailed in Table 2.

TABLE 2. Calculating the Consistency Score for a Hypothetical Example

Category: Airplanes

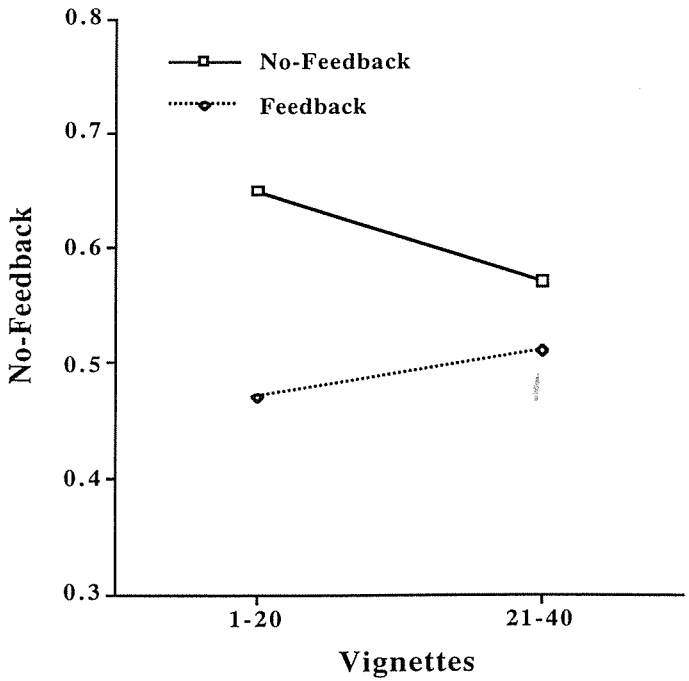
Exemplar	Handshape	Proportion
1. Biplane	X	X: 2/4 = .50
2. Military	X	
3. Paper	Y	Y: 1/4 = .25
4. Leggo	Z	Z: 1/4 = .25
		.33

Figure 4 presents the overall consistency score, based on this method of calculation, for subjects in the Feedback and No-Feedback condition, averaged across the five different object categories. In this figure, the data from each condition are compared as a function of time, that is, for the first set of 20 vignettes and for the second set. Contrary to our initial prediction, we found that subjects without a Listener were more consistent in their handshapes than subjects with a Listener. This result suggests that the availability of a language partner actually caused adults in our study to be less consistent, at least initially, rather than more consistent. But why should this be?

One interpretation is that the presence of a Listener exerts constant pressure on the Gesturer to be semantically clear -- that is, to only produce gestures that are transparent in form and thus easy to understand. The pressure to be clear might account for the variability in handshape we find early in the study for subjects in the Feedback condition. This variability may then serve as a catalyst for a reorganization into a more consistent language system. The sheer quantity and complexity of gestures may increase the demands placed on the Gesturer, thereby setting the stage for reorganization into a coherent system. If so, we might expect to see consistency emerge in the Gesturer's handshapes after a period of continued use.

Although we found subjects in the No-Feedback condition had higher consistency scores than those in the Feedback condition, this increased consistency did not in all cases reflect a system of contrasts. That is, subjects in the No-Feedback condition often used the same handshape indiscriminately, (i.e., the same handshape was used for planes, trees, vehicles, etc.)yielding high within-category consistency scores even though the handshape supplied little information. New measures thus are needed that can capture both within-category consistency and between-category variability. Ideally, such a measure would allow finer discriminations to be made across conditions.

FIGURE 4. Consistency of Handshape as a Function of Time



4. Conclusion

The primary aim of this research was to investigate the development of a spontaneous gestural language system by hearing adults who have no prior experience with conventional language under controlled experimental conditions. Specifically, we asked how the availability of a communication partner influences the lexical, syntactic, and morphological properties that emerge -- or fail to emerge -- in the language systems they create.

We found that the opportunity for negotiation and feedback between users of a language has important consequences for how individuals convey information about objects, their locations, and movement. At the lexical level, we found that when communication involves the two-way transmission of information, not only does the absolute number of gesture tokens increase rapidly but the number of different kinds of gestures increases as well. Thus the size of the gesture vocabulary that individuals produce together will be considerably larger than the vocabularies of individuals who do not enter into a shared language system.

Our conjecture is that this difference may have a direct bearing on the development of a productive morphology within such systems. That is, we suggest that, as the number of different gesture kinds amassed by individuals becomes too unwieldy to easily process, there should be a strong push to reorganize those gestures into an internally consistent and contrastive linguistic system. Prior to that point, however, a preference for semantic transparency may override concern for the establishment of internal standards of form.

Finally, at the syntactic level, we found an immediate tendency toward a stable gesture order regardless of whether or not the system was shared by another person. Moreover, the preferred order did not mirror the ordering pattern that the Gesturer commonly used in speech but was emergent in the act of gesturing itself. It may be, then, that aspects of the gestural modality, for example, the use of space or movement, are introducing structure into the gestures individuals produce.

The paradigm employed in this research allows us to trace the evolution of new linguistic forms as they emerge in a laboratory setting. Although this procedure is, by necessity, artificial in nature, what we hope to gain is the ability to control some of the factors that might be relevant to the process of language creation and language change. Additional studies are now in progress to test individual pairs of adults, as well as children, over a span of several months. Those data so far indicate increasing complexity of gesture form with continued expansion of time. We also are studying what effect the introduction of a third, novice gesturer has on an existing language system. Will the presence of a learner, rather than a creator of gesture, lead to greater demands for structural elaboration? These questions and many others should provide us with the basis for exploring the complex forces that shape the emergence of language.

Endnotes

* We thank Veronica Sansing for assistance in data collection and coding and David McNeill for helpful comments and discussion. This research was supported in part by NICDC Grant No. F32 DC00264 to L. Gershkoff-Stowe and NICDC Grant No. RO1 DC00491 to S. Goldin-Meadow.

References

Barber, E.J.W. & Peters, A.M.W. (1992). What child language and archaeology have to say to each other. In J.A. Hawkins and M. Gell- Mann (Eds.), *The evolution of human languages*. Proceedings of the Workshop on the evolution of human languages. CA: Addison-Wesley.

Bates, E. & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157-193). N.J.: Erlbaum.

Givón, T. (1979). *Discourse and syntax: Syntax and semantics*. N.Y.: Academic Press.

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.

Harris, Z.S. (1991). *A theory of language and information: A mathematical approach.* N.Y.: Oxford University Press.

Pinker, S. (1994). *The language instinct: How the mind creates language.* N.Y.: William Morrow.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707-784.

Slobin, D. (1977). Language change in childhood and history. In J. Macnamara (Ed.), *Language learning and thought.* N.Y.: Academic Press.

Studdert-Kennedy, M. (1991). Language development from an evolutionary perspective. In N.A. Krasnegor, D.M. Rumbaugh, R.L. Schiefelbusch, and M. Studdert-Kennedy (Eds.), *Biological and behavioral determinants of language development* (pp. 5-28). N.J.: Erlbaum.

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