The development of constructions through gesture use

Barbara F. Kelly
Stanford University

1. Introduction

Children's earliest utterances containing a verb can be considered a forerunner to the development of argument structure constructions. In these early utterances verbs such as *make*, *give, put, look*, and *want* are often used with nouns such as *ball* and *book* and pronouns such as *that, this* etc. When children combine verbs and nouns in these utterances, it is clear that they already have some notion of how the words go together and how the early constructions are likely to be construed by an interlocutor. However, researchers have not yet established where this knowledge comes from. Since it occurs early in their multi-word speech development it seems that young children's early gesture and gesture + speech patterns may be an important place to begin examining the basis for the acquisition of constructions. If children use gestures prior to the development of constructions then it is crucial that we look at these early communications as a potential locus for the early development of constructions. Analyzing constructions once children are already producing them does not account for how children *reach* the stage of using their first constructions. The question that remains is whether early construction development of English-learning children is influenced by their gesture and gesture + speech development. In the study reported here, I address this question through a novel analysis of children's construction acquisition, focusing on the role of gesture in children's earliest construction acquisition.

2. Background

Children's early utterances occurring as single-word speech are often combined with non-verbal communicative acts, such as pointing and reaching gestures (Bates 1976, Carter 1978, Greenfield & Smith 1976, Morford & Goldin-Meadow 1992). In combinations of speech and gesture, the two modalities may express associations between two different entities about which a child wants to communicate with relation to a single event (Butcher & Goldin-Meadow 2000) such as a child pointing toward a ball and uttering her name "Katie". Studies to date indicate a close association between gesture and speech in development. Although there has recently been a renewed interest in young children's gesture development and its role in communication, this research has not yet been linked to research examining children's argument structure development or the development of constructions. While we cannot examine differences in a child's intentions when they use only a gesture or a word versus a word plus gesture, we can examine the communicative affect of using two modalities rather than just one. A useful way to determine what information is being communicated is to look at the caregiver response to a child's communication and determine how caregivers treat children's gesture and gesture + speech communications.

In using more than one modality to communicate, children can communicate the same types of information in their multi-nodal communication that they later convey through two-word speech (Goldin-Meadow & Butcher 2003). This suggests that their multi-modal interactions may be a locus for learning about how they can combine communications about different elements such as actions and objects. Researchers across a range of different theoretical perspectives have shown that from their earliest utterances (and potentially earlier, see Gordon 2003) children have some knowledge of how to combine verbs and nouns that go with them (Hyams 1986, Pinker 1989, Tomasello 2003). However, research position differ in explanations regarding how children come to have such knowledge.

One current proposal that is particularly pertinent to construction development, is Tomasello's (1992, 2003) *verb-island hypothesis*. Tomasello suggests that children learn about constructions on the basis of their knowledge of verbs. He suggests that children learn the syntactic frame for verbs on a verb-by-verb basis. Only after learning several high frequency verbs do children then generalize from these to less frequent verbs. While Tomasello makes this claim only with respect to verbs, others have shown that this hypothesis regarding verb islands can be extended to other high frequency words, including pronouns such as *she* and *I*, and proper nouns such as *Mummy* and the child's name (Pine, Lieven & Rowland 1988).

The existing research on the acquisition of constructions together with the research in this volume, has gone a long way to explaining how children learn constructions across individual words and extend them to new uses. A fundamental remaining issue not addressed in the research to date relates to the origins of children's construction knowledge. If, from their earliest multi-word utterances, children have knowledge of some constructions, such as verb argument structure, where does this knowledge come from? Does it have its basis in children's innate linguistic knowledge? Is it cognitively motivated? Or is there another, altogether different explanation for the patterns found in children's early multi-word speech? More specifically, this study explores whether the early construction development of English-learning children influenced by their gesture and gesture + speech development. In order to approach this question two specific research questions will be addressed in this study:

i.   *Do caregiver responses differ across children's different gesture communications?*
ii.   *Does caregiver scaffolding in dialogic interaction play a role children's construction development?*

3. Method
3.1 Subjects
Data for this study are taken from the Very Young Children corpus (Bugenthal, Clancy, Kyratzis, Lerner, and Zimmerman 1999). The corpus is a collection of longitudinal video-taped data of spontaneous interactions of children aged about 12 months to 30 months in a daycare center. Recordings of five monolingual children were carried out at roughly weekly intervals. The five children focused on in this research are one boy, Brailey and four girls: Chera, Lette, Caitlin, and Fiona.

3.2 Data coding
Data were transcribed and coded for two communicative strategies used by a child to maintain or initiate a communication. These were gestures or vocalizations directed toward another individual. For each strategy coded this involved the child looking at the caregiver and then moving their gaze toward a target object and using the strategy. The data were coded for two distinct communicative strategies. These are:

- *Vocalizations* which may be words or proto-words (these are children's attempts at producing adult words, for example one child in the corpus uses [ma] to indicate that she wants a drink).
- *Gestures*, which are gestures directed at another individual – three gesture types were coded:
  i) POINT: a movement of the index finger or outstretched hand toward an object while focusing a simultaneous gaze on the object or a caregiver.
  ii) ATTENTION FOCUS: holding an object out while looking at the caregiver or holding it out and then throwing it down while looking at the caregiver.

iii) GIMME*:* holding out an open palm as if to receive an object while gaze is on the caregiver or the object.

Children's gestures and the subsequent caregiver responses were coded across each of the five children. For POINT and ATTENTION FOCUS gestures, a caregiver response was coded if a child used the gesture and the caregiver then altered the direction of her attention, for example, by moving her head and looking in the direction being indicated by the gesture. For a GIMME gesture a response was coded if the caregiver responded by giving or denying the child the target object. After having moved attention to the indicated direction the caregiver often acted upon the object indexed by the gesture. There were no instances in which a caregiver manipulated an object being indexed by a gesture without having first looked in the direction of that object. A non-response was coded if a caregiver ignored a child's gesture, either intentionally or because of distraction from another source.

4. Results and Discussion

Results indicate that children's gestures used both alone and in combination with speech play a role in the development of constructions for young English-learning children. Caregiver response differed consistently across children's different gesture communications. These responses were used by the child as a type of scaffolding in dialogic interaction prior to their use of multi-word speech. Results to the first research question are presented and discussed below.

4.1 Caregiver response across different gestures
In examining caregiver response to children's gestures, several issues arise both relating to the gestures used and the specific caregiver responses. If children receive caregiver response when they gesture, what types of elements are they indexing with those gestures, and what types of responses do specific gestures bring about? This is addressed in the first research question of the study: *Do caregiver responses differ across children's different gesture communications?*

Children's communicative gestures and the related caregiver responses are presented in Table 1 below. The interpretations of children's communicative gestures are presented in terms of both the caregiver's interpretation of the child's viewpoint and the caregiver's viewpoint regarding the action. I have coded the function of each gesture in two ways: 1) with the child as the agent or experiencer (*hypothetical child viewpoint*) and 2) with the analyst as agent (*caregiver viewpoint*). I want to be clear that when I use the term *child's viewpoint*, it is short for *potential interpretation of the child's viewpoint*. It is never based on an assumption about the child's actual viewpoint or the child's intentions in using a gesture, which we cannot know. We also cannot know the actual viewpoint of the caregiver. However we can observe her responses to the child's gestures (and the child's subsequent response to these responses) as a means of determining the caregiver viewpoint through her words and actions. The caregiver gesture interpretations and caregiver responses to children's gestures are presented together in the following table:

GESTURE	CAREGIVER INTERPRETATION		CAREGIVER RESPONSE
	hypothetical child viewpoint	*caregiver viewpoint*	
POINT	look	look	**look** at object 88% (n=97/111)
GIMME	want/get	give	**give** child object 79% (n=60/76)
ATTENTION FOCUS	show	look	**look** at object[1] 82% (n=42/52)
	give	take	**take** object 41% (n=21/52)

Table 1.   *Caregiver responses to child gestures*

The results in Table 1 indicate that when a child gestures, the caregiver overwhelmingly responds by carrying out an action, suggesting that children's early gestures serve as a means for eliciting action. Overwhelmingly, when children gesture caregivers undertake some kind of action. Children come to learn that the use of different gestures results in different outcomes in terms of caregiver action response. Caregivers consistently respond differently to different arm and hand configurations and this variance in response then socializes children into using these different gestures to obtain different outcomes. The variance across the different gestures is further reinforced by the consistency in response across a single gesture over time. For example, one child at 12 months says "ba" while pointing to a book, and at 16 months she points to another child's shoes and says "shoes". In both instances the caregiver turns her gaze toward the shoes suggesting that she assumes, in both cases, that these are the target of the gesture and of the word. From the earliest child-initiated communications in the corpus, when children used different communicative gestures, they received consistent caregiver responses. This can be seen below when we look at each of the gesture uses in detail.

4.1.1 POINT
When a child points at an object, in 88% of cases the caregiver responds by directing her gaze toward the object the child is pointing at. She treats the gesture as though it is a call to action for her to look in the direction toward which the child is pointing. For example, when Fiona pointed to an ambulance she heard on the street the caregiver directed her attention to where Fiona was pointing. When the children points the caregivers responded in this way regardless of whether the point was at the early stages of development and involved an extended index finger with barely retracted fingers and thumb, or later in the child's development when the gesture involved an extended index finger with a fully retracted hand. From the children's earliest recorded uses of a pointing gesture caregivers responded by looking in the direction of the point, an action that they later used in response to children's use of the verb "look", whether it was accompanied by a gesture or used alone. Turning now to the GIMME gesture uses, we see a very different caregiver response.

---

[1] In each response where a caregiver takes the object being proffered she also looks at the object, and this accounts for the two different figures for responses to the ATTENTION FOCUS gesture.

4.1.2 GIMME

When a child reaches out a hand in a GIMME gesture, in 79% of instances the caregiver gives the child the object, interpreting this action as though the child wants the object. For example, during a meal Lotte used a GIMME gesture toward an the caregiver was cutting, and the caregiver responded by handing a piece to her. However, using a GIMME gesture did not always result in the child being given the object toward which they were gesturing. This is evident in the majority of the remaining 21% of GIMME gesture uses which were instances where the caregiver did not want the child to have an object. Eleven out of 16 (69%) of the non-giving responses were accompanied by a clear indication that the caregiver interprets the gesture as a request that she refuses. For example, in one instance Brailey reached for another child's food and was told "That's Fiona's, is not for you". If these non-giving responses are added to the total responses having a *want/give* interpretation, the caregiver response to GIMME would be 71/76 (93%).

4.1.3 ATTENTION FOCUS

When a child holds an object up towards a caregiver in an ATTENTION FOCUS gesture, in 82% of instances the caregiver looks at the object. The caregiver interprets this action as though the child wants her to focus her attention on the object being proffered. Interestingly, the percentage of responses to ATTENTION FOCUS and POINT gestures is higher than responses to GIMME. It may be the case that because the action of looking is easier than the action of giving something, caregivers respond to POINT and ATTENTION FOCUS gestures more readily that to GIMME gestures. In 41% of ATTENTION FOCUS responses, the caregiver then takes the object from the child (either as soon as the child holds out the object or after the object has been held out for some time), interpreting the gesture as though the child is offering to give the object. These are not two different interpretations of the ATTENTION FOCUS gesture. Rather, taking the object is a secondary response that only occurs after a caregiver has first looked at the object. For example, Caitlin held a box up to a caregiver who looked at it and then said "yes". When Caitlin remained standing with the box held out for a further 3 seconds, the caregiver then assumed she was giving the box and took it from her. This type of use differed from GIMME gestures occurring when the adult expected to be given an object. In another instance, Brailey was lying on the change table about to have his diaper changed and was playing with a clean diaper. When he held it out to the caregiver she immediately took it from him, assuming this was a 'giving' action. There was no pause between the caregiver looking at the diaper and taking it from him, suggesting that was construed on as a 'giving' gesture and never a 'showing' gesture.

Although we have now established that caregivers treat early communicative gestures as though they are requests for action to the recipients of the gesture, we have yet to establish whether this occurs across a range of target referents. Further, it is not clear whether the elements indexed by children's gestures differ across developmental stages or whether they are relatively consistent from when children begin gesturing to when they combine their gestures with words which either match or differ from the target of the gesture. For the remainder of this paper I will focus primarily on the development of constructions as evidenced in the communications of one child, Chera. In the following table, I show the range of different targets indexed by Chera when she uses each of the gestures. These are presented according to the following developmental stages: *Stage 1* - gesture is used alone; *Stage 2* - word used alone (this stage is omitted since the focus here is on gesture use); *Stage 3* - gesture used asynchronously with a proto-word or word, indicating the same element; *Stage 4* - gesture used in synchrony with a word or proto-word indicating the same element; *Stage 5* - gesture used with a word that indicates a different element. In this table, I present a sample of the shared targets used by Chera across the different developmental stages identified with each of the gestures (see Kelly 2002, 2003 for further

discussion). When I label a target as *food,* the target is an unidentified type of food; if the food was recognizable on the video it is labeled as the specific food or drink, e.g. *banana* or *juice*.

GESTURE	STAGE	SAMPLE TARGETS
POINT	*Stage* 1	food, milk, toy
	Stage 3	cracker, banana, picture
	Stage 4	book, juice, washcloth
	Stage 5	food, juice, other child
GIMME	*Stage 1*	food, bottle, diaper
	Stage 3	bottle, washcloth, diaper
	Stage 4	food, cracker, cup (filled with drink)
	Stage 5	bib, milk, food
ATTENTION FOCUS	*Stage 1*	food, toy, cup
	Stage 3	spoon, cup, diaper
	Stage 4	food, cup, toy
	Stage 5	cup, banana, sock

Table 2. *Chera's targets indexed by gesture across developmental stages*

As can be seen from Table 2 above, across the different stages of development, even though Chera has different abilities to communicate when she is at different stages, she still uses the gestures to refer to the same type of targets. For example, in Stage 1 when she was using gesture without speech, she used GIMME gestures to index food and continued to use this gesture to index food at Stage 5. The objects listed in Table 2 are indicative of the types of referents indexed by all the children across the database.

The majority of the objects that are the target of Chera's gestures are static objects. In all instances gestures have been interpreted as indexing referents rather than indicating actions. If the child gestured in the direction of a person acting on an object, the object was coded as the target. If the child gestured toward a person performing an action, for example, a child rocking or jumping, the actor was encoded as the target. When the target of a gesture is in motion, although it is impossible to be certain whether the child is gesturing toward the action or the target, on the basis of the caregiver response, which always assumes the child is indexing the target, I code the gesture as indexing the target. This decision regarding the coding of an object rather than an action as the target is only relevant in 6% of instances (14/237) across the database. These include examples in which a plane is moving overhead, water is running from a hose, and a child is riding in a toy car. In other words, in 94% (223/237) of cases there was no potential event or action target when children used gestures – overwhelmingly gestures occurred with static objects as their targets. The only systematic difference in targets indicated by the three gestures across the stages is that there are no human, or even animate targets used with the ATTENTION FOCUS or GIMME gestures. Use of a POINT toward a non-speech act participant is the only type of gesture toward a human target - the interlocutors are never indexed.

As we have seen, caregiver responses differ across different gestures but remain consistent across the same gesture over different developmental stages. In the following section I look at how children begin to develop constructions around their pre-verbal gestures by

examining the second research question: *Does caregiver scaffolding in dialogic interaction play a role children's construction development?*

4.2 Caregiver scaffolding in dialogic interaction

Tomasello (1992, 2003) argues that when children produce their earliest verb-islands, they do not have noun or verb categories (see also Nino 1988). While children may not categorize a lexical class of nouns and verbs, it is clear that they can differentiate between gestures that can be used to elicit actions and words that encode specific objects. In looking at children's gestures combined with speech, we can begin to see how argument structure constructions, such as those predicted by the verb-island hypothesis, develop. The multi-modal analysis presented here establishes that long before children can produce verbs and arguments, their gestures and noun combinations have been interpreted as requests for actions on objects. These gesture and noun combinations will ultimately end up as being encoded as predicates and arguments.

Caregivers play an important role in the transition from gesture + word, to word + word communications. When a child uses a GIMME gesture combined with a noun that encodes a concrete object, the caregiver often responds by re-coding the handshape of a gesture using words. It is safe to assume, given the evidence presented in Table 1, that when children begin using verb-islands, for some specific verbs they are already communicating similar information through their gestures. Tomasello (1992, 2003) argues that when children build on verb-islands they are building on established constructions. In fact, when we look at the representation given in Table 3 below, it becomes clear that prior to the use of their first verbs, children already have a schema in which the outcome of a gesture + word communication is a caregiver response. How then do we relate children's early gesture and word communications to Tomasello's stage at which children have an established construction? An analysis of one child's (Chera) use of *want* can illustrate how the gap is bridged. I suggest that between children's earliest gesture and word uses occurring at Stage 3, and Tomasello's early construction uses, a child will internalize information through dialogic interaction in which the caregiver expands the child's gesture + word communications, rephrasing them as word + word communications. In Table 3, I present the acquisition pattern for Chera's use of the verb *want* across modalities. Her age (in months) is given at each use, to the left of the column headed 'Age'. A similar pattern to this also occurs with the verbs *give*, and *look*, although *give* is most often realized as *gimme*.

AGE	ARGUMENT	VERB		ARGUMENT	CAREGIVER RESPONSE
		gesture	*word*		
13		GIMME			*want more pears?*
14		GIMME		milk	*want more milk?*
17		GIMME +	want		*You want more?*
17			more		*more milk?*
18	I	GIMME +	want	that	no response
25	I		want more	cheese	hands her cheese

Table 3.   *Development of the verb 'want' for one child from 13-25 months*

The above table illustrates one path by which a child develops from using a gesture alone to using a gesture combined with a noun argument. At 13 months, Chera used only a gesture and the caregiver responded by giving her the target she was indicating by her GIMME gesture (pears), while reformulating the apparent request in words "want more pears?". By 14 months Chera had added the word *milk* to her request, and the caregiver response is the same except that

milk has slotted in for *pears*. At 17 months Chera reaches toward some food and says *want*, while indexing the target through her gesture, and the caregiver expands this utterance, asking "you want more". Chera then picks up on the use of the word *more* and uses it here when she says "more milk?". With no other cues present, such as a GIMME gesture, the caregiver is not clear about the request and she replies by first checking it. By 18 months Chera has combined the use of the GIMME gesture + *want* + *that*. Although she receives no response to this request (the caregiver's attention was on another child), she can clearly make maximal use of all the strategies at her disposal in indicating the elements she is making a request about. At 25 months she no longer needs to use a gesture to specify the element she is requesting or to request the caregiver's action. At this stage she can rely on words to get what she wants, for example, when she says, "I want more cheese", to which the caregiver responds by putting some cheese in her bowl.

By the time Chera is combining words she is already using gesture to indicate the elements that will be direct object arguments. She communicates without using A role referents (subjects of transitive verbs) as these are often herself and the caregiver, and are already part of the scene. When she begins combining speech and gesture, the word can indicate either the target (encoded with a noun) or the action (encoded with a verb). Chera requests by using a gesture with what will eventually become a noun argument, with the verb *want*, and also with *more*, which appears to be interchangeable with *want* in Chera's uses. The caregiver also uses *want* in each of her replies (aside from the response in which she asks for confirmation). The caregiver maps the child's handshape such as GIMME to the verb *want*, taking the child's point of view when she displays a mapping of handshape onto words and actions.

I have shown elsewhere that when children begin to communicate about wanting something, they use a GIMME gesture initially and then combine it with the label for the object toward which they are gesturing. At this stage the children occasionally include the name of the object they are gesturing toward (Kelly 2003). Children also occasionally include an action request by using a verb together with their gesture. The verbs used with each of the gestures were consistent with the caregiver's responses to the use of the gestures alone, and consistent with the verbs used in caregiver expansions of children's gestures. For example, when a child uses GIMME and the adult responds with a verb the verb is always either *want* or *give*. This is the case for all of the verbs used in the study that have a meaning that was originally expressed through a gesture; POINT and initially ATTENTION FOCUS are construed by the caregiver as a request to *look*, and in the few early instances when they are combined with a verb the verb used is *look*. The data suggest that in these instances, children retain the gesture until they have acquired the verb and until they realize that it can be used to do what they earlier did for the verbs presented in Table 1, with a gesture (Kelly 2003). Although a more detailed analysis is necessary, the data suggest that children will only use the verb alone when they have already used it with a gesture (gesture + *look*), and received the same response from a caregiver that they received when they used the gesture combined with a word (gesture + *book*). This at least indicates to the child that the request for action can vary in form (it can be a gesture or a verb) in the same way that the target of a request for action can vary, i.e. a gesture or a noun.

By looking at Chera's communications as one example of a child's path to construction development, we can see that these early gestures and multi-modal gesture-speech communications are precursors to the development of constructions. Through dialogic interaction with a caregiver, children learn to build upon the basis of their communicative gestures and their early words, and eventually combine these modalities as a means of conveying multiple pieces of information at once. The cognitive ability to convey multiple pieces of information about a single event is, in turn, the basis for children's ability to combine multiple words with respect to a single

event and to extend these constructions by consistently combining a relatively small number of verbs with a larger number of arguments.

5. Conclusions

I have established throughout this paper that the early construction development of young English-learning children is influenced by their gesture uses. Further, we've seen how caregiver responses to children's gestures and gesture + speech communications provides scaffolding for the child by modeling of a mapping from gesture + word to word + word. These caregiver responses help children to learn to build constructions around gestures. In claiming that children's constructions are built around gestures, I do not mean to claim here that this is the only path children take to learning constructions. However, I do assert that this is one important way they can learn to use early constructions before moving on to other more complex constructions. Tomasello (2003) and Goldberg (1995) among others argue that children's constructions are built around single words or 'islands'. In fact, children's constructions begin developing prior to their first words and are centered around their pre-verbal gestures. The knowledge that they have about constructions when they begin to talk comes from their early gesture-based interactions. As this paper has shown, constructions arise long before the multi-word utterances examined in the majority of child language studies. Overall, these findings provide an insight into what information children may use as they transition from using gesture as a primary form of communication to using multi-word constructions.

References

Bates, E. (1976). *Language and context*. New York: Academic Press.

Bugenthal, D., Clancy, P., Kyratzis, A., Lerner, G., & Zimmerman, D. (1999). The very young children project. Corpus archived in Dept. Sociology, University of California, Santa Barbara.

Butcher, C. & Goldin-Meadow, S. (2000). Gesture and transition from one- to two- word speech: When hand and mouth come together. In: D. McNeill (ed.), *Language and gesture*. Cambridge: Cambridge University Press. pp. 235-58.

Carter, A. (1978). The development of systematic vocalizations prior to words: A case study. In: N.Waterson and C.E. Snow (eds.) *The development of communication*. Chicester: Wiley. pp. 127-138.

Goldberg, A. E. (1995). *Constructions.* Chicago: Chicago University Press.

Goldin-Meadow, S. & Butcher (2003). Pointing toward two-word speech in young children. In S. Kita (ed.). *Where language and culture meet*. New Jersey. Lawrence Erlbaum Associates.

Gordon, P. (2003) The origin of argument structure in infant event representations. *Proceedings of the 26th Boston University Conference on Language Development.* Somerville, Mass: Cascadilla Press.

Greenfield, P. & Smith, J. (1976). *The structure of communication in early development*. New York: Academic Press.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Kelly, B. F. (2002) The development of gesture, speech, and action, as communicative strategies. In *Language and Gesture: Proceedings of 27th Berkeley Linguistics Society*: 1-10.

(2003). The emergence of argument structure from gesture to speech. Unpublished PhD dissertation, University of California, Santa Barbara.

Leung, H. & Rheingold, J. (1981). Development of pointing as a social gesture. *Developmental Psychology, 17*: 215-220.

Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language, 19*(3): 559-580.

Ninio, A. (1998). Acquiring a dependency grammar: The first three stages in the acquisition of multiword combinations in Hebrew-speaking children. In G. Makiello-Jarza, J. Kaiser & M. Smoczynska (Eds.), *Language acquisition and developmental psychology.* Cracow: Universitas.

Pine, J. M., Lieven, E. V. M. & Rowland, C.F. (1997). Stylistic variation at the 'single-word stage': relations between maternal speech characteristics and children's vocabulary composition and usage. *Child Development, 64*: 807-819.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: MIT Press.

Tomasello, M. (1992). *First verbs*. Cambridge: Cambridge University Press.

 (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge MA: Harvard University Press.