

The organization of children's pointing stroke endpoints

Running head suggestion: Children's pointing stroke endpoints

Author: Mats Andrén, Centre for Cognitive Semiotics, Lund University

mats.andren@ling.lu.se

Abstract

The timing of index finger pointing gestures of three Swedish children (recorded longitudinally between 18 and 28 months) was analyzed. 63% of the pointing strokes ended in direct association with the child's own spoken utterance. This is in line with standard descriptions of gesture timing. However, 35% of the pointing strokes were sustained for a longer time—until a response was received from an interlocutor. It is shown here that parents give significantly more elaborated responses when children's pointing strokes are sustained and that the children work actively to achieve this result. The timing of such pointing gestures is thus a matter of interactive coordination between child and interlocutor. Finally, these findings are used as the basis for a discussion of different types of descriptions of gesture timing in the literature and how these may relate to each other.

1. *Introduction*

In the context of ethology, Hinde (1957, p. 118) stated that: “The mechanisms underlying behaviour are diverse, and a given pattern of behaviour may be brought to an end in many different ways. Nothing is gained by grouping all ‘causes of endings’ under one heading”. In this paper, I will try to make very much the same point, but more specifically with respect to children’s pointing stroke endpoints. Most research on gesture timing during the last 20 years has been devoted to the formulation of gesture- and speech-production models, and models of ‘thinking for speaking’. This line of research has yielded important insights into the ways in which gestures are usually coordinated with spoken utterances, especially with regard to the onset of strokes and mainly regarding iconic gestures (cf. Nobe, 2000).¹ However, not all aspects of gesture timing are of this kind. The present study focuses on aspects of gesture timing whose logic is primarily interactive; that is, it is continuously adjusted during the course of its performance with respect to the behavior of the Other. Also, this study concerns the endpoints of 18-28 month-old children’s pointing strokes, rather than the onsets of strokes in the iconic gestures of adults, which have been much more studied.

At issue is how responses from interlocutors vary in relation to two types of pointing stroke durations: (a) strokes that end in direct association with the utterance in which the pointing stroke also started and (b) strokes that end some time after *another utterance* has been delivered (in most cases by the interlocutor). The two main hypotheses are (H1) that

¹ Strictly speaking, Nobe (2000) use the term ‘representational gestures’, as including ‘iconics’, ‘metaphorics’, and ‘abstract deictics’.

parents give more elaborated responses during children's sustained pointing gesture strokes and (H2) that the children fine tune this type of sustained strokes adaptively in relation to the responses they get from their parents: the less elaborated responses they receive, the more they tend to work for such a response.

Similar thoughts about the interactive functions of sustained pointing gestures have been expressed in the literature, based on observations of adults (Sidnell, 2005; Clark, 2005) and children with Down's syndrome (Wootton, 1990). Bavelas (1994, pp. 203, citing personal communication with Adam Kendon in 1988) writes: "when a gesture is held longer than would be needed simply to convey information, it becomes a *kinetically held question*, that is, a request for response from the addressee." The aim of the present study is to make a more systematic evaluation of these claims.

2. Method

2.1 Data

Recordings of three Swedish children from the Strömquist-Richthoff corpus (Richthoff, 2000) were used. Each child, two girls and one boy, was recorded at home at least once a month between the ages of 18 and 28 months as they were interacting with a parent. Throughout the recordings, the participants sat by a table, interacting side-by-side. Activities included book reading, eating, playing with toys, and general conversation.

All instances of index finger pointing performed by the children in the first five minutes of each recording were coded according to the categories described in the following two sections. A total of 393 instances of index finger pointing in the children were found.

2.2 Explicit exclusion of some instances from analysis

There is much variation in children's performance of pointing gestures. Not all of them are performed with an extended index finger and the other fingers curled. Furthermore, in some cases the stroke itself has a movement structure, giving it a kind of inherent temporal extension (in contrast to punctual strokes), and in other cases, the pointing gesture is part of a series of gestures rather than being used on its own. Such "additional" features of pointing gestures can be expected to affect their timing characteristics, but since the aim was to study variation in pointing only along one single dimension, namely differences in parental response when pointing gestures were sustained or not, it was decided to focus strictly on cases where the pointing gesture is as "pure" and prototypical as possible. To be sure, the different ways in which children perform pointing gestures is an interesting topic of systematic study in its own right (Andrén, in preparation), but accounting for the precise timing in relation to all of these features would necessarily involve a much more complex analysis. Therefore, explicit criteria were formulated of when to *exclude* instances of index finger pointing from analysis. Instances were excluded under the following circumstances:

- a. The child was simultaneously holding an object in the hand that performed the pointing gesture (cf. Andrén, in press) (n=7).

- b. The pointing stroke also exhibited iconic features such as displaying form or motion (n=9).
- c. The pointing gesture was performed without speech (n=8).
- d. The pointing gesture seemed to be monologic (private) rather than directed to the interlocutor (n=6).
- e. The parent and the child started talking simultaneously (n=21).
- f. The pointing gesture was part of a series of gestures (n=13).
- g. The pointing gesture was very diffuse and did not seem to orient to a specific target (n=8).
- h. The gesture appeared to be combined with haptic exploration of a material being pointed to, using the index finger (n=5).
- i. The pointing gesture was affected by practical problems, such as the child pointing to a book that the parent was simultaneously moving, or the child and the parent collided physically during the action (n=13).

In all of these cases, the timing characteristics are potentially at least slightly different in terms of the child's coordination of the gesture with the own utterance and/or in terms of interactive coordination. After the exclusion of these instances, 303 of the 393 instances remained.

2.3 Coding of the data

The concept of a *stroke endpoint* is used in this paper to refer to the moment in time when a pointing gesture to a certain target eventually turns into a full retraction or into a

preparation phase of a new gesture. In cases where there were, for example, repeated tapping of the target during the performance of the pointing gesture, possibly with some pause between some of the taps, this was considered to be a single stroke, rather than being a series of strokes, since it is part of the same overall pointing to a single target. For each index finger pointing gesture, three *types of stroke endpoints* were distinguished:

1. *short*: Endpoint in direct association with the child's own utterance—at the very end of, or during, the utterance.
2. *between turns*: Endpoint shortly after the utterance was finished, but still before the first transition relevance place (TRP) in the next utterance. The first TRP in an utterance corresponds to the first point in the utterance where the turn-so-far may be perceived as a complete turn, although the utterance need not necessarily end at this point (Sacks et al. 1974). The turn-so-far can be a full grammatical clause, but also, for example, a response morpheme.
3. *sustained*: Endpoint after the first TRP in the next utterance, produced either by the parent or the child.

Sustained strokes were considered single gestures even if they exhibited features of *stroke renewal*, such as tapping the target once more after a hold, as long as there was no proper retraction of the pointing gesture in between. This is why the term sustained gestures is used here, in contrast to the more technical and narrow sense of a stroke hold, which would not be said to last across such renewals of a stroke. It should also be noted that the same

sustained pointing gesture was sometimes sustained over several further utterances from both of the participants.

Second, all parental responses to the children's utterances that included pointing gestures (sustained or not) were coded for *degrees of response*, arranged on an ordinal scale, ranging from least responsive to most responsive. The categories were defined relative to what the child was talking about and pointing to, and they were defined in the following way:

1. *no response* – Neither responding nor initiating features: when parents did not say anything, or did not perform a certain act if an act was requested, or when the parent indeed did say something, but initiated a new sequence rather than replying to the child's utterance.
2. *minimal response* – Responding, but no initiating features: when the response only contained short response morphemes such as “yes”, “mm”, “no” or simple repetitions of what the child said (Child: “a ball”, Parent: “yes, a ball”) mainly serving as an acknowledgement of what the child said, rather than adding any new content.
3. *expanded response* – Both responding and initiating features: when there were not only aspects of acknowledgement in the reply, but also initiating aspects, saying something more or new about what the child was pointing to (Child: “there”, Parent: “I wouldn't wanna taste that”) or when the parent performs an act in compliance with a child's request for such an act.

This is, of course, a rather harsh simplification of the intricacies of interactive coordination of response and initiative in dialogue. For example, some utterances, such as *wh*-questions, are more response demanding than others. Nevertheless, it was judged sufficiently detailed to provide a foundation for testing the hypotheses. For a considerably richer treatment of initiative and response in dialogue, see Linell and Gustavsson (1987).

Third, in cases where the children produced further utterances during a *sustained* stroke, often with parental responses “inserted” in between, all these subsequent child utterances were coded using a distinction between three levels of *communicative effort*. This distinction was intended to capture the type of effort the child puts into the gesture stroke itself as well as other means of drawing attention to the multimodal utterance as a whole, such as using a stronger voice than in the child’s previous utterance. The three levels are defined as:

1. *plain hold*: a continuous hold of the gesture with a new spoken utterance which was essentially a repetition of the previous utterance at the same level of intensity.
2. *renewed stroke*: a renewal of the stroke such as tapping the target again, but without a full retraction, repeating a similar utterance at the same level of intensity as in the previous utterance.
3. *upgraded renewal*: not just renewing the gesture and repeating the previous utterance as in the previous category, but also adding intensifiers such as performing the gesture in a more intense and salient way, using a stronger voice, turning to the parent to establish

mutual gaze, or providing a markedly more elaborated version of the previous utterance.

3. Analysis

3.1 The existence of two main types of timing

Most of the analyzed pointing stroke endpoints were of the *short* type (63%), ending in direct association with the child's own utterance in line with standard descriptions of gesture timing. However, there was also a substantial amount of stroke endpoints of the *sustained* type (35%), where the stroke was sustained until at least one more utterance has been delivered (90% of those utterances come from a parent). The *between turns* type of endpoints was very rare in comparison (2%), which means that the children used two almost categorically distinct ways of placing the endpoint of the pointing stroke either within their own utterance or after a next utterance has been delivered, but seldom in between. Moreover, 5 of the 7 instances of the *between turns* type had a stroke where the target was tapped repeatedly with the index finger. This implies that the *between turns* type may often occur in strokes that have an inherent temporal extension due to having a complex movement structure (such as tapping repetition), which may compete to some extent with detailed coordination of gesture and speech. The rest of the analysis concerns only the two most common types of stroke endpoints (*short* and *sustained*). Even though development was not the focus of this study, it may be pointed out that the relative frequencies of *short* strokes and *sustained* pointing strokes for the group as a whole

remained constant over the investigated period (correcting for an overall reduced tendency for pointing from around 23 months and onwards). The three children were fairly similar to each other, with one child performing slightly more *sustained* strokes than the other two in the analyzed data.

3.2 Parental responses in relation to short/sustained strokes

According to the first hypothesis (H1), it was expected that *sustained* pointing strokes should be associated with more elaborated responses from the parents than *short* strokes, and vice versa. A Pearson Chi-square test confirmed this hypothesis: χ^2 (df=2, n=296) = 22.34 (p<0.01). Raw frequencies are presented in Figure 1. In cases of *sustained* strokes, *expanded response* was significantly more common, whereas *no response* and *minimal response* were significantly less common, compared to when strokes were of the *short* type. When strokes were *short*, the pattern was the opposite (also significant). In sum, children received significantly more response in cases where the stroke was sustained into further utterances. There is only one instance in the data where a child abandons a sustained pointing gesture apparently without having received any response.

@@ INSERT FIGURE 1 HERE. Caption: "Stroke endpoints and types of response"

It should also be remembered that many means are available for eliciting responses in communication. The claim here is certainly not that the duration of pointing strokes is the

primary means for eliciting responses. *Expanded response* was the most common response type in both conditions (*short* and *sustained* stroke, see Figure 1), and the crucial finding was that parents gave *even more* response when children performed sustained pointing strokes than they did in the context of short strokes.

3.3 *The internal dynamics of sustained pointing sequences*

In 69% of the cases where a stroke was *sustained*, the stroke ended directly after an adjacent response given by a parent. This, in itself, is evidence in favor of the second hypothesis (H2), that getting a (satisfactory) response from the parent is indeed the primary stopping condition of the sustained strokes. In 31% of the cases even more utterances were exchanged, including a few instances where the stroke was sustained although an *expanded response* had in fact been received. All utterances in such exchanges from both participants were always about the referent being pointed to. It could be argued that the general principle of gestures being coordinated with co-expressive speech remains true here too, even across turns and partly even across speakers.

To further test the second hypothesis, children's reactions to their parents' responses were investigated in cases where the parent responded during a sustained child pointing gesture. More specifically, if the child kept the pointing gesture for at least one more of his/her own utterances after a parent's response, the nature of the child's pointing gesture as it was performed with the new utterance was coded according to communicative effort (*plain hold*, *renewed stroke*, or *upgraded renewal*). There was a strong negative relation

(Spearman Rank=-0.70, $p<0.01$) between communicative effort and the degree of response from the parent, which is strong evidence in favor of the second hypothesis. The less response the children received to their sustained pointing gestures, the more communicative effort they mobilized in the next turn, and vice versa.

4. Discussion and conclusions

According to the findings of the present study, parents gave significantly more response when children performed *sustained* index finger pointing gestures. Furthermore, the children were shown to orient to the content of these parental responses. This was visible in two ways. First, in most cases the children immediately withdrew their sustained pointing gestures when a parental response was given, and second, in cases where there were no such immediate withdrawal, there was a significant inverse relationship between the degree of response from the parent and the communicative effort invested by the child in the child's subsequent utterance(s). That is, the less response a parent gave to a child utterance with a sustained pointing gesture, the more likely the children were to upgrade their "demand" for a response through various sorts of intensifying resources. In short, the children did not perform such sustained pointing gestures at random. They seemed instead to be part of the children's established and typified repertoires of methods for eliciting parental responses. The fact that the children were less satisfied with minimal responses than with extended responses is interesting since it means that the "goal" of their pointing was not only to achieve intersubjectively shared reference to certain targets (joint

attention), in which case simple acknowledgements would have been sufficient. The children rather seemed to aim for receiving various sorts of evaluations and comments on the target being pointed to, i.e. an active form of *social referencing*. It remains unclear when the use of these different types of pointing stroke endpoints emerge, since they are already present in the first observations here, where the children are 18 months old. As pointed out in the analysis, the relative frequencies of *short* and *sustained* strokes remained constant, for the group as a whole, during the period studied here (correcting for the fact that there was an overall decline of pointing gestures around 23 months).

Since sustained pointing gestures and the responses from the parents are happening simultaneously, it may not be appropriate to interpret one as cause and the other as effect. As was shown in this study, both parent and child orient to such sustained pointing gestures actively and mutually. This can be seen as a demonstration of the utility of multimodal resources in dialogue since gesture and speech do not interfere with each other in the way that simultaneous speech does. As mentioned in the introduction, adults sometimes make use of sustained gestures too for the very same general purposes (and in several different cultures). However, it would be an overgeneralization to conclude that all kinds of sustained gestures are to be interpreted as requests for response under all circumstances. For example, at the final moment of a theatre act actors commonly freeze and sustain postures and gestures, but in the turn-taking system used in this setting the audience is supposed to wait until *after* a sustained gesture is retracted before providing a response in the form of applause (Broth, 2002). In conversational interaction, though, gestures sustained in this way may well serve similar functions across most contexts. A common

denominator between the theatre example and the conversational situations studied here seems to be that a sustained gesture is markedly “in play”, whatever that implies in a given context.

Regarding the general theoretical issue of gesture timing, it is interesting to note that different researchers use different terms to talk about such phenomena. In interactionally oriented research gesture timing is often described as a kind of skillful *achievement* or in terms of *recipient design* (Kendon, 2004; Goodwin, 2000), whereas psychologically oriented research tend to use the term *synchrony* (McNeill, 2005). Although interactional and psychological interests need not be mutually exclusive, it is clear that these differences in terminology highlight very different aspects of gesture timing. The term synchrony tends to evoke descriptions of gesture timing primarily as a matter of neural mechanisms in relation to utterance formulation, whereas the terms achievement and recipient design tend to highlight continuous processes of mutual orientation that are not understandable without reference to two or more co-present bodies, and their contextual embedding in a shared field of activity. Whereas psychological research on gesture timing has mainly focused on onsets of strokes (and preparations) as indications of mental activity such as “motor planning”, the interactional perspective seems to be the one that has recognized the importance of stroke endpoints as orderly, visibly recognizable and interactionally consequential (Wootton, 1990; Sidnell, 2005; Clark, 2005; Bavelas, 1994; Goodwin, 2000).

A crucial difference seems to lie in what is considered the ‘starting point’ of behavior (explicitly and/or implicitly). Does action originate in the mind, or in situations?

Obviously, the question is ill-formed, since this is not best understood as an either-or question. Still, the emphasis of acts—especially communicative acts—as creations of the mind in psychological research has a tendency to downplay how action is systematically sensitive and responsive to the contingencies of social situations. Interactional accounts of action are often more explicitly geared at treating action in a way that relates it to the situation and previous acts; that is, treating action as a kind of ongoing balance act between initiating and responsive/responding aspects of action. In line with this description, Linell (1998, p. 211) writes about a change of emphasis from intentionality to responsibility within dialogical frameworks. In this vein, I would like to argue that it is important to think about agency of action in a way that pays proper attention *both* to its initiating and responsive qualities. It seems clear that the sustained pointing gestures described in this paper are not delivered as readymade wholes, as they also exhibit features of responsiveness with respect to the behavior of the Other during their very performance. To be clear, the point of this argument is certainly not to claim that one or the other approach to language and communication is ‘wrong’ *per se* or that there are approaches that do not have limitations. It is also questionable to push talk about perspectives too far as if they were uniform. The point is rather that the perspectives tend to *complement* each other due to their relatively systematic highlighting of different aspects—in this case regarding gesture timing. Still, I would like to conclude by suggesting that there exists no moment in time such that the process of communication is entirely your own (cf. Schutz, 1962).

5. Acknowledgements

Thanks to Gale Stam, Jordan Zlatev, Simon Harrison and Ulrika Nettelbladt for valuable comments on previous drafts of this paper. I would also like to acknowledge the Centre for Cognitive Semiotics (CCS) at Lund University for providing a productive interdisciplinary environment for research on the relations between different semiotic resources.

6. References

- Andrén, Mats (in press). Gesture and object manipulation. In Chris Sinha, Cíntia Rodríguez & Jes Vang (Eds.), *A companion to Cognitive Semiotics 1: Meaning and materiality*.
- Andrén, Mats (in preparation). Forms of pointing in young children.
- Broth, Mathias (2002). *Agents secrets. Le public dans la construction interactive de la représentation théâtrale*. PhD thesis, Uppsala University.
- Bavelas, Janet B. (1994). Gestures as Part of Speech: Methodological Implications. *Research on Language and Social Interaction*, 27(3), 201-221.
- Clark, Herbert H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7(4-5), 507-525.
- Goodwin, Charles (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489-1522.
- Hinde, Robert (1957). Consequences and goals: Some issues raised by dr. A. Kortlandt's paper on "aspects and prospects of the concept of instinct". *British Journal of Animal Behaviour*, 5(4), 116-118.
- Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

- Linell, Per & Lennart Gustavsson (1987). *Initiativ och Respons: Om dialogens dynamik, dominans och koherens*. Linköping University: Department of Communication Studies.
- Linell, Per (1998). *Approaching Dialogue*. Amsterdam: John Benjamins.
- McNeill, David (2005). *Gesture and Thought*. Chicago: Chicago University Press.
- Nobe, Shuichi (2000). Where do most spontaneous representational gestures actually occur with respect to speech? In David McNeill (Ed.), *Language and Gesture* (pp. 186-198). Chicago: Chicago University Press.
- Richthoff, Ulla (2000). *En svensk barnspråkskorpus: Uppbyggnad och analyser*. PhD thesis. Göteborg University.
- Sacks, Harvey, Emanuel A. Schegloff, & Gail Jefferson (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.
- Schutz, Alfred (1962). *Collected Papers I: The Problem of Social Reality*. The Hague: Martinus Nijhoff.
- Sidnell, Jack (2005). Gesture in the pursuit and display of recognition: A Caribbean case study. *Semiotica*, 156(1), 55-87.
- Wootton, Anthony J. (1990). Pointing and interaction initiation: The behaviour of young children with Down's syndrome when looking at books. *Journal of Child Language*, 17(3), 565-589.