Discourse Processes. 16, 363-386.

COHESION AND GESTURE

David McNeill, University of Chicago
Elena T. Levy, University of Connecticut and Haskins Laboratories

Abstract

Discourse cohesion is viewed from the perspective of a speech/gesture synthesis. Based on narrative and conversational data, cross-cultural evidence is presented of cohesive elements marked by repetitive gestures that maintain continuity with respect to their location in space, the hand with they are produced, and/or their form. The data show the joint contribution made by speech and gesture to the process of creating and maintaining discourse topics. It is claimed that an approach to discourse that focuses on events taking pace at the moment of speaking, unlike approaches that assume the prior existence of planned discourse units, can account for the impact of speech and gesture on thought; for example, that the execution of a gesture helps the speaker to track presupposed background information, and so provides a basis for the production of the communicatively dynamic part of an utterance. The proposed model of discourse production is a dialectic in which gesture and linguistic form are two non-redundant interacting `voices,' with the gestural `voice' commenting on, supporting, contradicting, bracketing, reinterpreting, etc., the linguistic `voice.'

COHESION AND GESTURE

David McNeill, University of Chicago

Elena T. Levy, University of Connecticut & Haskins Laboratories

In an earlier paper (Levy & McNeill, 1992) we argued that gestures and referring expressions perform similar discourse functions: gestures and speech cooperate in creating new themes or continuing old ones. Specifically, both types of surface referring forms accumulate when communicative dynamism is high. Thus Givón's (1985) principle of more quantity of expression—that the less predictable/accessible/continuous a topic is, the more coding material is used to represent it in language—applies to gestures as well as speech. Focusing on the temporal dimension of speaking, we identified processes through which speech and gesture contribute to the thematic structure that is created at a given point in a discourse. From a temporal point of view, thematic structure is built up utterance by utterance, as new elements are linked to old ones, creating either continuity or discontinuity with the prior context. Elements that are discontinuous carry a high degree of communicative dynamism (Firbas, 1964, 1971; Danes, 1974), while those that are continuous are cohesive, carrying "the presupposition of something that has gone before" (Halliday & Hasan, 1976).

A basic explanation for the cooperation of speech and gesture in creating thematic structure is an underlying unity of speech and gesture, that speech and gesture are nonredundant manifestations of one process in utterance and discourse generation. Gesture in this view is actually *part of language*—not a subsystem, either distinct or overlapping. It is the analogic, imagistic, global and synthetic part of language that is processed *together* with the familiar hierarchical, linear, segmented and `linguistic' part. Language is a system in which the analogic and the linear-segmented combine. Thus Givón's principle applies to gestures as well as to speech because it directly applies to this underlying dualistic process.[1]

In this paper we focus on the surface marking of continuous referential information. The general principle is that, all things being equal, information about references that is more continuous will be marked with fewer and less complex surface devices. For example, in one speaker's narration, as the "surprise element" of constituents decreased—as continuity with the context increased—fewer and less elaborate gestures tended to occur. However, linguistic elements that are continuous with their context are at times accompanied by gesture. In these cases, cohesive elements tend to be marked by repetitive gestures that maintain continuity with respect to their location in space, handedness and/or form. In this paper we explore gestural cohesion of space, handedness, and form.

Approach to Referential Cohesion

From the point of view of speech as it unfolds in time, the speaker's choice of a surface referring form at any given moment is dependent on communicative intent and earlier context, at times on context that stretches back to distant parts of the discourse. On-line choices are made in a matrix of continuous discourse connections (Levy & McNeill, 1992). Consider the following choice of referring expression:

> Charles$_i$ gives the guy$_j$ a tip
> and Ø$_i$ says that
> "I$_i$'m sorry"
> the guy$_j$'s muttering
> so <u>he</u>$_j$ goes out.

The underlined pronoun in this clause is not coreferential with the nominal in the immediately preceding clause ("the guy"), but rather it indexes the referent that has been most frequently mentioned in the preceding "paragraphs" of text. Relative to the immediately preceding clause, a fuller referring expression, for example, a proper name, would have been more appropriate. In the broader context of the ongoing narration, however, the pronoun is in fact appropriate since its referent has been built up as the topic over the course of time, in part through the relative density with which it has been mentioned. The dense chain of references to "Charles" has created a "tightness" of cohesion, which them contributes to the topic structure that is created in the discourse. As the discourse unfolds, each successive pronominal reference to "Charles" tightens the cohesive relationship that is being formed, and in so doing strengthens the referent's status as the discourse topic. In the final sentence, the speaker chose a pronoun to index the referent that had been established as the most topical at the global level.

From this point of view, less explicit referring terms both reflect a cohesive link that has already been created, and strengthen cohesion for future discourse. In general, this double connection of less explicit referring terms suggests that, all other things being equal, complex surface forms are reserved for those parts of the discourse that will become new topic units, and that less complex forms are used for the duration of the topic . (The "all other things being equal" argument is necessary since surface forms are multifunctional, so that the same form can mark different functions when it occurs at different points in the discourse. For example, explicit referring forms are used both to start a new topic unit, as well as to mark contrast, or emphasis.)

For gestures, as well, one would expect that patterns of occurrence and non-occurrence and continuity vs. discontinuity of form, space, and manner of execution would both create and reflect topicality. There is evidence that gestures tend to occur at points of topic shift, such as new narrative episodes or new conversational themes (when communicative dynamism is high) (Levy & McNeill, 1992). Conversely, one would expect that highly presupposed linguistic elements would either lack gestures entirely, or would be accompanied by gestures that are specialized for their cohesive function of

form—less complex internally on some appropriate absolute scale, less complex than other gestures that do signal topic shifts, or repeating essential features of earlier topic-initiating gestures.

The following example, from one speaker's narration, supports the general principle that less "quantity" of gestures occurs with presupposed information. For highly presupposing constituents, such as zero and unstressed pronoun references, no gestures occurred at all. On 3 occasions, gestures occurred in sentences with unstressed pronouns but precisely excluded the pronoun. With modified noun phrases, there were iconic gestures with the reference (n=13); they were simple in form. All the complex gestures were coextensive with clauses and verb phrases; that is, coextensive not with referring forms but with predications. Only with these largest constituents did complex two-handed iconics take place. Thus, the probability and the complexity of the gestures increased when the quantity of the linguistic material in the accompanying speech also increased.

In a second example, Kendon (1972) found that discourse "paragraphs", or "locution clusters" (defined by "what the speaker is talking about or how is treating the topic of discussion"), were introduced by shifts in the speaker's use of hands and arms in gesticulation. For example, during one locution cluster, the speaker primarily used his right arm, in the following cluster he used his left arm, and in the next bused both arms. Thus, changes in gestural handedness occurred at the start of a new topic. These qualities were then sustained for the duration of the topic. At a higher level of discourse organization, speaker turns tended to be accompanied by postural shifts, sustained for the duration of the turn. The continued use of the same gesture hand or posture may help the speaker keep alive the discourse theme while the problem of new information is being dealt with. In these ways, the speaker is able to maintain and strengthen topicality by marking change and then maintain cohesion via the gestural system.

In this paper, we examine how the speaker's unwitting spontaneous gestures contribute to cohesion that is demanded by the construction of interpretable discourse. Together with our earlier paper, where the focus was on the role of gesture in creating thematic discontinuities, we paint a more complete picture of the mind of the speaker at the moment of speaking.

Sources of Data and Coding Methods

We present our subjects with an animated color cartoon or full-length film "stimulus," which the subject then immediately retells from memory to a listener. The listener is a genuine interlocutor who has not seen the cartoon or film and is free to question the speaker. The instructions emphasize that the listener himself will have to retell the story (although he or she is not always asked to do so) from the narrator's original presentation, and this is to encourage completeness and clarity; no mention is made of gestures and no guidelines are given for how the narrative should be structured. The entire performance is videotaped, and this videotape is the primary data of the investigation.

We have collected narratives from adult speakers of different languages (several Indo-European languages, including English, and non-Indo-European languages: Georgian, Swahili, Chinese, and Japanese). We report examples in this paper from the English and Georgian narratives. In addition, we have studied the gestures produced during videotaped conversations, and some of our examples below come from this source.

We list here the major types of gesture recognized in our coding scheme (see McNeill & Levy, 1982; McNeill, 1985). Following the methods described in McNeill (1992), the four types—iconic, metaphoric, beat, and (abstract) deictic—can be coded with excellent reliability (between 0.85 and 0.95).

Iconics. Iconic gestures bear a close formal relationship to the semantic content of speech. That is, in their form and manner of execution they exhibit aspects of the action or events described by the accompanying narrative discourse. In storytelling, iconic gestures appear chiefly at the narrative level—the level of story events (McNeill & Levy, 1982; Cassell & McNeill, 1991).

Metaphorics. Metaphoric gestures are like iconics in that they are representational. The imagistic content of a metaphoric gesture, however, presents an abstract idea, not a concrete object or event. A metaphoric gesture displays the vehicle of a metaphor (Richards, 1936). Metaphoric gestures occur chiefly at the metanarrative level of the storytelling process—the references to the story proper and its structure (Cassell & McNeill, 1991). Some gestures involve oppositions of space imbued with abstract meanings to outline, for instance, the plot of the story or the contrast between the moral values of the characters. One specific type of metaphoric gesture that plays an important role in the framing of the discourse is the "conduit" metaphoric gesture (named after a similar verbal metaphor; see Reddy, 1979; Lakoff & Johnson, 1980). In narrative discourse, the conduit metaphor often appears only in gesture form. Conduit metaphoric gestures often appear as a cupped hand which seems to contain the narrative itself and offer it to the listener. An example, is a speaker stating that he has just seen a cartoon and is about to recount this kind of story to his listener. The speaker creates and supports with cupped hands an "object" that is metaphorically the cartoon and his upcoming narration in this genre. The gesture displays the conduit image, speech here does not. Gesture thus complements speech—adding its own imagery of the narrative event.

Beats. Of all the gesture types, beats are the least significant looking; but appearances can be deceptive, for beats are among the most revealing of gestures for uncovering the speaker's construction of the narrative discourse. In performing beats, the hand moves with a rhythmical pulse that tends to line up with the stress peaks of speech (McClave, 1991). Unlike iconics and metaphorics, beats tend to have the same form regardless of the content (McNeill & Levy, 1982). The typical beat is a simple flick of the hand or fingers up and down, or back and forth; the movement is short and fast. The semiotic value of the beat lies in the fact that it indexes the word or phrase it accompanies as being significant not purely for its semantic content but also for its discourse pragmatic content. The beat is sensitive to the momentary relevance of the

larger discourse structure or narrative situation as a whole. Examples are marking the introduction of new characters, summarizing the action, introducing new props, etc. Thus beats may accompany information that does not advance the plot line. With beats these events on the metalevel of discourse can be inserted in to the narrative level, signaling the fact that they depart from the chain of events comprising the plot line (Cassell & McNeill, 1991).

Abstract pointing. Deictic gestures, or points, have the obvious function of indicating objects and places around the speaker, but they also play a part in narrations and other discourse situations where there is nothing objectively present to point at. Although the gesture space may look empty to a camera, to the speaker and listener it is being filled with real entities in the discourse. In storytelling, deictics establish characters in space the characters and mark the occurrence of narrative boundaries. An example is a speaker who said "the artist and Alice are walking by," pointing first to his right and then making an iconic gesture for walking by; the pointing gesture set up the characters in the scene, the iconic gesture depicted their action. Abstract pointing often occurs at the beginnings of new episodes and scenes, and here it is the dominant gesture type (Levy, 1984; Levy & McNeill, 1992). In this context, pointing may be the most common way of marking the establishment of a new focus space in Grosz's sense (Grosz, 1981).

In narrations of cartoon stories, about three-quarters of all clauses are accompanied by gestures of one kind or another; of these, about 40% are iconic, 40% are beats, and test are divided between deictic and metaphoric gestures. In narrating films, where the emphasis is more on narrative structure, the proportion of metaphoric and deictic gestures increases at the expense of iconic gestures (statistics from McNeill & Levy, 1982, and McNeill, 1992).

Cohesion of Space

We have found that narrators often use space to create cohesive linkages across narrative texts. The speaker assigns a distinct part of space to separate characters. A chain of gestures performed in the same place can convey continuities and previously established thematic roles. The following examples show that shifts of space co-occur with initial references in a scene and continued use of space with continued references; at the same time, the referring forms themselves are conditioned to decline in specificity by continuing use of the same spatial assignments.

A. The first example is from a full-length film narration (an early Hitchcock, *Blackmail*). The speaker was setting up a scene in which there are two characters (this was not the first mention of the characters in the entire narrative, but their fist appearance in this scene). The references to the two characters are thoroughly interleaved. However, gesturally, the references were separated into spaces:

Center Space (male character)	Right Space (female character)
	she picks up a palette
	and Ø attempts to do a painting
then what he does is	
he steadies her hand	
	then she signs the painting
and he gets her a drink	
and Ø starts to play the piano	

Speech and gesture are co-expressive in this passage in two ways. The linguistic forms referring to the two characters show a progression from more explicit to less explicit, and this progression takes place separately in each space. Moreover, entering or reentering a space was the occasion for upping the explicitness of the referring form; conversely, remaining in a space was the condition for not upping explicitness or for reducing it. In the female character's space, the referring forms were "she"—"Ø"—"she". The first two references directly followed one another, without intervening reference to the male character. The zero subject was then followed by "she," an increase of explicitness, and this was after the male character's space had intervened. Similarly for references in the male character's space. The sequence was "then what he does is"—"he"—"he" "Ø", that is, decline of explicitness. The highly overmarked reference, "then what he does is," where the referring form is a full clause, created a major discontinuity at the first mention of the male character and the first time his space was used. This was just after the double use of the female character's space. The speaker was introducing the spatial dichotomy itself as well as the new character. The highly explicit marking of the referring form for the male character, which at first seems excessive, is in fact appropriate for the new space after the well established other character's space. The rest of the references to the male character followed the pattern of decreasing explicitness. The second "he" came after an excursion into the female character's space, and was thus a reintroduction, and was directly followed by the least explicit reference to the male character, zero.

B. The Georgian narrator made a similar cohesive use of space. In the following example from McNeill (1992), the narrator was describing how the film maker had cleverly linked two scenes through the use of form, space, action, and time. One character (a young woman) has killed another (a none too savory male). In horror after the deed, the female character wanders the streets of London and comes across repeated visual reminders of what she had done. The final one of these is a drunk lying on the street with his hand stretched out in the just same posture as the dead man's. Hitchcock cleverly transforms this image into a second image of the victim's landlady at the precise moment she discovers the dead man; the impression is that the landlady made her gruesome discovery just as the young woman came across the drunk. The cinematic transformation was described by the narrator as follows (English translation by Kevin Tuite): "At this point a scream is heard from her, but this scream changes suddenly. The scream continues but the face changes and you see that landlady, who is standing exactly

the same way as this woman stood in the street, and she is looking at a hand, but this is really the dead person's hand."

The narrator's gestures alternate between the left side of space, which she sets up as the scene in the street with the female character, and the right space, which is the scene at the flat with the landlady. The gesture space was thus divided into old scene/new scene and the transformation of the old into the new was visualized as a piece-by-piece shifting of the left gesture space over to the right gesture space. The narrator first shifted the scream, next the face, then the posture, and finally the hand itself. The narrator first sets off the scream on the left (brackets show the approximate—since this is a translation—lexical accompaniments of the gesture stroke; Kevin Tuite's transcription and translation):

at this point [a scream is heard coming from her]
Left hand points = scream.

The narrator holds the scream on the left, and contrasts it to the right space, which is to be the transformation of the scream:

but [this scream] [changes suddenly]
Left hand holds in midair.
Right hand swings forward and to the right.

Now the narrator reactivates the scream on the left, while holding the new scream on the right:

[the scream continues]
Right hand holds in midair.
Left hand beats.

Next she starts to transform the face, again moving between left and right:

but [the face changes]
Left hand holds in midair.
Right hand swings forward and right again.
as [this woman stood in the street]
Right hand held in upright position.
Left hand loops down and to the right.

From now own, in perfect accord with the left-right division of the space, all gestures with both hands are performed in the right space, the space of the new scene. The point of view has become that of the landlady, no longer that of the young woman, and the space shifts to the right. The right hand is shown looking down, conveyed by the pointing left hand, all in the right space:

[and she's looking] at a hand

Right hand held in upright position.
Left hand points down in right side space.
but this is[really the dead person's] hand
Right hand still held in upright position.
Left hand again points down in right space.

The gestural solution to the problem of describing a complex cinematic transformation was utterly spontaneous. It was in no way modeled by the film itself. The successive shifts took place smoothly as the links between the scenes came up in the narration. The effect was a cohesive joining of references to the two scenes even as one of them was being transformed into the other.

C. We find cohesive and discontinuous uses of space at the metanarrative—story about the story—level as well. The following excerpt comes from another narrator's retelling (in English) of the Hitchcock film:

[it- it's sort of a fade-out]
Metaphoric: fade-out (curly & uncurl hand) as hand moves from left side to center space.
y'know Frank obviously made stalks off
and then the next time we see [anyone] [involved ...]
Deictics: point down and to the right.
Frank and Al- [not Frank]
Metaphoric: negation (closing hand).
[the artist] [and Alice] are [walking by]
(1) and (2) Deictics: point first to right and then to center space.
(3) Iconic: walking from center to left (opposite direction).

This illustrates how space is used to generate discontinuity on the metanarrative level. The metaphoric gesture for the fade-out of the preceding scene was made by the hand moving from the left to the front space; this is a metanarrative reference to the film's structure. By contrast, the new scene is introduced to the narrator's right; this is a second metanarrative level reference. The metanarrative space thus involved a contrast between front and right. The new space on the right, meaning the new scene, was then used for the next narrative level description of the characters walking by. Discontinuity on the metanarrative level was marked linguistically, as well as gesturally ("and then next time we see anyone involved ..."), confirming the gestural marking.

D. An example of cohesive use of space in a different narrative context is taken from a video recorded conversation, where one participant uses space to differentiate himself from an unwelcome topic introduced by the other participant. The participants were previously unacquainted male graduate students at the University of Chicago. The tape was made as part of a research project on face-to-face interaction by Starkey Duncan and Donald Fiske (Duncan & Fiske, 1977). The participants were simply brought together before the video camera and told to have a conversation. This command typically elicits a series of maneuvers during which the speakers try to find a common

theme about which they can talk. The example we are considering is from this exploratory phase; it ends with the participants discovering their workable theme, but only after a cat-and-mouse game that is the substance of our example. In conversations, no less than narrations, the predominant gesture is pointing when a new theme is being sought (Levy, 1982, 1984); and so it was here. One of the participants, H, was rather frantically trying to find something that could serve as a topic of conversation. In the way of students, he asked where the other participant, O, had gone to school before. O, however, mysteriously held back; he seemed to be avoiding the question of where he had been an undergraduate. His evasiveness led to the cat-and-mouse game where H tried to find out, and O tried to hide, this biographical datum. Space was an important factor throughout this whole episode. H returned over and over to the space he established for the topic, "where O had gone to school." O for his part avoided this space, first pointing above it, later pointing in the opposite direction. Clearly O was responding to H's use of the space, by steering clear of it at all costs. In this sense, through their proffering and avoiding of the same space, the two participants jointly created a cohesive space that governed their individual gestural performances.

The excerpt begins with two rapid questions from H:

H: is this your first year [here]?
Points down to own space.
or [where did you] come from before?
Points forward into shared space with O and draws small circles.

The first point marks "here" as the very place where H is located. It could mean this room but probably means the University of Chicago, where the experiment was being conducted and where H and O were students. The second point contrasts this "here" with a new space that introduces the space of "before." The significant fact about this new space is that it is the interaction space, shared by H and O (viz., the space where their gazes intersect, the participants facing each other at an angle). It thus helps to constitute H's idea of a shared topic of conversation. However, O replies noncommittally:

O: um Iowa. I lived in Iowa

This sends the conversation into a brief exploration of Iowa as a possible topic, but that theme quickly peters out. H then returns to the idea of O's past:

H: did you [go to] school [there] or uh?
Points both times into shared interaction space.

Thus H is consistently using the shared space for the meaning, "where you came from before." O answers:

O: I did go to school [there]
*Points **above** shared interaction space.*

O gives in to H's siege to the extent of pointing in the direction of the interaction space but points above it, as if to signal that while his "there" was indeed "where he came from before," this is not his idea of the topic of conversation. This was elaborated upon in O's next gesture:

> **O**: [I went to school] here also
> *Points to **left** periphery (i.e., **away** from H and the shared space).*

Clearly, this "here" is not H's "here." Indeed it is as far removed from H's "here" as possible. We don't suppose that O was aware of tipping his hand (so to speak), but his gesture reveals that he had divided the space differently from H: the *left* side was to be his "here" and this excluded the "here" of H. The drama had a happy ending: shortly after this critical moment O capitulated and revealed that he had been an undergraduate at a Jesuit college in Chicago (so his "here" was the *city* not the *university*, and it was this distinction that the left space had been assigned to carry). It then marvelously turned out that H also had been a student at a different Jesuit college, and so they had their theme after all—Jesuit education, a topic about which, as it turned out, O was the more enthusiastic!

E. Our final example of cohesion through space is a "repair" of spatial cohesion, which leads to an otherwise unmotivated linguistic choice. The example thus demonstrates that verbal output is influenced by gesture space. Initially, gestures establish a locus for Sylvester (the cat) to the right and for Tweety (the bird) to the left:

> OK [cat] [trying to catch the bird]
> *(1) Hand moves right (= the cat).*
> *(2) hand moves left (= the bird).*

The left-right separation is preserved in the next clause, even though the order in which the characters are mentioned is reversed:

> [Tweety Pie] and [Sylvester]
> *Moves left (= the bird).*
> *Moves right (= the cat).*

Next, Granny is introduced and she also is on the left (Tweety's side, as is appropriate as Granny is Tweety's owner and ally against Sylvester):

> with [his owner] an older woman
> *Moves left (= the bird).*

These spatial divisions may have appeared initially because of the way that space was represented in the cartoon itself (McCullough, 1992). But the gesture space quickly loses the meaning of the concrete cartoon space and takes on more abstract referential values. Thus, during the second and third clauses above, left vs. right had become noniconic and functioned solely to distinguish the characters. The main point of the example is the

following repair. The speaker next explains about Sylvester's locus (his "operation") and accompanies this by a gesture to the *left* (i.e., Tweety's side):

and uh Sylvester the cat's [operation]
Moves left (= the bird).

This space shift does not mean a thematic discontinuity; it is treated as an error and is immediately repaired both gesturally and linguistically:

[base of] [operations] is across the street
*Moves **right**.*
*Moves **far to right**.*

The linguistic repair of "operation" with "base of operations" seems trivial and unmotivated until we take into account the gesture space. It is the reversal of the spatial arrangement of the characters that required the repair, not the choice of word. This is confirmed by the fact that the second gesture in the clause above, accompanying the repetition of "operations," was an exaggerated rightward movement. Thus, cohesion in space was repaired and this led to an otherwise unmotivated linguistic choice.

Cohesion of Handedness

A. In a well constructed discourse, subordinate clauses that precede main clauses should be more cohesive with their context than the main clauses they accompany (these subordinate clauses should reflect information that is presupposed from the earlier discourse). We therefore expect that more complex gestures should occur with main clauses than with this type of subordinate clause. The following example from one of the cartoon narrations indeed shows higher gestural complexity with the main clause:

and ... [as he's com]ing up and
One handed iconic: right hand rises straight up from lap into center space, for Sylvester coming up.

the [bowling] ball's coming down
One handed iconic: left hand drops down from head level into center space, for the bowling ball coming down.

he [swallows it]
Two coordinated hands iconic: left hand moves into opening formed with right hand, for the bowling ball entering Sylvester.

That is, during the two subordinate adverbial clauses ("he's coming up," "the bowling ball's coming down"), there were one-handed iconics. During the main clause ("he swallows it"), there was a coordinate two-handed iconic where one hand changed its shape. The subordinate clauses refer to information carried forward from previous clauses but the main clause doesn't convey presupposed information, and plausibly this

clause does more to "push the communication forward" (Firbas, 1964, 1971). It was indeed accompanied by a change of gesture handedness—from one hand, which had been the norm, to two. The very complexity of this gesture could add to the communicative dynamism of the discourse at this point.

B. A second example illustrates the same point. The source is the same conversation cited earlier. The atmosphere has by now cleared, the participants are comparing impressions of the University of Chicago to other institutions they know about, and wondering what effect the large number of graduate students has on school spirit. First, there is a string of right handed gestures:

I wonder if one of th- [the] [wonder wa- ya know one of the reasons why it
(1) *(2)*
may be] like this is that- wouldn't [be like this] if your were [doing] graduate work
(3) *(4)*
at say [Madison or Champaign or someplace like that]
(5)

(1) —(5): Right handed gestures all in right periphery.

The extended subordinate clause, "if one of th- ... or someplace like that," was accompanied by 5 right-handed gestures all made at the periphery. This is followed by another extended series of clauses making up the main clause, "because ... the students." The gestures in this case were made with the left hand in center space. Note that handedness and space work together. The consistent use of one hand/space followed by use of the other hand/space, keeps central and peripheral themes apart; successive gestures with the same hand and space links clauses to other clauses having the same thematic status. The main clauses are the following, and all but one was made with the left hand in the central space:

um [because] [like] ['bout eighty] per[cent of the people who're in]
(6) *(7)* *(8)* *(9)*
[the area] [graduate students] [or seventy] five percent ya know
(10) *(11)* *(12)*
[an' ya don' have that] [undergraduate element] [which is] [you know like a lot of]
(13) *(14)* *(15)* *(16)*
the students

(6)—(16): Left handed gestures all in center, except for (8), which was a right-handed gesture in the center, and (16), made with both hands in the center.

The right-hand "exception" at (8) in fact proves the left-hand rule. The meaning of "'bout eighty" is that of approximation, a more peripheral notion for which the right hand, here the hand of the presupposed, may have seemed the more appropriate choice. A similar difference might explain the activation of the right hand at (16). The "which is like a lot of the students" is a non-novel gloss on the information given earlier, and again the right hand may have seemed essential for the meaning.

C. A final example of cohesion by handedness is again taken from the conversation. The example comes from a single speaker and was accompanied by a continuous tattoo of beats made with the right hand. The beats continued until there was a 25 sentence interlude, during which both speakers contributed and our target speaker did not gesture at all. After this break the speaker returned to his theme, which was Chicago's grimness, and the tattoo of beats with the right hand likewise resumed. Beats thus marked the presentation of the thematic idea and may have helped to keep it alive as a reference point The theme begins in the following extended stretch:

[I don't know it jus' seems at a place like this there's so much concentration
(1)

on one] [a- area] [well just on yer] [yer mind you know yer] [intellectual
(2) *(3)* *(4)* *(5)*
development] [an' there's there's so little emphasis and ya lose so much- I think as
(6)
a- I guess I] sound like a [Jesuit product but] [you lose so]
(7) *(8)*
(Interlocutor: the whole man)

[much so much of y- ya know i- i- in yer yer whole being ...]
(9)

(25 sentence interlude)

it jus' [seems everything's so serious] [and so concentrated]
(10) *(11)*
(1)—(11): Right hand continuously beats.

The only exception to the stream of beats was a retraction just prior to (7), as the speaker said "sound like a Jesuit product." This statement in fact fits the rule set up for this passage, since the statement was a side comment on the speaker himself, and departed from the beat-marked theme of the University of Chicago.

Cohesion of Form

We have thus far shown how gesture space and gesture handedness can mark cohesion in narration; we now do the same with gesture form. Gesture forms recur, often across considerable stretches of discourse text, and mark the reappearance of a particular plot element, character, or narrative value.

A. A striking example of cohesion of this kind is the following, in which the narrator highlighted what, to him, was the "leitmotif," "crux," or "essence" of the full-length Hitchcock film cited earlier. (This is yet another speaker.) The gestures in

question have an unusual form, an 8-hand: the thumb and middle finger in a ring, the other fingers extended. Chance repetitions of this unusual handshape are highly unlikely. The first appearance of the 8-hand was well into the narration, while the speaker was describing the psychological consequences of the murder for the young woman, her confusion and terror as she hears people talk about the crime, although she is not suspected of any connection with it. The following shows all occurrences of the 8-hand (curly brackets show the clause numbers; the transcript and gesture analysis are the work of Kevin Tuite):

{239} [it echoes] ... [in Alice's] mind
*(1) Left hand moves to front and **forms 8**.*
(2) Left hand closes to a point and points right; then closes to a fist.
{240} [it goes knife knife] knife knife
*Left hand **reforms 8**; there are small beats on the 1st and 2d "knife."*
{242} Alice would you [... cut me a s-] [cut me a] slice of bread?
*(1) Left hand moves left and forward, **forms 8**; then closes to a fist.*
(2) Left hand uncurls to left and spreads open.

The narrator continues to describe the aftermath:

{250} everyone is asking for [news] [of] the murder
(1) Left hand retracts slightly and opens to form the OK shape.
*(2) Left hand moves forward and down, **opens to 8**, then closes.*
Next, the blackmailer is introduced who proceeds to carry out his nefarious scheme:
{267} [on the pretext] [of making a tele]phone call to New Scotland Yard ...
(1) Left hand retracts to right and closes.
*(2) Left hand **forms 8**, palm right, moving to left; a beat on "tele," and then hold until the pause.*
{268} and uh [takes Alice into the phone] [booth ...]
*(1) Left hand **opens to 8** .*
(2) Left hand opens and fingers extend forward.

The blackmailer, in the phone booth, shows Alice her own glove, which she had left behind in the artist's flat and which the blackmailer, poking around, has found:

{286} [this- I found] this is the murders man- a- murdered man's apartment
*Left hand moves forward and down, and **opens to 8**.*
{287} [this clear][ly ...] [is yours]
*(1) Left hand **forms 8**, then closes to a fist.*
(2) Left hand opens slightly, then retracts to adjust glasses.
(3) Left hand moves forward and left, then closes to a fist.

The next two clauses describe Alice's boyfriend, Frank, a police detective who has determined to frame the blackmailer (called "this murderer" by the narrator—a different murder for which the blackmailer has already been punished; being an ex-felon, he is vulnerable to Frank's frame-up):

{366} and Frank is pursuing this murderer with the intent [of making it] [look like]
*(1) Left hand closes to OK, then the lower 3 fingers uncurl to make a **variant of 8**.*
(2) Left hand closes and points.
{367} he was the man who killed the artist ...
{368} now ... let's see [okay]
*Left hand moves up from lap and **forms 8**.*

Frank's frame-up succeeds brilliantly: the blackmailer bolts, is chased, escapes into the British Museum, and,

{375} [then he falls] [backwards through] [the glass dome pfft!]
(1) Right hand points and arcs to the left.
(2) Right hand points and loops up and right, then left and down.
*(3) Right hand **forms 8** with palm facing down, the down stroke on "pfft."*

Afterward, Alice, tormented by conscience, is about to confess her crime:

{394} that [she is in fact] the murderer
*Left hand points and swings up and left; **opens to 8** and points to left with palm forward.*

But events interrupt her and she never does.

So, what is the crux, essence, or leitmotif of the film, according to this narrator? It seems to be the blackmailing of Alice and the framing of the blackmailer: that is the moral crux. First, the 8-hand marks the aftermath of the murder in Alice's mind—her vulnerability, guilt and terror; then the blackmailer's presentation of the glove, the only thing that links Alice to the crime; then Frank's plan to turn the tables on the blackmailer. The crux expands at the end to include the consequences of all this wickedness: at {375}, the moral irony of the blackmailer's demise and, at {394}, Alice's final getting off of the hook, the film's ultimate moral irony. The 8-gesture connects clauses to this crux across a huge span—more than 150 clauses. The form of the gesture is not inappropriate to the concept of a crux, because it suggests precision but differs from the conventional OK sign (made with the forefinger contacting the thumb); this fits with the meaning being somehow the "essence" but lacking any approbation. The implicit contrast of the 8-shape with the OK sign suggests that the latter, although never actually appearing, had an influence on the formation of the 8-shape by the narrator (for emblems in general and the specific interpretation of the OK sign in terms of precision, see Morris, et al., 1979; for extensive discussion of emblems, including the Morris, et al. survey, see Kendon, 1981).

B. A similar case of cohesion achieved through form is shown in the Georgian language narration, where the speaker restarted her narration after interpolating bits of quoted dialogue and a lengthy side comment. Each of the restartings was accompanied by the same words and same gesture, the H-shape (first 2 fingers extended, the other fingers curled in). During the quoted dialogue, during the first interruption, a different gesture form was introduced (the conduit metaphor cup-of-meaning shape) and it was

also repeated. Thus, two cohesive chains were intertwined, held together internally and discriminated externally by gesture form (again, the English translation by Kevin Tuite):

[... and they are talking to each other]
Right hand in H-shape moves left-to-right twice: shows reciprocal activity and dual number.

Now the interpolated quoted dialogue commences:

this man [asks her]
Right hand uncurls into conduit shape: shows the man's question.
uh [come down- come on over] to my place
Right hand in conduit shape.
[have you ever seen] a painter's studio?
Right hand in conduit shape.
[this] woman says: no I haven't seen one
Right hand in conduit shape.
[well, come and see!]
Right hand in conduit shape.

Next the narrator resumes the scene description, by returning to the earlier words and H-gesture:

[while they are talking to each other]
Right hand resumes H-shape for the two people talking.

The H-shape chain went much further. The same words and gesture for the idea of reciprocity and duality occurred again after a 25 sentence break, during which the speaker did many other things, including introducing another character and discoursing on the social class membership of the various characters:

... um [at this time these] two are talking with each other
Right hand takes on H-shape again.

C. The Georgian narrator provides one last example of cohesion by form in her use of moving her two hands back and forth in antiparallel directions. She used this movement to depict the idea of moral opposition. The movement seems to have been a metaphor for the idea of opposition: two sides, face-to-face, good vs. evil. The narrator introduces the concept of opposition along with antiparallel movements in the first example below; then she presents successively the two sides of the opposition and each side is accompanied by just one of the two antiparallel movements. Finally after 6 clauses, during which the narrator describes the hero's and villain's physiognomies but doesn't refer to the conflict of good vs. evil, she returns to the opposition theme and again performs the original antiparallel movement. This movement, including its factored versions, thus covers some 10 clauses:

{53} here they are [standing on two sides, face to face]
Metaphoric: both hands move back and forth in antiparallel directions.

This introduces the concept of opposition and the antiparallel movement. Next, she highlights one side of the opposition and moves her left hand through just one of the two components of antiparallel movement:

{54} [on the one side]: the police, that is, justice, logic, order
***Left hand** moves back and forth, through one side of the antiparallel movement.*

Next, she mentions the other side of the opposition and moves her right hand through the other antiparallel component:

{55} [and on the other side]: the criminal, thief, killer
***Right hand** moves back and forth, the other side of the antiparallel movement.*

Now the six other clauses intervene, including one clause contributed by the listener. After these side comments, the speaker returns to the theme of opposition together with the same antiparallel movement introduced at the start:

{62} that is, [justice and guilt]
Resumes original metaphoric antiparallel movement.

Summary

These examples illustrate the gestural marking of continuous information in discourse. In all of the examples, gestures first mark the existence of a referent, and later help to maintain reference to it. In so doing, they help to differentiate the initial referent from other possible referents in the context. Reference is maintained through cohesiveness of space, handedness, and/or form, including style of movement. In some cases, gestures serve a second discourse pragmatic function as well: at times they contribute to the marking of information as central or peripheral with respect to the larger goals of speaking.

Discussion

On-line construction of discourse. An advantage of an approach to discourse context that focuses on events taking place at the moment of speaking—as opposed to hierarchical approaches in which surface devices are assumed merely to reflect discourse units that are set up prior to the act of speaking—is that the former applies to discourse production that is not planned in advance, as when a topic is negotiated by two or more conversational partners (see Ochs, 1979, on planned vs. unplanned discourse). From the perspective of the real-time processes involved in the production of speech, an emphasis on the moment of speaking applies to those instances of monologic speech where the topic is negotiated by a single speaker in the course of speaking itself (Bakhtin, 1981; Wertsch, 1985; Levy, 1987; Özyürek, 1992).

Earlier we suggested that the continued use of gesture may help the speaker keep the discourse theme alive. This process of continuing the theme may help the speaker him/herself keep track of what has already been said, and to contribute a sense of nonclosure to stretches of discourse that are cohesively tied together. That is, the execution/maintenance of a gesture may help the speaker to track the presupposed background information; this then provides a basis for the construction of the "communicatively dynamic" or new part of the utterance that we discussed in our earlier paper (Levy & McNeill, 1992). In the example of a scream serving as a scene-transition, the narrator held her left hand in midair to maintain the reference to one character, while her right hand began to move to represent the shift to the second character.

This perspective—the use of gestures as place-markers for the speaker—ties in with Vygotsky's (1934) view that thinking and speaking are interdependent activities. Not only does thinking influence the emergence of speech, but speaking influences the development of thinking as well. While arguing against the view that speech is a mere reflection or external expression of thinking, Vygotsky remarked that meaning does not exist and develop in complete isolation from its "material carrier." Rather, speaking itself is used to construct meaning. More generally, it follows that discourse is used in the process of constructing a theme or point of view.

Our argument in this paper extends Vygotsky's concept to the speech/gesture synthesis. Not just speech, but *speech and gesture* are used jointly to work out a plan and goal; that is, far from discourse being planned in advance, a plan can be created, in significant cases, by speaking and gesturing itself (Levy, 1987).

As the stream of discourse goes on, later points are influenced by earlier points. Informational continuity requires such influence. We propose to analyze the ongoing discourse as a kind of internal dialogue within one speaker, the "voices" of which are given distinct material carriers—speech and gesture. Bakhtin (1981) conceptualized the internal dialogue in terms of two semantic systems that play out in each sentence. One is the voice of the author, the other the voice of the character, the narrator, or other voice that helps constitute the dialogue inherent in the text. The interplay of the voices is an essential quality of the narrative, that is, what constitutes it as a narrative. One voice affirms, evaluates, or contradicts the other (Özyürek, 1992). For example, one of our narrators began to describe a scene from the cartoon, then interrupted herself: "the last one he tries is [to walk] across ... part of the problem is that Tweety [Bird is inaccessible]." The first gesture, before the interruption, was an iconic that depicted the "surface" on which Sylvester was attempt to walk (namely, overhead trolley wires), and the second, during the interruption, was a metaphoric gesture depicting the concept of inaccessibility as it applies to the situation of the story (one hand held up and to the rear, for Tweety on an upper story in his cage, the other hand held below and near the lap, for Sylvester earthbound). In Özyürek's analysis, there are two voices; before the interruption was the narrator, the interruption itself was the author or some other omniscient observer. The latter was evaluating and explaining the former. What is significant here for our purpose is that the shift of voice was accompanied by a shift of

gesture. The gesture's contribution is a voice of the Bakhtinian type. In keeping with Vygotsky's remark above, each voice has its own material carrier, both the utterance and the contrasting gestures. The words are framed by the gesture in a different way. That is, the gesture of the interrupting voice contributes uniquely to the meaning of the discourse. Conversely, when the speaker uses gesture for cohesion, and maintains the same gesture space, handedness, or form, we say that the same voice continues to be heard over time. The Bahktinian concept of dialogic voices, joined to the Vygotskian concept of the material carrier, thus provides an interpretation of the role of gesture in the creation of discourse structure. The gesture can provide one voice and speech another voice, and the relationship between them—the commenting on, affirming, evaluating or rejecting—moves the discourse forward.

We can apply this form of analysis to the examples cited earlier. In each case, we find an interplay of voices each with a separate material carrier, speech or gesture.

When the 8-hand marked the crux of the film, this precision image (distinguished, we speculate, from approbation) was part of the narrator's discourse meaning; the other parts were the various verbalizations that in themselves do not normally convey precision but yet, joined to the gesture, have the discourse meaning of the essence of the story (it echoes—it goes knife knife—cut me a slice of bread, etc.). The image of precision constantly reprised evaluates the verbal references—they are the crux of the film whereas, taken alone, they only recount events from the film but do not mark them as the crux. Thus, considered with the gesture shape, they provide the discourse meaning the speaker was seeking. This is an excellent example of what we (and, we believe, Bakhtin) mean by a dialogue within a single speaker with each voice having its own material carrier.

When another narrator reserved one part of space for the female character and another part for the male character, this spatial separation became part of the thematic structure. As residence in each space continued, the thematic identification of its particular character grew stronger, and the explicitness of the referring forms went down, in keeping with the increasing thematic status (Levy, 1984). In terms of voices, we have here mutual reinforcement, gesture space and the dimension of linguistic explicitness ratifying each other.

When one of the conversation participants shifted from the right hand to the left hand at the boundary between a long subordinate clause and an equally long main clause, we see how the gesture system is able to keep alive an extended relationship of topic and comment, while the speech system is free to roll out detailed referential specification. The words were framed by the hand choices and the blanket shift of hands at the subordinate/main clause boundary bracketed the rambling fragmented speech stream into clearly demarcated clumps of thematic material. This is yet another example of an internal dialogue, in this case one voice (the gesture) providing the continuity, and the other voice (the spoken utterance) the differentiation.

In these ways and many others (examples could be multiplied indefinitely), speech and gesture work together to create continuity and variation of information. We propose to represent the relationship of speech and gesture as a dialogue of voices. The voices have different carriers. One is global, synthetic, and imagistic; the other is linear, segmented, and linguistic (McNeill, 1992). The voices help constitute the speaker's thought processes at the moment of speaking. Together with our earlier paper (Levy & McNeill, 1992), we have attempted to show something of how gesture and speech collaborate to construct the ebb and flow of continuity and discontinuity in discourse. As discourse proceeds, the two forms of meaning presentation together create and maintain topics. As Vygotsky (1934) remarked , the speaker may need surface forms, material carriers, to help master the flow or his or her own mental processes.

Other forms of cohesion. In this paper we have concentrated on continuity of reference; that is, on the continuity of characters and objects in the real or fictional world. Other forms of cohesion also exist and elsewhere we have analyzed the gestural marking of `event lines' or `participant frames' (cf. Hanks, 1990). Event lines establish possibilities of cohesion on different organizational levels (Cassell & McNeill, 1991), and these too can be interpreted as an dialogue of voices each with its own material carrier. Metaphoric gestures, for example, are used selectively to indicate the metanarrative level—the story about the primary story—and often little else in the spoken form of discourse carries this information. Each event line thus has its own voice and its own universe to add to the overall cohesion of the discourse.

Acknowledgments

Preparation of this paper was supported by grants from the Spencer Foundation and the National Institute on Deafness and Other Communication Disorders. Our research on gesture and language has been supported in the past by the Linguistics Program of the National Science Foundation. We would like to acknowledge the support of the Haskins Laboratories, New Haven, CT. We thank Adam Kendon and Talmy Givón for comments on an earlier draft.

Note

(1) The concept that images (or the analogic representations behind them) are actually part of language, processed *together* with the hierarchical, linear, segmented parts, suggests that, over historical time, there would be produced a certain amount of linguistic iconicity. By `linguistic' we mean those elements of language that are socially constituted and maintained in standards of well formedness for the language. Constituting and preserving standards would presumably be easier if the standards don't fly in the face of major patterns of thought at the moment of speaking. This convergence of standards with the moment-by-moment course of processing could take place if, for example, language systems try to anticipate certain salient and highly regular imagistic thought patterns, and this is our interpretation of at least some aspects of the `iconicity in syntax' argument (Haiman, 1985). For example, to maintain local proximity conditions, such that closely related semantic elements tend to be sequentially close in surface syntax

(Givón, 1991; Bybee, 1985), could be seen as a socio-cultural regularization of predictable mental processes based on imagistic spatial representations

References

Bakhtin, M. M. (1981). *The dialogic imagination* (M. Holquist, ed.). Austin, TX: University of Texas Press.

Bybee, J. (1985). *Morphology.* Amsterdam: John Benjamins.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Cassell, J. & McNeill, D. (1991). Gesture and the poesis of prose. *Poetics Today*, **12**, 375-404.

Danes, F. (1974). Functional sentence perspective and the organization of the text. In F. Danes (ed.), *Papers on functional sentence perspective* (pp. 106-128). The Hague: Mouton.

Duncan, S. & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory.* Hillsdale, NJ: Erlbaum.

Firbas, J. (1964). On defining the theme in functional sentence analysis. *Travaux linguistiques de Prague*, **1**, 267-280.

Firbas, J. (1971). On the concept of communicative dynamism in the theory of functional sentence perspective. *Philogica Pragensia,* **8**, 135-144.

Firbas, J. (1974). Some aspects of the Czechoslovak approach to problems of functional sentence perspective. In F. Danes (ed.), *Papers on functional sentence perspective* (pp. 11-37). The Hague: Mouton.

Givón, T. (1985). Iconicity, isomorphism and non-arbitrary coding in syntax. In J. Haiman (ed.), *Iconicity in syntax* (pp. 187-219). Amsterdam: John Benjamins.

Givón, T. (1991). Isomorphism in the grammatical code: Cognitive and biological considerations. *Studies in Language,* **15**, 85-114.

Grosz, B. J. (1981). Focusing description in natural language dialogues. In A. Joshi, B. L. Webber, & I. Sag (eds.), *Elements of discourse understanding* (pp. 84-105). Cambridge: Cambridge University Press.

Haiman, J. (1985). *Iconicity in syntax.* Amsterdam: John Benjamins.

Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hanks, W. F. (1990). *Referential practice: Language and lived space among the Maya.* Chicago: University of Chicago Press.

Kendon, A. (1972). Some relationships between body motion and speech. In A. W. Siegman & B. Pope (eds.), *Studies in dyadic communication* (pp. 177-213). New York: Pergamon Press.

Kendon, A. (1981). Geography of gesture. *Semiotica*, **37**, 129-163.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Levy, E. T. (1982). Toward an objective definition of "discourse topic." In K. Tuite, R. Schneider, & R. Chametsky (eds.), *Papers from the 18th Regional Meeting of the Chicago Linguistic Society* (pp. 295-304). Chicago: Chicago Linguistic Society.

Levy, E. T. (1984). *Communicating thematic structure in narrative discourse: The use of referring terms and gestures.* Doctoral dissertation, Department of Behavioral Sciences, The University of Chicago.

Levy, E. T. (1987). A Vygotskian perspective on discourse: From complex to concept. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition* (University of California at San Diego), **9**, 100-105.

Levy, E. T. & McNeill, D. (1992). Speech, gesture and discourse. *Discourse Processes*, **15**, 277-301.

McClave, E. Z. (1991). *Intonation and gesture.* Doctoral dissertation, Department of Linguistics, Georgetown University.

McCullough, K.-E. (1992). Visual imagery in language and gesture. Paper presented at the Annual Conference of the Linguistic Society of Belgium on the theme of the relationship between linguistic and conceptual representation (November, 1992).

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, **92**, 350-371.

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

McNeill, D. & Levy, E. T. (1982). Conceptual representations in language activity and gesture. In R. Jarvella & W. Klein (eds.), *Speech, place, and action* (pp. 271-295). Chichester: Wiley & Sons.

Morris, D., Collett, P., Marsh, P. & O'Shaughnessy, M. (1979). *Gestures: Their origins and distribution.* New York: Stein & Day.

Ochs, E. (1979). Planned and unplanned discourse. In T. Givón (ed.), *Syntax and semantics: Discourse and syntax*. (pp. 51-80). New York: Academic Press.

Özyürek, A. (1992). Seeing the `dialogic angle' in gesture. Unpublished MS, Department of Psychology, University of Chicago.

Reddy, M. (1979). The conduit metaphor—A case of frame conflict in our language about language. In A. Ortony (ed.), *Metaphor and thought* (pp. 284-324). Cambridge: Cambridge University Press.

Richards, I. A. (1936). *The philosophy of rhetoric*. New York: Oxford University Press.

Vygotsky, L. S. (1934). *Myshlenie i rech': Psikhologicheskie issledovaniya* (Thinking and speech: Psychological investigations [MS], N. Minick, trans.).

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, transs.). Cambridge, MA: MIT Press.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.