

Gesture and Ground

Author(s): Justine Cassell and David McNeill

Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society (1990), pp. 57-68

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

GESTURE AND GROUND¹

Justine Cassell & David McNeill
University of Chicago

I. Introduction

Since the foreground / background distinction was first described (Hopper, 1979; Hopper & Thompson, 1980), it has retained widespread acceptance as a way of analyzing written and spoken texts. Much of the literature since then has taken the contrast more or less for granted and has refined it, changed the terminology, or simply added supporting evidence from additional languages and language families. For English discourse, however, adequate criteria for foregrounding vs. backgrounding have never been described.

We believe that (a) that the dyadic notion of foregrounding / backgrounding misses much of what is going on in narrative texts; (b) part of narrative structure may be carried in the non-verbal channel, as well as the verbal; (c) for English narrative the majority of ground distinctions may be carried by gesture, thus explaining the difficulties encountered in trying to isolate one linguistic system that carries information structure in narrative.

In this paper we will first discuss the distinction between foregrounding and backgrounding as it has been described in the literature. Then we will go on to replace the notion of ground with one of participation frameworks or **event lines** in information, and **narrative level** in speech. And finally we will demonstrate the role that gesture plays in the structuring of narrative.

II. The Concept of Ground

In extended narrative texts, it has been claimed, a distinction is always made between the 'story line' and the 'supporting material'. This distinction has alternately been stated in terms of the material that a narrative is about (sometimes known as the **fabula** in Russian Formalist literature), and in terms of the language in which the narrative is couched (sometimes referred to as the **syuzhet**). That is, Hopper & Thompson (1980) say:

That part of a discourse which does not immediately and crucially contribute to the speaker's goal, but which merely assists, amplifies, or comments on it, is referred to as **background**. By contrast, the material which supplies the main points of the discourse is known as **foreground**. (1980:280)

and Hopper (1979) says:

It is evidently a universal of narrative discourse that in any extended text an overt distinction is made between the language of the actual

story line and the language of supportive material which does not itself narrate the main events. (1979:213)

The sort of **material** which is considered foregrounded is events which occur in a temporal order -- that is, punctual non-overlapping events -- and which advance the listener's knowledge of the plot. This resembles Labov's definition of a minimal narrative text as being "a sequence of two clauses which are temporally ordered" (1972: 360). Backgrounded material can be states which do not engage in relations of sequence with events, and which comment on or amplify the main events.

As far as ground distinctions in **language**, the paradigm example of the sort of linguistic distinctions meant is Swahili. Swahili (Hopper, 1979) marks verbs denoting events on the main story line with the prefix *ka-* while events marked as subsidiary receive the prefix *ki-*. Thus, in a story analyzed by Hopper (1979) those events that are punctual, occur in the narrative in the temporal order in which they occurred in the world, and are not simultaneous with or subsumed by other events, are marked with the foreground marker *ka-*.

It is not entirely clear in the work by Hopper, and Hopper & Thompson, whether the concept of ground should be looked for in language or in the story events that exist independently of the discourse which discusses them. In this paper we are going to assume that it is possible to talk about events existing independently of the narrative that talks about them -- although of course these events may be fictive -- and we would like to separate information structure in the story from information structure in the discourse. This issue becomes most necessary when languages without clear morphosyntactic systems for marking information structure are examined.

III. Ground in English

English has repeatedly been the subject of discussions of ground, but never with much success at identifying a stable heuristic for separating foregrounded and backgrounded material.

Hopper & Thompson claimed that transitivity might mark the foreground-background distinction in English and other languages without overt aspectual systems. Their definition of transitivity, however, is not a strictly morphosyntactic one, but also includes semantic contrasts -- that is, it includes a distinction in the kinds of things talked about, and how those things exist in the discourse world. Transitivity involves a number of components: the aspect of the verb, the conscious activity of the agent, and the referentiality and degree of affectedness of the object. For this reason it is impossible to correlate transitivity as a linguistic distinction with a contrast in the kinds of things talked about -- the content criterion for ground. In addition, in practice, Hopper & Thompson's analysis of English texts does not show that transitivity distinguishes ground in this language -- a large percentage of background clauses in the discourse that they analyse have high transitivity values (as mentioned by Delancey, 1987).

IV. Content vs. Linguistic Criteria for Ground

In much of the literature, content criteria for ground have been confused with linguistic criteria (although see Reinhart, 1984). That is, it is unclear whether

what is foregrounded is information that will be referred to by a given clause, or the linguistic nature of the clause itself. In addition, content criteria for ground depend on notions such as 'pivotal events' or 'central and significant information' which contain a covert definition of narrative as concentrating on events in a narrow sense of the term -- excluded are the events of the storytelling, for example.

If, however, we are to look for linguistic devices that express the notion of ground -- which is, after all, the enterprise originally undertaken by Hopper -- then we must have both a non-linguistic distinction and linguistic devices that correlate with that contrast, and the notion of story events (outside of their linguistic instantiation) must be able to handle **all** of the kinds of events narrated in a story.

V. Event Lines

We substitute for the notion of foreground and background, the concept of 'event line' (Cassell & McNeill, in press). The distinction between foregrounding and backgrounding only makes sense in the context of a definition of narrative where single protagonist, single viewpoint events are primary. Events can be seen as foregrounded only if there is one narrator telling one story, and the plot of the **story** is identical to the plot of the **narrating event**. Ground implies a narrating event where the storyteller has only one goal in relating the story. If there are other event chains (looking at a cartoon, recounting the cartoon, representing the organization of the narrative, interacting with the interlocutor, etc), then there cannot be a single or simple distinction between 'foreground' and background'. In contrast, we use 'storytelling' or 'narrating' to refer to the entire set of events that make up the conveying of a story by one person to another. Each of these events is grist for the storytelling mill and may be referred to by the storyteller.

We study stories that are told about a cartoon or feature film just seen. For these narrators the following distinctions can be drawn: stories generally refer to what may be called *emplotted events*: incidents or occurrences that follow one another in a real or fictive world (e.g. a character climbs up a drainpipe to reach an upper story).

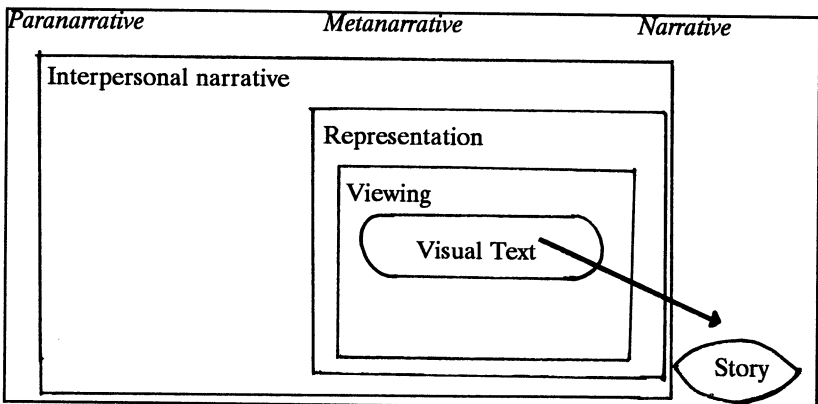


Figure 1: EVENT LINES AND NARRATIVE LEVEL

In addition to emplotted events, though, the stories we are concerned with also commonly contain references to the event of observing the visual text, or cartoon story ('it was a Sylvester and Tweety cartoon', or 'it was an old movie, a very bad print'), and/or to the event of the storytelling ('I'm going to tell you about a cartoon I just saw'). Reference to these 'metanarrative events' often acts in actual storytelling to create junctures between the parts of a story. The five event lines are listed in Figure 1.

As mentioned above, all of these event lines may be referred to in storytelling. Reference to the visual text, viewing, and speaker's own representation form the **metanarrative** level of the discourse: the part of the narrative that is about narrating. The interpersonal narrative is what we are referring to as the **paranarrative** level of the discourse: the part of the story where the narrator steps out and speaks in his/her own voice to the listener. These notions resemble the three kinds of mimesis described by Ricoeur (1984:52-87).

1. um have you seen any of the uh Bugs Bunny cartoons?	[PARA]
{yeah like}	
2. right, ok this one actually wasn't a Bugs Bunny cartoon	[META]
3. it was one of the- the series	[META]
{oh, ok}	
4. and it had Tweety Bird and Sylvester	[META]
{alright (laugh)}	
5. so so so you know	[PARA]
{the cat right?}	
6. right uh huh	
{ok}	
7. and uh the first scene you see is uh	[META]
8. this this window with birdwatcher's society underneath it	
9. and there's Sylvester peeking around the window	[NARR]

Figure 2: NARRATIVE LEVELS IN SPEECH

They are also listed in Figure 1. The beginning of one speaker's narrative reproduced in Figure 2 gives an example of each narrative level in actual storytelling.

What we call event lines comprise a clear content criterion for ground -- not in terms of the importance of an event but in terms of what aspect of the story is referred to -- what participant frameworks are evoked. Our narrative level, on the other hand, is a way of looking at the linguistic instantiation of ground. It has to do with discourse -- with the clauses that refer to a given event -- and they may be marked in one of a number of ways.

In the next section we are going to talk about how narrative level is marked in gesture.

VI. Gesture and Narrative Level

Narrators make reference to all three levels, and all 5 event lines during a typical story. Adults also demonstrate a typical pattern of gesturing correlated with reference to these different kinds of events (McNeill & Levy, 1982). It is this gesturing which may serve as the indication of ground, or narrative level. We're **not** talking about emblematic gestures (such as those that mean "a-ok", or "V for Victory"). Those emblematic gestures carry meaning in the absence of speech. The four types of gestures that we will describe, on the other hand, occur with speech. They occur most often with the stress peak of a clause, and they carry meaning insofar as they are understood in the context of the speech that they are produced alongside of.

We are claiming that when people tell stories they speak not only of the events of the story per se but also of the events of the representation of the story, and of the events of the narrating. And these categories, we claim, can be differentiated by the gestures that accompany them. In Figure 3 is reproduced the same narrative fragment as in Figure 2 but with the gestures noted². We will be making reference to these gestures in the next section.

(A) Gestures at the Narrative Level

The most common kind of gesture produced overall in narratives is the **iconic** gesture. Iconic gestures depict some feature of the action or event being described. Iconics are recognizable by virtue of their form and content. An example is in clause (9) in Figure 3 where the speaker rests her head on her hand and mimes looking left and right while saying "and there's Sylvester peeking around the window. Iconics occur most often with narrative speech. There are several kinds of iconic gestures, corresponding to viewpoint and perspective in narrative, and we will describe these further below.

So, narrative clauses are accompanied by iconics, while meta- and paranarrative clauses are not.

(B) Gestures at the Metanarrative Level

Storytellers tell about the act of narrating by making overt reference to the structure of the story being told, or to the structure of the narrating event. They accompany these references by another kind of gesture, the **metaphoric**. Metaphoric gestures, where the concept being depicted has no physical form, accompany

metanarrative speech, as do deictic gestures. Metaphoric gestures appear with references to the narrative story structure where the cartoon narration may be objectified and commented on in the verbal channel, and presented as an object in the gestural channel. Figure 3 gives an example of a metaphoric gesture: the speaker's hands are rolling over each other while she says, 'it was one of the series.'

2. [right], [ok this one] [actually wasn't] a Bugs Bunny cartoon beat beat beat	[META]
3. [it was one of the- the series] metaphoric	[META]
{oh, ok}	
4. and it had [Tweetie Bird and Sylvester] beat	[META]
...	
7. and uh [the first scene you see is uh] metaphoric-deictic	[META]
8. [this this window] [with birdwatcher's society underneath it] iconic iconic	
9. and [there's Sylvester peeking] around the window iconic	[NARR]

Figure 3: GESTURE AND NARRATIVE LEVEL

Deictics at this level spatialize, or locate in the physical space in front of the narrator, aspects of the story being narrated. An example in Figure 3 is the hand pointing at empty center space while the speaker says, "the first scene you see is uh".

Another kind of gesture accompanies metanarrative speech and other clauses with a metapragmatic function, and that is the beat gesture. Beats are small baton like movements that do not change in form with the content of the accompanying speech. As well as occurring with metanarrative speech, beats also accompany speech repairs and reported speech. The semiotic value of a 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. An example in a metanarrative context are, in Figure 3, the beats that accompany each segment in 'right], [ok this one] [actually wasn't a Bugs Bunny cartoon.' Beats can be seen as signalling some kind of move away from the main temporal axis of the story. They reveal, however, that this move can be in the context of a single word -- as in a

repair -- or an entire segment of the discourse -- as for metanarrative speech.

The narrative excerpt in Figure 4 is from the beginning of another cartoon narration by an adult speaker. It demonstrates how gestures function to indicate the kind of background narrative information that is given in the clause they accompany.

The narrator here forms a largish bowl shape with his two hands and raises this bowl from his lap towards the listener. This is a conduit metaphoric gesture: the speaker presents information as if contained in a bounded object that can be passed to the listener (cf. Lakoff & Johnson, 1980; Reddy, 1979). In this example, speech and gesture work together to make clear the metanarrative level at which the narration begins. The speech presents the cartoon as an example of a type, while the gesture represents the narrator, not as an observer, but as the conveyor of this object which is the cartoon. The narrator also performs beats to accompany the introduction of the genre ("it was one of the cat and bird cartoons").

<p>Well ++ [it was one of the ahm Tweetie Pie and ahm] ++ the beats</p> <p>cat ++ [cartoons] metaphoric: hands present cartoon to listener.</p>

Figure 4: CONDUIT METAPHORIC GESTURE

Conduit gestures like the one just described are not found in narrations in all languages. Whereas we have found excellent examples of conduits in a Georgian language narrative, we have none in Chinese or Swahili narratives. These narratives do contain abundant metaphoric gestures of other kinds, but do not depict abstract ideas as bounded containers. The metaphoric use of space, however, appears in all narratives regardless of the language spoken.

(C) Gestures at the Paranarrative Level

To go on to the third narrative level, when narrators speak as themselves outside of a narrative situation but adopting the role of a participant in a socially defined situation of speaker and hearer, they make only a small number of gestures of a restricted kind. Iconic gestures are virtually absent, as are metaphoric gestures. Deictic gestures are found, to point out the participants of the event: in this case the narrator and listener. This is similar to the findings of Goodwin & Goodwin (to appear), who have observed that gaze plays an important role in structuring the participation of a speaker and hearer in narrating events. Beat gestures are also found, when they mark the **inception** of a paranarrative portion of the discourse, or when they signal repairs or other metalinguistic work. An example of a deictic at this level is pointing towards the listener while saying '[right] uh huh.' Pointing at the listener in paranarrative contexts seems to be a specific clue for this level of narrative organization, opposing it to the narrative and metanarrative levels. The paranarrative point indexes shared knowledge, and the shared activity of narration.

Thus we find a distinctive gestural pattern corresponding to each of the narrative levels: for narrative level, iconics; for metanarrative level, beats, metaphoric, and pointing at empty gesture space; for paranarrative level, pointing

This is manifested directly in the form of the concurrent gestures. The character's focalization narrates events in close-up, as it were, and the C-VPT (character viewpoint) gestures depict this directly. The more peripheral events are narrated at a distance, and the O-VPT (observer viewpoint) gestures present this. Gestural focalization is inferred from the form and space of the iconic gesture. We infer the character as the focalizer when the depiction enacts the character and his movements: the narrator's hand plays the part of a character's hand, her body the part of this character's body, etc. The gesture space includes the speaker's own body as an integral part. An example from Figure 3 is the gesture previously discussed: "and there's Sylvester peeking around the window"

Conversely, an O-VPT gesture concentrates the entire event in the speaker's hand. The hand equals the entire character, its movement is the motion of this character, not just one part of him, and the gesture space is localized in front of the speaker, excluding her own body. The speaker is still part of this gesture, but now as an observer of the event. An example is given in Figure 6.

"he tries going [up the inside of the drainpipe and]"

O-VPT iconic: right hand forms a fist with forefinger extended and moves straight up. Narrator's hand depicts the character's body.

Figure 6: OBSERVER FOCALIZATION

We may use gesture as a clue to the function of accompanying speech. The character voice tends to appear with single-clause sentences that use, where possible, active, transitive verbs and minimal subjects: An illustration is 'and θ / drops it down the drainpipe' which was accompanied by a C-VPT gesture. The O-VPT conversely tends to appear with multiple-clause sentences. Corresponding to the distance built into the O-VPT gesture, distance is introduced in these sentences since one clause describes the action and the other clause may express a narrative observer. Another example with an O-VPT gesture is, 'and we see him go up the drainpipe,' where one clause indexes the observer's focalization. Thus, even though in English the narrative foreground is said to be not specifically indexed linguistically, one kind of layering occurs within the narrative foreground in gestural form, and this leads us to note an unsuspected parallel in syntax.

(b) Perspective

The observer may stand either outside the event as an onlooker on the event, or may stand inside the event but not as a participant in the action. An example is using a hand to show a figure swinging past the narrator herself, from back to front, while saying, 'and you see him swinging down [across a rope].' Unlike a C-VPT gesture, the gesture depicts the character as a whole entity, not just the hands of the character. However, as in the C-VPT, the gesture space extends backward to include the narrator. In an outside perspective, in contrast, the space is in front of the narrator. We have not found parallels between gestural perspective and the syntax of accompanying speech. This gesture is shown in Figure 7.

1. Internal Perspective:

"and you see him swinging down [across a rope]"
iconic: fist swings from behind head towards listener

2. External Perspective:

"and he puts a [board on a box]"
both hands describe box and then board in center space

Figure 7: PERSPECTIVE WITHIN NARRATIVE

Within the foreground, then, we find a rich structure of subdivisions. The speaker's sense of the perceived centrality vs. peripherality of the events in a narrative takes on concrete form in gesture focalization and perspective. Narrative distance may be represented by the actual physical distance of a narrator from the narrated. These images do not only come out of, but also can have real ongoing consequences for the speaker's thoughts, and therefore for narrative and memory. Thus, for both the speaker, and for the listener, gestures help to build a representation of the narration, at all of its levels, and play an important part in the 'telementation' of the story (Harris & Taylor, 1989).

(B) Layering within 'Extranarrative'

As mentioned above, within the background we find a distinction between reference to the structure of the story (metanarrative), and reference to the structure of the narrative event itself (paranarrative). We can see this distinction as another example of perspective in gesture.

A. METANARRATIVE

"and it had [Tweetie Bird and Sylvester]"
metaphoric: indicates left then right to spatialize characters

"and uh [the first scene you see is uh]"
deictic: indicates center as locale for new scene

B. PARANARRATIVE

"[right] it's one of the series"
deictic: point at listener

"[so so so you know]"
deictic: point at listener

Figure 8: PERSPECTIVE WITHIN NON-NARRATIVE

Although the same space may be pointed out in gesture, depending on the narrative level of the accompanying speech, the gesture will have a different value.

Thus in example (A), Figure 8, the first deictic spatializes an opposition among characters, the second deictic spatializes an opposition among scenes. At the paranarrative level, on the other hand, as under (B) in Figure 8, the deictic spatializes an opposition between speaker and hearer: more exactly it indexes both of them as participating in the same speech event.

V. Conclusions

In this paper we have argued for the following major conclusions: (a) the theory of ground contains a covert naive definition of story that conflates the different kinds of events, and different kinds of goals that a speaker may talk about. (b) ground is not capable of describing all of the narrative phenomena that have to do with the relationships between different kinds of events in narrative. (c) narrative structure, which has seemed problematic for English speech, is more clearly marked by nonverbal means. Once we can discover from gestures where the background and foreground are distinguished, we can sometimes observe hitherto unsuspected linguistic contrasts, as in the case of simplex/complex sentence structure being used for different focalizations. Observation of the gestures that accompany narrative discourse illustrates the potential for non-spoken resources of narrative structure that in English may be richer than those available to speakers in the verbal channel.

Footnotes

1. Preparation of this article was supported by grants BNS 8211440 and BNS 8518324 from the National Science Foundation, by grants in 1981 and 1989 from the Spencer Foundation, by a 1989-90 dissertation research grant from the National Science Foundation, and by a 1989-90 dissertation grant from the Spencer Foundation. We wish to thank Amy Dahlstrom for commenting on the manuscript, and many others for especially helpful comments at the presentation. Correspondance can be sent to David McNeill, Department of Psychology, University of Chicago, 5848 South University Ave., Chicago, Il. 60637, USA.

2. Square brackets mark the extent of a gesture. Plus signs (++) indicate the duration of a pause.

3. C-VPT stands for 'character viewpoint' and O-VPT stands for 'observer viewpoint'.

Bibliography

Cassell, Justine & David McNeill. (in press). Gesture and the Poesis of Prose. Poetics Today.

Delancey, Scott. 1987. Transitivity in Grammar and Cognition. Coherence and Grounding in Discourse. Amsterdam: John Benjamins. pp.53-68.

- Goodwin, Charles and Marjorie Harness Goodwin. (in press). Context, Activity and Participation. The Contextualization of Language, ed. by Peter Auer & Aldo di Luzo. Amsterdam: Benjamins.
- Harris, R. & T.J. Taylor. 1989. Landmarks in Linguistic Thought: the Western Tradition from Socrates to Saussure. London: Routledge.
- Hopper, Paul. 1979. Aspect and Foregrounding in Discourse. Syntax and Semantics (Vol. 12): Discourse and Syntax. New York: Academic Press. pp. 213-241.
- Hopper, Paul & Sandra Thompson. 1980. Transitivity in Grammar and Discourse. Language Vol 56:2. pp. 251-299.
- Labov, William. 1972. The Transformation of Experience in Narrative Syntax. Language in the Inner City. Philadelphia: University of Pennsylvania Press.
- Lakoff, George & Mark Johnson. 1980. Metaphors We Live By. Chicago: University of Chicago Press.
- McNeill, D. & Levy, E. (1982). Conceptual Representations in Language Activity and Gesture. Speech, Place, and Action, ed. by R. Jarvella & W. Klein. Chichester: John Wiley & Sons.
- Reddy, Mark. 1979. The Conduit Metaphor -- A Case of Frame Conflict in our Language about Language. Metaphor and Thought, ed. by A. Ortony. Cambridge: Cambridge University Press. pp.284-324.
- Reinhart, Tanya. 1980. Principles of Gestalt Perception in the Temporal Organization of Narrative Texts. Linguistics 22. pp.779-809.
- Ricoeur, Paul. 1984. Time and Narrative (vol. 1). Translated by Kathleen McLaughlin & David Pellauer. Chicago: University of Chicago Press