

Toward a Cognitive Theory of Narrative Acts

EDITED BY FREDERICK LUIS ALDAMA

Katja Mellmann:

Voice and perception

An evolutionary approach to the basic functions
of narrative

p. 119-140

University of Texas Press  *Austin*

2010

CHAPTER FIVE

Voice and Perception: An Evolutionary Approach to the Basic Functions of Narrative

KATJA MELLMANN

The distinction of *voice* (who speaks?) and *perception* (who sees/hears/smells?) (Genette 1986:186; 1988:64) can be said to be the egg of Columbus in Gérard Genette's analysis of narrative discourse. Whereas in traditional models of literary narrative we had to deal with typologies mainly (for instance, of "narrative situations"; see Stanzel 1971, 1984; Fludernik and Margolin 2004; Genette 1980), we now possess a systematic description of the imagination evoked by a text, which takes into account the quasi-ontological (see Bortolussi and Dixon 2003) status of its constituents. In this chapter I search for the *cognitive functions* that correlate with the text features of "voice" and "perception" and for how they bring about such a "layered" imagination in the reader. The aim is to explain how and why literary narratives can run properly in the human mind—which is another way of asking how humans could develop narrative discourse as a way of communication at all.

To make more graspable what I am trying to do, let me start with a consideration by Genette that I regard as crucial for a profound understanding of the interplay between narrative texts and human cognition. Genette says: "Unlike the director of a movie, the novelist is not compelled to put his camera somewhere; he has no camera" (Genette 1988:73). Vice versa, it can be said: Unlike the novelist, the director of a movie is not compelled to talk to the audience; he has no voice. If these propositions are true (I shall discuss them later), it can be stated that there are *two distinct narrative functions*, as I will call them, which can occur independent from one another in principle, although they used to occur in combination very often.

Literary narratives—especially since the age of realism, but also in the Homeric epics and all through history—*can* (but need not) possess longer or shorter "focalized" passages, and motion pictures *can* (but

need not) employ title cards, subtitles, or voiceovers, and thus establish an enunciating instance for a moment. But the typical literary narrative is understood to be a (wholly or dominantly) *telling* medium (“voice”), and the typical cinematic narrative is understood to be a (wholly or dominantly) *showing* medium (“perception”). How can it be that one can work without the other? My proposition is that they rely on different strategies of information gathering which evolved separately in the history, and prehistory, of human evolution. Different modes of presentation, different media, make stronger use of the one information system or the other.

In the first section below, I analyze selected passages from Virginia Woolf’s *Flush* (1933) in order to illustrate the distinction of voice and perception and explain the concept of focalization. The second section attempts a first linkage between the textual features of “voice” and “perception” and distinct cognitive programs, as well as a rough assessment of their historical realizations. Taking up again Genette’s comparison of literature and film in the third section, I shall then discuss the question of whether one can plausibly speak of filmic “narratives,” although there is no obligatory voice instance in cinematographic representations.

Voice and Perception as Two Distinct Functions in Literary Narrative

Unfortunately—or fortunately—Genette did not care too much about giving a holistic system of narrative theory. Rather, he formed occasional concepts for what he needed to analyze Proust’s *Recherche*, and there are slight shifts between his concepts in *Discours du récit* (1972) and *Nouveau discours du récit* (1983).¹ This is why I do not attempt to give a scholastic exegesis of the “real” Genette here and by this establish a narratological “catechism” which Genette himself refused (1988:74). I just pick up what I conceive as one of his basic ideas: the distinction of voice and perception and the concept of focalization. The structuralist approach was the best one we ever had to extract persistent patterns from a diversity of works of art, and persistent patterns we require when correlating text structures with cognitive functions. One such pattern in literary narrative is that “someone talks to us” (“voice”), another one that fictional events cannot be shown but from a particular angle or position (“perception”).² In order to avoid “the too specifically visual connotations” of terms like *point of view* and *perspective*, Genette has suggested “the slightly more abstract term *focalization*” (Genette 1986:189; 1988:64). I doubt whether the visual metaphor in terms like *point of view*

or *perspective* has ever led to serious problems. I appreciate the term *focalization* rather for omitting connotations of “attitude” which the former terms were charged with and which, in the first instance, refer to the concept of “narrator.”

Preliminary Note on the Concept of “Narrator”

The mental attitude (set of values and convictions, intentions, cultural knowledge, etc.) we conceive by reading a story is due to what I am calling a “psycho-poetic effect” (Mellmann 2006:99–103; forthcoming) on the part of the reader, that is, the effect of “making psyche” (where there is none) by imagining any kind of “subject” in response to certain text structures. Because *Homo sapiens* is an eminently social animal, our social cognition systems are eager to extract relevant information wherever possible. There are many kinds of text features which are well suited to serve as input to particular social algorithms (for examples see Mellmann forthcoming), and once triggered, those social algorithms entail, as a kind of by-product, mental images of quasi-personal unities (like “character,” “narrator,” “author,” or Stanzel’s “figural” narrator). Such implied subjective entities can *combine* with the speaker instance or the perceptual instance³ in the imagination of the reader, but they are not identical with them. However, in literary narratives the sustained “voice” provokes a permanent psycho-poetic effect of “narrator.” What has often been criticized in narrative theory under the title of “anthropomorphic fallacy” (see Bortolussi and Dixon 2003:174) is but the symptom of a universal human bias, “a metaphor we live by” in Lakoff and Johnson’s term (see Smith 2000:159). The fact that “someone talks to us” simply entails the image of *somebody* talking to us (see Carroll 1995:157). This is why, in literary narratives, the default position to ascribe certain attitudes to is located on the side of the “voice” rather than that of “perception.” I give an example from Woolf’s *Flush* as an illustration:

Flush darted to the sofa.

“Oh, Flush!” said Miss Barrett. For the first time she looked him in the face. For the first time Flush looked at the lady lying on the sofa.

Each was surprised. Heavy curls hung down on either side of Miss Barrett’s face; large bright eyes shone out; a large mouth smiled. Heavy ears hung down on either side of Flush’s face; his eyes, too, were large and bright; his mouth was wide. There was a likeness between them. As they gazed at each other each felt: Here I am—and then each felt: But

how different! Hers was the pale worn face of an invalid, cut off from air, light, freedom. His was the warm ruddy face of a young animal; instinct with health and energy. Broken asunder, yet made in the same mould, could it be that each completed what was dormant in the other? She might have been—all that; and he—but no. Between them lay the widest gulf that can separate one being from another. She spoke. He was dumb. She was woman; he was dog. Thus closely united, thus immensely divided, they gazed at each other. Then with one bound Flush sprang on to the sofa and laid himself where he was to lie for ever after—on the rug at Miss Barrett's feet. (Woolf 1983:20)

The charm of this scene is due to its ironic parallelism with romance. The narrator, by his diction, cites romantic discourse and thereby reveals a humorous attitude towards this quite ordinary occurrence between a human and a dog. That is, we draw conclusions about what the narrator's attitude may be from looking at his way of telling the scene. Word choice, imagery, manner of speaking, and their aptness or inaptness in respect of the events, are the clues by which we learn about the narrator's attitude. Independent from this mental "standpoint"—the amusement about the scene as a whole—we are guided through multiple perspectivalizations of this scene which the narrator employs to show the event itself. For instance, when we are told about the curls/ears, the bright eyes, and the large mouth twice, these seem to be seen through the eyes of Flush/Miss Barrett—in other words, from in front, as in a face-to-face communication. Only from this perspective can we see the curls/ears "*on either side,*" *both* eyes, and that the mouth is "*large.*" These perspectives "through the eyes of a character" (Genette's "internal focalization") are not those of the narrator, if we imagine the narrator as a person external to the story (Genette's "heterodiegetic narrator").

I will return to the above quoted passage in more detail below. For the moment it is important to notice that "attitude" in literary narrative mostly is a matter of "voice" (word choice, imagery, manner of speaking), and that it should not be interfused with "perspective," which is a matter of "perception" (of what is perceived before it is told).⁴ So what does "perspective," or "focalization," as a matter of "perception," mean exactly?

Focalization

Genette explicated his concept of focalization as "a restriction of 'field'—actually, that is, a selection of narrative information [. . .]. The

instrument of this possible selection is a *situated focus*, a sort of information-conveying pipe that allows passage only of information that is authorized by the situation” (Genette 1988:74). Whenever a text passage suggests such a “situated focus” as viewpoint, it is, in Genette’s terms, “focalized,” which means that there is some restriction of information due to situational conditions within the fictional world. The opposite would be an “unfocalized” passage, showing (theoretically) unlimited information about the narrated world. For an example let us have a look at the very first sentences of the novel:

It is universally admitted that the family from which the subject of this memoir claims descent is one of the greatest antiquity. Therefore it is not strange that the origin of the name itself is lost in obscurity. Many million years ago the country which is now called Spain seethed uneasily in the ferment of creation. Ages passed; vegetation appeared; where there is vegetation the law of Nature has decreed that there shall be rabbits; where there are rabbits, Providence has ordained there shall be dogs. (Woolf 1983:7)

The narrator begins with telling something out of his general *knowledge* (“It is universally admitted . . . Therefore . . .”), not something which is actually *perceived*. One may argue that there is some restriction of information too, for the narrator’s knowledge may be not unlimited either. But this would be no “restriction of ‘field’” in the sense suggested above. The limits of the narrator’s knowledge are due to reasons external to the storyworld, not to a particular “focus” situated in it.

A slight change takes place with the sentence “Many million years ago . . .” At this point, a kind of story begins (as with “Once upon a time . . .”), and we start imagining a specific fictional world (wherein a future Spain seethes uneasily in the ferment of creation, and so on). We have the impression of a camera recording the spectacle of creation, which is due to the spatio-temporal data in the text. If we ask how the situation of perception is defined, we can answer: somewhere in the “ferment of creation,” where Spain is about to emerge (*where?*), and: many million years ago (*when?*). As soon as we are given spatio-temporal coordinates of a fictional situation we face the germ of focalization.

However, in this example the spatio-temporal scope is very wide; a larger spatial frame than that of the tohubohu and a larger temporal frame than that of several ages, in comparison to which these could pass as a “restriction of field,” are hard to imagine. Thus, the grade of focalization is very low here. It is much higher in the above-quoted

sentences: “Flush darted to the sofa. ‘Oh, Flush!’ said Miss Barrett.” Here the events are told as they might be heard and seen by someone who is present in the room. This is a focalization strategy so frequent in modern literature that Franz K. Stanzel (1971; 1984) made it one of his “prototypes” of “narrative situations.” It evokes the impression of an invisible witness, always walking about with the characters (Stanzel’s “figural” narrator, or Henry James’s “reflector person”), who functions as the “information-conveying pipe” between the fictional world and the storyteller. I call it *anthropomorphic focalization* for its humanlike dimension of spatio-temporal access.⁵ However, the spectrum of possible focalizations is still much broader: Between the scope of wideness of a camera set at such a great distance that it can record the universe, and of such a permanence that it can record millions of years, and the scope of wideness of a camera set among characters, changing places with them, *any gradation is possible*. Think for instance of the initial zooming (e.g., from country/age across town/season to house/day) in the beginnings of many a novel. And also *below* the level of anthropomorphic focalization one can go down endlessly into microscopic time and space. Moreover, an author can combine narrow time scopes with wide space scopes, and vice versa. There are *infinite forms of focalization*.⁶

The best way to describe a particular focalization is to determine the “situated focus,” i.e., the “point from which the narrative is perceived as being presented at any given moment” (O’Neill 1992:333). Its spatio-temporal coordinates define the perceptual situation. Normally, the situation of the perceiving instance (camera) is not identical with the situation of the voice instance. There are some particular cases (like interior monologue or stream of consciousness) where both coincide and the narrator appears to be the perceiver at the same time. Another particular case is the autobiographical schema, where there are several years between the perceptual situation and the situation of utterance, but the (anthropomorphic) focalization can be attributed to the “I” of the narrator and his/her presence in the fictional world. Yet the default, or classical, case of modern literary narrative is that of many realist novels, where the situation of enunciation remains more or less unspecified, whereas perception varies between several particular situations within the fictional world.

So far, I have illustrated the difference between focalized and unfocalized narration, and the continuum between them, which can be described as different grades of focalization. But the grade of focalization (scope wideness) is not all which needs to be taken into account; we have

to consider also the grade of autonomy of voice and perception, that is, their specific interplay at work in a particular passage.

Different Grades of Autonomy of Voice and Perception

My conception of voice and perception as two more or less autonomous narrative functions attempts to integrate the classical distinction of showing and telling (or mimesis and diegesis) as two different modes of narration into Genette's model.⁷ Let us assume, as default position, an integrated narrative system in which the functions of voice and perception are well balanced and tightly intertwined, so that the reader is not forced to reflect on one of them in particular. Then the showing mode can be reformulated as a deviation which is marked by the dominance of the perceptual function, the telling mode as a deviation with dominant voice function. I initially provide two examples that shall illustrate the extreme poles of this opposition.

The first extreme case is akin to what Roland Barthes described as "reality effects." Those effects are brought about by detailed descriptions, as for instance in a text by Flaubert, which comprise even apparently useless details like a barometer hanging on the wall. If the perceptual data given in a text thus exceed the demands of telling a consistent story and the voice appears to will-lessly follow the input from the perceptual system by just reflecting what is perceived, then the perception system differentiates from the integrated default system and develops into a distinct system of its own, which works in accordance to its own logic (indiscriminately recording everything), independent from external needs and intentions; to show what is there seems to be an end in itself. The voice function, in contrast, is subordinated to the perceptual function in that it does not comment, add, select, or stylize anything, but confines itself to a mere reflector function. This is why in passages with dominant perception and subsidiary voice function we easily forget that the perceived events are mediated through a voice instance. The fictional events seem to "tell themselves" (Genette 1986:164) and "finally say nothing but this: *we are the real*" (Barthes 1989:148).

The other extreme pole can be illustrated by a passage from Diderot's *Jacques le Fataliste*:

Vous voyez, lecteur, que je suis en beau chemin, et qu'il ne tiendrait qu'à moi de vous faire attendre un an, deux ans, trois ans, le récit des amours de Jacques, en le séparant de son maître et en leur faisant courir

à chacun tous les hasards qu'il me plairait. Qu'est-ce qui m'empêcherait de marier le maître et de le faire cocu? d'embarquer Jacques pour les îles? d'y conduire son maître? de les ramener tous les deux en France sur le même vaisseau? Qu'il est facile de faire des contes!

The narrator here leaves no doubt about his power to tell what he wants to. That is, by saying that something is that way, he would make the story go that way; the perceiving instance would have to follow every turn of the voice instance. Passages like those fancied by Diderot's narrator, with dominant voice and subsidiary perception function, make the reader particularly aware of the artifact character of a narrative, the fictiveness of its content, and thus entail the very opposite of reality effects.

Now let us have a look at more subtle combinations between voice and perception and turn back to the above-quoted passage about the first encounter of Flush and Miss Barrett. In the first two sentences, we face the default version of narrative texts:

Flush darted to the sofa.

"Oh, Flush!" said Miss Barrett.

Our impression of the passage is that of a homogenous narrative act. The two narrative functions of voice and perception together constitute one integrative system. It would be hard to tell if the words "Flush darted to the sofa" determine the perceptual image of the event or if, vice versa, the event as it is perceived determines the wording. The two narrative functions of voice and perception seem quite equivalent here.

Not so in the next two sentences:

For the first time she looked him in the face. For the first time Flush looked at the lady lying on the sofa.

The perceived event ("looked") is not the only information given in these sentences (like "darted" and "said" were); in addition we are told that this happens "for the first time." This is something which cannot be seen, something which the narrator simply *knows*. One could say that in these two sentences different sources of information are blended on the discourse level. Imagine the text would go like this: "She gazed at his face for a while. It was the first time that she had a look at him." Then the presentation of the perceived event ("gazed") and the additional comment that "it was the first time" would be more separated

than in the original sentences. By interdigitating one with the other through a particular way of wording, the voice somewhat dominates the perception, that is, the latter cannot fully unfold as an autonomous dimension of the text (with stronger, activity-indicating words such as *gaze* instead of *look* perhaps, or with proper spatio-temporal data such as “for a while”); the perceived event is reduced to an only shortly introduced fact in order to fit together with the cardinal information that it was the first time. Similarly, the beginning of the novel (the creation scene): I said the spatio-temporal data would provide only the *germ* of a focalized passage, because beside the all too wide scope of focalization we have here a voice instance so dominant again that we can hardly assume a strong perceptual function. If the text went like this: “Where there is vegetation the law of Nature *decreed* that there shall be rabbits; where there are rabbits, Providence *ordained* there shall be dogs”—i.e., if the epic preterit of “seethed uneasily” was kept on, the illusion of ongoing events (Nature and Providence entering the stage, so to speak) would be maintained. But by switching to the present perfect (“has decreed”/“has ordained”), Woolf evokes the impression of a narrator talking out of his general knowledge (about accomplished laws of nature) again. If perceptual elements are throttled down like this to a mere mention by the voice, which seems to have control over how much of the perceived world gets on through to the discourse level, the perception system has not differentiated into a distinct system of its own but in a way pertains to, or subsidiarily subserves, the voice system. In terms of system theory such non-differentiated systems can be called “trivial machines” (see von Foerster 1984:9f.), which means that they have not yet established themselves as operatively closed systems and developed their own logic, but function after the rules of another and produce a rather predictable output.

The above-quoted passage then goes on as follows:

Each was surprised. Heavy curls hung down on either side of Miss Barrett’s face; large bright eyes shone out; a large mouth smiled. Heavy ears hung down on either side of Flush’s face; his eyes, too, were large and bright; his mouth was wide. There was a likeness between them.

The first sentence affiliates properly with the antecedent in that it retains the subsidiary perceptual function: Although we feel that the fact that “each was surprised” must have been perceived somehow (by looking into Flush and Miss Barrett) before it is told, we would settle for the abridged version that it simply is that way and would do without

a detailed description of how this surprise unfolds in Flush's and Miss Barrett's minds. That is, we would follow the claim of dominance of the voice instance. But then we are guided in this internal perspective exactly, when we gain insight of how Flush and Miss Barrett see one another, so that the perceptual function now gains autonomy and the voice function is throttled down to a subsidiary trivial machine (reflector function). The following sentence, saying that there was a likeness between them, now does not seem to be a comment on the part of the narrator, evaluating these perceptions, but a thought Flush and Miss Barrett (at least she) might have had and which is just replicated (by means of free indirect discourse) by the voice instance.

The dominance of perception, on the one hand, seems to be maintained until the end of the passage. On the other hand, we more and more feel that, even if it was nothing but Miss Barrett's thoughts that is reflected here, it is not done in accordance with her attitude:

As they gazed at each other each felt: Here I am—and then each felt: But how different! . . . Broken asunder, yet made in the same mould, could it be that each completed what was dormant in the other? She might have been—all that; and he—but no. Between them lay the widest gulf that can separate one being from another. She spoke. He was dumb. She was woman; he was dog. Thus closely united, thus immensely divided, they gazed at each other.

We clearly feel shining through the spirit of the narrator, his attitude, his own thoughts and feelings toward the event, expressed by a particular way of phrasing. Thus, also the voice function establishes a differentiated system here. This autonomy of the voice instance goes without stopping the above-declared autonomy of the perceiving instance. You can catch its persistence by the voice's real-time reaction to ever new events in the perceived scene, indicated by the dashes: "She might have been—all that" indicates that the voice ties itself down to the real-life time scale of Miss Barrett's thoughts by performing the same pause of reflection she makes; the same with "and he—but no," which follows her self-interruption of thought. Thus it is granted enough room to the perceptual function to unfold completely, though we are quite aware that the definite way of utterance (employing the vocabulary of romantic love encounter) is due to the narrator rather than to Miss Barrett (whose thoughts may be of a rather nonverbal nature).

I have now typified four ways of combining voice and perception in literary narratives:

1. integrated narrative default system with balanced voice and perceptual function;
2. autonomous voice instance, with competence to discretionarily abridge the output of the perceptual device;
3. autonomous perceptual system, coming to the fore as the inner film of imagination in the reader, and thus pushing back the voice to a mere reflector function;
4. autonomous voice *and* perception, producing a truth of its own each.

In options (2) and (3) one of the two functions shows clear dominance and differentiates out of the narrative default system as an autonomous system itself, while the other function is reduced to a subserving trivial machine. In option (4) we face a kind of double-coding brought about by two fully developed narrative systems, which might irritate one another because of their “structural coupling”⁸ but for the rest work autonomously.

I used these rather abstract terms from system theory in order to make more transparent the very complex design of narrative discourse by reconstructing its underlying principle. Moreover, the conception of voice and perception as of potentially autonomous systems—together with Genette’s comparison of novels and films which I introduced above—can help to firm up the supposition that these two different dimensions of narrative texts correspond with two equally different faculties of the human mind, which are able to be performed independently from one another in principle.

Voice and Perception as Inputs for Two Different Systems of Information Gathering

John Tooby and Leda Cosmides, two exponents of evolutionary psychology, have given a notable explanation of narrative by suggesting that socially communicated information is often formatted in a way that mimics firsthand experience, in order to cooperate with our evolutionarily elder, more basic mechanisms of information gathering:

Indeed, we evolved not so long ago from organisms whose sole source of (non-innate) information was the individual’s own experience. Therefore, even now our richest systems for information extraction and learning are designed to operate on our own experience. It seems therefore inevitable, now that we can receive information through communica-

tion from others, that we should still process it more deeply when we receive it in a form that resembles individual experience, even though there is no extrinsic reason why communicated information needs to be formatted in such a way. That is, we extract more information from inputs structured in such a form. What form is this? People prefer to receive information in the form of stories. Textbooks, which are full of true information, but which typically lack a narrative structure, are almost never read for pleasure. We prefer accounts to have one or more persons from whose perspective we can vicariously experience the unfolding receipt of information, expressed in terms of temporally sequenced events (as experience actually comes to us), with an agent's actions causing and caused by events (as we experience ourselves), in pursuit of intelligible purposes. (Cosmides and Tooby 2001, 24)

What Tooby and Cosmides refer to as quasi-experiential structure or form is, in terms of narratology, the focalization of a text, especially what I called anthropomorphic focalization above. It is this situated focus, adhering to human dimensions of time frame and spatial movement more or less, which evokes the impression of vicarious experience. Hence, we can state that the narrative function of perception correlates with our ability to extract information from environment, that is, it makes use of basically the same information systems that are employed in normal environmental experience of all organisms capable of autonomous movement. Whereas the narrative function of voice—the fact that someone talks to us—correlates with information systems that evolved together with human language, that is, in a much later stage of evolution. The voice aspect of narratives thus should involve mainly those cognitive mechanisms which are associated with verbal communication (semiotic faculties, memory, syntax logic, and the like), whereas the perception aspect should run a simulation on the perceptual system of our brains and focalized passages should be processed by the same second-order circuits as are involved with processing sensory inputs.

So far, the one has nothing to do with the other, and, as Tooby and Cosmides put it, “there is no extrinsic reason” to combine these two systems of information gathering by telling a story. But it is true that, in ancient times, most of our knowledge was bound and passed on in the form of stories. Think of the Gilgamesh epic, the Homeric epics, Greek myths, the Old Testament, medieval novels, or simple folktales. They contain a good deal of knowledge about our ancestors, foreign countries, fabulous animals, the beginning of the world, and, moreover,

about vice and virtue, love and hate, bliss and sorrow, danger and salvation, and so on. There are several studies on narrative that focus on its capacity for storing adaptive information (Sugiyama 1996; 2001; 2006; Eibl 2004:257–272), but, as far as I can see, the particular role that focalization plays in this way of storing information has not been analyzed yet. This may be due to the fact that in most ancient myths and simple folktales the voice function preponderates and focalized passages are the exceptional case.⁹ A persistent autonomous perception system seems to be more typical for modern literature. But even a subsidiary perceptual function can be a good device for preserving information in the form of stories. Imagine I was a Pleistocene hunter-gatherer saying to you:

1. There is a waterhole a half day's hike away from here in this direction.

You would be informed well enough for an immediate departure to the waterhole. But imagine I tell you about it as follows:

2. You take the path we went yesterday to the stone pit. When you see the big oak tree, you climb down the scarp at the right-hand side and cross the plain toward its leafy end. There you will find a waterhole.

This instruction maps out an imaginary way, that is, a little (proto-) *narrative* of you (as situated focus) going through a particular landscape. This imaginative map of the story makes you independent of a particular direction that I show you from where we are standing in the moment of our conversation. So if you decide not to go for the waterhole until the next day, you will find it from wherever you might be then.

Now imagine you want all your kin to know about this very good waterhole in case you will not be with them when coming to this place again the next year. Then you might tell a story like:

3. Once upon a time there was a little boy looking for a waterhole. He walked on a rocky hill and looked out as far as the eye could see. Beside the hill there was a great plain extended to the horizon. At its end, the boy noticed a small stripe of greenness, and he said to himself: where there is grass there must be some water. He climbed down the scarp and walked ahead toward the green stripe. He marched for half a day, the sun was shining hot, and he was getting tired and more and more thirsty. But coming nearer he could clearly see now

the green grass. He had not deceived himself. When he reached the grassy ground he had not to search long ere he found a wonderful waterhole.

Now we have a classical third person narrative that suits for information storage over longer periods of time and can be communicated from one individual to another several times. (You can adorn it further with several adventures the little boy has to get through, obstacles he must overcome, and so on, to make it more interesting¹⁰ and easier to keep in mind.)

What is particular for human language is that the reference to facts (Karl Bühler's "representation function" of language) can be isolated from the reference to persons (Bühler's functions of "expression" and "appeal"), that is, from its situational anchorage; whereas animal languages, and presumably the proto-language of the early hominid species, are undifferentiated "tri-functional" languages (Eibl 2004:209–275). To flesh out a message (1) by a conversational first- or second-person proto-narration (2) or to transform such proto-narratives into a purely representational third-person story (3) might have been important steps in this process of linguistic sophistication. The narrative format of a message deliberates its descriptive content from the original context of utterance and thus makes it a quasi-objectified, transportable thing. In this sense, the isolation of the descriptive function of human language was described as a "Vergegenständlichung" (reification) of information by Karl Eibl (2004:209–275). What Cosmides and Tooby refer to as the "text book" type of communication is nothing but a radicalization of the same principle: an absolute detachment from any situational meaning, even from vicarious situations. I give a possible textbook version of our story:

4. Waterholes are often found amid leafy places.

This proposition contains the most generic information essence of the story, omitting any additional data about a particular time or area or person. However, successful application of information detached from any situational context is difficult, because the individual has to know by herself under what conditions such an abstract proposition is true. But obviously we are adapted to this problem: We are able to handle abstract information because of a cognitive "scope syntax, that tag[s] and track[s] the boundaries within which a given set of representations can safely be

used for inference and action” (Cosmides and Tooby 2001:20). This is how humans came to “live with and within large new libraries of representations” (Cosmides and Tooby 2001:20), which allowed them to store and standardize a greater amount of information than one single individual would be able to gain in a life span, and to pass it on to the next generation. That is, every new generation can build on the knowledge achieved by the last one, which leads to the specifically human, “cumulative” or “cascading” type of cultural evolution (cf. Eibl 2004:236).

However, these great “libraries of representation” have to be literal libraries most of the time. Except for some proverbs and similarly small packages of abstract information, oral cultures are mainly *mythical* cultures, which rely on *storied* lore on a much larger scale than on abstractly coded knowledge. Mere textbook types of cultural patrimony appear rather late in human history and depend on the existence of social elites possessing the required education and media.

The fact that an *autonomous* perception system in literary narrative occurs only seldom before modern times may also be due to cultural evolution and the evolution of media. I would not preclude completely the possibility that high differentiation grades of the narrative perceptual function have always been there in *oral* storytelling. Imagine a gifted storyteller at the campfire of a hunter-gatherer tribe, doing the best he can to entertain his audience, to thrill them, to move them. He would be well advised to flesh out the perceptual situation of his narrative with concrete detail to the utmost. But when oral traditions are recorded in media, this should be done as economically as possible, for most of the time new media are very expensive in the beginning. On this basis, it is not surprising that most of the narrative literature written down until one or two hundred years after the invention of letterpress does not make great use of a differentiated perceptual function. Furthermore, written stories in former times often were not designed to be *read*, least of all alone, silent, and in an immersive attitude of reading, but they served as an aid to memory for an ultimately oral performance again. Conditions were completely different in the realist age, when a most sophisticated art of narrative book-fiction emerged, tapping the full potential of the narrative perceptual function.

With regard to that rough historical sketch, it should have become palpable that the novelist is not compelled to put his camera somewhere in the strong sense of a high grade of focalization or a fully differentiated perceptual system. Low grades of focalization as in a bird’s-eye view or through frequent switches (Genette’s “multiple focalization”)

and a merely subsidiary perceptual function (i.e., merely mentioned perceptions) are sufficient to evoke the feeling of narration. In novels, we even tolerate longer voice-only passages (e.g., reflexive commentaries), although they, strictly speaking, interrupt the narration. Thus, storytelling per se is not particularly different from standard communication, where little proto-narrations like that of text (2), or conversational first-person narrations, with low differentiation grade of the perceptual function, occur now and then. To have the impression not only that somebody talks to us but that, moreover, we really see the told events, as if we were there ourselves, is an additional effect brought about by artful storytelling, which is not limited to a particular age or media but can emerge everywhere along the histo-cultural continuum under various conditions.

This seems quite plausible if we assume that the cognitive program which supports the perceptual function of narrative texts is much older than our linguistic faculties and was already there when humans first invented verbal storytelling. But then the question arises whether there is—if perception is what distinguishes narration from standard communication—something like a non- or preverbal storytelling making use only of the characteristic one of the two narrative functions. H. Porter Abbott (2000:248–252), for instance, suggests that the gift of imitation and a thereupon based “mimetic culture” (Donald 1990) gave way to an at least rudimentary form of narrative. Michelle Scalise Sugiyama (2005:181–183), on the other hand, prefers to confine the notion of narrative to verbal accounts, because “nonverbal expressive media (for example, visual art, dance, music) are actually quite inefficient narrative devices,” which fail, especially, “in representing the thoughts, beliefs, and motives” of literary characters and in communicating necessary background information (182). Scalise Sugiyama does not, among her examples, discuss dramatic art, which Abbott seems to consider in the first place and which, in the form of film, has indeed become the most powerful media for fiction in the twentieth century. As it clearly is *stories* that millions of people consume every night watching TV, it should be worthwhile to analyze the specific combination of narrative functions discussed here which is at work in films. The immense popularity of that medium, outrunning even novels by far, may be seen as a clue that the dominance of the perceptual function perhaps meets a universal human bias. It is quite possible that it simply feels more convenient if the evolutionarily elder, more basic system for information gathering is

served directly by audio-visual stimulation, without the (maybe richer, maybe not) mediation through language.

Voice and Perception in Nonverbal Narrative

Some narratologists hold that it is not correct to speak of filmic narratives, as the Latin word *narrare* means “declare, enounce, tell,” and there is no obligatory voice instance in cinematographic representations. However, some film theorists have tried to determine the implicit discourse level of filmic narrative. For instance, François Jost (1987:15) argued that the picture of a house does not signify “house,” but rather “look, there is a house”—that is, a voice-like deictic communication. This sentence, however, is just a factitious verbalization of the autonomous perceptual function (showing). It does not even cover the reflector mode of a subsidiary voice function, which I would rather indicate through “There was a house,” or “It is a house,” i.e., a clearly declaring kind of sentence. So one question is: how does the demonstrator of events become a declaratory instance (narrator) in the mind of the spectator? Another even more basic question, which shall be treated first, is: how can the observer of events become a demonstrator at all? Jost’s suggestion of an implicit deictic speech act implies agency. Correspondingly, the recent debates about whether films have narrators (for surveys see Lothe 2000; Smith 2000) center on narration as agency. I do think that films produce a psycho-poetic mental concept of narrator in the viewer, but I do not believe that this mental construction arises directly from the perceptual instance, because the personified I of the camera primarily is an observer, not a demonstrator. To understand the perceiving instance as an authorial instance of demonstrator already *presupposes* the psycho-poetic effect of a para-social instance rather than being the source of it.

As I said above, the psycho-poetic effect of a narrator, as a subject of attitudes, is only loosely connected to the literary voice instance (that is, only because of its continuous presence), but is not identical with it. So why preclude the theoretical possibility that in films it is the continuous presence of the perceiving instance which occupies the default position of a narrator? In fact, even the literary perceiving instance can sometimes give rise to anthropomorphic mental constructions. But are these anthropomorphic instances communicative instances at the same time?

Vocal calls have a communicative function in all animals, and that is why the linkage of voice with communicative functions, and a thusly implied social agency, is entirely intuitively plausible. Perception, however, as a simulation of individual experience, does not per se imply communicative functions or social agency. Rather the opposite: the illusionary “reality effect” of a differentiated perceptual system is much stronger in cinematographic representations than it could ever be in literary ones. Consequently, the authorial instance of a mediator is much easier to forget in filmic representations. The mental representation of an authorial instance would depend on a rather sophisticated awareness of the artifact character of a film. While the cognitive scope syntax (“This is what someone told me to be true”) is continuously busy in verbal communication, there is no intrinsic need to activate this faculty toward experiential simulations, except that the filmic representation overtly deviates from our perceptual habits. Naïve spectators, I would say, do not reflect on the authorial aspect of camera unless they are urged to do so; the unconditioned presupposition that the represented perceptions are shown by someone would rather be a culturally learned mental attitude toward filmic representations.¹¹ Yet if the supposition of an authorial instance is already conditioned by other means, it should be quite normal to understand the perceptual function as a communicative device of showing. So what other means are there to precondition the psycho-poetic effect of an authorial instance?

Narrative fictions often involve real world references and, by this, interfere with our own general knowledge about the world. The filmic depiction of real world elements thus might entail an “implied hypothesis” of the kind “it’s like this, isn’t it?” as Murray Smith (2000:164) suggests for his especially clear example of Patrick Keiller’s depiction of the city of London (in *London*, 1993). This effect of understanding implied statements about the world may be reduced in more fictitious genres such as for instance romantic love comedies, fantasy films, or horror movies, which often do not refer to historically real times, places, or persons; but even those genres bear implicit statements about “that’s what humans are like” and thus give rise to a psycho-poetic effect of the authorial kind.

Another important device to imply authorial instances is the plot. So far, I have treated narrative fictions (literary or filmic ones) as more or less pointless representations of events in a temporal sequence. Yet recent attempts to give a definition of narrative note that it consists in a representation of a *non-contingent* sequence of events (Eibl

2004:255). The non-contingency of narrative sequences is often ensured by simple plot schemata, like, for example, that of “separation and reunion” (see Eibl 2008), “searching and finding,” or “punishing the vicious/rewarding the good ones,” and we tend to derive meaning from these plots, that is, to infer messages from narrative sequences. The principle of poetical justness is especially clear in this respect: failure signifies wrongness, success signifies rightness. That is, the sheer sequence of events implies hypothetical moral issues. Or, more generally speaking: also the non-contingency of narrations implies hypothetical propositions about the world and thus refers to someone making such a statement.

From interpreting represented perceptions as transmitting a certain kind of knowledge, it is just a small step to other communicative acts such as persuading, suggesting, deducing, arguing, and whatever the genuine functions of verbal communications might be. This is how the common idea of a filmic language could have come into being. Though the metaphor of a language of film does not provide a solid ground for an articulated theory of film, as Gregory Currie (1993) and others have argued very plausibly, it does make sense in that it signifies the communicative quality of films. The voice instance implied by the metaphor does not literally perform a verbal speech act, which could be analyzed with linguistic instruments, but it expresses the feeling of intentionality which also voice-less fictions induce in the recipient, and thus makes it possible to speak of cinematographically represented stories as of narratives.

That noted, we can settle both literary and filmic fiction somewhere on the same continuum of narrative. Both commit declarative acts, the one directly by continuously verbalizing the perceived events, the other indirectly by depicting and arranging them in a way that appeals to our general knowledge about the world. The fact that they make different use of the two systems of information gathering (verbal communication, individual experience) does not determine their capacity to install an authorial narrative instance communicating with the recipient.

Taking film as a representative example for nonverbal narrative, it can be regarded as quite possible that the human species developed narrative skills even before the gift of language. Displaying real world elements in any kind of representational system—be it pictorial, dramatic, musical, or anything—and intentionally arranging them might correspond to what Abbott (2000:250) described as the “literature effect” of (nonverbal) storytelling: the ritualized staging of mimetic rep-

resentations as replicable units. And as such ritualizations are another way of reifying information by isolating it from its primary situational contexts, even the linguistic sophistication of third-person narratives seems to be nothing but a later affiliation of an already nonverbal cognitive faculty. What is significant for this faculty of reification is not its particular medial condition but the fact that reified information seems to need a social instance of sender to whom we ascribe a certain message and attitude. Storytelling thus seems to intrinsically be a form of social interaction, and Scalise Sugiyama (2005), though neglecting nonverbal forms of storytelling, seems to be right in assuming that storytelling evolved, in the first instance, to transmit information. Literary voice and perception as observed by Genette represent only the instruments to serve that purpose; they provide apt material for cognitive programs that are designed to extract and restore information, but they at the same time only answer to the need of sharing common information among group members.

Notes

1. This has led to a broad controversy on his notion of “focalization” (among others). For a recent survey on some major discussions see Bortolussi and Dixon 2003:166–178; see also Genette’s own responses to early critics in Genette 1988:64–78.

2. Bortolussi and Dixon (2003:172, 174–176) argue that Genette’s technical distinction between “voice” and “perception” does not find its equivalent in the mind of most real readers, who tend to synthesize them in their imagination. While I principally join Bortolussi and Dixon’s endeavor toward a reader-oriented theory of narrative, my conception of the reader in this theory is not an empirical one, but rather a theoretical one. I posit an “anthropological model reader” (Mellmann 2006:21) that can be understood as a compound of several psychic mechanisms postulated by evolutionary psychology. This conception is not determined by the question of what of the complex cognitive operations is accessible to the introspection of the reader himself, so I cannot agree with Bortolussi and Dixon’s conclusion that the analytic distinction of enunciator and perceiver “loses . . . relevance” (172) if not mirrored by interviews with real readers.

3. Terms like that of “perceiving instance”/“enunciating instance,” “perceptual situation”/“situation of enunciation,” and the like, and what I will introduce in the following passages, were developed in Mellmann 2006:164–204.

4. Chatman’s (1990:139–149) distinction between “slant” (in respect of the “reporter” of a story) and “filter” (in respect of “observer” of a story) represents a similar endeavor as mine is here. However, I think he is mischaracterizing Genette when he treats his term *focalization* as just another synonym for *point of*

view, pertaining to both voice and perception. Genette (1980) in fact develops questions of voice and narrator in a separate chapter, while focalization affiliates to the chapter “Mood.”

5. I avoid Genette’s term *external focalization*, because it makes no sense but in the presence of characters, and then it is rather an antonym to “internal focalization” than a stand-alone type of focalization. Both internal and external focalization show the same (humanlike) degree of scope wideness. “External” focalization can be said to be the default case of this spatio-temporal scope, whereas “internal” denotes the particular case when “the focus coincides with a character, who then becomes the fictive ‘subject’ of all the perceptions” (Genette 1988:74). But imagine a phrase like: “Leaves tussled by the wind outside the window of Miss Barrett’s empty room.” It shows the same “situated focus,” but there is no character aboard which the focalization could be “external” to. (For further examples see Mellmann, forthcoming.)

6. I emphasize this because it is often said that, after Genette, there are three forms of focalization: “zero,” “internal,” and “external focalization” (see for instance Martinez and Scheffel 1999:64). First, as I mentioned above, Genette did not aim to give a complete system (wherein three would make a whole), but, in the respective passages (Genette 1980:189–194), built on Stanzel’s typology in order to introduce his own concept of focalization. Second, “zero focalization” strictly spoken is no *type of* focalization, but the opposite; unless one takes very wide scopes (like that of “tohubohu/ages”) as omniscient views, like Genette (*ibid.*; 1988:74) himself was prone to do.

7. Genette (1980:162–164) himself integrated this distinction in terms of different grades of “distance”—a concept which then has to be applied to different narrated objects (events, speech, etc.) individually. I hope to give a more economical, systematic solution of the problem. Moreover, I principally avoid the spatial metaphor of “distance” because of its vagueness. Beside its literal meaning of near/far it assembles a great deal of figurative meanings like temporally near/far, directly accessed/mediated, pure/commented, true/stylized, sincere/ironic, positively/negatively evaluated, or even elevated/prosaic. Though all the aspects indicated by these oppositions might bear some relation to the question of mimesis and diegesis, they do not seem very useful for a clear-cut definition.

8. In Niklas Luhmann’s system theory, “structural coupling” (after Humberto R. Maturana and Francisco J. Varela) signifies the interrelation between two autonomous (“self-referent”) systems which represent “environment” to one another (i.e., do not overlap or interpenetrate), yet can be “irritated” by one another. Applied to our narrative model: The perception system cannot produce attitudes, opinions, or anything that would require more than a sensual system (seeing, hearing, smelling . . .). But it can change its orientation in reaction to an irritation from the voice system. Take for instance the change from external to internal focalization, which can be interpreted as reaction to the proposition “each was surprised”; as if the perceiving instance wanted to focus on what the voice instance had just picked as the central information. The fact that the voice system chose this of all information is due to its own logic: a particular intention of depiction. The voice system itself cannot produce percep-

tions (events), but only words, but it can by this produce a particular selection and appearance of events, which can entail particular focalizations on the part of the perception system again, and so on.

9. See Fludernik (1996; 2003) who observed a dominance of the “teller frame” in pre-modern literature.

10. For a detailed study on emotional triggers in literature see Mellmann 2006.

11. However, once one has taken this attitude, it is again quite natural to follow the social implication of showing. In doing so, we rely on a universal human predisposition, which in developmental psychology is known as the ability for “joint attention.” Humans, on a much larger scale than other primates, follow another one’s gaze (perhaps the higher percentage of the white in the eye co-evolved together with this endowment), listen if another one seems to be listening, and so on. That is why, once we assume that the filmically mediated perceptions are not our own but those of another, we are still interested in them and are quite prepared to pay to them as much attention as we would to our own perceptions.

#####

Abbott, H. Porter. “The Evolutionary Origins of the Storied Mind: Modeling the Prehistory of Narrative Consciousness and Its Discontents.” *Narrative* 8, no. 3 (2000): 247–256.

Barthes, Roland. “The Reality Effect.” In Roland Barthes, *The Rustle of Language*. Trans. Richard Howard. Berkeley: University of California Press, 1989. 141–148.

Bortolussi, Marisa, and Peter Dixon. *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge and New York: Cambridge University Press, 2003.

Carroll, Joseph. *Evolution and Literary Theory*. Columbia and London: University of Missouri Press, 1995.

Chatman, Seymour. *Coming to Terms: The Rhetoric of Narrative in Fiction and Film*. Ithaca: Cornell University Press, 1990.

Cosmides, Leda, and John Tooby. “Does Beauty Build Adapted Minds? Toward an Evolutionary Theory of Aesthetics, Fiction and the Arts.” *SubStance* 94/95 (2001): 6–27.

Currie, Gregory. “The Long Goodbye: The Imaginary Language of Film.” *British Journal of Aesthetics* 33, no. 3 (1993): 207–219.

Diderot, Denis. *Œuvres. Texte établi et annoté par André Billy*. Paris: Gallimard, 1951.

Donald, Merlin. *The Origins of the Modern Mind*. Cambridge: Harvard University Press, 1990.

- Eibl, Karl. *Animal poeta. Bausteine der biologischen Kultur- und Literaturtheorie*. Paderborn: Mentis (Poetogenesis, 1: 2004). English abstract at http://www.lrz-muenchen.de/~eibl/animaeng_htm.html
- . “Epische Triaden. Über eine stammesgeschichtlich verwurzelte Gestalt des Erzählens.” *Journal of Literary Theory* 2, no. 2 (2008): 197–208.
- Fludernik, Monika. *Towards a “Natural” Narratology*. London: Routledge, 1996.
- Fludernik, Monika, and Uri Margolin. “Introduction.” In *German Narratology*. Ed. Monika Fludernik and Uri Margolin. *Style* 38, nos. 2–3 (2004): 148–187.
- von Foerster, Heinz. *Principles of Self-Organization—in a Socio-Managerial Context*. In *Self-Organization and Management of Social Systems: Insights, Promises, Doubts, and Questions*. Ed. Hans Ulrich and Gilbert J. B. Propst. Berlin: Springer, 1984. 2–24.
- Genette, Gérard. *Narrative Discourse: An Essay in Method*. Trans. Jane E. Lewin. Ithaca, N.Y.: Cornell University Press, 1980.
- . *Narrative Discourse Revisited*. Trans. Jane E. Lewin. Ithaca, New York: Cornell University Press, 1988.
- Jost, François. *L’œil-caméra. Entre film et roman*. Lyon: Presses Universitaires de Lyon, 1987.
- Lothe, Jakob. *Narrative in Fiction and Film: An Introduction*. Oxford: Oxford University Press, 2000.
- Martinez, Matias, and Michael Scheffel. *Einführung in die Erzähltheorie*. München: Verlag C. H. Beck, 1999.
- Mellmann, Katja. *Emotionalisierung—Von der Nebenstundenpoesie zum Buch als Freund. Eine emotionspsychologische Studie zur Literatur der Aufklärungsepoche*. Paderborn: Mentis-Verlag, 2006.
- . “Objects of Empathy. Characters (and Other Such Things) as Psycho-poetic Effects.” In *Characters in Fictional Worlds*. Ed. Jens Eder, Fotis Jannidis, and Ralf Schneider. Forthcoming.
- O’Neill, Patrick. “Points of Origin: On Focalization in Narrative.” *Canadian Review of Comparative Literature* 19 (1992): 331–350.
- Smith, Murray. “Aesthetics and the Rhetorical Power of Narrative.” In *Moving Images, Culture and the Mind*. Ed. Ib Bondebjerg. Bedfordshire: University of Luton Press, 2000. 157–166.
- Stanzel, F. K. *Narrative Situation in the Novel: Tom Jones, Moby-Dick, The Ambassadors, Ulysses*. Trans. James P. Pusack. Bloomington: Indiana University Press, 1971.
- . *A Theory of Narrative*. Trans. Charlotte Goedsche. Cambridge and New York: University of Cambridge Press, 1984.
- Sugiyama, Michelle Scalise. “On the Origins of Narrative: Storyteller Bias as a Fitness-Enhancing Strategy.” In *Human Nature* 7 (1996): 403–425.
- . “Food, Foragers, and Folklore: The Role of Narrative in Human Subsistence.” *Evolution and Human Behavior* 22, no. 4 (2001): 221–240.
- . “Lions and Tigers and Bears: Predators as a Folklore Universal.” In *Heuristiken der Literaturwissenschaft. Disziplinexterne Perspektiven auf Literatur*. Ed. Uta Klein, Katja Mellmann, and Steffanie Metzger. Paderborn: Mentis, 2006. 319–331.
- Woolf, Virginia. *Flush. A Biography*. London: Hogarth Press, 1983.