Constructing speaker images: The problem of enunciation in discourse analysis

PEKKA SULKUNEN and JUKKA TÖRÖNEN

Introduction

The social world is constructed in communication. Since Berger and Luckman (1966) this has become a widely accepted slogan for a variety of sociological approaches, but the obvious corollary that the construction is an act of a subject is usually forgotten. Even constructivist social problems research, which is heavily oriented to the study of rhetoric (Spector and Kitsuse 1987 [1977], Ibarra and Kitsuse 1993), has difficulty in recognizing this and giving answers to what kind of subject it is and what the relationship is between the subject and the world that is being constructed (Agger 1993).

Constructing, identifying, or reconstructing the subject of action and the subject of discourse on action is a general problem for any semiotic, phenomenological, or hermeneutic sociological theory, but in this article we approach it more narrowly from the perspective of discourse analysis, and call it the problem of enunciation. Any speech or text inevitably constructs an image of somebody who issues it, and by the same token also of those who are supposed to receive it. These images are more or less fictive and often hidden or left in the shadow of what we usually believe to be the ‘message’. In scientific texts, for example, the author is often thought to disappear behind the objective facts or results that are being reported. In popular narrative fiction the author may be irrelevant and invisible. In everyday speech, subjects are apparently present and visible but often ignored by sociologists who are looking for opinions, attitudes, or facts unless they are specifically interested in interaction, as in conversational analysis.

Nevertheless, some image of a speaker or an author is always implied if not made explicit in any text or speech. In scientific texts, the author is usually presented as someone whose knowledge is relevant (to the assumed reader) and can be trusted. In literary texts authors or speakers often appear in complicated networks of communication, and in everyday
speech the positions or the points of view from which people observe and comment upon the world may be more relevant than the way people conceptualize the world they are observing. For example, in qualitative interviews on peoples’ opinions on some political matter it may be less interesting to look at their approval or rejection of a policy than to pay attention to whether they are talking from the point of view of a ‘client’ or an ‘object’ of the policy, or from the point of view of policymakers (Sulkinen 1992).

This article aims at conceptual clarity in this area, in order to identify some mechanisms by which speaker images are constructed and to point out certain regularities in their operation. Two devices, we suggest, are of particular importance for sociological analysis: projections and enunciative modalities.

**Speaker images and the functions of enunciation**

In sociologically oriented discourse analysis the image of the issuing subject has been central in critical linguistics (Kress and Hodge 1979; Fowler et al. 1979; Fowler 1991), sociosemiotics (Hodge and Kress 1988) and critical discourse analysis (van Dijk 1993). For them, the focus is on asymmetrical relationships of communication, in which one participant (usually the 'sender') has more authority than the other.

For us, enunciation is in a more general sense the dimension of discourse where speaker images are created, and relationships of authority are only one albeit an important aspect of them. In terms of intertextual communication, the structures of enunciation have four major functions. First, they are important in arousing emotions; second, they are vital in building up motivation of addressees (audience, readers) to follow the argumentation; third, they are used in establishing a contract of confidence between the text or speech and its addressees; and finally, they operate to produce positive (but more or less imaginary) self definitions of the speakers or authors.

However, it is a complicated dimension for several reasons. First, speaker and receiver images tend to get confused with the concrete situation of interaction in which the speech or text is issued. For the sake of conceptual and analytic clarity these two should be distinguished. Speaker images – for example, a certain way of authorizing a scientific text – may be cultural facts in the same way as a myth or a conventional or even ritualized way of speaking. But they are also reflections of the concrete interactional setup in which they are issued and received. It is not always easy to conceptualize the difference, as we shall show below.

The second problem is that the dimension of enunciation consists of a web of embedded structures that may be very complicated. For example, a television news report is embedded, first, in the general framework of a news broadcast, then in an introduction by the anchor, then in a close-up of a reporter in a visual background relevant to the report. Finally, several people may be interviewed and they may refer to a number of statements by others, etc. How should we conceptually structure this embeddedness, and which speaker image are we interested in anyway?

Third, the devices by which speaker images are constructed are complex. Simple devices of personal, temporal, and spatial disengagements and re-engagements (Greimas and Courtés 1979) may be involved to distinguish the time, place, and actors from the here and now of telling a story, as when someone says, “Would you believe what happened this morning when John was driving his car out of the garage...!” But also more invisible modal devices often imply a speaker image, like when a politician says: “Our children must be protected from street violence”.

**‘Narratology’**

The dimension of enunciation has been an essential element in the theory of rhetories since Aristotle. The three classical forms of persuasion — ethos (ethical appeal), pathos (emotional appeal), and logos (rational appeal) — are strategies of convincing the audience of the trustworthiness of the speaker, of committing the addressees to the cause, and of grounding their trust in reason (Aristotle 1946; Noth 1990: 340). The New Rhetoric has widened the scope of rhetorical analysis to stress that all argumentation, however logical, employs implied and informal techniques of convincing an ‘audience’ that cannot be boiled down to formal analytic reasoning (Perelman and Olbrechts-Tyteca 1971; Perelman 1982).

The rhetorical perspective is looking at speaker-receiver images from the point of view of convincing or persuading. This is an important function of the enunciative dimension, but not the only one. The problem was revived from a different perspective in structuralist literary criticism and semiology towards the end of the 1960s (Communications 8 [1981 (1966)]. In French semiotic and linguistic theory the issue about the subjects of communication became increasingly relevant in the course of the 1980s, partly influenced by pragmatic developments in Anglo-Saxon linguistics and language theory (Forest 1986).

In literary studies the problem became so important that Tzvetan Todorov (1969) gave it a name: ‘narratology’. A literary text is seen to consist of two different and hierarchically organized elements: a story
and a discourse. The story represents what is being told; discourse consists of all the structures of telling (Chatman 1978; Culler 1981; Prince 1988: 21). The story is usually thought to consist of actors (characters) and a temporal series of events that is independent of the situational interaction of telling it. Discourse, on the other hand, refers to techniques of telling the story (past or present tense, active or passive narrator, where the story takes place, etc.). This same distinction is made in different terminologies but they correspond quite precisely. Early Russian formalists talked about ‘fabula’ and ‘sposo’. Todorov himself talked about ‘histoire’ and ‘discours’, and Claude Bremond (1964, 1973) in his theory of folktales distinguished a ‘récit raconté’ from the ‘récit racontant’. At some point there was a tendency to believe that underneath apparent verbal expressions an invisible meaning structure could be detected, and that structures of enunciation (who is speaking to whom from what perspective) would belong to the former. This idea was reflected in Kristeva’s (1969a) terms ‘genotext’ and ‘phenotext’ and in Greimas’s (1966) scheme of the ‘paracours générique’ with its immanent and manifest levels of meaning (Neith 1990).

Sometimes three different levels or dimensions have been distinguished. Barthes (1966), for example, thought that the ‘lowest’ level consists of functions and indices that form a ‘text’ which is governed by structures that refer to the interaction between the (image of an) author and the (image of a) receiver. Genette (1980) talks about ‘histoire’, ‘récit’, and ‘narration’, a tripartition that now in standard English terminology is referred to as ‘story’, ‘text’, and ‘narration’. This tripartition resembles Halliday’s (1978) three functions of language: mimetic, textual, and interpersonal.

A literary text is an aesthetic object, and the organization of its content as a whole that begins somewhere and proceeds to an end is a key element of its structure. However, the structures of narration may be quite difficult to disentangle from other elements of textual organization, as will be shown below.

For our purposes the textual structure is less important, because sociological texts sometimes have no organization at all (for example, qualitative interviews), or because the textual structure is of little interest. Therefore we prefer to speak of the enunciative dimension and the dimension of utterance (Sulkinen and Törnänen 1997) instead of using the standard terminology. The choice of this terminology implies that the meaning content of a text cannot be separated from the effects of meaning that it acquires from the structures of enunciation. On the other hand, it implies that the structures of enunciation are not always unambiguously distinct from the content of a story, and therefore looking at texts from this perspective is an analytic strategy rather than an effort to develop a general model.

The narratological way of looking at the enunciative dimension can be illustrated by a simple model (adapted from Boeth 1961; Chatman 1978; Tamm 1985, and Chatman 1990; see Fig. 1).

A literary text is a medium of communication between the real author (RA) and real readers (RR), but this communicative interaction is of course not visible in it. Structuralist theory of literature explicitly excluded the biographical author from its horizon of interpretation, but could not deny the importance of intratextual subjective structures. Any literary text includes a number of internal narrators (N) and their narratees (Ne), for example, in dialogues. Often, but not always, the whole text is organized as a story told by a stable narrator (N) who reports all the settings, events, actions, and feelings of the persons in the story. Sometimes, as in the novels of Toni Morrison (e.g., ‘Jazz’) the role of the first person Narrator may change from one character to another so that the same events will be told from different perspectives. Whatever its narrative composition, all well-structured texts have an implied author (IA) (Iser 1978) who is responsible for the whole. As Chatman (1990; 84) suggests, an alternative term for the implied author could be the inferred author, because as a subject of the whole text it cannot appear in the text but can only be inferred by the analytic reader.

As soon as an author appears (as in the following sentence: ‘The author of this book will attempt to present a theory . . .’) it is — in terms of this model — either in the position of a narrator or even an internal narrator (who will, from now on, be quoted by the narrator in this example). As pointed out above, the system of embedded structures can be very com-

![Text diagram](image)

Figure 1. The narratological model
plex even in the most ordinary texts, such as television broadcasts. In the simplest case, at least one narrator (N), however invisible, is always necessarily apparent in all texts.

It is the concept of the implied author in this model which has most aroused debate. Even accepting the exile of the biographical author imposed by the structuralist program, the idea of the implied author is filled with difficulties. If such an author is never apparent but must be detected by an analytic reader, what is its reference? Some have argued (Pietäiä 1992) that the implied author in fact refers to conventions of the genre, or to a personal style (image) of the author (Booth 1961) or to the communicative structure of the text as a whole (Chatman 1990); while some others (Genette 1988) believe that the concept is unnecessary and confusing.

From our point of view the narratological model has limited applicability in three respects. First, the assumed symmetry of embedded structures may be misleading. Many texts have complicated structures of internal narrators, quoting each other or giving the floor to successive subnarrators that may not be hierarchically organized and may not have corresponding internal addressees. A journalist, for example, may quote or interview several persons who communicate directly with the assumed receiver (compare Goffman 1981: overhearers). Besides, as is often the case in television news reporting, the viewer is depicted both as the 'sender' who expects the journalist to find out information and as the 'receiver' or beneficiary of that information. In general, it might be said that in most sociologically relevant texts the image of the narrators, authors, or speakers is much more interesting than that of the receivers. The second limitation is related to the idea of the implied author. Whereas in literature research the global embedded composition of the text may be inherently interesting, in sociological analyses of interviews, for example, the global structure of the interview may be of very little relevance. And thirdly, the devices that produce speaker images are often much more subtle than explicit structures of embedding. Modal relationships, for example, can as such have this function.

Voice and focalization

Gerard Genette's (1980) distinction between voice and focalization is an important contribution that essentially widens the scope of the narratological model. In every text one or several voices can be heard telling something, but the same voice can report things from several points of view. In a romantic novel, for example, the narrator - who is relatively invisible and reports on settings, actions, and feelings firmly as if possessing absolute first-hand knowledge — may tell the reader about the feelings of the heroine but can switch in the next sentence to report on the feelings of the hero.

Following Bal (1977: 29), Rimmon-Kenan (1983) adds to this the distinction between the subject and the object of focalization. The subject of focalization fixes the attention of the reader on the object's point of view. The same subject may change the object of focalization, and vice versa, different subjects may focalize the reporting on the same object.

Paula straightened her blouse nervously. The curve of her breast did not escape Juan's attention. He didn't try to conceal his interest in Paula anyway. An investigating gaze swept over her from top to bottom and stopped at her face. She tried to stand it but she had to turn her face away. (Ker 1990 [1988]: 13. Translation ours)

In this example the narrator stays the same (the 'author' who reports on actions and feelings of characters) but the focalization changes. In the first sentence, the subject and object of focalization is Paula, as if she were looking at herself in a mirror. In the next two sentences the object of focalization is still Paula, but its subject is now Juan. In the last two sentences, the subject of focalization is again Paula, first interpreting Juan's (object) way of looking at her, and in the last sentence observing her own reactions (object) to it.

Focalization may be oriented to feelings or other mental states (internal focalization) or it may be limited to observing external actions or states of things (external focalization). In the latter case, also called zero focalization (Rimmon-Kenan 1983), the reader or audience is not explicitly informed about the inner worlds of the persons described, but may still be led to make conclusions about them (Prince 1988: 32).

One of the big questions of literary theory is to what extent focalization and voice are necessarily borne by different (internal) narrators. The concept of voice was originally developed by Bakhtin (1981), and later Ducrot (1984) used it for his theory of irony. Even the same narrator can use simultaneously multiple voices, and when one denies what the other says, ironic effects are produced, as in the utterance: 'this is the happiest day of my life' [spoken by an entrepreneur who has just learned that her undertaking has become bankrupt]. In principle, voice and focalization may be situated in one narrator, but this is a particular case. Therefore, as Chatman (1990) suggests, the two concepts should be kept separate: '... separate mental behaviors, stances, attitudes and interests of narrators and characters require separate terms' (1990: 141). The narrator reports but (usually) does not experience or observe things in
the same sense as the characters of a story unless the narrator is one of them. Still, the narrator can have attitudes (sympathy, hostility, disgust, belief, etc.) to what is being told, and Chatman suggests the term 'slant' for them. Instead, the mental acts of characters are said to be 'filtered' by the narrator who may report on them selectively from the point of view of one character instead of others.

In the theory of literature the concepts of voice and focalization are important elements in the study of style. For sociological uses, similar ideas are relevant, for example, in the analysis of the points of view taken by interviewees, the mechanisms of persuasion in polemical texts and the strategies of acquiring legitimacy and competence for authors of scientific texts. However, we are going to suggest a slightly different way of conceptualizing the problem.

Direct, indirect, and free indirect narration

Focalization is therefore a very subtle aspect of enunciation. But even detecting the voice — i.e., who is actually speaking — is not always a simple matter. In direct and indirect speech, two voices can be clearly heard as in (a) and (b) in the following examples:

(a) Peter said: ‘John was shopping there’
(b) Peter stated that John had been shopping there.
(c) He (John) was shopping here.

The narrator in (a) reports that character 'Peter' said something, quoting his exact words, but in indirect speech (b) the narrator reports on an observation of a speech act. Direct and indirect speech are fairly clearcut cases governed by grammatical conventions such as using quotation marks, the hierarchical order of principal and subordinate clause, tenses, etc. In both cases several levels of embedding may be involved, but it is always clear whose voice is speaking. Instead, free indirect speech (c) has caused headaches for theorists in this respect. Who is speaking, the narrator or the character? Toolan (1988; see also McHale 1978) suggests that both speak simultaneously, and says that free indirect speech is a double discourse of the two. Often, as in the example (c), the narrator communicates an experience or an intention of a character on the basis of intimate knowledge that is indicated by the deictic expression ‘here’.

In a television news report such intimacy is often obvious, because the camera can show a journalist on the spot reporting on what people are doing, the reporting can be in the present tense, and the journalist may confirm his or her interpretation by interviewing the people. In literary texts strange effects can be produced by combining spatial proximity of the (hidden) narrator and the character with their temporal separateness (the past tense in [c]).

Free indirect speech is not governed by clear-cut grammatical conventions, and it is sometimes used to explicitly transgress conventional uses of language. As the term suggests, it is stylistically free and frequently employed in 'experimental' or 'avantgarde' literature (Toolan 1988). In that role, its function is to deepen the polyphony of the text by expressing the multiplicity of speakers and attitudes. It can do this, then, when it is possible to identify the speakers of each text passage but especially when this cannot be done (McHale 1978). However, free indirect speech is also frequently used in 'the realist tradition of narration' to help readers to identify themselves with the characters (Rimmon-Kenan 1983).

The dimensions of enunciation and utterance

The research on embedded structures, voice, focalization, and especially free indirect speech, leads us to conclude that the dimension of enunciation cannot be pressed into a comprehensive and logical model. This is perhaps a conclusion that applies to language use generally, we can identify problems and issues related to meaning effects, then study linguistic devices that may be used to produce them, but never achieve a point where we could tell exactly which devices produce which effects under what conditions. In the end, language is and will remain 'unknown' (Kristeva 1969b).

Therefore, what we are proposing as a 'model' is not meant to simulate language use but to formulate sociologically relevant questions about discourse (and of social action). For this purpose it is useful to retain the distinction between discourse and story, but for reasons explained above we prefer to call them the dimensions of enunciation and utterance. Utterance refers to that which is told about the world, how it is organized into structures, and how relationships between its elements are defined. Enunciation refers to the positions from which this world is described. The textual level, and with that the notion of the implied author, is of lesser interest unless we are specifically dealing with conventions of a genre like certain types of television soap operas, newspaper reporting, or even scientific texts.

Both dimensions may be more or less coherently organized into any of the three textual configurations we have distinguished elsewhere: narratives, classifications, or explanations (and of course their combinations) (Sulkunen 1992). A televised news report often tells not only the story
that is the subject of the news but also a story about how the journalist has found out about it. Scientific texts also often report their findings within a story about earlier research that has failed and then telling another story about the present study that succeeded in discovering true knowledge (Griegas 1979; Swales 1981; Pratt 1986). Classifications can be involved in phrases such as ‘We are the kind of people who couldn’t care less whether alcohol is available in grocery stores or not’ (‘we’ contrasted with others). The enunciative dimension can also be dressed in the form of an explanation, such as ‘What makes people believe that [...] is that it makes them feel better than others.’

Already these simple examples show that there is hardly any kind of socially relevant speech or text in which the dimension of enunciation is not apparent somehow. There are cases where this dimension is very poorly developed. Typical examples are popular novels such as romances (Radway 1984) and adventure stories or news reports in newspapers (Petilä 1992), where the narrator may remain almost completely invisible. This we call zero-degree enunciation (cf. Kiefer 1987: 80, 92). It corresponds approximately to the case that Greimas and Courtès call expressions of utterance (énoncé énoncé), in contrast to expressions of enunciation (énonciation énoncée). The latter term refers to cases where the dimension of enunciation is made explicit.

In zero-degree enunciation the invisibility of the narrator does not mean its inexistence. On the contrary, in these cases the narrator is omniscient, knowing and reporting exactly and firmly everything relevant that happens in the story. Especially in reception studies, it may be relevant to ask how readers react to such reporting in terms of how much they have confidence in the narrator (usually identified with the named author of the text) (Törnönen 1995).

It should be emphasized once again that the dimension of enunciation in the text is never an exact image of the real act of issuing it. To what extent it is a representation, a means of reproduction, or even production of the conditions of its own production (Mainguenaï 1987) is a complex epistemological issue that cannot be discussed here. For now we are only interested in what appears in the text itself.

The mechanisms that regulate the enunciative dimension of the text and its relationship to the real act of enunciating it are called disengagement (désenclavage) and re-engagement (ré-enclavage) in Greimasian semiotics (Greimas and Courtès 1979). The subject of enunciation must first (in the logical sense) be distanced from the subject issuing the text in the here and now. This can take place on two levels. In the example (a) above, Peter is first distanced from John simply by naming differently the subject of enunciation and the subject of the utterance. This kind of disengagement is called désenclavage énoncé; the speaker is not the person who is spoken about. This is the only possibility when no trace of the dimension of enunciation is visible (énoncé énoncé). On the second level, the subject of the enunciation is distanced from the (hidden) narrator who reports that it is Peter who is saying something, not him or herself. This is called désenclavage énonciatif, because it brings to light explicitly the (intratextual) structure of enunciation (énonciation énoncée). Disengagement can be actorial, as in (a), but also temporal or spatial, distanciating the subject of the enunciation from the real speaker/author as different actors, acting in different time and space, or any of the three together.

Re-engagement refers to procedures that create illusions of the presence of the real speaker/author in the discourse. A good case is free indirect speech in (c) above, where the actorial and temporal disengagements first inform us that the narrator is not the subject of the utterance. Then the spatial re-engagement (‘here’) creates the illusion of the simultaneous presence of both the narrator and the reader with the subject of the utterance (John shopping). Greimas and Courtès (1979: 119–121) emphasize that re-engagements can only produce effects of illusion, parallel to those of the referential illusion: the real act of issuing the speech and text remains always outside the horizon of the (intratextual) analysis of discourse, as the real world will always remain outside the use of language that refers to it.

The simultaneous account of disengaging and re-engaging mechanisms is capable of seizing infinitely fine nuances that may be of great importance in the analysis of poetic language. For example, the discordance between actorial and temporal disengagement with spatial re-engagement in free indirect speech in (c) illustrates that the play in the enunciative dimension may produce quite strange effects of simultaneous distance and presence of the narrator and the characters that are subjects of narration.

Projections

In sociological use, awareness of such mechanisms is helpful but is not the objective in itself. For our purposes the most important function of disengagements and re-engagements is that they connect the enunciatore (narrator or subnarrator) with the utterance. In doing this, they define the point of view of the subject to the world that is being talked about. ‘We are the kind of people who couldn’t care less whether alcohol is sold in grocery stores or not, but young people and the down-and-outs should
not be permitted access to alcohol". This statement expresses an opinion on a policy matter, but it also divides the social world into 'us' (who are competent self-controllers) and 'others' (who are not competent), and furthermore accepts that the competent peoples' point of view should dominate in policy making.

Here we have a very simple case of connecting characters of the enunciative dimension with characters of the dimension of utterance. The re-engagement (*en narration énoncé*) ('we are') creates an illusion of immediate presence and vests these opinions and positions with the form of a factual statement.

Such connections between the dimensions of enunciation and utterance will be called projections, because they transmit properties of actors in the world to characters in the enunciative dimension and vice versa. In the example above, the division of the social world into competent and incompetent people (as if an objective fact) is transmitted to a positive self-definition of the speaker as a member of a collectivity. In the semiotic sense the example is trivial, but nevertheless worth some attention because it is a common and important way of constructing positive identities.

Much more interesting cases are those where the enunciative dimension is made explicit and elaborated (*énonciation énoncée*). A case in point is a televised news report on the Finnish Federation of the Left (Finnish Public Television, Channel 2, February 18, 1992) that we have analyzed in detail (see Fig. 2). The report consists of two stories, one in the dimension of utterance and the other in the dimension of enunciation. The story in the dimension of the utterance reports how the Federation, recently formed on the ruins of the Soviet-oriented Communist party and its cover organization, the Democratic Peoples' Union, has set itself the goal of becoming a modern, nondogmatic and European-oriented left-wing political party. It is torn with contradictions between modern Euro-Communists and the old-style Soviet Communist elements. The focus of the story is on who is actually in power in the new Federation and by what means. The answer suggested is that old-style Soviet Communists control the finances of the Federation because the old Communist party had secret financial sources from the Soviet Union. These sources remain secret and interviews of the President and the Secretary General reveal their inability and unwillingness to discuss the matter. The story ends with a shot of a ski jumper falling injuriously after a bad leap.

The story of enunciation first sets the journalist (subject) in motion to find out the truth (object) of the Federation's financial sources, in the interest of the telespectators (both senders of the journalist's mission and receivers of the object). The best efforts of the journalist fail because of

![Diagram](image)

Figure 2. Federation of the Left

the tenacity of the anti-subject, the old-style Communists, in refusing to divulge any information on the matter.

The style and structure of the report appear to conform to the requirements of neutrality and objectivity in political journalism, yet it creates anger in the viewers against old-style Communists and sympathy for the modern wing. The reason is the projection of the story of the enunciation onto the story of the utterance. The anti-subject of the former (old-style Communists) is also the anti-subject of the latter. Therefore the telespectators who are in the actantial role of sender-receiver in the enunciation story, frustrated by the anti-subject, are placed on the side of the subject (the modernists) in the story of the utterance. Even if the latter is told
in an apparently neutral way, not assigning bad qualities or intentions... evidently bad qualities or intentions viewers if the modernists were to gain the upper hand, the projection... projection device of localization that in Chatman’s terms should be called a slant... the narrator’s (the news studio, represented by the anchor and the camera) attitude to what is being told, while the internal worlds (emotions) of the actors are almost completely ignored.

A different but interesting type of projection was identified in the lyrics of Finnish juke-box songs that were popular in working-class pubs in the early eighties:

Tarja, at times I spy your eyes in someone who isn’t you, they take me back to memories so real that they seem true. But then I awake and come to with a start.

Tarja, no one can replace you in my heart. (Sulkunen 1996; trans. by Andrew McCafferty)

The speaker ‘I’ reports his own feelings and addresses himself to ‘you’. Tarja, as if in a dialogue. The dialogical form of the enunciation in many of these songs simulates kitchen table conversations, conversations in the pub or, as in this example, a conversation that a man can imagine having with a lost spouse or girlfriend and that a woman could wish to participate in. The function of this kind of simulation in the pub situation is that it helps the listeners to identify their feelings with those of the characters of the lyrics, to the extent that the lyrics expresses their feelings for them. In fact, the pub-goers did indeed occasionally use the song lyrics as elements of their conversations (Sulkunen et al. 1985: 102–139, Sulkunen 1996).

As these examples show, projections may be important elements in the analysis of emotions, and this point will be developed in a subsequent speaker images, and in this they function in several ways. First, projections (“We are the kind of people...”) or to listeners (“Tarja”). Second, they example, a typical rhetorical figure in political speeches is like the following: “We all know that A... But we also know that B... Should we not therefore try to do our best that X...?” The projection of the ‘us’ who here communicate binds the audience to the point of view of the speaker, who also is an actor in the world that is being talked about.

Third, projections are important elements in establishing an alliance of faith between the speaker/author and the assumed receiver of the message. One type of such alliance is typical of journalistic discourse exemplified by the Left Federation report discussed above. Another wide area is the contract of confidence between the author/speaker and the reader/audience, assigning legitimacy and competence to the former. Here the role of enunciative modalities is essential.

Enunciative modalities

Especially in persuasive text or speech it is essential that the enunciator has authority over the subject matter and the right and credibility to influence whoever receives the speech or text. It is unconvincing if somebody just says: ‘Alcohol and tobacco advertising should be banned in the name of public health’ or ‘Social structure is a simultaneous process of differentiation and homogenization’. Even if the argumentation in itself would be clear, we need to know on the basis of whose authority these statements are made and for what reasons we should be committed to them. Whether it is to cajole readers into believing (as in scientific texts, transmitting knowledge) or into doing (as in polemical texts, transmitting willingness to act), the one who persuades must have a relationship of competence and of legitimacy to the subject matter and to the assumed reader. In both respects the modalities defining the degree of truth and certainty of utterances are of special importance.

We have elsewhere (Sulkunen and Törnönen 1997; compare Greimas and Courtès 1989) divided them into two groups, those evolving around the axis of being and appearing and those evolving around the axis of knowing and believing. Both of these groups are called enunciative modalities, because they make the enunciative dimension explicit. In the former case, the enunciator reports and comments on what is (not) and appears (not) to be the case also from the enunciatee’s point of view. It deals with ontological truth in the sense that it asserts what the world is (not) like to anyone observing it. Therefore we call this the group of verdictory modalities (see Fig. 3).

In the second case, the enunciator reports and comments on received knowledge and beliefs. Instead of evaluating appearance in terms of what truly is, the enunciator evaluates the degree of certainty of statements. This requires expertise or epistemic competence that is superior to that of the enunciatee. Therefore we call this the group of epistemic modalities (see Fig. 4).

The contrast of certainty of a statement P, it should be noted, is not certainty that the statement P is false. It is, rather, absolute uncertainty, and we prefer to call it imagination, because assumption and doubt are
leads to increasing problems] during the era of middle strength beer [in the period
when medium beer has been liberalised from the alcohol monopoly].

(b) imagination: 'Nobody knows for sure how high the costs caused by these
harms are for the society. In any case, the amounts circulate around billions of
marks, on top of which come human suffering and losses that are immeasurable
by money'.

As in other modal groups, the 'meta-terms' on the sides of the squares are
logical and not descriptive. They define the enunciator's position vis-
à-vis the world's appearance to anyone, or vis-à-vis received knowledge
or beliefs, and they can be formulated in infinitely different semantical
ways. 'Assumption' in the epistemic square, for example, stands for
positions as different as falsity, ignorance, prejudice, guessing, or convic-
tion. These squares are not semantic models but ways of asking questions
about the position of the speaker image in relation to the enunciator and
to the world that is being talked about.

In the construction of speaker images, the difference between veridic-
tory and epistemic modalities is essential. Veridictory modalities place
the enunciator and the enunciatee in the same footing, looking at reality
from the same point of view, and therefore their use creates an alliance
of solidarity between the two. In the preceding example of the Left
Federation, the solidarity between the journalist and the telespectators is
further enhanced because the story of enunciation is a story about veridic-
tion: The financial sources of the Federation are known to exist but the
anti-subject prevents the journalist to make them apparent, to transform
a secret into plain truth.

In scientific texts the veridictory modalities may have the function of
soliciting the motivation and solidarity of readers. The following example
is the opening paragraph of a book on macrosociological theory:

Attempts to define precisely the concept of community and its subconcepts —
the group and the society — may seem unnecessary and boring. Both in sociology
and in everyday language one can in normal cases safely speak of communities
and of societies without exact definitions of them. People would know anyway
what we are talking about. ... However, if one attempts to find a system of
description for community, a definition is needed. It not only restricts the scope
of the area we are studying but also informs us of certain variables that can be
used to compare different kinds of communities. The variables will describe
certain basic processes in all communities. (Allardt 1964: 1)

The narrator first recognizes the apparent futility of academic definitions
of commonsensical things and then states that what thus appears is an
illusion (veridictory modality), as the reader, too, will readily recognize.
The sameness of the readers' and the narrator's perspective is further
underscored by the distinction between (ordinary) ‘people’ and ‘us’ (sociologists).

Epistemic modalities are commentaries on the validity of received beliefs or knowledge, and whether they are affirmative or negative they raise the enunciator above the level of the enunciatee, into an intermediary position. Their function is to establish an alliance of legitimacy between the enunciator and the enunciatee on the basis of the former’s competence. Sometimes a newspaper editorial may affirmatively refer to scientific research (that is assumed to enjoy the reader’s confidence) to back up an argument, but this is a relatively weak strategy of persuasion, as in the following quotation from a newspaper editorial (Kansan Tahto 1993; Torrønen 1995):

An interesting contribution to this debate comes from researchers of forensic medicine of the University of Helsinki. Their study shows that medium beer has clearly increased mortality from cirrhosis of the liver during the last couple of decades.

Much stronger effects are achieved when the subject of enunciation first reports on received knowledge (that readers might not have had access to), then evaluates it as more or less justified and finally promises new knowledge.

The sociological text quoted above continues with 22 pages of veridictory statements on how various appearances in social life are or are not what they seem, developing an alliance of solidarity with the sociological readership and mounting up its motivation to acquaint itself with the argument that ‘the author of this book’ is going to make. However, before the presentation of the argument begins, the narrator turns to earlier research, listing a number of misleading results and contradictions in treating social differentiation and homogeneity as one single dimension, differentiation leading to conflicts and change while homogeneity is thought to be the foundation of stability and order. As a consequence of such difficulties, authors like Ralph Dahrendorf believe that theories of differentiation and theories of homogenization should be kept separate.

His starting point is that there are phenomena that only the theory of integration (homogeneity) can explain, and others that can only be explained by theories of coercion (differentiation). These theories explain some aspects of the same reality, and, according to Dahrendorf, for the present time they cannot be combined. As a programme of theoretical research, Dahrendorf’s position is surely too pessimistic. In any case, this book will attempt ... to present a theory that combines elements from both approaches ... The intention here is to show that strong pressure to homogeneity has very different consequences depending on other structural characteristics of society. Respectively, attempts will be made to show that conflicts and cross-presures have different effects in different social structures, especially in societies with different degrees of differentiation in the division of labour. (Allardt 1964: 26)

The last four pages of the introductory chapter are devoted to accumulating legitimacy for ‘the author of this book’ as a competent commentator of scientific knowledge who will here suggest some new solutions to previously known problems.

Allardt’s text is academic and educating. It dates back to the heyday of positivism, and although not committed to any strict epistemological doctrine, it respects the ideals of objectivity and impartiality in scientific discourse. It does this with a strong emphasis on the contract of confidence between the author and the presumed academic readership, and makes claims for legitimacy by emphatic use of not only veridictory but also epistemic modalities.

For comparison, let us look at another type of scientific text where a different relationship between the author, the presumed reader and society is established in a strongly emotional way by the enunciative devices discussed above.

The promise of the Enlightenment: The Sociological Imagination

The Sociological Imagination by C. Wright Mills (1970 [1959]) is a brilliant example of how the enunciative dimension can be used to develop a commitment of the reader to a battle on the scientific field for a moral cause higher than the simple advancement of knowledge. As is well-known, the book is a eulogy of Enlightenment values: freedom and reason, in the midst of a postmodern mass society that is descending from the overripe ‘political age’, with disastrous consequences for mankind, in the form of massive economic crises and the threat of a nuclear war, and for the individual in the form of lost self-determination and destruction of free intelligence. The first chapter of the book, ‘The promise’, sets out the programme and defines the respective positions of ‘men’, ‘the intellectual community’, ‘the reader’, and the ‘author of this book’ in this world.

Norman Denzin, one of the leading American affirmative postmodernists in sociology, finds Mills’s nostalgia for the modern values deluding, even oppressing. For him, The Sociological Imagination is a hypocritical text, pretending that people are worse off than they really are, and wanting to move the reader ‘by altruistic moral indignation about the
way sociologists like Parsons and Lazarsfeld do sociology'. Therefore, 'His manipulation of me, flesh-and-blood reader, to his own ends erases my trust. His book is unethical, totalitarian, manipulating me and all "nowadays" persons to his own ends' (Denzin 1991: 58).

Denzin's negative reaction is not only a sign of disagreement about the values of Enlightenment as promoted by Mills. It is also a way of taking a distance from a text that is constructing social reality as a contradiction between these values and the reality of contemporary society and sociology, formulated in a rhetoric that derives its power from the use of the enunciative modalities and projections.

'The promise' is a typical polemical text in that it is arranged in the form of a qualifying test in a narrative scheme (Sulkunen and Torronen 1997). It is an actantual structure that does not 'move'. The subject and its helpers object, counter-subject and its helpers (opponents to the subject) are placed in their positions, the task is defined and motivated, but the subject never sets out to encounter the anti-subject or to complete the task. The readers remain in a state of suspension, waiting and hoping to see the hero's victory, and willing to lend him their support.

The hero is 'the sociological imagination', helped by the classical sociological tradition. The object is true knowledge of structural social factors that underlie peoples' private troubles. The adversary is false sociology, represented by three 'tendencies' that distort the classical sociological tradition. To the first belong the historical prophetic views by Arnold Toynbee or Edward Spengler. The second is Parsonsian systematic theory of 'the nature of man and society', the third, the empiricist research technology that has developed into a bureaucratic ethos with no insight to the structural realities of American mass society.

Nowadays men often feel that their private lives are a series of traps. They sense that within their everyday worlds, they cannot overcome their troubles, and in this feeling, they are often quite correct: what ordinary men are directly aware of and what they try to do are bounded by the private orbits in which they live; their visions and their powers are limited to the close-up scenes of job, family, neighborhood; in other milieux, they move vicariously and remain spectators. And the more aware they become, however vaguely, of ambitions and of threats which transcend their immediate locales, the more traps they seem to feel. (Mills 1970 [1959]: 9)

In the famous opening paragraph of The Sociological Imagination a neutral narrator observes what appears to be the sense of being trapped to ordinary men and confirms that they indeed are. The veridictory statement builds an alliance from the very start between the narrator and the people who 'need, and feel they need a quality of mind that will help them to use information and to develop reason in order to achieve lucid summations of what is going on in the world and of what may be happening within themselves' (Mills 1970 [1959]: 11). That quality of mind, 'the sociological imagination', is offered to them in this book by the narrator.

Its task is to solve a problem: to reveal a secret that is but does not appear to be the reality in peoples' lives. The secret is the way private troubles — unemployment, urban life, broken marriages, and a general psychological malaise — depend on historical social structures, and the promise is that awareness of this dependency helps people to turn their private troubles into public issues and political action (Mills 1970 [1959]: 18). The secret that is most particularly troubling in contemporary society is the unawareness of values and threats to them, which produces indifference as a mental reaction. Another is anxiety, resulting from awareness of a threat but unclarity of the values that are being threatened.

The clarity of this narrative structure is further enhanced by a list of false heroes and their false solutions: psychologism (Mills 1970 [1959]: 19–20) and natural science (1970 [1959]: 20–23) as well as substitute heroes and their substitute solutions: art and fiction (1970 [1959]: 24–25) that for their best efforts have 'not the intellectual clarity required for their understanding and relief today'. The real hero is adequate social science called the 'sociological imagination' and offered to the people by the narrator.

In contrast to the text by Allardt analyzed above, 'The promise' employs only weak epistemic modalities of belief and knowledge in renouncing the false solutions.

It is true, as psychoanalysts continually point out [=believe], that people do often have 'the increasing sense of being moved by obscure forces within themselves which are unable to define'. But it is not true [=knowledge], as Ernest Jones asserted, that 'man's chief enemy and danger is his own untruly nature ...'. On the contrary, 'man's chief danger today lies in the untruly forces of contemporaneous society itself ...' (Mills 1970 [1959]: 19–20)

The contract of confidence is not very strongly based on the narrator's scientific competence. Instead, two projections focalize the narration to the point of view of ordinary people and incite them emotionally to support the author's mission. First, the narrator is identified as one of the people who need the sociological imagination: 'We have come to know that every individual lives ... in some society ... The sociological imagination enables us to grasp history and biography and the relations between the two ...' (Mills 1970 [1959]; emphasis added) and 'What we experience in various and specific milieux ... is often caused by structural
changes' (1970 [1959]: 17; emphasis added). The task of formulating the sociological imagination is felt to be necessary and beneficial not only for intellectual reasons of science, but because it has an emancipatory mission that is common to the reader, to other people, and to the author of this text alike. Mills is a 'populist' in the precise sense that he writes in the name of the people, not only of experts.

However, Mills is also a heroic writer. Towards the end of the chapter he switches to first person, projecting the public person, the celebrated sociologist C. Wright Mills, to the position of the narrator: ‘It is my aim in this book to define the meaning of the sciences for the cultural tasks of our time. I want to project the kinds of effort that lie behind the development of the sociological imagination ... I want to make clear the nature and the uses of the scientific disciplines today, and to give a limited account of their contemporary condition in the United States’ (1970 [1959]: 25). The first person singular develops an image of a tragic hero that fights against the corrupted but dominating tendencies in American social science (1970 [1959]: 27) sketched out earlier. The result is anger, in a way quite similar to the anger in the news report analyzed earlier. The narrator-hero (journalist: Mills) are struggling to arrive at true knowledge but their work is obstructed by forces that are allied with the power elite, the cause of the many private troubles in which the people appear to be trapped, and indeed are.

No wonder, then, that Denzin among others reacts to this text with emotion and rage. The populist radical Mills engages the reader in anger against his adversaries, and those who have lost faith in a political society based on Enlightenment values of modernity might easily see in his mission nothing but a private strategy in the battlefield of scientific prestige and fame. For us, Mills’s text is one example, perhaps extreme of its kind, of how even a scientific text can mobilize emotions and how it does this by the use of enunciative modalities and projections in the enunciative dimension.

Conclusions

From the constructionist point of view (Miller and Holstein 1993) the idea of a speaker image is intuitively appealing but so far undertheorized. Social reality is constructed in discourse about the world but the subjects of discourse are inevitably part of the world they are talking about. However, as the semiotic and literature research that we have discussed above has pointed out, the construction of speaker images is a very complicated matter.

In summary, we should like to stress two methodologically vital points. First, the images of the subjects of discourse are always images and not ‘reality’ as such. In what ways they are representations of the real world is a complicated issue that will need a separate discussion. Second, the devices of constructing such images are reflections of a triple relationship between (a) the subject of enunciation and its addressee(s), (b) the relationship between both of them and the world that is constructed in discourse, and (c) between the subject of enunciation and prior knowledge about the world.

We have identified two especially interesting devices - projections and enunciative modalities - that are relevant here. Projections are important in the analysis of emotions and in focalization. Enunciative modalities are important in establishing a contract of confidence between the enunciator and the enunciatee. Of these, veridictory modalities help to build up the enunciator’s motivation to follow the argument and to create a contract of alliance with the enunciator, while epistemic modalities have as their basic function to establish the enunciator’s competence and legitimacy.

Notes

1. As in all semiotic models of discourse analysis, terminology poses problems. We employ the many available terms referring to subjects of discourse in the following manner.

   Enunciator is a general term for the agent who presents the contents of a text. Its counterpart is enunciatee. In the same sense we speak of subjects of enunciation and their addressees. When we have verbal data it is natural to use the terms speaker/audience. In the case of written data, in turn, we can use either the terms author/reader or narrator/narratee. The narrator/narratee pair is appropriate when the text has a clear structure and the voice who narrates is coalescing from a relatively stable position. When the text has a complicated structure and is polyphonic it is better to use the terms author/reader.

2. In practice, difficulties may arise in distinguishing them whenever it is unclear whose belief or knowledge is being evaluated or from whose point of view. The essential difference is the enunciator’s competence. Even if there is a re-engagement that creates the illusion of identifying the enunciators with the enunciator, we would still speak of epistemic modalities, as in the following: ‘So far it has been believed that ... but as we all know by now, this belief is a sign of ignorance’.

References

Pekka Sulkimen (b. 1948) is a Docent in the Department of Sociology at the University of Helsinki as well as a Senior Research Fellow at the Social Research Institute for Alcohol Studies (pekka@sriaks.ifi). His research interests include semiotic theory, semiotics, and sociology of consumption. His recent publications include The European New Middle Class (1992), Keys to Sociology Perspectives on Society (in Finnish, 1995), Constructing the New Consumer Society (ed. with J. Hohnwood, 1996), and 'The production of values: The concept of modality in textual discourse' (with J. Torrønen, 1997).

Jukka Torrønen (b. 1964) is a Lecturer in the Department of Sociology at the University of Helsinki as well as a contract researcher at the Finnish Foundation for Alcohol Studies. He is interested in semiotic theory, semiotics, reception study, cultural studies, mass media theory, and welfare state theory. He has previously published 'The researcher reading and speaker image' (in Finnish, 1994), 'Analysing newspaper editorials on alcohol politics' (in Finnish, 1995), 'The genuine love in masculine world' (in Finnish, 1996), and 'The production of values: The concept of modality in textual discourse' (with P. Sulkimen, 1997).