

# *Other voices, other sources*<sup>1</sup>

JAN ANWARD

I miss your broken-china voice  
How I wish you were still  
here with me  
(Tom Waits: 'Hold on')

## 1.

"There was a knock at the door and I jerked involuntarily, adrenaline shooting through me. I glanced at my watch: 12.25. My heart was thumping so hard it made my hands shake. I crossed to the door and bent my head.

"Yes?"

"It's me," Charlie said. "Can I come in?"

I opened the door. Charlie was leaning against the frame."<sup>2</sup>

How does Kinsey Millhone know that "me" is Charlie Scorsoni? Benveniste's famous formula "Est 'ego' qui *dit* 'ego'" (Benveniste 1966: 260) tells us that the referent of "me" is the one who speaks that very "me". And the one who speaks "me" is, in the simplest case, the one from which the utterance comes forth, the source of the utterance. This is why we can't help attributing utterances to a ventriloquist's dummy.

---

<sup>1</sup> This article uses material from a project which I buried in 1993. I am grateful to Viveka Adelswärd, Jens Allwood, Chuck Goodwin, Linda Jönsson, Per Linell, Ulrika Nettelbladt, Ragnar Rommetveit, and John Swedenmark for bolts of lightning which helped kick this present incarnation to life.

<sup>2</sup> Sue Grafton: *'A' is for alibi*. Bantam Books, 1987, p. 175.

In this case, though, the source of the spoken "me" is hidden by a door, through which only his voice travels. But that voice is sufficient. Est 'ego' dont la voix dit 'ego', 'ego' is the one whose voice says 'ego'. Between the utterance and the speaker falls the voice.

## 2.

Humans, like other mammals, have both individual voices and the capacity to recognize individual voices. Although the nature of the capacity to recognize individual voices still remains something of a mystery (Nolan 1991), as does the precise nature of the properties which constitute an individual voice, it is clear that differences in i) mean fundamental frequency ( $F_0$ ), ii) mean pitch range of  $F_0$ , iii) volume, iv) voice quality, and v) articulatory settings may be heard as a difference in voice (Laver 1980), and that a combination of values on these dimensions can be identified, memorized and recognized as a particular voice, for example a high (i), monotonic (ii), thin (iii), hoarse (iv), and nasal (v) voice. Such a combination is most likely not a simple list of values, but rather some kind of Gestalt which remains invariant under transformation. Moreover, other factors, such as tempo, rhythm, and typical segmental qualities, also play a role.

Constellations of values on the dimensions (i – v) carry a lot of other information, as well. Speakers are individuated with a number of attributes attached.

These attributes may concern linguistic and regional background. The old notion of *Artikulationsbasis* and its phenomenological counterpart: the distinctive way in which a particular language sounds, involve these dimensions, and Elert & Hammarberg (1990) show that regional differences typically involve simultaneous contrasts in several of these dimensions.

The attributes may also be fairly stable social attributes, such as gender, age, and class. Mean  $F_0$  is a reliable indicator of gender and age (Karlsson 1990, Traunmüller, Branderud & Bigestans 1989), while voice quality is a good indicator of social class (Giles & Powesland 1975).

The attributes may also be rather transient ones, such as physiological state, emotional state, attitude, and footing (Goffman 1981). A change in mean pitch range of  $F_0$  may signal a change in attitude or emotional state (Williams & Stevens 1972, Bruce 1982), as may a change in volume. A change in voice quality may signal a change of physiological state, or a change in footing. A change in articulatory setting might indicate a change in attitude or footing (Gazdar 1980). However, even here, the contrasts are probably multidimensional. For example, House (1990: 115) concludes that the salient cues to emotional state (or mood, as he calls it) "probably lie in the interplay between fundamental frequency, intensity dynamics, spectral characteristics and voice quality."

The individuality carried by a voice is a socially and contextually constituted individuality.

### 3.

The range of meanings which an individual voice makes hearable, qua voice, all belong to the realm of social or indexical meaning (Dahlstedt 1987, Scherer & Giles 1979, Ochs 1992)<sup>3</sup>. Meanings in that realm are directly geared to the dynamics of the speech event, indexing continuity and change in the identity of its components, their properties and their relations to each other. Identity of speaker, addressees, and embedding activity, properties of speaker, addressees, and activity, and relations of various kinds between speaker, addressees, activity and current topic are straightforward examples of indexical information signalled by talk (a comprehensive overview can be found in Saville-Troike 1989).

Social meanings typically belong to what is shown by talk, not to what is said by it, to use Wittgenstein's distinction (Wittgenstein

---

3 As Levinson (1979) points out, Grice's (1975) notion of conventional implicature, inferences which are non-cancellable (i.e. can not be cancelled by a following meaning of the same kind), irrelevant to truth-conditions, and detachable (i.e. conventionally tied to particular expressions) also covers precisely this realm of meaning.

1921). They are meanings outside of topic, which need a shift in topic and footing to become the focus of explicit agreement or disagreement. In (1), for example,

- (1) Adam: snygg barre du har lilla gumman  
Beata: nää  
(Adam: nice shack thou've got baby  
Beata: oh no)

the *nää* can not be used by Beata to protest the *lilla gumman* relation that Adam wants to involve her in. In order to do that, Beata would need to change topic and footing and say something like (2).

- (2) Beata: ja e inte din lilla gumma, gå å tvätta dej i munnen  
(Beata: I'm not your baby, go wash your mouth)

In Anward (1997), I claimed that social meanings are always shown, never said, but I am now convinced that that statement is too strong. Turns made up simply by interjections, such as (3) and (4), for example, say rather than just show social meanings. They may be responded to by a simple agreement or disagreement, which concerns precisely the social meanings involved.

- (3) Beata: usch vale  
(Beata: ugh)
- (4) Beata: mygghjärna  
(Beata: mosquito brain)

What led me astray was the fact certain expressive interjections, such as *ouch*, are awkward to disagree with. However, this has most likely nothing to do with the distinction between saying and showing. Expressions such as *I have a cold* are equally awkward to disagree with (unless, of course, you are a physician who has just examined the speaker). What we observe in such examples as *ouch* and *I have a cold*

is rather a difficulty or unwillingness to challenge assertions which are crucially grounded in purely private experiences.

Meanings produced on the basis of voices, though, are robustly on the side of that which is never said, but only shown.

#### 4.

Social meanings are not easy to handle. Suppose I say

- (5)    hej de e ja  
       (hi it's me)

in a voice which an outside observer, with access to a reasonable sample of Swedish speakers, would interpret as carrying the meanings in (6).

- (6)    I am a man  
       I am middle-aged  
       I am upper middle class

In any claim that by saying (5) in my voice I meant, among other things, (6), there are at least three problems buried. First of all, meaning presupposes a choice (Lyons 1968).<sup>4</sup> If the basic social meaning which I intended to convey by (5) was (7),

- (7)    I am Jan Anward

and I succeeded in conveying that meaning, then the items of (6) are simply fully predictable inferences from (7) and not independently selectable meanings. In that case, these items are not different in kind from the items of (8).

---

<sup>4</sup> This is not to say that more choices generate more meaning. See Nørretranders 1993, ch. 2.

- (8) I have lungs  
I have a larynx  
I have a mouth

Secondly, full-blown meaning requires awareness, intention and secondary intentions (Allwood 1976). Suppose Curt yawns in the presence of Dora. The yawn may be intentional or not, and Curt may be aware of it or not.<sup>5</sup> If the yawn is unintentional and Curt is unaware of what he is doing, then his yawn is just a natural sign (Grice 1957), interpretable by Dora as a sign of tiredness or boredom, but being communicated only in the way dark clouds communicate a coming rain. That Curt yawned meant something to Dora, but Curt didn't mean anything by yawning. If the yawn is unintentional but Curt is aware of it, then it is still a natural sign, but this time interpretable to Curt himself as well. It is only if the yawn is deliberate that we may safely claim that Curt meant something by yawning, for example that he wanted to go home. That meaning, however, is not properly communicated unless Curt intends his yawn to be apprehended by Dora in a particular way, and also intends that she apprehend his intention that she apprehend the yawn. Only then is Curt and Dora involved in act of full-blown communication.

Obviously, the meanings of (6), (7), and (8) are open to the same range of interpretations as is the meaning 'I am tired' conveyed by Curt's yawn. These meanings may be located anywhere on a scale between natural meaning and meaning communicated in a full-blown manner. Because of that, another dimension of impreciseness is added to the notion of social meaning.

Thirdly, finally, labels. All the labels used in (6) are disputed, both in social theory and in social practice. While they may seem the correct labels to use for an outside observer trained in a particular framework of sociological theory, their relevance to a flow of

---

5 The existence of intentional behavior of which one is unaware is of course controversial. See further discussion below.

interaction may easily be denied by other observers, not to speak of the parties to that interaction themselves (for pertinent discussion, see Antaki & Widdicombe 1998).

In summary, social meanings, such as those in (6), are deeply problematic, both with respect to what is communicated and with respect to how it is communicated.

## 5.

Voices, their sources, and the speaker identities they index are however inescapable aspects of spoken interaction.<sup>6</sup> In any conversation, I must be aware of my own voice, the voices of others and their sources, and the way these voices and sources 'populate' the conversational room. Otherwise, I can't log into the turn-taking system under way, and hence will be excluded from the conversation.

While other social meanings are context-dependent and even in a relevant context elusive and negotiable, the basic meaning of individuality indexed by voices and sources appears to be secured by the very nature of conversation.

Rolf Lindgren kindly drew my attention to the following pertinent example and described his and other peoples' experiences to me. Divers inhale heliox, a mixture of helium and oxygen, to make diving easier. If you inhale some heliox and then start talking, a curious effect appears. Since sound travels faster in helium than in air, what you hear is a speeded up version of your own voice. You sound somewhat like Donald Duck<sup>7</sup>. This is of course startling, but what is really unpleasant about the experience is that you can't switch off that strange voice coming from you. The gas must leave your respiratory system before your voice can revert to normal again.

For a fleeting moment you sense that someone else is speaking from within you.

---

6 Extending the notion of voice to signed interaction and written interaction (cf. Engdahl 1994) is not an easy matter and must be dealt with separately.

7 A concrete example can be found in the movie 'Sphere'.

This element of horror, of the Freudian *Unheimliche*, in the experience stems from the fundamental role of the voice for our self-images. In any conversation, I can't help hearing at least my own voice and being deeply affected by what I hear. My voice is both one of my most intimate parts and a crucial ingredient of one of the protecting environments of the self described in detail by Anzieu (1995, 2000): the 'sound envelope' of the self (*enveloppe sonore de soi*; see also Anzieu 1979). This is the sound environment in which a person feels relaxed and at home. It is composed of that person's own voice, the voices of significant others, and other familiar and reassuring sounds. Anzieu notes that this might be the earliest protecting environment of the self that there is, constituted already in utero. This notion is strongly corroborated by research on early speech perception (Juszyk 1997, especially p. 208 - 211), which shows that from the very start, our remembered experience of language is structured in terms of voices and occasions.

It is not hard to imagine what a disruption of your sound envelope would mean. What if your husband/wife/partner woke up one morning with a new voice, or with your voice; what if you found yourself late at night with a new voice, or somebody else's voice<sup>8</sup>. How would this affect you? What would be worst, the discovery of a new nose on you or the discovery of a new voice from within you?

## 6.

What happens in the heliox experience is of course that voice, source and individuality come apart. Someone else's voice is suddenly coming from the place where a few moments ago only you were living. A change in voice thus introduces another party to the ongoing conversation, another individual, or, for short, another I. However, in the heliox case, this change of I is not linked to a change of place.

---

8 Cf. the following haiku by Otsuji (1881 – 1919): I do not recognize / my own voice / this winter night.



The voice of the new I is coming from the place of the old I, thus creating a problem.

We can make sense of this in terms of the following principles, each of which has two parts, one dealing with change (a), and the other with continuity (b).

- AV. a. Another voice introduces another I
- b. Same voice maintains same I

- AS. a. Another source introduces another I
- b. Same source maintains same I

By AVa, a new I is introduced by the heliox voice. By ASb, however, this I is the same as the old I who was speaking before the heliox voice came on, since the source has not changed.

There are several ways in which this paradox can be resolved. But first some further conceptual labor.

## 7.

For Bakhtin, individuality is grounded precisely in the unique place that each person occupies (two persons can of course not occupy the same place at the same time) and the unique perspective this place affords (Bakhtin 1986, Clark & Holquist 1984, ch. 3).

However, a crucial insight by Bakhtin is that individuality must be recognized to come into existence. Only when the uniqueness of a person's place and perspective is confirmed by another person, is individuality realized. That is why dialogue and answer are so essential to Bakhtin. It is only through an action that is uniquely issued from a place and perspective and a proper answer to that action that a person in that place and with that perspective attains social reality as an individual. A proper answer to an action guarantees the uniqueness and indispensability of that action, because the answer could not have occurred, if the action had not occurred. If the

action furthermore is uniquely issued from the place and perspective of a certain person, then individuality of that person is secured.

The place which confers individuality on a person is thus not an absolute place, but a relative place, in our terms, a source from which a flowing of alterity is experienced by another individual.

Likewise, as Miller (1989) argues, there is no such thing as an absolute voice. A voice only exists as heard by someone else and it carries individuality only by being recognized by someone else as the voice of a separate individual.

This relational nature of voices and places fits into a learning scenario where a learner models her/his speech on that of a significant (m)other, but with a difference, which, once it is accepted, secures both the individuality of the learner and the individuality of the model (Anward 1990). In that kind of scenario, the I:s indexed by a voice or a source, in accordance with AV and AS, are from the very beginning positions in a field of relationships.

This learning scenario also combines well with the exemplar theory of Jusczyk (1997) and Anzieu (1995, 2000), and its natural consequence, a conception of linguistic experience as a network of token utterances (Anward & Lindblom 2000). In such a conception of linguistic experience, utterances do not have autonomous properties. Rather, each member of the network is structured by the whole network. Similarities with other members of the network, as well as crucial points of contrast, affords segmentation of utterances into units with syntagmatic and paradigmatic relations to other units licensed by the network.

What emerges then is a conception of linguistic competence as access to a network of sequentially organized, authored, and voiced utterances, which may be reused and recombined at new occasions for talk.

## 8.

The nature of this linguistic competence is such that a speaker can produce a new turn at talk in a voice which is heard as another voice

than her/his own. In this, speakers have a choice. A speaker can choose not to retain her/his normal voice and instead tune in another voice, as in the following example, from Nordberg (1984: 16), where a boy – call him Frans – shifts from his normal voice to a girlish voice (in *italics*), when describing the attitude of the girls he is talking about:

- (9) Frans: ... ta tag i killarna så här bara å rycka opp dom  
*kom här å dansa annars får du stryk*  
vet du dom va helt vilda på rom  
(Frans: ... grab the guys like this and just pull them to their  
feet  
*come here and dance or you'll get hit*  
you know they were completely wild at them)

Moreover, this shift of voice and the change of speaker identity it indexes are clearly intended and these intentions are also intended to be apprehended, otherwise the contrast in reference between the first and the second token of *du* (you) would not be recognizable. Thus, a new I can be introduced into an interaction by means a voice shift, in accordance with AV, in a manner of full-blown communication.

But from where does this new I speak? Again, we seem to be facing the problem that a change in voice is not accompanied by a change of source. However, in this case, unlike in the heliox case, there is a straightforward way out. As Adelswärd (2000) points out, neither addressees nor speakers need actually be present in an ongoing speech event. A stretch of speech populates a conversational space not only with participants in the speech event, as we have seen, but also with topical events and their participants, and such topical participants can be cast in the roles of speaker and addressee, as demonstrated in detail by, among others, Tannen (1989) and Günthner (1998). In (9), Frans sets into motion not only himself and his addressees but also the girls and the boys he is talking about, and in the stretch of speech in *italics* these topical participants are cast in the roles of speaker and addressee. When that stretch of speech is

spoken in the voice of one (any) of the girls, a new I is introduced, who is anchored in a topical participant, and who is speaking from the topical place of that participant, addressing another topical participant, one (any) of the boys.

## 9.

What happens to 'real' participants in the speech event, when the roles of speaker and addressee are attributed to topical participants, when, in Adelswärd's terms, topical participants are made into virtual participants in the speech event?

There is a long and multifaceted tradition which appears to claim that in a situation where a real speaker shares a speech event with a virtual speaker, the virtual speaker deprives the real speaker of parts of her/his speakerhood (Banfield 1982, Goffman 1981, Sells 1987, Levinson 1989). A monolithic real speaker of an utterance is typically attributed the roles of (i) source (Sells 1987), the one who actually speaks, a.k.a. speaker (Banfield 1982), animator (Goffman 1981), and transmitter (Levinson 1989); (ii) author (Goffman 1981), the one who chooses the words of the utterance, a.k.a. formulator (Levinson 1989); (iii) pivot (Sells 1987), the one in whose perspective the utterance is formulated; (iv) self (Banfield 1982), the one whose thoughts and feelings are expressed by the utterance, a.k.a. motivator (Levinson 1989); and (v) principal (Goffman 1981), the one who is responsible for the contribution the utterance makes to the ongoing conversation. The tradition I am referring to now seems to hold that the relationship between real speaker and virtual speaker with respect to these roles is a zero-sum game, that whatever roles the real speaker loses are gained by the virtual speaker. In (9), for example, this would mean that Frans, in the *italicized* stretch, would be totally eclipsed by the girl he lets appear as speaker in that stretch.

I believe that this view of the relationship between real and virtual speakers is too simple. Consider, for example, a problem which appears in acting and which can't seem to even receive a proper

formulation unless we allow for a more complex relationship between real and virtual speech event participants.

Suppose you play, say, Hamlet and adopt a voice that fits the role. Then, by AV, it is Hamlet, not you, who says "to be or not to be", and all the rest. And by the method of reconciling AV and AS which I have already described, it is Hamlet, not you, who is the source of these utterances. In the theatrical context, you ARE Hamlet. But are you also still yourself in that context?

There seem to be several answers to this question. One way is chosen by actors such as the late Peter Sellers. Such actors indeed tend to eclipse their out-of-role individualities and equip all their role characters with distinct individualities, complete with distinctive voices and accents. In the case of Peter Sellers, just think of inspector Closeau or dr Strangelove! This is a risky way, though, precisely because the actor's own individuality is eclipsed in all fictional contexts, which turns his out-of-role individuality into just another individuality in a huge repertoire of individualities. Every display of individuality will then involve a choice among a number of available individualities, each one of which will have to be strongly framed in order to be correctly inhibited or selected in a particular context. If framing for some reason is weakened beyond a certain point, there will be contexts where several individualities are available, and it will be impossible to tell which one is the appropriate one in that context. Indeed, this appears to be what actually happened to Peter Sellers. During the last years of his life, he complained that he no longer knew which voice was his.<sup>9</sup>

Bengt Nordberg has pointed out (p.c.) that another way is chosen by actors with highly distinctive voices, voices which are heard through every role they play. The late Swedish actors Anders Ek and Ernst-Hugo Järegård are cases in point. Nordberg proposes that this could be interpreted as a defense against fictional voices. By making your out-of-role voice heard through the voice of your roles, you retain your presence in all contexts, reducing effectively the salience

---

9 According to a Swedish newspaper article (which I forgot to save).

of all your fictional voices. By speaking in two voices simultaneously, you are neither Hamlet nor Anders Ek, but Anders Ek playing Hamlet.

A third way may be to withdraw individuality from your role characters, by adopting an impersonal voice that can not be mistaken for anyone's individual voice, for example the Swedish "routine theatre voice", described so vividly by Katarina Frostenson as "a not infrequently slightly chuckling tone of voice with many, too many, emphases and peaks". No actor who adopts this kind of voice will have any problems of keeping it apart from her/his own voice.

The problem of enacting someone else and retaining (or not retaining) yourself is not limited to professional acting, but arises in all cases where a real participant lets a virtual speaker appear, cases such as (9). A conjecture I would like to make is that people normally do not in such cases go for a Peter Sellers strategy but are content with an Anders Ek strategy or even a 'bad acting' strategy.<sup>10</sup> Another voice would, in other words, be a more optimal strategy than another's voice.

However, there could be no choice among these strategies unless real participants retain their presence in the speech event, even when the speaker role is taken over by a virtual participant. In the zero-sum theory, which only allows for one distributed I, it is not even clear where the dilemma would arise.

## 10.

Once we have allowed for real participants to co-exist with virtual participants in a speech event, it becomes possible to ask which social relations among them are introduced and maintained by speech. A further drawback with the zero-sum theory of speaker roles is that the only social relation imaginable in that theory between a virtual speaker and a real participant is that of (partial or complete) posses-

---

<sup>10</sup> For some suggestions what would constitute a non-theatrical routine acting voice, see Klewitz & Couper-Kuhlen 1999.

sion. This makes the theory peculiarly asocial. Moreover, the relational nature of voices and sources which I have insisted on throughout this article is completely lost.

A simple alternative theory goes like this. Whenever a new I is introduced, virtual or not, by a new voice or a new source, old participants do not melt away, partially or completely, but just stay on. What typically changes is instead participant roles. A new speaker will often turn a previous speaker into an addressee or an overhearer. Thus, in my analysis, the girlish voice in (9), by introducing both a virtual participant and a new speaker, also introduces a new addressee, one of the boys, and turns the previous speaker, Frans, and his addressees into overhearers. Thus, in keeping with the relational nature of voices and sources, I take AV and AS to be principles which modify entire participation frameworks.

Another example making the same point is (10).

(10) *Collum*

- Eva:       å den vårdarn han sa att de e ju för jäv:lit.  
              hon skulle ha dött för,  
Ivar:       ja just  
Eva       ett år sen.  
(Eva:       and that intern he said it is just too bad.  
              she should have died,  
Ivar:       absolutely  
Eva       a year ago.)

Example (10) is a stock example of reported speech. We have only one voice involved here, that of Eva, but there is a reporting stretch at the beginning of Eva's turn (*den vårdarn han sa att*) which sets up another source, a topical source, for the stretch of speech following *att*.

However, note that Ivar, who is the addressee of Eva, not of the intern (the intern is addressing Eva), responds positively to what the intern says. How can we understand this? Again, we need the

assumption that what changes with the introduction of a virtual participant is the entire participation framework. What Eva in fact does in (10) is deliver the moral of a story she has been telling. But instead of formulating it herself she puts it in the intern's mouth, addressing her topical counterpart, i.e. herself in the topical world in which she shares with the intern. And since the topical Eva aligns with the intern and the speech event Eva aligns with the topical Eva, it is entirely relevant for her speech event addressee, Ivar, to hear himself being addressed by the intern.

In this case, then, the speaker roles *i* – *v* are not split up into two sets, where one set is attributed to Eva and one set is attributed to the intern. Instead, each of these roles is distributed over a number of participants, both real and virtual. Both the intern and the speech event Eva are source, author and pivot of the stretch of reported speech, and all four participants are self and principal of that stretch of speech.

In (9), in contrast, Frans and his addressees are turned into overhearers by the appearance of the virtual speaker and have thus no responsibility for what is going between the girl and the boy. The collusion established between the girls and the boys (the boys obviously have some responsibility for the way the situation is finally defined) is not shared by Frans and his addressees, who rather can form a team of 'shocked' bystanders.

In both cases, there is no need for split speakers. A fully worked out dialogical analysis, across turns and within turns (Linell 1998), reveals both the team formations characteristic of egalitarian divisions of saying (Anward 1997) and the more monolithic speaker positions of monological discourse, without having to introduce participant roles beyond the ordinary ones of speaker, addressee, and overhearer, and without having to reduce the rather complex social relations established and maintained by speech to some kind of spiritual possession.



## 11.

We are now ready to get back to the horror in the heliox experience. There are two moments to this aspect of the experience. First, there is the unexpected change of voice, creating a moment where the I does not recognize itself any longer. This is not an uncommon experience. Using Mead's (Mead 1934, ch. 22) distinction between the 'I', the spontaneous agentive self, and the 'me', the 'I' of a moment ago, the object of the I's self-reflection, we can say that in such an experience, an aspect of the I is revealed in the me which the current I had not expected. I am surprised by my me.<sup>11</sup>

There are several ways of handling such an experience. One way is simply to accept whatever comes from the I and gets manifest in the me, without bothering too much about how it got there. I think this is the attitude taken by Rimbaud in his famous seer letter to Paul Demeny (Rimbaud 1871: 219): "Car JE est un autre. Si le cuivre s'éveille clairon, il n'y a rien de sa faute. Cela m'est évident: j'assiste à l'éclosion de ma pensée: je la regarde, je l'écoute: je lance un coup d'archet: la symphonie fait son remuement dans les profondeurs, ou vient d'un bond sur la scène." If we want to go dull again and relate this way to the notions of linguistic experience and linguistic competence I have just elucidated, we may say that the possibilities implicated in a network of stored utterances are so vast that linguistic surprises are bound to come to everyone.<sup>12</sup>

Another way is to attribute unexpected aspects of the I, typically slips of the tongue and other 'mistakes', to momentary mechanical malfunctions, temporary lapses, which can be framed and repaired in such a way that the competence and normality of the self are not in danger (Goffman 1981). Such a mechanization of the body seems to be a basic defensive strategy when individual integrity is threatened.<sup>13</sup>

---

11 For a modern variation on Mead's theme, see Nørretranders 1993.

12 For a famous idealization of such surprises and a corresponding devaluation of ordinary self-maintaining talk, see Novalis 1798.

13 For an extreme case, Joey the mechanical boy, who believed himself to be run by a machine, see Bettelheim 1976, and also Strang 1995.

A third way is to alienate the unexpected from the I, by attributing it to another I which somehow is speaking 'in' or 'through' the old I, from a separate place 'inside' the place of the old I or on a channel being somehow relayed through the place of the old I. The Freudian Unconscious, spiritual possession (our old acquaintance), and divine inspiration belong in this category.

In the case of a foreign, strange voice, as in the heliox experience, all of these options are available. It may be your 'real' voice finally rising to the surface (Odrischinsky 1989), it may be the result of a mechanical malfunction (in this case transparently related to the inhalation of heliox), or it may be someone else or something else speaking in or through you. Of these options, the first and the second fail to introduce a new I, either because the new voice 'really' belongs to the old I or because the new voice is just the result of a temporal disturbance of the old I.

Alienation, the third option, however, opens the doors for a new I and the next moment of horror in the experience, which is the question we have already posed: from where, more precisely, is the new I speaking?

We have seen how trouble arising because a new voice is apparently coming from an old source can be resolved by anchoring the new voice in a topical participant and a topical place. This holds not only for overt voices but also for voices you hear in your mind. As long as they can be anchored in remembered people and places, voices and sources still line up as they should.

Omnipresent aspects of the I, such as the Unconscious<sup>14</sup> or one's conscience, and omnipresent participants in the daily life, such as God and other spiritual powers, provide other possible anchors for new voices, both overt and in your mind. Provided an appropriate world-view, your integrity will not be threatened by spirits speaking

---

14 Miller 1989 argues though – on dubious grounds, I believe – that the Unconscious can have neither gaze nor voice, since, for him and for Lacan, the Unconscious is essentially non-relational.

through you or God revealing his plans to you in the privacy of your soul.

It is only when voices become free-floating, devoid of possible anchors, as in psychosis (Sechehaye 1951) and in dementia, that the horror becomes unbearable, for in that kind of experience, the only rational explanation is of course that your self has been taken over by someone unknown or by something unknown, speaking from a place you really can't locate.

## 12.

The principles AV and AS, in conjunction with a richly articulated dialogical approach to linguistic experience and practice, allow us to map in rather great detail those fascinating domains of human linguistic experience which are connected with the many ways of the human voice – common objects of wonder for Viveka and me.

## References

- Adelswärd, Viveka (2000). "'Who is talking?' Some thoughts on speakers, voices and virtual participants'. In *Creative Crossroads -Electronic Honorary Publication Dedicated to Yvonne Wærn on Her Retirement*. <http://www.tema.liu.se/people/chrga/yw65/festschrift/contributions.html>
- Allwood, Jens (1976). *Linguistic communication as action and cooperation*. Gothenburg Monographs in Theoretical Linguistics 1. Gothenburg: Gothenburg University, Department of Linguistics.
- Antaki, Charles & Sue Widdicombe, eds. (1998). *Identities in talk*. London, etc.: Sage.
- Anward, Jan (1990). 'Språkspel i språkcirklar'. In U. Nettelbladt och G. Håkansson, (eds): *Samtal och språkundervisning*. Linköping: Linköping Studies in Art and Science 60. 147-168.
- Anward, Jan (1997). 'Parameters of Institutional Discourse'. In Britt-Louise Gunnarsson, Per Linell & Bengt Nordberg (eds): *The Construction of Professional Discourse*, London: Longmans. 127-150.

- Anward, Jan & Björn Lindblom (2000). 'On the rapid perceptual processing of speech: From signal information to phonetic and grammatical knowledge'. <http://www.isp.su.se> via Forskning, Projekt IIS, Artiklar.
- Anzieu, Didier (1979). 'The Sound Image of the Self'. *International Review of Psycho-Analysis* 6. 23-36.
- Anzieu, Didier (1995). *Le Moi-peau*. 2<sup>nd</sup> ed. Paris: Dunod.
- Anzieu, Didier (2000). *Hudjaget*. Translation of Anzieu 1995 by John Swedenmark. Stockholm: Natur och Kultur.
- Bakhtin, Mikhail (1986). *Speech Genres and Other Late Essays*, Austin: University of Texas Press
- Banfield, Ann (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction*, London: Routledge & Kegan Paul.
- Benveniste, Émile (1966). *Problèmes de linguistique générale, I*, Paris: Gallimard.
- Benveniste, Émile (1974). *Problèmes de linguistique générale, II*, Paris: Gallimard.
- Bettelheim, Bruno (1976). *The Empty Fortress: Infantile Autism and the Birth of the Self*. The Free Press.
- Bruce, Gösta (1982). 'Developing the Swedish Intonation Model'. *Working papers* 22. Lund University, Department of Linguistics. 51-116.
- Clark, Katherine & Michael Holquist (1984). *Mikhail Bakhtin*, Cambridge, Mass: Belknap Press
- Dahlstedt, Karl-Hampus (1987). *Från Quartier Latin till Grisbacka*. Umeå: Två förläggare.
- Elert, Claes-Christian & Britta Hammarberg (1990). 'Regional voice quality variation in Swedish'. *Phonum* 1. 117-119.
- Engdahl, Horace (1994). *Beröringens ABC: en essä om rösten i litteraturen*. Stockholm: Bonnier.
- Gazdar, Gerald (1980). Pragmatic Constraints on Linguistic Production. In B. Butterworth (ed): *Language Production 1: Speech and Talk*. New York etc.: Academic Press. 49-68.
- Giles, Howard & Peter F. Powesland (1975). *Speech style and social evaluation*. London: Academic Press.

- Goffman, Erving (1981). *Forms of Talk*. London: Blackwell.
- Grice, H. P. (1957). 'Meaning'. *Philosophical Review* lxvi
- Grice, H. P. (1975). 'Logic and Conversation'. In P. Cole & J. L. Morgan (eds): *Syntax and Semantics 3: Speech Acts*. New York etc: Academic Press. 41-58.
- Günthner, Susanne (1998). 'Polyphony and the "layering of voices" in reported dialogues'. InLiSt No. 3. <http://inlist.uni-konstanz.de/index.htm>
- House, David (1990). 'On the perception of mood in speech'. *Phonum* 1. 113-116.
- Juszyk, Peter W. (1997). *The Discovery of Spoken Language*. Cambridge, Mass.: MIT Press.
- Karlsson, Inger (1990). 'Synthesis of normal female voices'. *Phonum* 1, 54-57.
- Klewitz, Gabriele & Elisabeth Couper-Kuhlen (1999). 'Quote – unquote? The role of prosody in the contextualization of reported speech sequences'. InLiSt No. 12. <http://inlist.uni-konstanz.de/index.htm>
- Kuno, Susumu & Kaburaki, E. (1977). 'Empathy and Syntax'. *Linguistic Inquiry* 8. 627-672.
- Lacan, Jacques (1977). *Écrits. A selection*, London: Tavistock Publications.
- Laver, John (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- Levinson, Stephen C. (1979). 'Pragmatics and social deixis: reclaiming the notion of conventional implicature'. *BLS* 5. 206-223.
- Levinson, Stephen C. (1988). 'Putting Linguistics on a Proper Footing: Explorations in Goffman's Concepts of Participation'. In Drew, P. & A. Wootton (eds): *Erving Goffman: Exploring the Interaction Order*. Oxford: Polity Press. 161-227.
- Linell, Per (1998). *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical perspectives*. Amsterdam: John Benjamins.
- Lyons, John (1968). *An introduction to theoretical linguistics*. Cambridge: Cambridge University Press.

- Mead, George Herbert (1934). *Mind, Self, and Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.
- Miller, Jacques-Alain (1989). 'Jacques Lacan et la voix'. In R. Lew & F. Sauvagnat, eds: *La Voix*. Paris. 184-193.
- Nolan, Francis (1991). 'Forensic phonetics'. *Journal of Linguistics* 27. 483-493.
- Nordberg, Bengt (1984). 'Om ungdomars samtalsstil. Några preliminära iakttagelser'. *Nysvenska studier* 64. 5-27.
- Novalis (1798). 'Monolog'. In *Dichtungen*. Rowohlt 1963. 5.
- Nørretranders, Tor (1993). *Märk världen*. Stockholm: Bonnier.
- Ochs, Elinor (1992). 'Indexing gender'. In A. Duranti & Ch. Goodwin, eds: *Rethinking context*. Cambridge: Cambridge University Press.
- Odrischinsky, Eva (1989). 'Rösten är själens muskel'. Radio P1. September 9, 1989.
- Rimbaud, Arthur (1871). 'Lettre à Paul Demeny, 15 mai 1871'. In *Poésies complètes*. Paris: Éditions Gallimard. 219-223.
- Saville-Troike, Muriel (1989). *The ethnography of communication: an introduction*. 2<sup>nd</sup> ed. Oxford: Basil Blackwell.
- Scherer, Klaus R. & Howard Giles, eds. (1979). *Social Markers in Speech*. Cambridge: Cambridge University Press.
- Sechehayé, Marguerite (1951). *Autobiography of a Schizophrenic Girl*. New York: Grune and Stratton, Inc.
- Sells, Peter (1987). 'Aspects of Logophoricity'. *Linguistic Inquiry* 18. 445-479
- Strang, Hilary (1995). 'Delusional circuitry'. <http://eserver.org/cultronix/hilary/>
- Tannen, Deborah (1989). *Talking Voices. Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Trautmüller, Hartmut, Peter Branderud & Aina Bigestans (1989). 'Paralinguistic speech signal transformations'. *Perilus* X. 47-64.
- Williams, C. E. & Stevens, K. L. (1972). 'Emotions and speech: Some acoustical correlates'. *JASA* 52. 1238-1250.
- Wittgenstein, Ludwig (1921). *Tractatus logico-philosophicus*. London: Routledge.