Abstract

*In direct face-to-face communication, we use a variety of different means of expression with different degrees of consciousness and intentionality (Allwood 1976). Communication is multimodal and embodied. In parallel to verbally formulating our conscious message, we also convey information with our voice, mimics, gaze, gestures and body posture. The aim of this paper is to analyse the use of spontaneous gestures accompanying spoken discourse on ‘others’. The first example shows that the ethnic references can be made in the gestural channel, keeping the spoken channel free from explicit ethnic reference. In the second example, a variety of gestures illustrate and emphasise the events referred to in the spoken discourse.*

**Keywords:** communication, embodiment, gesture, ethnic contrast
Multimodal embodied communication

In direct face-to-face communication, we use a variety of different means of expression with different degrees of consciousness and intentionality (Allwood 1976). Communication is multimodal. In parallel to verbally formulating our conscious message, we also convey information with our voice, mimics, gaze, gestures and body posture. Our communication is also embodied: we understand the meanings of words on the basis of our bodily everyday perceptual experience of what we see, hear and feel and we use our ‘expressive bodies’ to communicate. In most cases, the non-verbal information is expressed with a much lower degree of conscious control but sometimes this is not the case. In general, it is very difficult to say how conscious and intentional these various features of our communicative behaviour are (Allwood 1986, 2002).

Speech, gesture and action co-evolve in parallel, in a certain tempo and rhythm (Kita 1990) and the partners have to co-ordinate their utterances and non-verbal actions. The utterances and non-verbal actions (such as drawing and gesturing) can be conceived of as instructions for the listeners for how to change the meaning, how something is perceived, how one thinks or feels or what one wants to do with something that is currently in the conscious focus (Linell 2005). By expressing ideas, the speaker evokes images in the minds of the listeners. It seems that speakers even use internal images in spoken discourse. Our recent eye tracking study showed that verbal description of a scene triggers the eye movements of the listener/observer and engages mental imagery (Johansson et al. 2006, Holsanova forthcoming).
The consciousness of the speaker and of the listeners are synchronised. In this way, they are able to create joint attention and co-construct meaning (Linell 1998, 2005).

**Drawing and gestures in conversation**

Apart from spontaneously using gestures, mimics and body postures along with utterances, speakers also draw pictures, maps and sketches when they communicate. Pictures, maps and sketches show how the described reality is conceptualised. The use of iconic and pictorial representations is useful in communication, since it helps the speaker and the listener to interactively adjust their visualisations and achieve understanding (Johansson et al. 2005). In conversational data with spontaneous drawing, we investigated the visual focus movements across the drawn picture and their relations to the verbal focus movements in spoken discourse (Holmqvist & Holsanova 1997). In short, drawing allows cognitive processes to be captured, shared and elaborated (Holsanova forthcoming). By pointing at the drawing, the interlocutors can locate objects in a shared visual world and they can both follow and influence the gaze direction and gaze shifts of the other interlocutors (Goodwin 2003). Pointing and gaze is thus contributing to create mutual spatial orientation and a joint visual and attentional focus.
Gestural deixis for ethnic contrast

In this paper, we narrow down the interplay of the different communication modalities and concentrate on the use of gestures accompanying discourse on ‘others’ (Holsanova & Nekvapiil, 1995, Holsanova 1998a,b, Wästerfors & Holsanova 2005). We present two examples of how speaker gestures express contrast between ethnic groups. The first example shows that when a symmetric bipolar contrast has been gesturally established, there is no longer a need to refer to its poles in the verbal channel. It will then be sufficient that the contrast is referred to by gestural means only.

The second example shows the use of gestural information in an asymmetric center-periphery contrast. The speaker accidentally reverses the contrast, by gestural means mainly. She notices the oddity of the reverse and backtracks to reset the symmetry.

Examples are drawn from a conversation on ‘others’ videorecorded with three Swedish partipants, all women in their 20s. The examples have been transcribed and translated into English. Gestures have been marked in the text by parenthesis. In addition to text, transcriptions also show still images of the gestures. Arrows have been included in the still images in order somewhat to compensate the lack of moving images in this paper (cf Example 1).

1 This conversation was recorded within the project Det egna och det främmande: Hur vi samtalar om etniska och sociala grupper (Us and Them, Talking about Social and Ethnic Groups), financed by the Swedish Council for Research in the Humanities and Social Sciences. Results from the project have been presented in Jörgensen & Holsanova (1995, 1996), Holsanova (1998) and Holsanova, Jörgensen & Holmqvist (1999).
Example 1

The procedure - a wheel rolling forward (four full rotations)

men dom skulle nog röra
va man [skulle säge proceduren]
te å fä stanna
skulle gå mycket snabbare

Ethnic categorisation is *contrastive* by its nature (Jørgensen & Holsanova 1995, Holmqvist & Holsanova 1996, Holsanova et al. 1999). There is one more positive pole (us) and one more negative (them). When one pole has been given a positive characterisation, the other pole is almost automatically attributed the corresponding negative property.

Even when the contrast is presented verbally, it is usually given a *spatial* mapping. The different aspects of the contrast are then laid out as bipolar dimensions in space. The positive pole (us) is given positive values for each dimension, and the negative pole (them) negative values. In fig 1, we see the dimensions of a spoken Czech characterisation of the Czech-Polish contrast laid out across the border.

![Fig 1: Five contrast dimensions in a Czech-Polish contrast (from Holmqvist & Holsanova 1996).](image)

The spatial layout of ethnic contrast becomes particularly evident when the speaker actually draws a polary contrast that corresponds to and illustrates the contrast presented in the discourse. Figure 2 is a spontaneous drawing,
illustrating the simultaneous discourse contrast depicted in figure 3.

Fig 2: Spontaneous drawing by speaker of fig. 3. Negative contrast pole is to the left, positive is to the right.

Fig 3: Dimensions presented in the spoken discourse by the drawer of fig 2, simultaneous to the drawing.

The drawer of fig 2 and speaker of fig 3, a Swedish-Canadian man, contrasts the cost and quality of Canadian construction companies with respect to the ethnic status of their staff. In the spoken discourse, the speaker presents the three contrast dimensions by progressively moving his focus from one contrast pole to the next. The three contrast dimensions of the discourse were thereby mentally and perceptually superimposed onto the single drawn contrast bar in fig 2 (for details, see Holmqvist & Holsanova 1996, 1997).

**Bipolar contrasting by means of gesture**

The gestures used by the speakers in this article are so-called spontaneous, co-verbal gestures. There were hardly any other types of gestures in the
material, and none of the others had a contrastive function. For an authoritative overview of gestures and sign, see Kendon (1983).

Furthermore, we have only studied gestures made by hands and arms, not the speaker’s use of head and face. My main reason for this omission is that hand gestures provide a sufficient material in themselves. Also, it would be very difficult to present examples of face gestures without revealing the identity of the speakers.

Gestural presentation of contrasts shares the basic properties with the same use of drawings. Both modalities map the contrast onto space. Once the contrastive poles are localised in space, the speaker can efficiently refer to them by gestural deictics. The main difference is that while drawings leave permanent marks, it is up to the interlocutors to keep visualising that which has been presented gesturally.

In fig 4, the speaker places two opposite camps on each side of herself using her thumbs. The brackets in [two camps] indicate that these words coincide with the gesture in the figure. This so-called deictic gesture (McNeill 1992) locates discourse referents, the two camps, in the surrounding space.

When there is a table in front of the speaker, she may use the surface of the table to map out the contrastive space. In figures 5.1 - 5.2, the speaker, a young Swedish woman, superimposes the north-south dimension of the
European map onto the table.

they usually come from let’s say from Yugoslavia where there’s where there’s murder and rape an everything

dom kommer oftast ifrån vi säger ifrån Jugoslavien där det är där det är mord och våldtäkt å alltihopa

[they leave because it has been so horrible]

[dom åker därifrån för att det har varit så hemskt]

Fig 5.1: (in) Yugoslavia.

[they come here an do the same thing themselves]

[dom kommer hit å gör samma sak själv]

Fig 5.2 (in) Sweden.

In fig 5.1, the combination of a deictic gesture and coinciding localisations in the spoken channel effectively makes this places on the table Yugoslavia. In fig 5.2, the same combination of spoken and gestural means places Sweden on the table.

From now on, the southern Yugoslavia (them) is to her right and the northern Sweden (us) is to her left. This mapping is not only understood by the speaker, but as we will argue later, also by the listeners. Therefore, once these contrastive poles have been positioned on the table, the speaker can make systematic use of their positions.
In fig. 5.3, the speaker verbally refers to a causal relation. Simultaneously, she keeps her right hand on Sweden, thus referring back to the crimes of 5.2, while pointing at Yugoslavia. In effect she is thus saying that immigrants behave so badly because of something that has to do with the Yugoslavia of 5.1. This explanation is further developed during 5.4 - 5.7.
Observe that the speaker never mentions anything about Sweden or Yugoslavia in the spoken discourse during 5.3 - 5.7. The deictic placement of events in space is explicit in the gestural channel only. Gestural deixis provides a sufficient means of reference, once the two poles have been placed on the table, as in 5.1 - 5.2.

It should be clear that the speaker herself thinks of the events of 5.4 - 5.7 as taking place in Yugoslavia and Sweden, respectively. It is thinking so that makes her hands move deictically in congruence with the corresponding events in the verbal channel.

However, it is an open question whether listeners would be completely unable to perceive the speakers intended localisations were they to be without access to the gestural channel. The semantics of the spoken discourse is such that the localisations could perhaps be inferred. Having shown the text without the gestures to a few colleagues, we have nonetheless noticed that they have considerable problems localising the events anywhere. The non-specific causal schemas in the spoken discourse are taken to be just non-specific. People do not place the beating of the child in 5.6 in Yugoslavia and the beating of children in 5.7 in Sweden, until they are shown the accompanying deictic gestures in the figures.

In their series of so called mis-match experiments, Cassell, McNeill & MacCullough (1998) show that the information in the speaker’s gestural channel is indeed perceived and does play an important role for the
listener’s understanding. There is, therefore, a good reason to believe that the gestural deixis in 5.4 - 5.7 is important in the listener’s understanding of the speaker.

**Asymmetrical contrasting by means of gesture**

In this second example, we shall see the speaker gesturally illustrate the entire process of immigration into and segregation within the new country. She does this in a way that reverses the original asymmetry: The immigrants get the benefits of being inside, while all the rest are closed out. In the preceding examples, contrasts has been made along a symmetrical bipolar barlike dimension. In this example, the contrast makes use of the asymmetry in center-periphery and container schemas (Lakoff 1987, Holmqvist 1993).

- In a *center-periphery* schema, the positive center contrasts against the open-ended, often multiplex or mass-like periphery.
- In a *container* schema, the inside contrasts against the outside. The inside is protected from the outside by a border.

Container schemas differ from center-periphery schemas by the presence of a border which protects inside from outside. In this example, the speaker combines both these contrast schemas: The center identifies with the inside, and the periphery with the outside.

When fig 6.1 starts, the topic under discussion is the principles for immigration.
Gesturally, the person in need of help is placed centrally (fig 6.1). In the spoken discourse, the preposition 'in' tells us that the central position is also on the inside (i.e. inside the container Sweden).

In fig. 6.2, the speaker points at the periphery while referring verbally to the remaining family. i.e. placing these people in the periphery.

In the following discourse, the speaker will describe a process which in fact reverses the center-periphery contrast, putting the original 'them' in the center and 'us' in the periphery. The first step in this process is the arrival of the hords of relatives to the center, which is gesturally fully illustrated (figs 6.3 - 6.5).
[well if we look here how it is now] then it’s families not families it’s all the relatives coming in

Fig 6.3 Here, with this situation.

[ja om vi tittar här hur de e nu] så de alltså familjer inte familjer de e släkten som kommer in

because [one has got his foot\textsuperscript{2} in]

Fig 6.4 One is here.

then [all the rest arrive]

Fig 6.5 All the rest arrive.

After the sweeping motion in fig 6.5, the central position on the table is full of people, all relatives to the original immigrant. Now, the process continues with group formation (figs 6.6.1 - 6.6.3). Three groups are placed onto the table.

\textsuperscript{2} The Swedish equivalent used here literally translates to ’get ones finger in’.
an this I think is wrong totally wrong because then [they form their groups]

Fig 6.6.1 First group.

Fig 6.6.2 Second group.

Fig 6.6.3 Third group.

There are now several groups of immigrants on the table. The third and final step is an application of the container schema: Enclosing the groups into a central unity that separates them from the outside. Gesturally, the speaker does this by a circular motion around the three groups on the table. In the spoken channel she makes a vague reference, perhaps to culture (fig 6.7).
In fig 6.8, the speaker concludes that this conservative group formation on the part of the immigrants closes the rest of us out (the term ‘all the rest’ including ‘us’). The immigrants now take up the central position, leaving ‘all the rest’ in the periphery, as is clearly indicated by her gestural deixis.

Placing the immigrants on the inside and us on the outside is an asymmetry with peculiar consequences. It makes it appear as if we on the outside want in, which is normal, since the inside is protected and the outside not. However, in 6.7, the speaker made the inside be the immigrant culture. In this case, therefore, the speaker makes it appear as if we are standing outside their culture, wanting to come in.

Not surprisingly, the speaker retracts to a symmetrical contrast, by saying that ‘we close them out (too)’ immediately after 6.8. This way, the shutting out and wanting to come in is again symmetrical.
Discussion

We have presented two examples where spontaneous gestures are used as part of ethnic discourse. The first example shows that the ethnic references can be made in the gestural channel, keeping the spoken channel free from explicit ethnic reference. In the second example, a variety of gestures illustrate and emphasise the events referred to in the spoken discourse. Allwood (2002:16) compares the contribution given by the body movements to contributions given by prosody concerning emotions and attitudes. We can say that the demonstrative function of bodily mimesis (displayed by auto-fixation of gestures) can be compared to quotations as demonstrations (displayed by prosody) (Adelswärd et al. 2002, Clark & Gerrig 1990, Holsanova 1998a,b, 2006). Speakers often spontaneously quote people from the ethnic groups described. Holsanova (1998) argues these spontaneous quotes are not only a way of characterising the quoted person or group. It is also less threatening to use quotations. ”From the perspective of face work, it seems to be less threatening for the speakers to compress their valuations in quotations than to use evaluative attributes about ’the others’” (Holsanova, 1998b:258).

Gestural information, which is not present in the spoken channel, similarly saves the speakers face from the dangers involved in explicit spoken characterisation. Gestures is perhaps an even safer means of presenting the ethnic contrast.
But the question is whether verbal and gestural communication leads to an alignment of visual attention between speaker and listener. Do listeners look at the gesters produced by speakers in conversation? We know that faces and eyes and mouth regions attract visual attention but what about bodily movements? Goodwin (1986) suggests that speakers intentionally use gestures to attract the attention of their interlocutors. Streeck (1993) maintains that gestures can be overtly marked as communicatively relevant by the speakers if they look at their own hands. Gullberg & Holmqvist (1999) conducted an eye tracking study of gestures in interaction and found that auto-fixated gestures are fixated more often by listeners than other gestures. Their conclusion is that gestural and visual deixis functions as a powerful director of attention. As we argued above, listeners appear to perceive and take in the gestural information.

However, we would not argue that putting the ethnic characterisation in the gestural channel is a conscious strategy. Allwood (1976, 2002) underlines that different means of expression are used with different degrees of consciousness and awareness. Speakers in our examples are often not even aware of their own gesticulation. They appear to gesticulate as part of their attempt to express the ideas they have active in their minds. The hands simply draw and point without much conscious control on the part of the speaker.

It is however likely, that when the speaker softens or modifies the presentation of her ethnic contrast, this softening may have less of an effect in the gestural than in the spoken channel. This is what may have taken place in the first example above.

McNeill, Levy and Pedelty (1990) make systematic comparisons between the informational content of gestures and the ideational content of images, visualisation and thought. In the two examples above, the contrast clearly
resemble the *link, container* and *center-periphery* image schemas of Lakoff (1987). If we realise that spontaneous gestures have this semantic content, studying their use in ethnic contrast may become a tool to reveal portions of the communication that is not present in the spoken channel.

**References**


from The XVth Scandinavian Conference of Linguistics, 264–268. Oslo: Department of Linguistics.


