To tell and to show:
The interplay of language and visualizations in communication
Jana Holsanova, Roger Johansson, and Kenneth Holmqvist

In: Gärdenfors, P. & Wallin, A. (eds.)
A Smorgasbord of Cognitive Science, Nya Doxa 2008

The use of imitation, gestures and pictures has played a very important role in the evolution of human cognition (Zlatev, Persson; & Gärdenfors, 2005; Persson, 2008). Another important strand of research has studied the human ability to communicate multimodally, by using a combination of language, mimics, gestures, pictures, and body movements (Allwood, 2002).

Let us have a closer look at the use of various types of visualization that accompany language in communication. Imagine a group of friends involved in a lively conversation about their holiday adventures. Whether they talk about a dangerous bungy-jumping experience, hiking in the mountains, or enjoying the beach and pub life, they behave in a similar way. They do not only use language to describe events and things; they also use gestures, draw sketches, imitate voices, and engage their whole body to re-enact events. In short, when narrating, they not only tell but also show their experiences, tales, or stories.

Since Plato and Aristotle, these two activities are called diegesis (narrators describe things) and mimesis (narrators show things). In the following, we will be concerned with mimesis in more detail. The term has been taken by Plato from music theory and applied to designate scenic performance where actions of persons are imitated and re-enacted. In other words, mimesis is a re-production based on imitation of an action. In communication, mimesis is important for both speakers and listeners. It is a means of showing actions to make them visible (and audible) to the listeners. By using voice, gestures, and drawings, the speakers create scenes, embody and dramatize events, and thereby involve their partners.

“Mimesis gives us the sense of reality in fiction, the illusion of access to the reality of personal experience” (Lodge, 1990, p. 144). The more effective the imagination, the better the possibility for the audience to visualize the speaker’s experience.

There is, however, an important distinction to be made between two ways of “showing”: on the one hand we can indicate something by pointing at an object, on the other hand we can demonstrate things with the help of our voice, hands, or body. These two senses go back to Charles Peirce (1960), who identified three forms of relations between a representation and an object: an index refers to its object by means of a physical connection to it (e.g., a footprint as
a representation of a bear), an *icon* resembles the object it depicts (a drawing of a bear), whereas a *symbol* bears no resemblance to what it stands for and is thus arbitrary (the word “bear”). The first sense of showing is thus connected to what Peirce calls indexes, the other sense what he calls icons. According to Clark (2004), it is only the iconic sense of showing that is actually mimetic. In this context, Clark talks about *demonstrations* as selective depictions that “enable others to experience (in part) what it is like to perceive the things depicted” (Clark & Gerrig, 1990, p. 765). They make it easier for the listeners to imagine objects, scenes and events, what it is like to see, hear, and feel them, etc.

Producing and understanding mimetic devices relies on imagination and pretence (Clark, 2004, p. 8). When showing (and imitating) a person or demonstrating (displaying) events, the speaker runs a sort of simulation: s/he pretends to be another person and animates this virtual person’s actions. But imagination and pretend play would obviously not be possible without the ability to separate real states of affairs from the pretended ones (Frith, 1996). Also, the listeners must be able to understand demonstrations as pretend play to finally reach “joint pretence” together with their partners.

In communication, we can observe a great variety of mimetic devices that stimulate the imagination and involvement of the listeners. They can take form of pictures and illustrations, sound symbolism, quotations, iconic gestures, etc. In the following, we will take a closer look at these imagistic elements of communication and at the interplay between language and visualizations. In particular, we will be concerned with visualizations in different formats of mimesis: both auditive demonstrations, i.e. depictions by sound symbolism and quotations, and visual demonstrations, i.e. depictions by drawing, gestures, and imagery. The outline of the chapter is as follows: We will proceed from quotations (section 1), gestures (2), drawings (3), to mental imagery (4) and implications for learning (5).

1 To tell and to show with your voice (quotations)

A frequent type of mimetic devices used in conversation is *quotations*. Quotations are audible demonstrations that mediate direct experience and stimulate listeners’ imagination. In face-to-face conversation, we often imitate others, play scenes, and integrate others’ voices in our presentation in order to achieve certain effects in the current situation. By voice quality and a particular prosody, we animate non-present (virtual) characters and stage their activity. Not only singular speakers are animated but also plural voices, hypothetical speech and thoughts of others (Tannen, 1989; Holsanova, 1998; Adelsvärd, Holsanova, & Wibeck, 2002).

Look at the following example where speaker B tells his friend M about a telling-off
that his girlfriend Mary got at her job. He uses quotations and enacts a dialogue between a father of a child and Mary (M) who is working at a kindergarten. Since speaker B imitates two audibly different protagonist voices, which are marked in the transcript (cf. Holsanova, 2006b, p. 253).

Example 1

B so this child has difficulties eating you know, cause she was the only child and they have spoilt her terribly with sweets and things, so that when they serve normal decent food’. she simply doesn’t want it . you know, so . they have finally succeeded in making her eat a little and at this moment . her daddy enters, just in the middle of the meal, his kid sitting and eating,

M mm’

B and he says something (VOICE 1) *we’ll go home now* right, and Mary said that (VOICE 2) *we are eating now*

M mm’

B (VOICE 1) *yes but we’ll go home now* (VOICE 2) *we are eating now and you know what problems we have had with this,* right, (VOICE 2) *and she is eating now and she has to sit here,* . (VOICE 1) *we’ll go home now,* and Mary (VOICE 2) *you’ll go out till she has finished,*

B (laughing) (VOICE 1) *I want to talk to you,* hehe, (VOICE 2) *sure!* and in front of the door, she’s got a damn telling-off, (VOICE 1) it’s me who decides about my children

Direct quotation can serve the purpose of both dramatization and documentation. In narratives, we rather show what happened in order to increase dramatic intensity and to mediate direct experience. This is the dramatic (mimetic) function of direct quotation. By contrast, in reports such as news texts, quotation is used to increase the perceived objectivity and accuracy of the account to the reader. Here the documenting (diegetic) function of quotation is prominent. The writer claims authenticity by implying that he has direct access to the original speech situation (Redeker, 1991). In argumentative discourse, virtual participants in the form of authoritative references are often used to reinforce chosen opinions or to present a mental opponent whose opinions are to be questioned and undermined (Adelsvärd et al., 2002). For listeners, quotations mark an attentional shift towards other real or imaginary scenes, characters, and events (Sanders & Redeker, 1996).

Quotations position those who are animated as speakers in an unflattering light and hide behind in case of sensitive topics (Holsanova, 1998, 2006b), but can also increase the distance between the speaker and the described characters: “with quotations speakers can partly or wholly detach themselves from what they depict” (Clark & Gerrig, 1990, p. 792).

Another way of depicting persons and events is to use onomatopoeia or sound
symbolism (his heart went ticktock ticktock). For instance, when characterizing and criticizing politicians the speaker can let them whine like helpless children (the whining politicians say uhuhu), or express anger with the help of growling rather than with an extensive verbal description (she thought mmhrrmnhrhr). Onomatopoeia is a very effective way to ascribe attitudes without having to describe them verbally.

2 To tell and to show with your hands (gestures)
Probably the most usual way to embody an action or manner of action is through body language, in particular gestures. According to McNeill (1992), planning of utterances involves the interplay of imaginistic thinking and linguistic thinking (manifested as gestures and speech). There are, however, a number of different opinions regarding the relation between the content of gestures and of concurrent speech (Kita & Özyürek, 2002). The free imagery hypothesis states that gestures are generated from imagery, independently of the language, whereas the lexical semantic hypothesis states that gestures are generated from the semantics of lexical items in the accompanying speech and therefore cannot encode what is not encoded in the concurrent speech. Finally, the interface hypothesis states that gestures originate from an interface representation between language and imagery. According to this last hypothesis, production of speech and gestures is interrelated, and gestures may encode information that is not expressed in speech.

As mentioned earlier, there are two senses of “showing”: the first is indicating (by pointing) and the other is demonstrating (by animation). Prototypically, pointing is understood in its deictic sense: Through juxtaposition, pointing gestures instruct the hearer to attend to something beyond the talk and to locate what is being indicated (Goodwin, 2003). On the other hand, pointing gestures can take a step towards iconicity (e.g. when a person points at an object and traces the shape of the object that is being pointed at (Streeck, 1996). In the latter sense, when tracing an iconic shape, pointing gestures can be understood as demonstrations.

The information provided by gesture is analogue and depictive (Özyürek & Kita, 1999). The demonstrative function of bodily mimesis can be compared to quotations as demonstrations. Sensitive topics can be made in the gestural channel, keeping the spoken channel free from explicit reference (Holmqvist & Holsanova, 2007). Thus, similarly as in quotations, the gestural information which is not present in the spoken channel saves the speaker’s face from the dangers involved in explicit spoken characterization. In case of
disabilities, e.g. in aphasia, gestures and other non-verbal means of communication may compensate for the restricted use of language (Ahlsén, 2006).

The question remains how verbal and gestural communication becomes integrated by speakers and listeners. Do listeners look at the gestures produced by speakers in conversation? Goodwin suggests that speakers intentionally use gestures to attract the attention of their interlocutors. Cassell, McNeill, and MacCullough (1998) conclude from their mismatch studies that the speaker’s gestural channel is indeed perceived by listeners and plays an important role for the understanding. Gullberg and Holmqvist (1999) found in an eye tracking study of gestures in interaction that auto-fixated gestures, i.e. gestures that the speakers themselves focus on, are attended more often by listeners than other gestures. This result confirms the assumption maintained by Streeck that gestures can be overtly marked as communicatively relevant by the speakers if they look at their own hands.

3 To tell and to draw
Apart from spontaneously using gestures, mimics, and body postures along with their utterances, speakers also draw pictures when they describe their experiences. Drawings – such as sketch maps of houses and areas, illustrative drawings of people and clothing, or explanatory diagrams – are often incrementally produced and modified as a part of a conversation. The use of iconic and pictorial representations is useful in communication, since it helps the speaker and the listener to interactively adjust their visualizations and achieve understanding. Pictures, maps, and sketches are visualizations that show how the described reality has been conceptualised (Tversky, 1999).

Drawings can represent (a) a concrete spatial domain, such as geography, sizes, shapes, spatial relations; (b) non-spatial domains, such as amounts of money, temporal relations; (c) abstract domains, such as intensity, contrast, and quality dimensions, and (d) dynamic processes, such as stages in a decision process, causal relations, development over time.

Apart from being a support for visualization and demonstration (specifying the location and shape of objects, spatial relations, and events), drawings also fulfil other, more abstract, functions in communication. They are a useful tool for the identification of “where are we now?”, serve as a storage of referents, as an external memory aid for the interlocutors, as an expressive way of underlining what is said, and as a representation of a whole problem discussed in the conversation (Holsanova, 2008).

Speakers’ discourse contains concepts that appear as schematic representations and establish patterns of understanding. When speakers want to draw attention to a complex visual
idea, e.g. a scene, a navigation route or an apartment layout, they have to organize the information so that their partner can understand it. By uttering ideas, speakers evoke images in the consciousness of the listeners, the minds of the speaker and the listeners get synchronized, and the listeners co-construct the meanings. The process of how the listener's internal image is constructed from spoken discourse and simultaneous drawing has been studied within image-oriented semantics (Holmqvist, 1993) where discourse understanding is described in terms of evolving mental images.

In spontaneous face-to-face conversation, this process has a more dynamic and cooperative character (Clark, 1996). The partners try to achieve joint attention, formulate complementary contributions, and interactively adjust their “visualizations”. Quite often, the partners draw simultaneously with their verbal descriptions. Sketches and drawings are external representations that reflect the conceptualization of reality and serve as an aid for memory (Tversky, 1999). The utterances and non-verbal actions (such as drawing and gesturing) can be conceived of as instructions for the listeners on 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).

Using a delimited, common source of references makes it easier for the speaker and the listener to coordinate and to understand each other. Drawing as an external representation allows cognitive processes to be captured, shared, and elaborated. In a naturally occurring conversation, external visualizations help the partners to achieve a joint focus of attention and to coordinate and adjust their mental images during meaning making.

4 To tell and to imagine

In the previous sections, we mentioned mimetic devices that the communicative partners use when showing and re-enacting events and actions for the listeners. It has, however, been shown that we use gestures even in a situation when our partners cannot see us. The question is thus whether we as speakers use mimetic devices and re-enactment for ourselves, in the form of mental imagery.

*Mental imagery* is “the mental invention or recreation of an experience that in at least some respects resembles the experience of actually perceiving an object or an event, either in conjunction with, or in the absence of, direct sensory stimulation” (Finke, 1989, p. 2). In popular terms, mental imagery is described as “visualising” or “seeing something in the mind’s eye”.

We use mental imagery when we mentally recreate personal experiences from the past,
retrieve information about physical properties of objects or about physical relationships among objects, read novels, plan future events or anticipate possible future experiences, imagine transformations by mental rotation and mental animation and when we solve problems (Finke, 1989; Hegarty, 1992; Yoon & Narayanan, 2004). In other words, imagery plays an important role in memory, planning, and visual-spatial reasoning, and is considered a central component of our thinking.

Since mental images are closely connected to visual perception, this mental invention or re-creation of experience almost always results in observable eye movements that can be traced by new technology. Eye tracking methodology has become a very important tool in the study of human cognition, and current research has found a close relation between eye movements and mental imagery (Holsanova, 2001, 2006a). In order to verify the assumption that we use our ability to create pictures in our minds, we conducted a series of studies on mental imagery during picture description. The results of these studies contribute to our understanding of how speakers connect spoken discourse to mental imagery.

Already in our first eye tracking study (Holsanova, Hedberg, & Nilsson, 1998), we found some striking similarities between the participants’ eye movement patterns when they looked at a complex picture and their eye movements when they later on looked at a white board and described the same picture from memory. We then conducted a number of new eye tracking studies where participants looked at a blank white board and visualized a scene they had previously either seen on a picture or heard as a spoken description (Johansson, Holsanova, & Holmqvist, 2006).

Look at the following example showing a comparison of one person’s eye movement patterns during picture viewing and during picture description from memory.

Example 2: The stimulus picture (top) and one and the same participant’s eye patterns (bottom) after the viewing phase (left) and the description phase (right).
The results of our studies clearly showed that when describing a scene from memory the participants to a high degree moved their eyes in a pattern that “painted” the imagined scene on the white board in front of them. Additionally, it was found that the effect was equally strong irrespective of whether the original elicitation was spoken or visual, and that it was also present in complete darkness. Our results support the hypothesis that mental scanning (Kosslyn, 1980) is used as an aid in recalling picture elements, especially when describing their visual and spatial properties. Mental imagery seems to play an important role even for the speakers involved in discourse. We represent and re-enact our previous experience when we describe a scene from memory.

5 To show and to learn
The findings described above have decisive implications for learning. For speakers and listeners, both mental imagery and re-enactment of events can play an important role as a memory aid. For learning purposes, for instance in argumentation, quotations can be used to stage another person’s talk or thought, to reinforce chosen opinions or to present a mental opponent whose opinions are to be questioned and undermined (Wästerfors & Holsanova, 2005).
In the context of collaborative learning, the use of sketches and drawings is advantageous. First, sketches and drawings show how the described reality has been conceptualized. Second, they allow revisions, regroupings, refinements, and reinterpretations and are therefore an important thinking tool (Suwa et al., 2001). Third, they help the speaker and the listener to interactively adjust their visualizations and achieve understanding. In sum, drawings and sketches allow cognitive processes to be captured, shared, and elaborated.

In the context of textbook illustrations and instructional materials, it is important to outline visualizations according to the users’ mental model, i.e. the way users conceptualize how everyday objects and situations are structured or how they work. Visualizations have world-like qualities resembling actual objects or events. By means of this analogy, they function as a substitute for the referent and evoke similar experience to the real-world referent. Visualizations in instructional materials can offer highly realistic impressions of objects and events, which might otherwise be too small, too large, too fast, too far away, or too dangerous to observe in reality. In that respect, visualizations do not only replace real-world experience, they may even improve this experience by providing information that would not have been accessible in the real world (Scheiter, Wiebe, & Holsanova, 2008).

Visualizations in their representational function depict objects and relations mentioned in a text in a way that the meaning of the text is made more easily accessible for learners by making a text more concrete. Visualizations with an organization function provide an organizational framework for a text (e.g., how-to-do-it diagrams) and thereby make the content more coherent by highlighting the argumentative or organizational structure of the text. Accordingly, visualizations are often introduced in textbooks and multimedia instructions to clarify difficult passages and abstract concepts.

However, it should not be taken for granted that learners will extract the information from a visualization that was intended by the instructor. Rather, students need to be supported in extracting the relevant information from the visualization and guided as to how best to deploy their perceptual and cognitive resources. This support can be provided either by guiding learners’ attention towards its relevant aspects (e.g., highlighting) or by improving students’ competencies in dealing with visualizations (Scheiter et al., 2008).

To sum up, our communication is multimodal and embodied. By combining mimetic devices such as quotations, drawing and gestures, communicative partners depict, demonstrate, embody, and re-enact objects and actions in order to create vivid scenarios for each other. The usage of these devices evokes images in the mind of both the speaker and the listener, which in some respects resemble the experience of actually perceiving a scene or an
event.

6 References


