

Toward Enhancing User Involvement via Empathy Channel in Human-Computer Interface Design

Masashi Okamoto¹, Yukiko I. Nakano², and Toyoaki Nishida³

¹ Graduate School of Information Science and Technology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
okamoto@kc.t.u-tokyo.ac.jp

² Research Institute of Science and Technology for Society,
Japan Science and Technology Agency, Atago Green Hills MORI Tower 18F,
2-5-1 Atago, Minato-ku, Tokyo, 105-6218, Japan
nakano@kc.t.u-tokyo.ac.jp

³ Graduate School of Informatics, Kyoto University, Yoshida-Honmachi,
Sakyo-ku, Kyoto, 606-8501, Japan
nishida@i.kyoto-u.ac.jp

Abstract. This paper reports our research toward enhancing the ‘User Involvement’ in human-computer interaction, which is introduced to grasp the idealised state in which a natural communication between a computer and its user should be established. Moreover, we also introduce a new idea of ‘Empathy Channel’, through which humans can interact with virtual agents and objects on computers with more sense of reality. The main issue of this paper is to clarify various methods for establishing an Empathy Channel in building a human-computer interface. We believe a good design of human-computer interface with an Empathy Channel is one of the best methods to enhance the User Involvement.

1 Introduction

This paper reports our research toward enhancing the “User Involvement” in human-computer interaction. As we describe in Section 2, the notion of the User Involvement is introduced to grasp the idealised state in which natural interaction between a computer and its user should be established. Moreover, we also introduce a new idea of ‘Empathy Channel’, through which humans can interact with virtual agents and objects on a computer screen with more sense of reality. The main issue of this paper is to clarify various methods for establishing an Empathy Channel in designing a human-computer interface. We believe a good design of human-computer interface with an Empathy Channel is one of the best methods to enhance the User Involvement.

In the following sections, first, we describe the requirements for establishing the User Involvement. Then, we insist that empathy is the key notion to comprehend cognitive abilities of a human to step into another world different from the real world he actually lives in. Secondly, we show some related works supporting our ideas. Thirdly, we introduce the reference-point ability to formulate the essential aspects of the User Involvement precisely, and then we discuss the possible designs of Empathy Channel in the human-agent communication environment. Finally, our future work on experiments to prove our ideas is shown.

2 User Involvement in Human-Computer Interaction

To establish a natural communication between humans and computers, the following issues have been discussed recently:

- The human-to-computer interaction should/could be equated to that of human-to-human. [1]
- Though many children nowadays have been deep into computer games, most of non-entertainment computer services have not attracted people as much as games have done.
- Non-verbal information must be taken into consideration since verbal information alone is not sufficient for establishing human-computer communication.

Behind these discussions there exists a strong belief that the attributes of human-to-human communication, where a huge variety of complicated information is conveyed in a natural way, can be equally applied to human-to-computer interaction. However, there is no persuasive theoretical frameworks yet which describe how a human as a cognitive subject can communicate with computers as ‘real’ partners.

In order to comprehend such issues appropriately, we introduce the notion of ‘User Involvement’ here. User Involvement is the cognitive way humans willingly engage in the interaction with computers, or the way in which humans are, on the contrary, forced to be involved in a virtual world which computers display or in a human-to-robot communication. We believe that designing a human-to-computer communication environment mainly concerns the management of the User Involvement.

2.1 Requirements of User Involvement

The main requirements to establish the User Involvement are considered as follows:

- **Cognitive/Communicative reality is achieved:** The user should feel the virtual world or the human-to-computer interaction as ‘real’.
- **Two (or more) cognitive spaces are linked:** When the user sees and steps into another world, there need to be at least two cognitive spaces, that is, his/her viewpoint (here) and what he/she sees (there) [2]. Moreover, there must be some link among those multiple cognitive spaces.

‘Reality’ is a difficult concept to define precisely. In a traditional philosophical context, reality refers to everything that actually exists and contrasted to non-existence or possibility. In our approach we use the term as it means “perceptions, senses, beliefs, and attitudes toward reality, that is, people’s ways of seeing reality”, as usually used in cognitive psychology, anthropology, or sociology. Then reality is classified into the following three dimensions:

- **Cognitive reality:** The way of seeing objects, events and their relations in the real/virtual world as real.
- **Communicative reality:** The sense of reality that is achieved through communication with others.
- **Social reality:** The collective and intersubjective sense of reality based on sharing thoughts or opinions with one another.

As social reality is concerned with a real or virtual community, we discuss cognitive and communicative reality in this paper.

When we feel the world ‘real’, we are not living in a single space. It is because we not only exist in an external world but also can ‘think’ in an internal world, namely in our mind. Seeing things, animals and other people outside, we often employ our thoughts on each of them inside. We are living in and cognitively shuttle between both of these two cognitive spaces, when our body is a vehicle or a channel to connect them. Moreover, if we can also find reality in a virtual world, it might be said that we already have stepped into another different space. This is why two or more cognitive spaces should be linked for the User Involvement to be established.

2.2 Multiple Modes of Involvement

When the User Involvement is considered as such, the way for users to interact with computers might be dual or multiple in itself. Norman [3] points out two types of cognitive involvement for humans to interact with the world: ‘experiential cognition’ and ‘reflective cognition’. The former is the data-driven information processing with reactivation of information patterns in human memory, while the latter concept-driven processing in deep reasoning such as decision-making or planning. People use each cognitive mode appropriately in the human-to-computer interaction.

On the other hand, Laurel [4][5] also made it clear that users engage in the human-computer interaction in two ways, that is, a ‘first-person’ engagement (as active participant), and a ‘third-person’ engagement (as a passive observer)¹. The former is seen in such activities as walking through woods, writing a letter and so on. Meanwhile, the latter engagement is observed in watching movies or reading novels, when the viewer or reader is outside the action and describes the events in the movies or novels using third-person pronouns. She insisted that first-personness is enhanced by an interface that enables inputs and outputs that are more nearly like their real-world referents, and predicted that the conventional menu- or command-based system will be replaced with systems that employ the natural language in ways that are mimetic of real-world activities.

The multiplex of involvement suggests that a human is usually involved not in a sole activity, either cognitive or behavioural, but in multiple cognitive spaces where he can smoothly switch his mode of involvement. Therefore, our research is focused on what factors enable a human to smoothly switch such multiple modes of involvement.

In the next section, we will show that ‘Empathy Channel’ becomes one of the links between cognitive spaces to enhance the User Involvement.

¹ Note that Laurel also mentioned *second-person* experiences, which frequently occur in operating a computer. However, it is not clear what characterises the second-person engagement, while the relation between the first-person and the third-person experiences might be compared to the binary opposition between ‘subjective’ and ‘objective’. We thus excluded the second-person engagement from our discussion here.

3 Empathy Channel

In the human interaction, a human not only interacts with others but also reflects on his own interaction. Therefore, a human has two viewpoints when interacting with others; one is an ‘object-level view, the other a ‘meta-level view’’. The two viewpoints are tightly linked through the self-identification of ‘ego’ (reflective subject) and ‘self’ (experiential subject)² On the contrary, in observing others interaction at a distance, such duality is not established in general. However, if the observer can empathise with one participant in that interaction, he can also acquire the virtual object-level view, so that he can experience the interaction as if it were his own. This dual viewpoint via empathy shows how the ‘first-person’ engagement can be achieved (See Fig. 1).

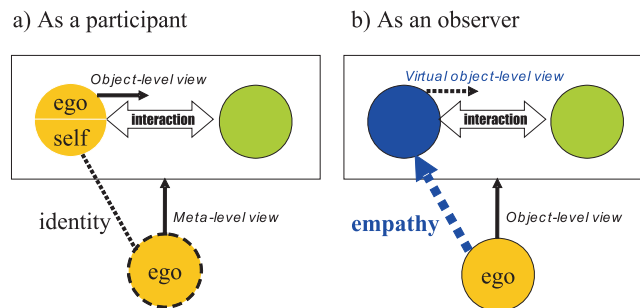


Fig. 1. The Empathy Channel: (a) involvement as participant in interaction, (b) involvement as observer outside interaction

In other words, empathising with others enables us to cross over two cognitive spaces (i.e. *here* and *there*), and enhances the psychological involvement with interaction in another world, which leads us to feel the world “real”. Since empathising with characters or objects in a virtual world in particular functions as a channel to connect the real and the virtual world, we call it ‘Empathy Channel’ here. As suggested in subsection 2.1, it is one of the conditions that maintain our cognitive and communicative reality, which will enhance the User Involvement in human-computer interaction.

3.1 Empathy Channels in Computer Games

The good examples in which the Empathy Channel does effectively function are computer games (Fig.2).

In the early stages of computer games, the bird’s-eye view is commonly used (seen in Fig. 2a). But, although such view excels in looking over the world where the user is involved, it has a disadvantage in feeling involved in the game playing itself. Later, in simulation games in particular, the first-person view was adopted in order to enhance

² This duality was originally provoked by G. H. Mead [6]. He distinguished Me and ‘I’ in that the former refers to the social aspect of self and the latter to the subjective aspect. In this paper, the ‘ego-self’ distinction roughly corresponds to the ‘I-Me’ distinction of Mead.

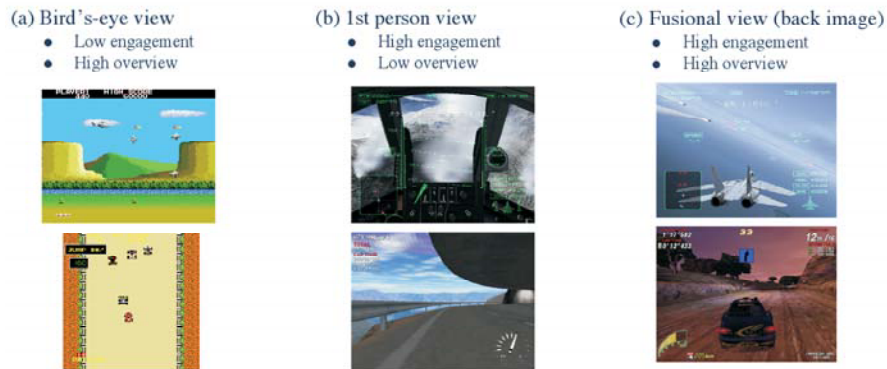


Fig. 2. The transition of user's view in computer games: (a) old-fashioned games with bird's-eye view, (b) simulation-type games with 1st person view, (c) popular games with fusional view

the User Involvement (Fig. 2b). However, it in turn loses an overview of playing fields. After that, most of today's popular games have adopted the fusional view using the back image of the player's vehicle or virtual ego (Fig. 2c). In our opinion, it is a typical example of establishing the User Involvement via Empathy Channel. Through an empathised object (which is a back image of user's vehicle here) a user can smoothly step into the virtual world and go back to the real world to comprehend what he is doing.

3.2 Related Works

The effect of establishing an Empathy Channel in the user involvement has been proved in some empirical researches.

For example, Miyazaki [7] examined the emphatic effects of a back image on readers. When the experimenter presented a juvenile story to subjects, half of the subjects were presented a storybook with pictures drawn from the observer view, while the rest was presented the same story with pictures from the fusional view featuring a back image of the protagonist. After the presentation, all the subjects were asked about the feelings of the protagonist and the contents of the story. As a result, those subjects presented with the back images could comprehend more precisely what the protagonist in the story is feeling and were more strongly involved in the story. In brief, the back images helped the readers to experience the virtual world as if it were their own.

On the other hand, Morikawa, et al. [8] reported that using the HyperMirror system, which was built by them in order to enable the users in different places to be present virtually on the same computer screen, established a smooth communication among them. In this experiment, the users strongly felt as if they had stepped into the HyperMirror space together without any eye contact. Compared with the video conferences or TV phones, it is noteworthy that lack of eye-contact does not prevent the communicative reality of the users. We assume that this experiment is an example of the Empathy Channel through which the self-image in a virtual world is naturally identified with the actual self of the user.

4 Empathy Channel as Reference Point

Toward building an effective and attractive human-computer interface it should be discussed how to establish an Empathy Channel to enhance the User Involvement. Take a human-agent communication environment for example. In such an environment which enables the user to communicate with virtual agents, he must be able to effortlessly step into the virtual world in advance. However, the virtual world on a computer screen feels not real for the user unless the screen is huge enough to surround him. It is thus necessary for either cognitive or communicative reality to be achieved.

In order to explore the candidates of Empathy Channel, we adopt the idea of the 'reference-point ability', which was originally provoked by Ronald W. Langacker [9], a cognitive linguist. This idea is also developed to formulate the model of the User Involvement in this section.

4.1 Reference-Point Ability

Imagine that you are asked by someone on a street about the location of the nearest post office. Then you would explain to her in the following way: "Can you see a traffic signal over there? Go there and turn left, then you'll reach the post office soon". In such an occasion, we cannot always access the target directly, so we use something that is easily accessible in our cognitive environment instead. Langacker [9] calls the clue for accessing the target as a 'reference point', and suggests that the reference-point phenomenon is so fundamental and ubiquitous that we often forget its effect. In our verbal activities it is often through some reference point we reach the target.

Langacker describes the essential aspects of the reference-point ability as shown in Fig. 3. The circle labelled C represents the *conceptualiser*, that is, a human as the cognitive subject. R is the *reference point* and T is the *target*, with which the conceptualiser establishes a mental contact using the reference point. The dashed arrows indicate the mental path which the conceptualiser follows in reaching the target, while the ellipse D represents an abstract entity, named the *dominion*, which can be defined as the class of potential targets. It is important that the reference point has a certain cognitive salience.

4.2 Reference Points in User Involvement

We believe that the reference-point ability also works in our non-verbal activities and can be applied to the User Involvement in a human-computer interaction as well. As

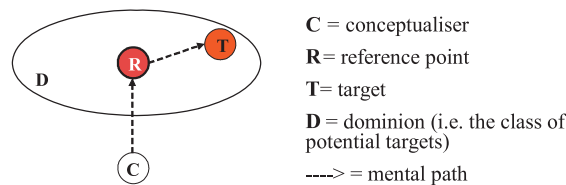


Fig. 3. The essential aspects of the reference-point ability (Langacker, 1993)

described in Section 3, the user of a computer steps into the virtual world via some Empathy Channel that links his living space and the virtual space. From the point of view of the reference-point ability, the Empathy Channel is established by something that is cognitively accessible and can be empathised by the user. Therefore, it is a reference point for accessing the objects or characters in the virtual world as a target. Moreover, as suggested in Section 2, we conceive the real world or its entities via ourselves, which lie both in the internal space (i.e. the mind) and in the external space (i.e. the real world). Thus it can be said that we use ourselves as a reference point to conceptualise the entities in the real world.

Some of the aspects of the User Involvement in the view of the reference-point ability are illustrated in Fig.4.

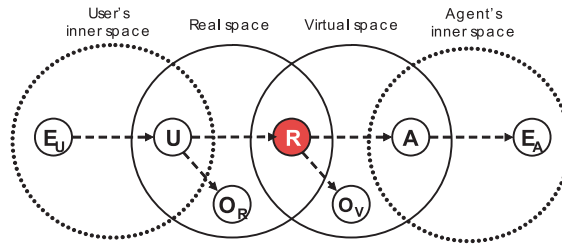


Fig. 4. The astigmatic model of User Involvement in human-to-computer interaction

In this model³, **U** represents a user or his self/body, **E_U** is the ego of the user or the cognitive subject, and **O_R** refers to an object/entity in the real space. **A** represents a virtual agent, **E_A** is the ego of the agent, **O_V** refers to an object in the virtual space, and **R** represents the empathised object as a reference point (i.e. the Empathy Channel).

In conceptualising something (**O_R**), a human uses his self/body (**U**) as a reference point in our everyday lives. If the user recognises the reality of the **O_V** in human-computer interaction environment, his reality is achieved through some Empathy Channel (**R**) as a reference point to access **O_V**. Additionally, if the virtual agent on a computer screen is animated or lively, and thus felt to be real for the user, he will assume the existence of something like the agent's intention or belief (in **E_A**) even though it is just an affectation.

5 What Is Needed for Establishing an Empathy Channel?

Our main goal is to clarify the conditions and the candidates for the Empathy Channel to be established in the design of a human-computer interface. The astigmatic model of the User Involvement suggests what will be required in the Empathy Channel:

³ We call the model as the *Astigmatic Model* because the overlapping among the cognitive spaces is essential for the user involvement to be established with each reference point. In our hypothesis, one reality is always constructed on those overlapped cognitive spaces like seeing an object with astigmatic eyes.

- (1) The high accessibility in the real space
- (2) The consistent functioning in the virtual space
- (3) The continuous correspondence of behaviours in both the real and the virtual spaces

The reason for (1) is that the empathised object should be highly accessible to the user in the real space to function as the reference point. Additionally, since the empathised object also lies in the virtual space, it should consistently function as a member of the virtual space to maintain the reality of the world, and it is the reason for (2). Moreover, the empathised object connects two cognitive spaces simultaneously. If its behaviours in each space are not corresponding to each other, the connection will be broken.

5.1 Empathy Channel for Human-Computer Interfaces

When we first think about a human-computer interface, a mouse or a keyboard is easily evoked. According to the astigmatic model of the User Involvement, a good human-computer interface should be (1) easily accessible or operational by the user in the real world, and (2) consistently functioning in the monitor, and should (3) maintain the coordination of each behaviour of the device in both worlds. From that view, the mouse and the keyboard are not so good in that using such devices contains no natural behaviours in our everyday lives, and that the correspondence of the user's action in the real space and the device's response in the virtual space is frequently broken. Therefore, many researchers are now trying to create a new interface device using our natural behaviours or movements to replace the mouse and the keyboard with. We believe that such a new interface should be an appropriate one to the User Involvement.

Similarly, a good human-computer interface in a human-agent communication environment should follow the theory of the User Involvement. Different from a mouse or a keyboard, the interface in the human-agent communication environment consists of a virtual world and virtual agents that resemble our real world and its habitants. Then, establishing a cognitive or communicative reality is necessary for the user to be smoothly and deeply involved in the virtual world.

5.2 Empathy Channel for Cognitive/Communicative Reality

The candidates for an Empathy Channel in a human-agent communication environment are shown as follows:

- Establishing cognitive reality
 - Self-image of the user in the virtual space
 - Matching age/sex/ethnicity
 - Smooth viewpoint switching
- Establishing communicative reality
 - Exchanging verbal information
 - Eye contact
 - Appropriate reactions of a virtual agent
 - Joint attention

Though we do not insist that all the possible candidates are fully cited above, these are considered to function well as establishing an Empathy Channel.

In order to achieve the cognitive reality, as shown in Section 3, self-images of the user in the virtual space work as an Empathy Channel. Not a full image but a partial body image of the user will be an Empathy Channel, like the image of a hand or a face of the user. In empathising with an agent in a monitor, the correspondence of sex, gender or ethnicity between the user and the agent will surely enhance the empathising effect. Furthermore, in movie-like contents, there needs to be a smooth viewpoint switching that would reduce the cognitive burden of the user experiencing the virtual world.

On the other hand, a user should feel the communication with the virtual agents as real. Since the prototype of the natural human-computer communication is a human-to-human communication, the characteristics of the human communication can be also applied to the human-agent communication environment, such as exchanging verbal information, an occasional eye contact, appropriate reactions or responses corresponding to the precedent actions, and joint attention. However, it is difficult for the current technologies to realise as sufficient communicative behaviours of the virtual agents as in a human-to-human communication. For instance, exchanging verbal information between a user and a computer requires highly intellectual technologies of natural language processing, speech synthesis, speech recognition and so on. Moreover, lack of eye contact between a user and an agent often makes the agent and the computer look like just a mock figure and a vacant box. Again, needless to say, the range of possible reactions that perfectly correspond to the user's actions is very limited. Those characteristics will surely work as the Empathy Channel for communicative reality if they are sufficiently realised, but we have to consider what is possible and effective in the limited actions the current computers can perform.

5.3 The Visual Settings Using Empathy Channel

Lastly we describe the possible settings of the human-agent communication environment using the Empathy Channel. See Fig.5 below.

A typical setting of the virtual agent is like (a), which features a conversational agent that faces up to the user and talks to him. In such a setting, the user gets less satisfied with the agent as its appearance becomes more human-like, because the user places more expectations on the agent. In particular, lack of eye contact will betray the expectations. That is why cartoonish agents are often used instead in interactive contents.



Fig. 5. The possible settings of human-computer communication environment: (a) the setting featuring an agent facing the user, (b) the setting featuring a back image as Empathy Channel, (c) the setting featuring joint attention as Empathy Channel

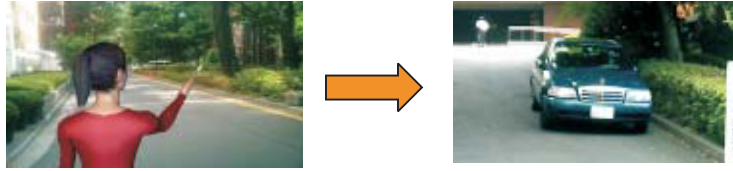


Fig. 6. The shot transition motivated by the agent's attention behaviour

Therefore, we suggest that using such a facing-up agent setting should be kept to the necessary minimum. Instead, the back image setting or the joint attention setting should be more used. As shown in (b), the user can easily empathise with the agent through its back image as an Empathy Channel. Presenting the landscape with the back image will enhance the User Involvement more than presenting the landscape alone. It is because the correspondence between the viewpoint of the agent and that of the user produces the sense of involvement as if the user were actually standing in the landscape.

The joint attention setting (c) evokes the attention of the user toward a target which the agent is watching. The difference between the back image setting and the joint attention setting is that the former makes the user empathise with the agent and share a mutual viewpoint while the latter does not. Instead, the joint attention setting makes the user and the agent virtually share the communicative environment including the target. Furthermore, not only the static visual settings but also the dynamic transition settings can enhance the User Involvement. See the following Fig. 6.

In movie-like contents featuring the human-agent communication, the shot or scene transition is very important for the User Involvement to be established. If some irrelevant shot is inserted in movie contents, the viewer will be embarrassed and will have to process the cognitive burden that the irrelevant shot lays on. Therefore, the smooth switching of the shots should be considered in constructing any movie-like contents. We suggest that one of the smooth shot transition types with less cognitive burden for the viewer is that which is motivated by the attention behaviours of the agent. The attention behaviours, such as turning a gaze, pointing by hand, and a verbal reference toward the target, will enable the viewer to smoothly accept the shot transition from an agent shot to a target shot. Once such an Empathy Channel as described here is established, the user will be inclined to communicate with conversational agents even in a small computer monitor.

6 Conclusion and Future Work

In this paper, we have presented our theory of the User Involvement. The main issues described so far were as follows:

- As well as in our everyday experiences, the user should also creatively use multi-involvement in a human-computer interaction.
- To enhance the User Involvement, it should be considered how to establish an Empathy Channel between the real world and the virtual world.
- Empathy Channel works as a reference point for the user to step into a virtual world.

Furthermore, we discussed the requirements of establishing an Empathy Channel in human-computer interfaces, and then the effective and empathetic visual settings of a human-agent communication environment were described in both static and dynamic settings in the view of the User Involvement.

We thus conclude that establishing an Empathy Channel will surely enhance the User Involvement in a human-to-computer interaction. However, the theory should be attested by building systems based on the User Involvement with the subsequent experiments on the psychological effects toward the users.

Now we are constructing the following two types of human-computer interaction environment based on the User Involvement theory. One of the ongoing researches is building a supporting system for the immersive CG contents creation, which is based on the analysis of the shot transition in actual TV programme. The other is establishing the natural human-robot communication environment using a listener robot, which can establish the joint attention. The reports of these ongoing researches will be done in the near future.

References

1. Reeves, B., Nass, C.: *The Media Equation: how people treat computers, television, and new media like real people and places*, CSLI (1996)
2. Miyazaki, K., Ueno, N.: *Viewpoint (in Japanese)*. University of Tokyo Press (1985)
3. Norman, D. A.: *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*, Addison-Wesley (1993)
4. Laurel, B.: Interface as mimesis, In: Norman, D. A. & Draper, S. (eds.): *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum (1986)
5. Laurel, B.: *Computers as Theatre*, Addison-Wesley (1991)
6. Mead, G. H. *Mind, Self, and Society: from the standpoint of a social behaviorist*. The University of Chicago Press (1934)
7. Miyazaki, K.: The Effects of Human Back-image as a Mooring Point in Empathetic Comprehension through Visual Images (in Japanese). *Japan Educational Psychology Association Proceedings* (1993) 35
8. Morikawa, O., Hashimoto, R., Yamashita, J.: Self Reflection Can Substitute Eye Contact. *CHI2003 Extended Abstracts* (2003) 944–945
9. Langacker, R. W.: Reference-Point Constructions. *Cognitive Linguistics* 4 (1993) 1–38