in: W. Schuler, J. Hannemann and N. Streitz (Eds.), Designing USer Interfaces for Hypermedia. Springer 1995, pp. 58-67.

Metaphor Engineering: a Participatory Approach

Matthias Rauterberg and Markus Hof
Work and Organisational Psychology Unit, ETH, Switzerland

Abstract. A method for 'metaphor engineering' is introduced in the context of participatory multimedia design. Our hypothesis, that adults talk to children more with a metaphorical language than to other adults, was empirically verified. Especially male adults do not qualify for the job as metaphor engineers. Metaphors are powerful, but not sufficient to come up with a good interface. Single, over-detailed metaphors can be too restrictive and unwieldy. Effective interface metaphors often evolve over time through design, evaluation and redesign. The approach presented in this paper guarantees that the starting point of the design cycles can be optimised.

Keywords. Metaphor, multimedia, hypertext, user interface, participatory design

1 The Interface Metaphor

The construction of hypertext systems needs several design solutions, "which requires dedicated methodological support". One important design problem is finding a metaphor for the interface architecture, which is the design basis for the screen layouts and the dialogue structure (Carroll 1982) (Carroll 1985) (Chauvet 1991). This approach leads directly to the usage of multimedia interface technology (Carroll 1988) (Eberleh 1991) (Shum 1990).

Waterworth (1992) distinguishes four different levels of analogy in interface design (cf. Hutchins 1989). The *conceptual model* is the overall view of the system, or part of the system, as conceptualised by the design team. This may comprise one overall metaphorical view or several metaphors and also non-metaphorical aspects. *Mental models* describe how users view the system, which of course will vary with the sophistication and experience of the user. This should hopefully be reasonably compatible with the conceptual model, but will not be identical in detail nor will it be complete. "A metaphor in this context is a mapping relation between aspects of the conceptual model and the world at large (e.g., a desktop). *Interaction modes* are details of user operations that are included in the conceptual model and may correlate with metaphors (clicking to open a file, dragging to move from one folder to another, uttering a word to place a marker, for example)" (Waterworth 1992, p. 91).

Waterworth (1992) illustrates furthermore the different levels at which system functionalities and metaphor characteristics correspond to each other in Table 1 (see also Hammond and Allison 1987). A metaphor need not be appropriate at all levels of description. Different metaphors at the same level of description may be appropriate and multiple metaphors may often be useful (Weyer 1985).

Level	Interaction events	Example metaphor	System perspective
Pragmatic	Tasks	Finding out about a topic	User session(s)
Semantic	Ways of doing tasks	Going on a library tour	Organised set of traversals
Syntactic	Combinations of moves/items	Selecting a book and 'opening' it	Displayed node
Lexical	Items/moves (icon, clicks, pointing, etc.)	Pointing at a picture	Input/output token recog-nised

Table 1. Levels of system-metaphor mapping (Waterworth 1992).

Metaphors are not 'right' or 'wrong' descriptions: rather they are 'stimulating' invitations to see a target domain in a new light (Richards 1936) (Haverkamp 1983). But, where do metaphors come from? As Carroll, Mack and Kellog (Carroll 1988) claimed, that "there is not now (and likely never will be) a discovery procedure for metaphors," we try to overcome this position. In the context of a participatory approach (Rauterberg 1992) we are looking for a method which generates a metaphor by the domain expert himself. This participatory approach ensures that the generated metaphor fits the application domain. If we try to design a multimedia information system for a new domain (e.g., proofs in philosophical theories), then we need a lot of different and unknown design concepts. We cannot continue to use well-known metaphors (see Figure 1) (Henderson 1986) (Carroll 1988) (Gould 1990). One way to discover these concepts is to ask domain experts. So, we are looking for a method, which stimulates the domain expert to describe his target domain in a metaphorical language.

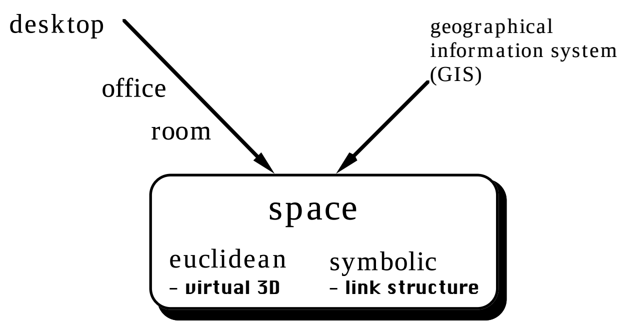

Figure 1. Historical influences for the Euclidean and symbolic space concept.

The domain knowledge of an expert is directly communicable only to other experts with the same domain knowledge. If we try to understand a foreign domain then we need a mapping or transformation of the unknown knowledge into our own knowledge structure. Incidentally, each teacher looks for these mappings (the arrows in Figure 2). Hammond and Allison (1987) make a distinction between mapping abstraction of primary metaphor entities and mapping invocation, of secondary metaphor entities. Primary metaphor entities are activated initially in an all-or-none fashion. The secondary metaphor entities are generated later, when a lack of knowledge is perceived. To avoid unnecessary problems Waterworth (1992) recommends to omit secondary metaphor entities.

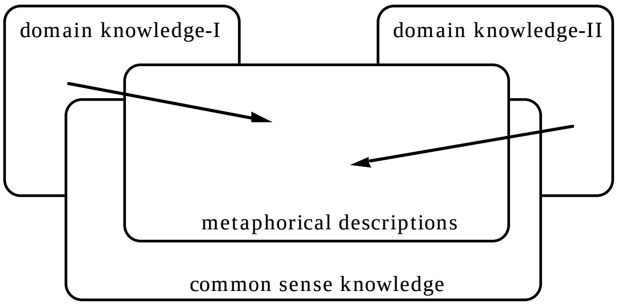

Figure 2. Communication of domain specific knowledge through mapping to a metaphorical description as part of the common sense knowledge.

'In the knowledge lies the power' has become a popular maxim in the context of artificial intelligence. 'In the metaphor lies the power' will become a popular maxim in the context of hypertext and multimedia system design. But, how can we discover suitable metaphors? The conventional way is to choose a metaphor from the set of known metaphors. Another design strategy is to develop a metaphor and look how it works (Carroll 1988). Only a domain expert himself, however, can correctly build all necessary mappings. He is the only one who has full control over his domain knowledge as well as the appropriate metaphorical descriptions of the common sense knowledge. This is the core of our approach: to look for a metaphor engineering method in the framework of a participatory design approach.

To optimise knowledge engineering methods, the following observation was quite interesting: if the knowledge engineer is a woman, then male experts do explain considerably more details compared to the situation where the knowledge engineer is a man. Initiated by this idea we looked for a social context in which a metaphorical language is normally used: this situation is given if adults explain something to children.

First we have to proof the 'children' idea to invent a 'metaphor engineering' method. Second, if the 'children' idea was successful, then we have to develop a practical metaphor engineering method.

2 Validation Procedure

To proof and validate our 'children' idea as a basis for a metaphor engineering method we carried out an empirical investigation. Domain experts explained parts of their domain knowledge to both adults and children. We recorded all explanations of the domain experts on video and analysed this material in terms of the amount of metaphorical descriptions.

2.1 Subjects and Experimental Setting

A total of 6 male students of physics participated as domain experts in this experiment. Twelve children were schoolboys (N=6) and schoolgirls (N=6) of a primary school in Basel (N=8 with age among 8 and 13 years, N=4 with age among 12 and 13 years). Twelve adults (male N=7, female N=5) were students of computer science or physics (N= 9, age 20-25 years) or people with an other educational background (N=3, age 40-50 years).

Each expert had about 10 min explanation time for two different domains. To control sequential effects each expert has to explain both domain problems twice to 2 children and 2 adults in two different orders ([D1 –> D2], [D2 –> D1]), so that each expert explained the two domains to 4 different 'metaphor engineers' (see Table 1). A total of about 40 min explanation time per expert was recorded on video.

2.2 Target Domains and Procedure

The domain experts were instructed to explain the following two problems (domains): (D1) "Why does the sound of a car change, if the car passes by?" (Doppler effect); (D2) "What happens in a bulb, if someone closes the circuit?" (Light generation).

We used a 3-factorial test design with the following three independent variables. Factor A is 'naivety': the presumed knowledge of the 'metaphor engineers' (child, adult). Factor B is 'sex': the sex of the 'metaphor engineers' (male, female). Factor C is 'domain': two domains "Doppler effect" and "light generation." The factor 'sex' was approximately balanced. The factor 'naivety' and 'order' were completely balanced.

To analyse the video sequences we derived categories from the literature (Haverkamp 1983) (Lakoff 1980) (Lieb 1967) (Richards 1936) (Wiedemann 1986). We used 3 fixed categories 'onomatopoeia', 'gesticulation', 'drawing', and open metaphorical categories, which were protocolled during analysing the videos.

A metaphor is defined as follows (Lakoff 1980) (Lieb 1967): (1) If we interpret a word or a syntactical structure of words of an explanation in their ordinary, context free sense ("common sense"), then the meaning of this part of the explanation is senseless or impossible. We call this first condition the "internal incompatibility." (2) If we relate the context free interpretation of a part of an explanation to the context of the whole explanation, then the meaning of this relation is senseless or impossible. We call this second condition the "external incompatibility." We classified parts of an explanation as metaphors, if we found an internal or an external incompatibility.

We discuss one example of a metaphor classification. One domain expert used the metaphorical description of "balls" for atoms to explain domain D2. He said: "All balls strike together." We have no internal incompatibility, but an external incompatibility in the context of electricity in a bulb. Metaphors with an internal incompatibility are often used in poems (e.g., "the green moon sang a sad song"). Sometimes we use this type of metaphor to describe special aspects of user interfaces (e.g., "the interface looks like a transparent house").

2.3 Dependent Variables and Results

Two trained raters (students of computer science at the ETH) analysed 48 video sequences - one for each experimental condition- in 15 seconds' intervals. In this paper we present the results of three dependent variables.

(1) The 'percental ratio of metaphors overall' used in each explanation trial is the sum of all 15 seconds' intervals, which included the usage of any type of metaphor. This measure is calculated over all metaphorical categories (fixed and open) corrected by the individual duration of explanation time.

(2) The 'percental ratio of different metaphors' is the sum of all 15 sec. intervals, which included the usage of different metaphors.

(3) The 'percental ratio of repetitions' of metaphors is the difference between the 'percental ratio of metaphors overall' and the 'percental ratio of different metaphors'.

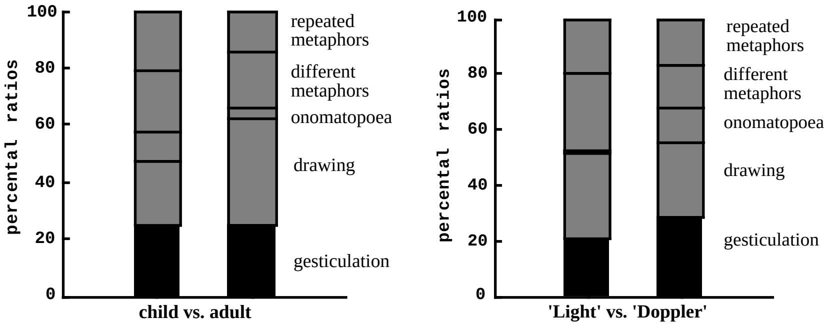

Figure 3. Percental ratios of the behavioural categories differentiated for the factor 'naivety' and the factor 'domain.'

We analysed the data with a 3-factorial analysis of variances. To give an overview of our results we present the percental ratios of our categories, which describe the most interesting aspects of the experts' behaviour. The portions of each category in Figure 3 give only the relative relations among the depicted categories and not the absolute values among all categories.

2.3.1 Percental Ratio of Metaphors Overall

The percental ratio of metaphors overall is an absolute measure to describe how much time each domain expert explains the problem to the 'metaphor engineer' with metaphorical terms. The main effect 'naivety' is significant ($p \leq .016$, see Table 2). We can observe a moderate difference between both domain problems ($p \leq .067$, see Table 2). The six experts tried to illustrate the domain "Doppler effect" using metaphorical terms on average 35.7% of their explanation time. To illustrate the domain "light generation" they used metaphors in 28.3% of their explanation time.

The domain experts used significantly more metaphors explaining to children (mean 37.8% of explanation time, N=24, see Table 2) compared to talking to adults (26.3%, N=24, see Table 2). We did not find a significant effect of the sex of the 'metaphor engineers,' but a significant interaction between the factors 'naivety' and 'sex' ($p \leq .017$, see Table 2). Basically, the experts used the maximum of metaphorical terms talking to boys and the minimum of metaphors talking to men. The domain experts differed in their use of a metaphorical language. So, we got a significant difference between the 'metaphorical' language of the 6 experts ($p \leq .001$; result of an additional analysis).

Source	df	F-test	prob
A naivety	1	6.226	.017
B sex	1	2.337	.134
C domain	1	3.539	.067
A x B	1	6.176	.017
A x C	1	1.288	.263
B x C	1	.015	.902
A x B x C	1	.213	.647

	male (%)	female(%)	total	N
child	39.7±13.7	35.8±12.7	37.8	24
adult	19.6±14.9	35.7±14.5	26.3	24
total	28.9±17.4	35.813.2	32.0	
N	26	22		48
(group means of metaphors overall)				

Table 2. ANOVA results of the dependent variable 'percental ratio of metaphors overall' with factors 'naivety', 'sex', and 'domain', and the absolute means for the factors 'naivety' and 'sex.' [Cell mean ± standard deviation]

2.3.2 Percental Ratio of Different Metaphors

The experts used significantly more different metaphors talking to a female (19.6%, see Table 3) than talking to a male person (14.7%, $p \leq .014$, see Table 3). The domain "light generation" evoked on average 20.7% metaphors, the domain "Doppler effect" only 13.3% ($p \leq .009$, see Table 3). We see a clear minimum of different metaphors in the cell 'male-adult' (significant interaction 'naivety' and 'sex', $p \leq .001$, see Table 3).

The experts differ in explaining both domain problems. Two experts (33%) prefer significantly the female 'metaphor engineers.' Also, to illustrate the 'light' problem the experts use significantly more different metaphorical categories than to illustrate the 'Doppler' problem. To demonstrate the Doppler effect most of the experts used onomatopoeia (Light generation: 1%, Doppler effect: 11%, $p \leq .001$, see Figure 3).

Source	df	F-test	prob
A naivety	1	2.726	.107
B sex	1	6.667	.014
C domain	1	16.928	.001
A x B	1	7.596	.009
A x C	1	.615	.438
B x C	1	.502	.483
A x B x C	1	.037	.847

	male (%)	female(%)	total	N
child	19.1± 5.6	18.7± 6.7	18.9	24
adult	11.0± 8.0	20.8± 8.4	15.0	24
total	14.7± 8.0	19.6± 7.4	17.0	
N	26	22		48
(group means of different metaphors)				

Table 3. ANOVA results of the dependent variable 'percental ratio of different metaphors' with factors 'naivety', 'sex', and 'domain', and the absolute means for the factors 'naivety' and 'sex.' [Cell mean ± standard deviation]

2.3.3 Percental Ratio of Repeated Metaphors

The experts used significantly more repetitions talking to a child (18.8%) than talking to an adult (11.3%, p≤ .020, see Table 4). We see a clear minimum of repeated metaphors in the cell 'male-adult' (8.6%, no significant interaction 'naivety' and 'sex', p≤ .095, see Table 4).

Source	df	F-test	prob
A naivety	1	5.884	.020
B sex	1	.241	.626
C domain	1	.001	.993
A x B	1	2.928	.095
A x C	1	1.171	.286
B x C	1	.076	.784
A x B x C	1	.271	.606

	male (%)	female(%)	total	N
child	20.6±11.6	17.1± 8.7	18.8	24
adult	8.6± 9.5	15.0± 8.5	11.3	24
total	14.2±12.0	16.2± 8.5	15.1	
N	26	22		48
(group means of repeated metaphors)				

Table 4. ANOVA results of the dependent variable 'percental ratio of repeated metaphors' with factors 'naivety', 'sex', and 'domain', and the absolute means for the factors 'naivety' and 'sex.' [Cell mean ± standard deviation]

2.3.4 Discussion

Our main hypothesis was confirmed (see Table 2). If the domain experts explain their domain problems to children, then we can observe that they repeat their introduced metaphors (see Table 4) and do not generate new metaphors within the context of a metaphorical category (see Table 3). In many cases the domain experts prefer to change the metaphorical category (see Table 3). So, children are suitably to play the role of a 'metaphor engineer', of course not exactly the same role as a 'knowledge engineer'; they function more like a 'catalyst' in a metaphor extraction session.

We can conclude from our results, that male adults are definitely not appropriate to play the role of a 'metaphor engineer' (see Table 3). The most appropriate persons are children or female adults (see Table 3 and Table 4).

What else do we need further on for a metaphor generation method? We have the verification that male domain experts use a more metaphorical language explaining something to children or women than talking to other men.

The final and till now unsolved step is to develop a guide, which describes the whole procedure: (1) criteria for selection of the domain expert, (2) criteria for selection of the 'metaphor engineer', (3) requirements for the interview session, (4) criteria for the analysis phase, and (5) a set of rules to convert narrative metaphors into design metaphors.

3 Multimedia Interface Design: an Example

What kinds of metaphors are generated by the domain experts? How can we transform these metaphors to multimedia design? We try to give preliminary answers to these questions in this section.

First, we present three examples, and then we discuss the transformation of these examples to multimedia design. One domain expert used the following description to explain the Doppler effect (D1): "Imagine you stand at a highway and you hear a car passing you: Iiiioouum!" The expert introduced the context 'highway with cars' and used an onomatopoeia to demonstrate the Doppler effect.

Figure 4. An interface design for the 'light generation' domain based on the 'people – road' metaphor.

Another domain expert described the light generation (D2) with the following words: "Imagine the circuit as a river or a brook. A power plant pumps water on a high level. At your home the water falls down and drives a water wheel." Another expert explained the same domain with the following metaphor: "Imagine a demonstration of people on a large road in a city. The head reaches a narrow lane. To avoid jams in the narrow lane the people have to hurry up. They slow down again when they reach a large road."

As one can see from these three examples, the transformations to the design of multimedia interfaces appear to be quite simple and straight forward. The first example leads us to the implementation of a sound track. The second example shows a more complex metaphor. We can transform water, power plant, house, and water wheel directly into a dynamic animated environment. With this metaphorical description, however, we can not explain the effect of resistance. From the third metaphorical description we can derive immediately an appropriate extension for the last metaphor. Consequently, we better change completely the design from passive 'water' to active 'people' (see Figure 4).

4 Conclusion

Let us summarise the main results of our study. We were interested in finding a method for metaphor generation, which is practicable in a participatory design model (Rauterberg 1992). According to our results, children and female adults are suitably to play the role of a 'catalyst' in a metaphor engineering session. We presented the following two practicable criteria, which are helpful to detect metaphors in a narrative interview (Lieb 1967): (1) If we interpret a word or a syntactical structure of words of an explanation in their ordinary, context free sense ("common sense"), then the meaning of this part of the explanation is senseless or impossible. (2) If we relate the context free interpretation of a part of an explanation to the context of the whole explanation, then the meaning of this relation is senseless or impossible. Finally, we illustrated the way from metaphor extraction to multimedia interface design. In this paper we presented the epistemological basis for this approach.

Metaphors are powerful but not sufficient to come up with a good interface. Single, over-detailed metaphors can be too restrictive and unwieldy (Waterworth 1992). Carroll (1988) is right that "effective interface metaphors often evolve over time through design, evaluation and redesign." We could show that the approach presented in this paper leads to a promising starting point of the cyclical design process.

5 References

(Carroll 1982) Carroll J. M. & Thomas J. C. (1982) Metaphor and the cognitive representation of computing systems. IEEE Transactions on Systems, Man, and Cybernetics SMC-12(2):107-115.

(Carroll 1985) Carroll J. M. & Mack R. L. (1985) Metaphor, computing systems, and active learning. International Journal of Man-Machine Studies 22:39-57.

(Carroll 1988) Carroll J. M., Mack R. L. & Kellog W. A. (1988) Interface metaphors and user interface design. In Handbook of Human-Computer Interaction. (M. Helander, ed., pp. 67-85) Amsterdam: North-Holland, 1988.

(Chauvet 1991) Chauvet J-M. (1991) Graphical user interfaces: metaphors we compute by. In Human Aspects in Computing: Design and Use of Interactive Systems and Work with Terminals. (H-J. Bullinger, ed., pp. 308-313) Amsterdam: Elsevier, 1991.

(Eberleh 1991) Eberleh E. (1991) Browsing cognitive task spaces instead of working on the desktop: an alternative metaphor. In Human Aspects in Computing: Design and Use of Interactive Systems and Work with Terminals. (H-J. Bullinger, ed., pp. 419-423) Amsterdam: Elsevier, 1991.

(Gould 1990) Gould M. D. & McGranaghan M. (1990) Metaphor in geographic information systems. In Proceedings of the 4th International Symposium on Spatial Data Handling. (K. Brassel & H. Kishimoto, eds., pp. 433-442) Columbus: Ohio State University, 1990.

(Hammond 1987) Hammond N. & Allison L. (1987) The Travel Metaphor as Design Principle and Training Aid for Navigating around Complex Systems. In People and Computers III. (D. Diaper & R. Winder, eds.) Cambridge: Cambridge University.

(Haverkamp 1983) Haverkamp A. (1983, ed.) Theorie der Metapher. Darmstadt: Wissenschaftliche Buchgesellschaft, 1983.

(Henderson 1986) Henderson D. A. & Card S. K. (1986) Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. ACM Transactions on Graphics 5(3):211-243.

(Hutchins 1989) Hutchins E. L. (1989) Metaphors for Interface Design. In The Structure of Multimodal Dialogue. (D.G. Bouwhis, M.M. Taylor & F. Neel, eds.) Amsterdam: North Holland.

(Lakoff 1980) Lakoff G. & Johnson M. (1980) Metaphors we live by. Chicago: The University Press, 1980.

(Lieb 1967) Lieb H-H. (1967) Was bezeichnet der herkömmliche Begriff Metapher? Muttersprache 77:43-52.

(Rauterberg 1992) Rauterberg M. (1992) An iterativ-cyclic software process model. In Proceedings of the 4th International Conference on Software Engineering and Knowledge Engineering. (pp. 600-607) Los Alamitos: IEEE Computer Society Press.

(Richards 1936) Richards I. A. (1936) The Philosophy of Rhetoric. Oxford: Oxford University Press, 1936.

(Shum 1990) Shum S. (1990) Real and virtual spaces: mapping from spatial cognition to hypertext. Hypermedia 2(2):133-158.

(Waterworth 1992) Waterworth J. A. (1992) Multimedia Interaction with Computers - Human Factors Issues. New York: Ellis Horwood.

(Weyer 1985) Weyer S. A. & Borning J. (1985) A Prototype Electronic Encyclopedia. ACM Transactions on Office Information Systems 3(1):63-88.

(Wiedemann 1986) Wiedemann P. M. (1986) Erzählte Wirklichkeit – zur Theorie und Auswertung narrativer Interviews. Weinheim München: Psychologie Verlags Union, 1986.

in:

W. Schuler, J. Hannemann, N. Streitz (Eds.)
Designing User Interfaces for Hypermedia

Springer-Verlag Berlin Heidelberg New York
(ISBN 3-540-58489-7)