Audio-visual Rhetoric:
Visualizing the Pattern Language of Film

Gesche Joost, Deutsche Telekom Laboratories, Berlin, Germany

Sandra Buchmüller, Deutsche Telekom Laboratories, Berlin, Germany

Roman Englert, DTAG Laboratories at Ben Gurion University, Beer Sheva, Israel

Abstract

Audio-visual Rhetoric is a knowledge domain for designers in theory and practice that is valid for all communicative actions through media that aim for persuasion. Within this domain, we introduce a framework for media analysis. We developed an Audio-Visual Pattern (AVP) language for film that is visualized within a notation system. This system shows auditory and visual parameters in order to reveal film's rhetorical structure. We discuss related theories from pattern language and rhetoric and apply the AVP method to analyze 10 commercials.

Keywords

Pattern Language, Film Analysis, Rhetoric, Emotion, Persuasion, Design Research

In different disciplines like software development and interface design, the concept of pattern languages is used in order to structure recurring design patterns. As one of the first, we find this concept in Christopher Alexander`s work on urban planning and architecture (1). In the 1970's, he defined a specific pattern language for buildings and environments that was based on the idea that every part of architecture can be defined as a solution for a design problem. One of those design problems could be: how can one design the entrance of a building so that people who enter are welcomed, given orientation, and shelter? According to Alexander, patterns are archetypal solutions to common design problems that aim to satisfy human needs in a certain context (1). Gui Bonsiepe introduced this concept to the design domain by investigating the nature of visual / verbal patterns in print advertisement (3). In the last years, he presented further studies on audio-visual rhetoric where he referred to audio-visual and interactive media as rhetorical structures (4)

In the field of Human Computer Interaction (HCI) Design, a pattern was defined more general as an abstraction from a concrete form which keeps recurring in specific non-arbitrary contexts (9). This definition is also valid for Jennifer Tidwell's (18) pattern language for interface design where she systematizes interface elements according to their function in HCI. Furthermore, Jan Borchers refers to Christopher Alexander when he describes a pattern language for interaction design where he adapts Alexander's

formal structure and terminology to define a pattern and its links to sub-patterns (5).

In this paper, we state that the concept of pattern languages can be defined as a new form of rhetorical figures and therefore, this concept refers to the ancient communication theory, rhetoric. The main difference between the two concepts is that the link from today's concepts of pattern language to rhetoric and the persuasive purpose got lost and therefore, it is a broken connection to the history of communications theory (3). We foster a stronger relation to rhetorical scholarship with its specific emotions-theory. Aim of all rhetorical actions is persuading an audience and addressing it on an emotional level. This plays a major role in all design activities (6). Therefore, audio-visual rhetoric with its pattern language is a key concept for media design and a core competence for designers. Hence, it is an objective for design theory and research today to elaborate on rhetorical knowledge for designers in theory and practice.

We suggest that the application of theoretical principles of rhetoric to audio-visual products can proceed in two steps: (a) to transfer the rhetorical framework from language to media analysis and (b) to visualize the pattern structure of audio-visual media within a notation system. Herewith, we re-establish the missing link of the pattern language approach to rhetorical scholarship, and focus on its potential to produce persuasion.

Our aim is to provide a method that eases media design and its analysis by referring to the set of AVPs. This kind of analysis furthermore allows an estimation of the potential level of emotional arousal because rhetorical scholarship links certain pattern structures to three different levels of arousal: *logos* being the lowest level of arousal using rational arguments for persuasion; *ethos* being the medium level of arousal based on entertainment and moderate positive values (pleasantness); and *pathos* being the highest level of arousal activating either very negative emotional responses or very positive ones (very unpleasant or very pleasant) (2).

Research Questions

In this paper, we address the following research questions:

- Can the concept of rhetorical patterns be transferred from rhetoric to design and its visual and auditory signs?

- Can an Audio-Visual Pattern language be introduced as a framework for media analysis and what is the benefit of such a framework?

- Does the visualization of the rhetorical structure of media enhance the cognitive metabolism of information?

The Theory of AVP

Rhetorical sources

Rhetoric is the scientific discipline that lies at the basis of this work. It is the theoretical background that provides systematic knowledge about communication patterns. The basic communication-principle of rhetoric is to address an audience in an appropriate way, using a medium in order to

persuade them about a specific message. Therefore, emotional arousal is used to foster persuasive power. Over the last 2500 years, rhetorical scholarship has developed a large body of knowledge about effective communication patterns for this purpose, rhetorical patterns being one example.

The rhetorical communication technique (*ars*) is traditionally divided into two parts: *rhetorica utens,* the applied part of scholarship and *rhetorica docens,* the theoretical and didactic part of it. Hence, there is an inherent link between both parts: in ancient Greek and Roman rhetoric scholars were taught to analyze best practice examples of oral talks in order to learn from practice. They had to investigate the composition of exemplary talks, learn about the topics and style that was used, and finally had to practice the oral presentation themselves. Therefore, the analysis of composition and structure of applied rhetoric as well as its oral presentation has always been part of the *ars rhetorica.* It makes sure that the theoretical system of rhetorical communication *(rhetorica docens)* is always updated according to its practical application and enhancements *(rhetorica utens),* and vice versa.

In this paper, the idea of linking the theoretical concept to its practical application is transferred to the design domain. The analysis of the rhetorical structure of audio-visual media should be part of all design education because scholars can learn from investigating best practice examples. Gui Bonsiepe (4) states that the analytical as well as practical skills of audio-visual rhetoric are a core competence for designers of the information age as they have to act in interdisciplinary and highly specified media contexts. This makes audio-visual rhetoric a skill to foster cognitive metabolism of complex media data.

Audio-Visual Patterns (AVPs)

The AVPs are structured in a pattern library for audio-visual media (see example AVPs in Table 2, for an extended list see (11)). It serves as a first systematic collection bringing together different sources from film, design and rhetorical scholarship, whereas it is also the basis for the notation protocols of this study. The patterns are assigned to different potential of emotional arousal: either to lower emotional impact (*logos* to *ethos*) or to higher emotional impact (*ethos* to *pathos*). Assigning patterns to a specific level of emotional arousal cannot be easily done in a distinct way – the assignment suggested in this paper is a proposal that is based on experience and reflects on criteria for rhetorical style. Nevertheless, single patterns can also appear in other emotional styles than they are commonly linked to – e.g., one can also find a metaphor in a commercial that overall uses low emotional involvement. This does not contradict the whole system at all, because patterns can be combined in a creative way and have to support the overall style of communication. From perspective of Audio-Visual Rhetoric, the texture has to be analyzed as a whole – regarding the quantity and quality of patterns as well as the composition – in order to reveal the level of emotional arousal it should evoke.

In rhetorical scholarship, the table of patterns was literally understood as a toolbox for *orators.* When planning a speech or presentation, orators of Cicero's times selected the appropriate figures from this source depending on

the topic they were talking about. For example, a topic of high pathos like peace and war, nation, or passion demanded patterns of high potential for emotional arousal like metaphor, climax, or exclamation. The doctrine of appropriateness, so called *aptum*, signifies specific patterns to be appropriate for specific communicative purposes like raising high, medium or low emotion in the audience. This body of knowledge is still valid today. Designers and film makers still search for appropriate patterns due to the communication purpose, although they usually do not reflect on the nature of patterns themselves. Today, the usage of patterns seems to be a tacit form of knowledge that is regarded as being only based on individual experience. The existing body of knowledge on this issue seems to be neglected, although we find in the different approaches to pattern languages (as indicated above) some first steps to recover it.

Selected list of AVPs	
1. Patterns with lower emotional impact (*logos* to *ethos*)	
a. Auditory visual patterns	
Analogy:	A verbally expressed issue is equivalent transferred to auditory visual signs
Parallelism:	Similarity of structure in a pair or series of elements of the auditory or visual track
Variation:	**auditory**: Melodic, harmonic or rhythmic variation of a musical motif
Variation	**visual**: Repetition of a visual motif with modification
Verbal anchor:	A visually expressed meaning is semantically specified by verbal signs
b. Auditory-visual correlations	
Congruency:	Auditory and visual sphere are corresponding (for example via illustrative, pseudo-realistic, realistic sound)
2. Patterns of higher emotional impact (*ethos* to *pathos*)	
a. Auditory visual patterns	

Antithesis:	A semantic contradiction between two issues (for example image and sound, image and text, image and music)
Emphasis:	Giving prominence to quality or trait by audiovisual sings
Hyperbole:	Exaggeration accomplished via visual or auditory signs beyond credibility, often with a pathetic effect
Climax:	An arrangement of shots in an order of increasing importance, where the single elements can be arranged in a loose order and substitutions with synonyms or tropes are optional
Metaphor:	Substitution of the "proper" image with an "improper" transfer
Metonymy:	An associated detail or associated image is used to evoke an idea or illustrate an issue
Paronomasia:	Repetition of shots or sequences with a minimal variation on the auditory or visual track
Polyptoton:	Repetition of shots or sequences with a formal variation (for example change of perspective, color, size)
Repetition:	Major rhetorical strategy for producing emphasis, clarity, amplification or emotional effect; repetition of auditory or visual structures as, for example, narrative, formal, chromatic repetitions
Synecdoche:	On the visual track a whole is represented by showing one of its parts or vice versa
b. Auditory-visual correlations	
Divergence:	Auditory and visual sphere are contradictory (for example as antithesis)
Complementary:	Auditory and visual sphere complement and boost one another (for example via metonymic or evocative sound, rhythmic parallels)

Table 1: Selection of AVPs in film

The Visual Notation System and its Iconography

The AVP Method for Film Analysis

Exploring audio-visual patterns of film or similar media such as commercials or animation is an analytical endeavour that is still in search for a useful method. In film theory and practice, there are different models to describe and analyze filmic structures. For film scholars, the most established method is using a written film protocol to elaborate on formal aspects as well as on narrative structures. This method was questioned by many scholars because of the lack of potential to reflect on the specific dynamic and audio-visual quality of film as such (15). Audio-visual rhetoric provides an approach to overcome this problem by introducing a notation system for film. In academic context, there are some examples of visual film protocols that were developed before, e.g. the notation by Dietrich Hahne (8) or the graphical representation of film by Wolfgang Ramsbott and Joachim Sauter (17). Nevertheless, none of the visual systems has been established for film analysis in a broader context. The existing models were single approaches that could not be developed to a further extend. Additionally, Hahne's iconographic system was quite complex and

not easy to apply for analytical purposes; hence, it did not have the character of a handy tool. A general issue is that there is no interdisciplinary collaboration between study of film, film production, rhetoric, and design research to come up with an applicable system. Bringing together all these competences, one could design a comprehensive system that could be used in various contexts – academic as well as applied. In this paper, we bring together at least the two disciplines mentioned last – knowing that this system has to be discussed in a broader context to be enhanced further.

Cognitive Benefits of Visual Diagrams

In the standard film protocol, the translation of audio-visual signs into a written text means to change the semiotic code in a radical way – from image and sound to text. This process involves a loss of information, particularly of the audio-visual and dynamic quality of the sign system. The idea of using a visual protocol as tool for film analysis is based on the hypothesis that a visual diagram can be processed cognitively much more efficient than language (4). This is especially true for the visual aspects of film and not for re-telling its storyline. It is not new to say that the tools and methods that are used for an investigation clearly influence the research results. This is also true for the AVP method we suggest in this paper. To set up a visual diagram of film focuses much more on visual and structural aspects than on the storyline. The aim is to reduce a loss of information that occurs when the audio-visual texture of film is transferred to written film protocol. Therefore, the audio-visual signs are translated into a visual structure. This method has additional advantages: the graphic displays information on one sight so that the recipient can process the data in parallel. Written text can communicate information only in a successive way – one word after the other. In the notation protocol, information about the whole clip in each of the channels can be visualized at the same time, allowing a parallel interpretation of data and of relationships among the various audio-visual elements. Here, the pattern structure can be easily identified on the basis of a graphical representation. For example, repetitions or climax patterns can be singled out quickly on the basis of their visual form. With this approach, large amounts of data from audio-visual media can be efficiently processed for analytical purposes. Furthermore, patterns and dynamic developments can be marked out visually, whereas critical points or changes in the succession can be displayed on one sight (see for example (13)).

The AVP Protocol

The output of the notation system is a pattern protocol that visualizes rhetorical structures in terms of audio-visual matching, emotional style (logos, ethos, pathos), climax, repetition and the like. On top of the analytical purpose, it reveals the style and structure that was built up during production process. These insights are relevant for film makers and designers in order to understand how media is created and how it affects the audience. The notation system itself consists of a set of visual icons (see Table 2, see also (12)). With this icon system, the identification of technical parameters of film is displayed in a cognitive efficient way. For example, each type of shot and each montage pattern have an iconographic representation. These icons are

part of a notation language that signifies precisely filmic phenomena without having to describe them word by word.

Icons of audio-visual correlation:	
Parallel/ contrapuntal/ cumulative	= ≠ >>
Shot icons:	
close-up	■
medium shot (closed/ open/ Over-the-Shoulder-Shot)	◖ ◖ ◖
full shot (closed/ open/ Over-the-Shoulder-Shot)	● ◖ ◖●
long shot (closed/ open)	• •
wide shot	•••
intercut	••
pan (example: from long shot to full shot)	• ●
intertitle	—
Angle-Views:	
high-angle view / bird`s-eye-view	⸱
low-angle view/ worm`s-eye-view	⸱
eye-level-view	⸱

Table 2. Extract of the set of notation icons

The icons are set within a notation field (see Figure 1). This field is defined by the axes of paradigm and syntagma that derive from the linguistic model Roman Jacobson introduced to describe language systems (10). This linguistic framework is easy to understand when we use the metaphor of a brick wall: the syntagma describes the succession of bricks in the horizontal layer, whereas the paradigm refers to the composition of bricks in the vertical layer. These bricks form pattern structures in both directions. Transferring this model to audio-visual media, we define the syntagma as the succession of different shots or expressions in time. In linguistic terms, it is called the axes of combination – combining different signs one after the other. The paradigm, however, refers to all the shots or expressions that were possible at this position but not selected. For example, the paradigm describes different camera angels that could be realized for one shot - from bird's eye view to worm's eye view. Therefore, it is the axes of selection describing the actual selected sign out of a whole range of possible ones. Christian Metz transferred this

theoretical model to film (16) in order to set a systematic framework describing general signification processes. For the notation system, these two axes form the theoretical framework that explains the pattern approach based on semiotics. It is useful in order to understand how AVPs actually create meaning and effect for the audience.
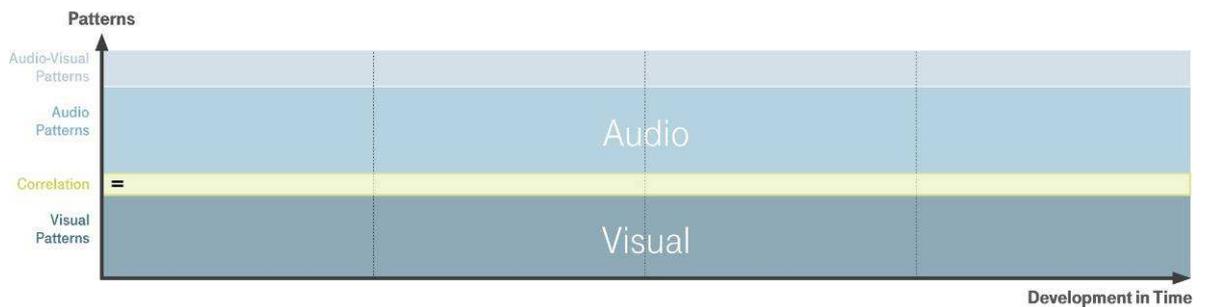

Figure 1. Notation field.

Furthermore, the notation system consists of 3 channels: the auditory, the visual and the correlation channel. In the first channel, all auditory parameters are described as being sound, dialogue, or music. The second one contains all visual parameters like single shot and montage figures, perspective, camera dynamics, colors and alike. The third channel specifies the relation between auditory and visual channel in three overall patterns: "parallel" (auditory and visual signs accord to each other), "contrapuntal" (auditory and visual signs are set in contrast to each other), or "cumulative" (auditory and visual signs intensify each other). With this correlation channel, the overall audio-visual pattern is described that gives insights into the role of auditory stimuli within the clip in relation to the visual succession. A contrapuntal AVP for example is a highly artificial element of style that is not often used in media, because it contradicts the viewer's expectation. Therefore, it is most often used to raise higher emotion in the audience by effects of surprise, confusion, or irony.

Notation Practice

We selected 10 commercials from public internet databases like Youtube. They can be accessed online at http://www.geschejoost.org/commercials/. Most of the selected clips did not have a high amount of spoken dialogue because we focused on image and sound elements. We choose those clips varying in the potential level of arousal from high to low. Furthermore, we selected one clip from the late 1980's because we observed that the communication style of commercials changed quite a lot until today: nowadays, you hardly find any commercial persuading with rational argumentation (logos), whereas most of them have high emotional impact and entertaining style. For each clip two of the authors made a pattern protocol using the AVP method.
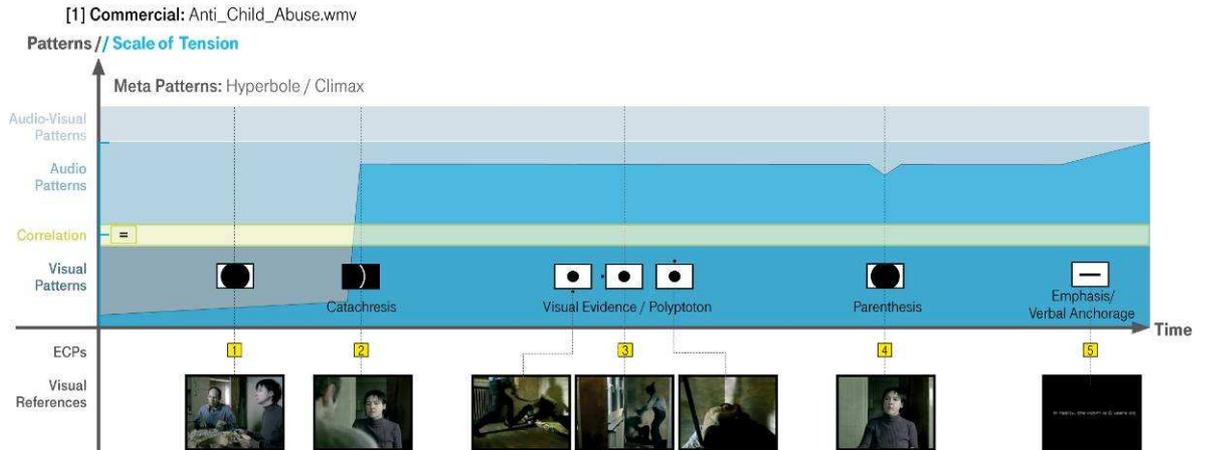
Figure 2. Notation protocol of the commercial "Anti child-abuse"

The structure and style of each of the 10 selected commercials are visualized in notation protocols (see examples Fig. 2 to Fig. 4). The development of tension is indicated by a curve whereas its devolution is depending on the pattern structure. In this respect, the vertical dimension becomes a scale of emotional intensity, while the horizontal dimension shows the development within the progressing plot. In the visual channel, significant patterns are indicated by the icons. For each clip, some special points of interest were marked, called emotional check points (ECPs). These ECPs seemed to be interesting in rhetorical respect as well as referring to the plot development because they marked specific AVPs like metonymies, emphasis or contrast. In future empirical evaluation, we want to test the viewer's emotional response at the ECPs.

To exemplify the benefit of a visual notation compared to a written film-protocol, we will give an example from one of the commercials. One of the patterns visualized in Figure 2 (commercial: Anti Child Abuse) would be described in a written film protocol as "3 times long shot with changing camera perspective, high arousal level, using the AVPs of visual evidence as well as polyptoton in the middle of the clip." It is much easier to express this complex information visually. The notation protocol indicates a heavy increase of arousal when the first AVP starts. It is a catachresis, being a harsh disruption of image, style or meaning. In this example, a typical scene showing a family having dinner together turns out to be the beginning of child abuse. Throughout the clip, the high tension is kept by using the AVPs of visual evidence and polyptoton (showing the same image from different perspectives). A parenthesis lowers the tension for a short time, being an insertion into the main plot. At the end of the clip, the arousal rises to its peak when the emphasis and verbal anchorage reveals the main message: that this violent scene on child abuse is even worse in reality.

Figure 3. Notation protocol of the commercial "Colgate toothbrush"

Figure 4. Notation protocol of the commercial "Lufthansa New York"

Discussion

The AVP method can serve developers of audio-visual media in various ways. First of all, it is a helpful tool to analyze and interpret film and understand its composition, as we have shown in 10 examples of commercials. In this respect, the only hurdle is that the set of icons of the notation system has to be learned once before actually using it. Second, using the notation system one gets a visual protocol comparable to music notations that can be used for reproduction purposes. The protocol stores information about the montage, the mise en scène, as well as the patterns that were used. Like the dance notation system made by Rudolf Laban (14), it stores this information for any kind of reproduction, which was in Laban's case another performance of dance. For film, it might not often happen that one piece will be reproduced in exactly the same way, but it is common to adapt a successful composition and pattern structure in a different context. This leads to the third point: the notation system is a tool for film design and planning in addition to the

technique of storyboarding, which is still a standard tool for the production process. With this visual aid, film makers can compose their texture beyond sketches of the scene and visual description. Using this tool, one can mark the montage patterns, the relation between sound and image, as well as the overall rhetorical development of film in a visual way before doing the actual shooting and editing. Sergej Eisenstein was one of the first film makers who used a visual language: it was for the movie *Aleksandr Nevsky* (1938), where Eisenstein had to arrange film music according to the visual part of the film (7). In this example, Eisenstein visualized the music notation composed by Sergej Prokofiev in relation to the film shots he edited – and invented the concept of "vertical montage".

But there are also weaknesses in the system as it is presented here. The current set of icons was developed for a specific kind of film – commercials. With this set it will be hard to analyze a whole movie because it is designed for a micro analysis of scenes and short sequences. In order to investigate the meta-level of a whole movie or to consider new technical features, the system has to be adapted and expanded, new icons and visual structures have to be added. The notation system should be understood as work in progress: it is a tool that the user can adapt due to his/her analytical purpose. The best way to enhance the quality of the system as well as the variety of possible application is to do a collaborative effort – to use the idea of open source development in an online community. This will be a next step for further research.

Conclusion and outlook

We introduce a new system of audio-visual patterns that visualizes the dynamic structure of media segments. As such, the system can serve as a meaningful aid for designers of media. The system is based on rhetoric theories that emphasize the persuasive goals of media messages. In the first part of the paper we described the system, its rational and notations. In the second part, we analyzed 10 commercials with this visual tool. This tool is meant to be applied in academic contexts like departments for studies of film and other media. With the list of AVPs and the 10 sample notation protocols we showed that the concept of rhetorical patterns can be transferred to audio-visual media as it provides a useful terminological framework for the analysis. The benefit of this analytical framework is that it reveals the rhetorical structure of media products a) in order to understand the product in better way, b) to use it for reproduction purpose, and c) to learn from the AVP protocols how to design successful media products by using rhetorical knowledge.

Next step in this research will be an empirical assessment to test the predictive power of this method. We will conduct experiments regarding the actual emotional response of recipients to proof whether it is similar to the system's prediction or not. Further work will be done in extending the list of AVPs for different kind of film as well as media in order to build a rhetorical knowledge base for audio-visual as well as interactive media.

References

1. Alexander, C., Ishikawa, S., Silverstein, M. (1977). *A Pattern Language. Towns, Buildings, Construction.* Oxford University Press, New York, USA.

2. Aristoteles (1989). *Rhetorik*, München, Germany.

3. Bonsiepe, G. (1961). Persuasive Communication: Towards a Visual Rhetoric. In Crosby, T. (Eds.) *Uppercase Nr. 5*, London, UK, 19–34.

4. Bonsiepe, G. (2008). Audiovisualistische Rhetorik in zeitbasierten Medien: Über die kognitive Relevanz diagrammatischer Visualisierungen. In: Joost, G., Scheuermann, A. *Design als Rhetorik*. Birkhäuser Verlag Basel. 217-232.

5. Borchers, J. (2001). *A Pattern Approach to Interaction Design.* Wiley & Sons, New York.

6. Buchanan, R. (1985): *Declaration by Design: Rhetoric, Argument, and Demonstration.* In Design Issues, Volume II, Number 1. 4-23; Reprint in: Margolin, V. (1989) (Ed.): *Design Discourse. History, Theory, Criticism*, Chicago. 91-109.

7. Eisenstein, S. (1969). *The Film Sense.* Ed. and translated by Jay Leyda, New York.

8. Hahne, D. (1992). *Komposition und Film: Projekt nach Motiven aus Camus „Der Abtrünnige" für Chor, Orchester und Spielfilm.* (Volkwang-Texte III. Bd. 5. Hrsg. v. Josef Fellsches).

9. Heidmann, F. (2007) Corporate Interaction: Interfacedesign als Teil der Corporate Identity. In Beyrow, M., Kiedaisch, P. (Eds.) *Corporate Identity und Corporate Design. Neues Kompendium.* av edition, Ludwigsburg, Germany, 126-135.

10. Jacobson, R., Halle, M. (1956) *Fundamentals of Language.* Den Haag, Netherlands.

11. Joost, G. (2008): Die rhetorische Pattern-Language des Films. In Joost, G., Scheuermann, A. (Eds.) *Design als Rhetorik*, Birkhäuser Verlag Basel, 233 – 249.

12. Joost, G. (2008) *Film-Sprache. Die audio-visuelle Rhetorik des Films.* Bielefeld.

13. http://www.geschejoost.org/AVRhetorik/

14. Laban, R. (1995). *Kinetografie – Labanotation: Einführung in die Grundbegriffe der Bewegungs- und Tanzschrift,* Ed. by Claude Perrottet, Wilhelmshaven.

15. Lehner, C. (1987). Einige zentrale Probleme der neueren Filmsemiotik. In Bauer, L., Ledig, E., Schaudig, M. (Eds.) *Diskurs film: Strategien der Filmanalyse.* Bd. 1 , München , Germany, pp. 59–72.

16. Metz, C. (1972). Probleme der Denotation im Spielfilm. In Metz, C. *Semiologie des Films*, München, Germany, pp. 154–161.

17. Ramsbott, W., Sauter, J. (1988). Visualisierung von Filmstrukturen mit rechnergestützen Mitteln. In: Helmut Korte, Werner Faulstich (Hrsg.):

Filmanalyse interdisziplinär. (Zeitschrift für Literaturwissenschaft und Linguistik. Hrsg. v. Helmut Kreuzer. Beiheft 15). Göttingen. pp. 156–165.

18. Tidwell, J. (2005). *Designing Interfaces*, O'Reilly.

Gesche Joost

Gesche Joost is the head of the Design Research Lab at Deutsche Telekom Laboratories in Berlin, Germany. Her current research topics are audio-visual rhetoric of media, interface and interaction design as well as gender aspects in design. She holds a PHD in Rhetoric and a Masters in Design. In 2007/2008, she was a visiting professor at the University of Applied Sciences in Hildesheim, Germany. In 2006, she has been elected one of the „100 masterminds of tomorrow" by the jury of the campaign "Germany - land of the ideas". This jury selected young people from Germany who will shape the future of their country.

Sandra Buchmüller

Sandra Buchmüller passed her exam at the Köln International School of Design in 2001. Then she has worked as a design researcher in collaboration with Prof Dr. Uta Brandes in the area of gender and design, as free designer for the "Entwicklungs-Gesellschaft Zollverein" at the World Heritage Site Zollverein, as well as User Experience Designer for Vodafone, T-Systems and T-Labs. Scope of her work is gendered aspects in design processes and artifacts, usability specifications and optimizations, as well as information and interaction design concepts.