Sonification as an Aid to Listening in Synthetic Speech

Robert E. Isaacson, Ph.D.
San Jose State University, One Washington Square, San Jose, CA
Ron Aust , Ph.D.
University of Kansas, 3061 Dole Building, Lawrence, KS
Brian Newberry, Ph.D.
California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA

Abstract: This paper presents the findings of a descriptive study that examined five dimensions of user satisfaction with the use of sonification in synthetic voice messages. Sonification is the use of non-speech audio cues to communicate information. Seventy-six participants responded to a seven point bi-polar semantic differential comparison rating scale questionnaire that asked them to compare messages without sonification to the same messages with sonification in terms of recruiting attention, aiding listening, helping to remember details, aesthetic appeal, and overall valuation. Overall mean response ratings for messages with sonification were all higher than messages without sonification across all five dimensions. The results of this study indicate that listeners assign higher valuation to synthetic voice messages with sonification. Future applications of sonification include using sound cues as a means to improve synthetic natural-language speech intelligibility and augment other sensory modalities for human-machine interface, data processing, instructional media, and to assist user populations with disabilities.

Introduction

The use of voice enabled user interfaces to access information has increased at an exponential rate during the past decade (Cohen, Giangola, Balough, 2004). Many voice user interfaces (VUIs) use synthetic speech technology to support communications interaction and provide users with access to information via the telephone, computer, and an increasingly vast array of portable electronic devices. This paper presents the findings of a descriptive study that examined five dimensions of user satisfaction with the use of sonification in synthetic voice messages. Participants were asked to compare messages without sonification to the same messages with sonification in terms of recruiting attention, aiding listening, helping to remember details, aesthetic appeal, and overall valuation. The purpose of the study was to explore whether listeners assign higher valuation to synthetic voice messages with sonification. The use of sonification and systematic sound design techniques to enhance synthetic voice messages is an area that has not been fully researched. Application development and design research is needed to create a systematic method for user-centered sound design for instructional media, software applications, and voice messaging systems. Human capabilities and limitations as well as other factors such as acoustics and psychoacoustics must be taken into consideration in order to improve human-machine systems. Future applications of sonification include using sound cues as a means to improve synthetic natural-language speech intelligibility and augment other sensory modalities for human-machine interface, data processing, instructional media, voice messaging systems, and to assist user populations with disabilities. Systematic sound design techniques include the use of audio signal processing effects and non-speech audio cues to convey information or facilitate communication or interpretation (Kramer and Santa Fe Institute (Santa Fe N.M.) 1994). Kramer (Kramer 1994) notes that other terms are often associated with sonification such as auralization, audiolization, acoustic display, virtual audio worlds, and virtual acoustic display. While there are many terms in use there is no common agreement or consensus over the use of the term sonification, so for purposes of this study in will be used to refer to any non-speech audio used to communicate or convey information.

Bishop and Cates (2001) among others (Marsh 1983; Spencer 1991; Tannenbaum 1998), have suggested that the theories from the parent disciplines/fields of cognitive psychology, information-processing and communications can be combined to provide the theoretical foundations for sounds systematic use in audio messages and instructional media. Information-processing theory is important because it provides a model for understanding the cognitive and perceptual processes involved in processing instructional messages. Communication theory is relevant because it provides a model for how to create more effective instructional messages. Patrick Marsh's (1983) theory of message design, provides a framework for the use of sonification in the design of instructional messages and audio messages in particular. Marsh suggests an approach to message design that is based upon classical rhetoric, information

theory, and human information processing. His model provides techniques for instructional designers to use that optimize the information load and message complexity dimensions of audio messages. Marsh (1983) observed that there is a need to add to our knowledge about analyzing audio messages so that they are more compatible with human perceptual and cognitive abilities. His theory provides two propositions that are useful constructs for this study. The first proposition is that cognitive load increases as message density increases. The second is that affection for instructional audio messages increases as diversity increases. This is possibly due to the fact that diversity in message design helps to reduce the negative effects of auditory habituation and saturation.

In a literature survey by Vroemen (1986), research was reviewed with respect to the design of color screen computer displays. He describes the following functions of color that can also be extended to the use of sonification:

- **Alerting:** Provide anticipatory sound cues that signal changes in message state.
- **Coding:** Create sound cues that are paired with key points that were stated or listed in a message.
- **Structuring:** Map sound cues that are equivalent to organizational and typographic features as a way to clarify the structure of the message.
- **Emphasizing:** Sound cues can be used to highlight message elements.
- **Ambience/Aesthetics:** Provide background sound in the form of ambient musical motifs or environmental sounds to make the message more aesthetically appealing. In addition, ambient environmental sounds can provide contextual clues that aid message interpretation.

The literature relating to sound design deals primarily with film and theater productions. Sound design is the process of creating the sonic character of a production (Altman 1992; Chion, Gorbman et al. 1994; Sonnenschein 2001; Alten 2002). Research related to the application of sound design techniques to instructional media is scarce. Research in the area of applied sound design for film and theater is also scarce; however, there has emerged over time a very sophisticated approach to the systematic use of sound.

The ability of sound to enhance visual elements in film is well established. Sound is a powerful link in the perceptual cycle (Neisser 1967) that has the ability to convey contextual cues which establish the emotional feeling and context within the physical environment thus reinforcing the visual image. Sound design has three primary domains which include speech, sound effects, and music (Chion, Gorbman et al. 1994; Alten 2002). These aural elements are powerful tools that can be used to convey both cognitive and affective information (Cook 1999).

A more systematic and structured approach to the use of sound and auditory messages in the computer interface called “earcons” was explored by Blattner (1990). Earcons are constructed from a sequence of tonal cues referred to as motives, rather than real world sounds as used by Gaver (1986) to create auditory icons. The advantage of using of tonal motives to form earcons is that the basic properties of sound such as pitch, timbre, rhythm, volume, can be precisely controlled, manipulated, and formed into more complex structures to convey information. Earcons are similar to Morse code in that unlike everyday sounds which have a known associated meaning, they must be learned.

Following is a partial listing of categories of sound cue affordances that form a basis for a systematic approach to user centered sound design. It provides examples of some of the possible ways the sonification and the use of sound cues can be conceptualized and developed.

• **Alerts:** Used to recruit attention.

• **Aesthetic Emotive:** The use of musical motifs or background music is similar to using graphic design elements to provide aesthetic appeal.

• **Ambient:** Designed to provide background environmental cues that provide contextual cues that can aid interpretation.

• **Typographic:** Designed to provide organizational, semantic, and syntactic affordances of document text structure and content.

• **Mnemonic Aids:** Designed to be used as a mnemonic device to aid storage of information into both short and long-term memory.

• **Parsers:** Designed to be used as spacers within an audio message in order to reduce message density.

• **Proximity:** Designed to provide information on distance to and from various persons, places, or things in the acoustic environment. Intended to provide feedback on "closeness" or distance too objects. Sonar detection is a good example of this.

• **Location:** Designed to be used as position reference acoustic space.

• **Temporal:** Designed to be indicative of location in time or position within an instructional piece, and sequence of events.

• **Intensity of force:** Energy in terms of volume level conveys level and intensity of force (loud volume sounds indicate greater energy or force).

• **Spatialization:** Enhanced audio signals that provide spatial cues that aid the listener to distinguish location and position of figure-background contrasts (reverberation, echo, phase shifting, pitch shifting, etc.).

The Study

The concept of user satisfaction is based upon subjective responses of the participants in the study. User satisfaction is a dimension that is frequently used in usability testing (Chin 1988; Bickford 1997; Shneiderman 1998). The effectiveness of an audio message will increase according to the degree to which it supports the tasks required of listeners and their information processing abilities (Marsh 1983). A more specific operational definition of user satisfaction can be stated in terms of the attitudes that users express towards a particular object or class of objects, i.e. elements of a human-machine interface system. Sarnoff (1960) defines an attitude as "a disposition to react favorably or unfavorably to an object or class of objects." The object or class of objects under consideration in this study was sonification in audio messages. User satisfaction was defined as the valuation listeners place on audio messages with sonification in comparison to audio messages without sonification.

Participants in the study included 76 graduate students enrolled in School of Education classes at the University of Kansas and California State University, San Bernardino. The participants then took part in the study in two parts. The first component of each instrument requested background demographic information. During the second component, a series of PowerPoint slides were presented to participants with a total of sixteen audio messages presented with and without sonification.. Participants were then asked to respond to the following series of five seven point bi-polar comparison rating scale items related to satisfaction with sonification in audio messages:

- In terms of gaining your attention, the message with sonification was...
- In terms of helping you listen, the message with sonification was...
- In terms of helping you remember details, the message with sonification was...
- In terms of aesthetic appeal, the message with sonification was...
- Overall, the message with sonification was...

Participants were also asked to provide general comments regarding their overall reactions to the use of sonification. Table 1 shows the seven point bi-polar comparison rating scale used to quantify and interpret participants' responses.

Table 1. Bi-polar Comparison Rating Scale

Much worse	Worse	Slightly worse	The same	Slightly better	Better	Much better
1	2	3	4	5	6	7

The sound cues that were used in the audio messages were a combination of auditory icons (real world sounds) and earcons (musical tones). Multiple variations of sound cues were used with each administration of the user satisfaction survey questionnaire and asked to provide general comments on the mappings of sound cues to the representation of various structural elements of the audio messages. A series

of sound cues both natural and synthetic were created and matched to various elements of a series of eight audio messages recorded with TextAloud MP3 text to speech software and ATT Naturally Speaking synthetic voices. The cues used included sound effects, synthesized tones, and short musical motifs cues intended to function both as perceptual enhancers and aids to auditory information processing. The messages used were representative of the types of messages that are found in various automated voice messaging systems that use synthetic speech. Following is a sample of the spotting sheet text transcript of the first of eight short audio messages which were converted to synthetic speech with indications of the type of sound cues which were used and where they were placed within the message.

Time Message Item
(3 tone framing cue - ascending pitch)
SYNTHETIC VOICE (female): The current local time is, 7:51 and, 5 seconds, on the morning of, Friday, September 1. Please say the name of another city, country, or state.
(3 tone framing cue - descending pitch)
USER: What time is it in France?
(3 tone framing cue - ascending pitch)
SYNTHETIC VOICE (female): The current time in Paris, France, is 1:52, and 24, seconds, on the afternoon, of Friday, September 1.
(3 tone framing cue - descending pitch)

Findings

Subjects

Seventy-six participants responded to the user satisfaction comparison rating scale questionnaire. All participants were graduate students in the School of Education program at the University of Kansas and California State University, San Bernardino. Of these, 43 were female (56.5%) and 33 were male (43.4%). In the age range categories, 14 (18.4%) were 21-25, 23 (30.2%) were 26-30, 12 (15.7%) were 31-35, 5 (6.5 %) were 36-40, 6 (7.8%) were 41-45, 6 (7.8%) were 46-50, 9 (11.8%) were 51-above.

Summary of Main Results

An analysis of participant mean responses to a series of comparison rating scale items was conducted to evaluate the relationship between user satisfaction and the use of sonification in audio messages. The research hypotheses were based upon the proposition that users will report greater affection for audio messages with sonification. The independent variable was sonification and the dependent variable was user satisfaction measured by the comparison rating scale survey questionnaire across five dimensions: attention, listening, remembering, aesthetics, and overall valuation. Eight audio messages were presented to participants and the total mean responses for each of the five dimensions of user satisfaction were averaged to determine the composite rating for all eight messages. In addition, a series of one-sample $\underline{t}$ tests were performed on each of the mean response scores for the items on the user satisfaction comparison rating scales to evaluate whether the means were significantly different from 4, the value that would be obtained if all participants had responded randomly or if the responses of all participants were spread equally across the seven point scale. Following is a summary of the results.

H1: Recruiting Attention

The first hypothesis, which predicted that subjects would assign higher affection values towards messages with sonification in terms of recruiting attention, was confirmed with the mean response indicating that the messages with sonification were rated higher ($\underline{M}$ = 5.16, $\underline{SD}$ = 1.35) than messages without sonification.

H2: Aids Listening

The second hypothesis, which predicted that subjects would assign higher affection values Towards messages with sonification in terms of aiding listening, was confirmed with the mean response indicating that the messages with sonification were rated higher ($\underline{M}$ = 4.97, $\underline{SD}$ = 1.34) then messages without sonification.

H3: Remembering Details

The third hypothesis, which predicted that subjects would assign higher affection values towards messages with sonification in terms of remembering details was confirmed with the mean response indicating that the messages with sonification were rated higher ($\underline{M}$ = 4.51, $\underline{SD}$ = 1.27) than messages without sonification.

H4: Aesthetic Appeal

The fourth hypothesis, which predicted that subjects would assign higher affection values towards messages with sonification in terms of aesthetic appeal, was confirmed with the mean response indicating that the messages with sonification were rated higher ($\underline{M}$ = 4.80, $\underline{SD}$ = 1.48) than messages without sonification.

H5: Overall Valuation

The fifth hypothesis, which predicted that subjects would assign higher affection values towards messages with sonification in terms of overall valuation, was confirmed with the mean response indicating that the messages with sonification were rated higher ($\underline{M}$ = 4.88, $\underline{SD}$ = 1.38) than messages without sonification.

Test of the Hypothesis Subscale Means

A series of one-sample $\underline{t}$ tests were conducted on the mean response scores to evaluate whether the means were significantly different from the midpoint rating (4) that indicated that the messages were rated as being the same. The results of the one-way sample $\underline{t}$ tests for subscale items related to gaining attention, aiding listening, remembering details, aesthetic appeal, and overall valuation produced observed $\underline{t}$ values which were significant ($\underline{p} < .05$) across all subscales for each item. The means for messages with sonification were significantly higher than the midpoint rating of 4 for messages without sonification on all subscales for each item.

Conclusions

The results of this research indicate that listeners assign higher valuation to audio messages which use sonification. Following is a summary analysis of participant comments related to the use of sonification in audio messages. Bradley (1981) identifies five distinctive characteristics of attitudes towards speech communication that are also useful for establishing construct validity for measuring user satisfaction with audio messages with sonification. The five characteristics of attitudes are direction, intensity, salience, dominancy, and stability. These characteristics of attitudes provide helpful terminology for the interpretation of participant comments and the comparison rating scale measures of user satisfaction used in the first part of the study.

First, user attitudes towards sonification are directional, either negative or positive. Participants were asked to respond to a directional bi-polar comparison rating scale for each of eight audio messages as being either much worse or much better with sonification. The participants mean response indicated that their attitudes and valuation ratings were positive for all five measures, attention, listening, remembering, aesthetics, and overall valuation of sonification. The overall valuations found in this study suggest that sonification is perceived to be an enhancement to audio messages rather than a distraction. However, reactions were mixed, some liked sonification and some participants did not find it useful. Participants react differently to the same sound cues. The following comments reflect this difference in reaction.

- At first it seemed a cute novelty. The birds and horns were annoying!
- The tweets were not annoying.
- The birds were annoying.
- The bird sound ok
- Some of the sounds are pleasing and others are not (car horn) ... bird sound and office noise ok ...
- I felt my reaction to the sonification overall really depended on the type of sonification used. The mechanical sounds were annoying to me whereas the bird calls were pleasant.
- There were too many sounds and it distracted me from the main message.
- The sonification was annoying to me in the message

Second user attitude towards sonification had varying levels of intensity or strength of emotional reaction. The reaction may be strong or mild, or varying degrees in between. A seven point bipolar semantic differential comparison rating scale was used to measure the intensity of participant attitudes towards each message from one to seven as being either much worse or much better with sonification. A midpoint rating of four indicated that the messages were the same or no different in terms of the attribute in question. Participant reaction was not strong in terms of intensity as the overall rating of the messages with sonification was only slightly better. Since the overall reaction was not at the extreme end of the rating scale, it suggests that users attitudes toward sonification in audio messages were mixed. This also may be the result of the specific sound cues used in the messages for the study. Individual tastes vary in music, so it is reasonable to assume that individuals may differ in the types of sound cues that they respond to most favorably. One possible way to increase user satisfaction levels would be to allow for personalization of sound cues. For example, some cell phones allow users to personalize the ring tones associated with their phones. The following participant comments express varying levels of intensity of reaction regarding sound cues.

- I found the choice of sonification in the directions message unpleasant
- The beeping was totally unnecessary and very annoying.
- The horns were annoying.
- I didn't like some of the noises used. I also didn't like it when the noises went on too long, the birds were good not the directions because the noises were harsh.
- In some cases the sound helped with concentration, with others it was more distracting soft sound vs. hard…
- My overall impression of the sonification was it was better than just listening to the voices.
- Really liked the oriental sonification - nice touch.

Third, user attitudes towards sonification may be influenced by salience or non-salience. Saliency or non-saliency is dependent upon context, need, relevance, and degree to which it supports the tasks required of the listener. For example, the use of sonification in quiet surroundings may be less important that if it is used in the presence of noisy environments where it helps to recruit attention to the message and more attentive to it. The sample message which provides users with directions from one location to another while driving a car is an example of a context where noise can be a distraction thus, the presence of sonification is more salient to the user in terms of recruiting and directing attention to the message in the presence of noise. Another use that participants found to support the task of listening was the use of sonification to convey system status. This is similar to the function that music on hold has while one is waiting for a phone call to be answered. It lets the user know that the system is functioning or is working on performing a task initiated by the user. This is an example of how sonification can support the user by providing awareness of system status. The following participant comments regarding the perceived usefulness of the sonification in the message illustrate the importance of salience as a factor in user satisfaction.

- It helped to know when new directions were going to be told.
- Sonification helped the person to know that the system was still working.
- Sonification helped you know that more directions were coming.
- At first the sound took away from the overall message; more a distraction. When the sonification related more to the message it was easier to listen to the messages and remember.
- The sonification helps when it provides a cue to a response required from the user.
- The more complex the material the better it is to have sound.
- I think the most important addition is a few tones before and after key information - it works like a highlighter.
- I did really like the sonification to fill in dead air spaces. It made me feel as if I was still connected and that something was happening to get results and answers to the questions that were asked.
- The sonification helped fill in dead air where I wasn't sure if were still connected
- The sonification helped to break up when a new thought was going to occur.
- The sonification helped fill in the boring blank space.

Additionally, the use of bird calls as a sonification element in the bird alert message is another example of how salience can be a factor in user response. The bird alerts message was the highest rated message in terms of overall valuation of sonification. This may be due to the fact that the bird alert message used recordings of bird calls to represent the bird being referred to in the message. This is an example an auditory cue that has a high level of iconicity, as the bird calls are recognizable in terms of the users' experience with them in the everyday world. The following participant comments regarding the bird alert message support this observation.

- The bird sounds were very helpful rather than annoying like some of the previous sounds.
- For a birder, especially, having the actual bird calls precede the message is great.
- Sonification helps you to remember the details only if they are the calls of the birds mentioned.
- Bird sounds added to content - made a big difference
- I could mentally divide the information better and the chirps supported listening like a visual would.

A fourth factor to understand is the dominancy of the user's overall attitude or belief system. This can influence a user's reaction towards sonification. For example, users who have more positive attitudes towards technology in general may value and have a greater appreciation for sonification as a user interface enhancement feature. Similarly, musicians or persons with a musical background may value the presence of sonification more than non-musicians as they are more experienced at communicating in non-speech sound and thus more sensitive to its use. The ability of users to personalize sounds used may improve a user's attitude towards sonification. One of the techniques that software developers have used to personalize application interfaces is to offer multiple options for the visual look and feel called skins. The use of thematic sonification sound cue collections that users could select from could provide a way to personalize sonification in audio messages in a similar manner to choosing different graphic skins for software applications. The rationale for this would be that the user would be able to have ownership of choice of the sound cues used, thus they would become more familiar with the associated meaning.

- Sonification seems better for conversations than instructional types of messages, on second thought, I thought it was easier, clearer to hear the information or narrative with sonification

A final factor to consider is that user attitudes towards sonification will become more favorable and tend to become fixed or more stable once the meaning of the sound cues are understood. This may be related to the fact that the meaning of sound cues that have low levels of iconicity must be learned. For example, auditory icons which are actual real world sounds have familiar meaning to users where as the meaning of earcons which are made of musical tones have to be learned within the context in which they are used. Once the meaning of sound cues are learned the attitudes of users towards earcons will tend to become stable or fixed because they now understand the meaning of the cues, thus their overall valuation of sonification may increase. To try to account for the learning curve associated with sound cues with low levels of iconicity some of the sound cues in the audio messages were repeated so that users would be able to learn the meaning by repeated exposure to the sound cues. Following are participant comments that support this observation.

- Since sonification cues are the same in both so far, 1 and 2, I "learned" what they mean.
- Knowing the routine will be helpful in attention and remembering information next time

Following are some suggested general rules of thumb regarding the development of sound cues for use as sonification elements based upon follow up discussion and comments offered by participants.

- Form should follow function as much as possible.
- Use sound cues in a consistent manner so that the meaning of cues can be learned.
- Use sound cues that have high levels of iconicity or have recognizable meanings when possible.
- Use sound cues that are perceptually distinct from listener's acoustic environment in terms of pitch, rhythm, and volume.
- Use sound cues with a gradual attack onset rather than a sudden attack to avoid startling users.

- Use the low volume levels rather than high volumes for sound cues.
- Avoid using sounds that are annoying or obtrusive.
- Use sound cues that are aesthetically pleasing.
- Use a reduced set of sound cues.

This study seemed to indicate that users were able to understand the meaning of sound cues that have iconicity, which is best obtained by the use of sounds which are referent in the real-world of everyday listening. A better understanding is needed of how sounds in everyday listening and synthetic musical tones can be used most effectively to convey information the listener. Future research should explore the design and selection of non-speech audio cues. Observing how individuals listen to and respond to music may provide a valuable information base to guide the design of more effective and aesthetically pleasing sound cues. The use of natural versus synthetic non-speech audio cues should be examined to determine which type of sound cues convey meaning that is immediately recognized by the listener. The use of Morse code as a source of non-speech audio cues for human-machine interaction should be explored as a way to provide sonification cues that have universal meaning across cultures and may be easier to learn and understand. In addition, it is suggested that future development of sonification elements for use in instructional audio messages should adhere to the National Braille Association (1979) requirements and specifications for readers for the visually impaired which provide a set of audio recording guidelines to follow for the description of images, mathematical symbols, and text structure. Message designers must be thoughtful when using sonification in audio messages. The inclusion of sound cues may not only be annoying to listeners, but can possibly add to the cognitive load of the message.

References

Alten, S. R. (2002). Audio in media. Belmont, CA, Wadsworth/Thomson Learning.

Altman, R. (1992). Sound theory, sound practice. New York, Routledge.

Bickford, P. (1997). Interface design: The art of developing easy-to-use software. Boston, AP Professional.

Bishop, M. J., Cates, W.M., (2001). "Theoretical Foundations for Sound's Use in Multimedia Instruction to Enhance Learning." Educational technology research and development : ETR & D. Volume 49 (Number 3): 5-22.

Blattner, M. M., D. A. Sumikawa, et al. (1990). Earcons and icons: Their structure and common design principles. Visual Programming Environments: Applications and Issues. E. P. Glinert. Los Alamitos, CA, IEEE Computer Society Press**:** 582-606.

Bradley, B. E. (1981). Fundamentals of speech communication : the credibility of ideas. Dubuque, Iowa, W.C. Brown.

Chin, J. P., Diehl, V. A. and Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of SIGCHI '88. New York, ACM/SIGCHI.**:** 213-218.

Chion, M., C. Gorbman, et al. (1994). Audio-vision : sound on screen. New York, Columbia University Press.

Cohen, M.H., Giangola, J.P., and Balough, J. (2004). Voice user interface design. Boston, MA. Addison-Wesley.

Cook, P. R. (1999). Music, cognition, and computerized sound : an introduction to psychoacoustics. Cambridge, Mass., MIT Press.

Gaver, W. (1986). "Auditory icons: Using sound in computer interfaces." Human Computer Interaction **2**: 167-177.

Kramer, G. and Santa Fe Institute (Santa Fe N.M.) (1994). Auditory display : sonification, audification, and auditory interfaces. Reading, Mass., Addison-Wesley.

Marsh, P. O. (1983). Messages that work : a guide to communication design. Englewood Cliffs, N.J., Educational Technology Publications.

National Braille Association. (1979). Tape recording manual. Midland Park, N.J. Washington, National Braille Association ; available from National Library Service for the Blind & Physically Handicapped, Library of Congress.

Neisser, U. (1967). Cognitive psychology. New York,, Appleton-Century-Crofts.

Sarnoff, C. W. (1960). "Psychoanalytic Theory and Social Attitudes." Public Opinion Quarterly(24): 261.

Shneiderman, B. (1998). Designing the user interface : strategies for effective human-computer-interaction. Reading, Mass, Addison Wesley Longman.

Sonnenschein, D. (2001). Sound design : the expressive power of music, voice, and sound effects in cinema. Seattle, WA, Michael Wiese Productions.

Spencer, K. (1991). The psychology of educational technology and instructional media. London ; New York, Routledge.

Tannenbaum, R. S. (1998). Theoretical foundations of multimedia. New York, Computer Science Press.

Vroemen, M. W. (1986). Colour in CRT-display design: Lay-out and coding. Utrecht, Universiteit van Utrecht, Psychologisch Laboratorium.