DOCUMENT RESUME

ED 388 265 IR 017 414

AUTHOR	King, Cynthia M.; And Others
TITLE	Digital Captioning: Effects of Color-Coding and Placement in Synchronized Text-Audio Presentations.
PUB DATE	94
NOTE	7p.; In: Educational Multimedia and Hypermedia, 1994. Proceedings of ED-MEDIA 94--World Conference on Educational Multimedia and Hypermedia (Vancouver, British Columbia, Canada, June 25-30, 1994); see IR 017 359.
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Assistive Devices (for Disabled); *Captions; Case Studies; College Preparation; *Comprehension; Computer Uses in Education; Films; *Hearing Impairments; Higher Education; *Special Needs Students
IDENTIFIERS	Black and White; *Color Coding; Digital Data; Gallaudet University DC

ABSTRACT

Captioning is the process of providing a synchronized written script (captions) to accompany auditory information. This article describes programs available for captioning digital media on computers, and discusses the results of a study on color-coding and placement of captions. Seventy-two students in the Preparatory Studies Program (PSP) at Gallaudet University (Washington, D.C.) participated in the study (PSP enrolls deaf and hard-of-hearing students and prepares them for college). A 15-minute segment from a Disney film was used in the study. Four versions of digital captions were prepared: (1) captions color-coded for speaker identification, centered at the bottom of the screen; (2) black and white captions, centered at the bottom of the screen; (3) color-coded captions with placement dependent on the location of the speaker; and (4) black and white captions with placement dependent on the speaker's location. Results indicate that comprehension is higher when captions are color-coded for speaker identification than when captions are black and white. There are no significant differences between centered captions and captions with variable placement dependent on location of the speaker. (AEF)

Digital Captioning: Effects of Color-Coding and Placement in Synchronized Text-Audio Presentations

CYNTHIA M. KING
Educational Foundations and Research
Gallaudet University, 800 Florida Avenue, NE, Washington, DC 20002-3695
email: cmking@gallua.gallaudet.edu

CAROL J. LASASSO
Education
Gallaudet University, 800 Florida Avenue, NE, Washington, DC 20002-3695
email: cjlasasso@gallua.gallaudet.edu

DOUGLAS D. SHORT
Institute for Academic Technology
2525 Meridian Parkway, Suite 400, Durham, NC 27713

Abstract: Captioning is the process of providing a synchronized written script (captions) to accompany auditory information. This article describes programs available for captioning digital media on computers and reports the results of a study on color-coding and placement of captions. Results indicate that comprehension is higher when captions are color-coded for speaker identification than when captions are black-and-white. There are no significant differences in comprehension between centered captions and captions with variable placement dependent on location of the speaker.

Captioning is the process of providing a synchronized written script (captions) to accompany auditory information. The most common form of captioning is that for *analog* television, where captions are embedded in Line 21 of the video signal and displayed only to those who have an external decoder or a television with a built-in decoder (Armon, Glisson, & Goldberg, 1992; Bess, 1993). New methods of captioning, however, are beginning to be developed for *digital* television and computer-based multimedia products (Armon et al, 1992; Hutchins, 1993; Short, 1992; Short & King, 1994).

King (1993) provides an extensive discussion of captioning and why multimedia developers should incorporate it into their products. Additional discussion of issues related to ensuring that deaf and hard-of-hearing people have access to information presented auditorially can be found in Jordan (1992) and Kaplan and De Witt (1994), as well as in electronic resources on the Internet, such as ADA-LAW (ada-law @vm1.nodak.edu) and Equal Access to Software and Information (easi@sjuvm.stjohns.edu).

The present paper addresses: (a) digital captioning, (b) captioning format research and standards, and (c) effects of color-coding and placement of captions used to represent speaker identification on the comprehension of deaf and hard-of-hearing viewers.

Digital Captioning

Synchronization of text and audio is possible in a wide array of computer-based environments. For example, IBM provides an option within the *OS/2* operating system by which captions, if they exist, can be turned on or off at the request of the end-user. IBM also provides a simple captioning tool, whereby developers can synchronize text files to accompany audio, as part of its Ultimedia Tools series. On both PC and Macintosh platforms, some video-editing programs support captioning via titling functions (e.g., ATI's *MediaMerge* and Adobe's *Premiere*; the titling function in *Premiere* is currently available only in the Macintosh version).

Multimedia programming environments, such as Apple's *Hypercard* and Asymetrix's *Toolbook*, can also be used for captioning. The research described in this paper involved the use of captioning tools developed, using Asymetrix's *Toolbook*, by Douglas Short of the Institute for Academic Technology. These tools were originally developed for providing translations for foreign language operas on laserdiscs (*SyncText*) and for providing notation and/or translation for music and lyrics on CD-Audio discs (*CD Time Liner*) (Institute for Academic Technology, 1992; Short, 1992).

The *SyncText* laserdisc captioning program was subsequently enhanced by Short and King (1994) to include tools for statistics calculation (e.g., words-per-minute) and other functions (e.g., a function to find words and phrases in the caption list) that are helpful during the captioning process. The enhanced tool, *CAP-Media LD*, also meets or exceeds the current specifications for television captioning. For example, the FCC Caption Decoder Standard of 1991 (as revised in 1992; Electronic Industries Association [EIA], 1992) mandates that seven background and foreground colors be available; *CAP-Media LD* provides twenty different colors. *CAP-Media LD* can align captions via all methods available for television (left alignment, eight tab-settings, and/or starting at any character in the 32-character line possible on television). It also provides automatic centering and right-alignment. Line length is not restricted in *CAP-Media LD*, and a full range of proportional and monospace fonts is available. The fonts include all characters in the required and optional font sets, as specified for television captioning (EIA, 1992). The maximum number of caption lines is set at four, identical to the limit in television captioning (in *CAP-Media LD*, this limit can be overridden manually, if desired).

CAP-Media LD, which works in conjunction with a set of tools called *Express Author* developed at the Institute for Academic Technology, permits the video and caption areas to be displayed full screen or sized to any portion of the computer screen. The program includes options for users to customize the target applications they develop using *CAP-Media LD*. For example, users can create lists of captions and graphic charts of video segments, which can then be used to control display of the video in non-linear order. Interactive captioning is also possible via hyperlinks between the captions and *Express Author*'s glossary tools. Thus, if a glossary and hyperlinks to the captions have been created, end-users are able to click on words or phrases as captions are displayed. The video is stopped automatically and the user is linked immediately to supporting materials--which can include text, video, audio, animation, and/or pictures--for clarification or elaboration. The glossary tools also permit end-users to add their own entries to the glossary.

Finally, *CAP-Media LD* is designed such that display of non-text objects (e.g., graphic symbols) and other *Toolbook* functions can be synchronized to audio-visual displays in the same way that text is handled. This design will allow, for example, synchronization between display of a laserdisc video clip and an audio file being used for descriptive video services (WGBH Educational Foundation, 1993) to provide access to non-visual information for blind and visually-impaired people.

Captioning Format Research and Standards

King and LaSasso (1993-1996) are using *CAP-Media LD* in a series of studies designed to assess the impact of four format features on the comprehension of content and affective response to captions. These features include: (a) methods of indicating speaker identification, (b) methods of indicating non-speech information such as background noise, music, and emotional tone, (c) placement of captions in relation to the video display area, and (d) timing, length, and duration of captions. These studies are part of a project, funded by the U.S. Office of Special Education and Rehabilitative Services (OSERS), which is designed to help set new standards for computer-based multimedia and digital television captioning. The results of these studies are also expected to be generalized to today's analog captioning system.

CAP-Media LD and computer-controlled laserdisc technology were selected for this research for several reasons. First, laserdisc technology permits the researchers to have precise control over the segments of video to be shown and permits non-linear access to the video (ensuring that all subjects see exactly the same portion of video and reducing costs associated with counter-balancing the order of video clips). Second, computer-based captioning is more flexible than Line21-analog captioning (e.g., more fonts and variations, such as placing the captions completely separate from the video, are possible). Third, computer-based captions are less costly and can be of higher quality than captions created in a television editing suite (when the captions involve features--such as different fonts--not possible with current Line21-analog captions).

This line of research grows out of a long history of research on the format features of captioning (e.g., Beldue, 1983; Braverman, 1981; Harkins, 1993-1994; King & LaSasso, 1992-1993). Recent research efforts have been motivated largely by the Television Decoder Circuitry Act of 1990, which mandated that, effective July 1, 1993, every television (13 inches or larger) manufactured in or imported to the United States must have built-in captioning decoder circuitry (DuBow, 1991) and by the work of the FCC and Electronic Industries Association (e.g., EIA, 1992, Hutchins, 1993) to set new technical standards for television captioning.

King and LaSasso (1992-1993) surveyed deaf and hard-of-hearing consumers to determine their opinions concerning captioning format and conducted video preference tests to determine their reactions to proposed new captioning features. Among the results were indications that caption consumers prefer (a) sans serif over serif fonts, (b) captions centered at the bottom of the screen rather than captions placed left-center-right (depending on where the speaker was on the screen), and (c) moderate movement of captions to avoid covering on-screen titles (e.g., names of people and sports scores). Consumers expressed strong preference for the existing Line21 monospace font and black background, when compared to character-generated proportional fonts and varied backgrounds. Whether such preferences were due to familiarity, resistance to change, and/or subtle differences in the quality of the Line21 and character-generated captions could not be determined. Consumers also indicated a desire for choice in features such as font style, font size, and color of the background box (e.g., LaSasso & King, submitted for publication).

In response to an open-ended question about captioning, caption consumers identified problems related to the amount of captioning, accuracy of captioning, obstruction of on-screen titles, and format features such as speaker identification. Regarding speaker identification, some wanted color used for different speakers (as is currently done in the Australian captioning system), some requested that the names of speakers be included with captions, and others wanted different emotional tones to be represented.

Effects of Color-Coding and Placement on Comprehension

Based on the reports of consumer preferences (e.g., LaSasso & King, submitted for publication) and the EIA-608 Report (EIA, 1992), which indicated that caption providers were exploring the use of color for speaker identification, a decision was made to conduct an experimental study of the effects of color and placement (used for speaker identification). [Note: the primary method used in current Line21-analog captioning for television is placement of the captions via eight-tab settings. For on-screen speakers, captions are placed in the location closest to the speaker. For off-screen speakers, captions are placed either in the same location as the last caption for that speaker or in a new location, if the speaker has moved.]

Method

Subjects. Seventy-two students in the Preparatory Studies Program (PSP) at Gallaudet University participated in this study. PSP enrolls deaf and hard-of-hearing students who are not quite ready for college and provides them with studies to improve English and mathematics skills, as well as other college-preparatory studies. Most of these students (82.6%) are profoundly deaf (90+ decibel loss in the better ear), with 5.8% with severe hearing losses (76-89 db), and 11.6% with moderate hearing losses (51-75 db). More than half (56.9%) are female. Only students with normal vision, with or without glasses, are included in the final sample (80 students participated in study, 8 were eliminated due to vision problems or additional handicaps).

Students were randomly assigned to one of four study conditions (described below). Analyses of variance indicate that there are no significant differences between the groups on (a) age, $F(3,68)=.91$, $p=.441$, $MS_e=6.75$; (b) hearing loss (measured in decibels), $F(3,65)=1.98$, $p=.126$, $MS_e=165.04$; or (c) reading level (measured by scores on the *Degrees of Reading Power* test, $F(3,64)=.81$, $p=.494$, $MS_e=54.95$.

Materials. A 15-minute segment from the beginning of Disney's *Sword in the Stone* was used in this study. Digital captions for the video segment were generated, using a 20-point, proportional, sans serif font and a full-screen display (the caption area covered the bottom of the video). Four versions were prepared: (a) captions color-coded for speaker identification, centered at the bottom of the screen, (b) black-and-white captions, centered at the bottom of the screen, (c) color-coded captions with placement dependent on the location of the speaker, and (d) black-and-white captions with placement dependent on the location of the speaker. The last two conditions emulated the system used by the caption provider who had prepared Line21-analog captions for the movie (i.e., captions were placed in the same place as would occur in television captioning).

The test materials for the study involved four separate *Toolbook* applications, using the digital captions (generated by *CAP-Media LD*) for the four conditions. The instructions, controls for playing each video clip, the test question pages, and the captions themselves were identical in all four applications--only the color and placement of the captions varied.

The 15-minute segment was divided into 22 video clips, with stopping points located at phrases where the identity of the speaker might be unclear to someone who could not hear the voices of the characters. Twenty of the stopping points--the actual test items--were selected from a larger set of potential items in a pilot-test conducted with hearing children and deaf and hearing adults. The other two stopping points were for a sample test question to ensure subjects understood the task and a final video clip (without a test question) to complete the video segment. Questions were in the form, "Who said ..." followed by the caption on the screen at the end of the preceding video clip. Answers were in the form of still-frames of the characters from the video, identified as A, B, or C.

Procedures. As previously noted, subjects were randomly assigned to one of the four conditions. Testing took place over three days in a given week, with each subject participating in a single test session of approximately 45 minutes on one day. Subjects watched the video on a 27" computer monitor in groups of 2-10 and recorded their answers in paper-based test booklets. Following the test, all subjects were asked to respond to a set of Likert-type items concerning their opinions of the captions they saw, and subjects in color-coded conditions were asked to name the color used for each character (black-and-white versions of still-frames from the video were used as prompts). The five-point, Likert-type scale had choices of "I like it. I like it a lot. I have no opinion. I don't like it. I don't like it a lot," with "it" being the feature of interest.

A factorial analysis of variance was conducted with test scores as the dependent variable and color and placement as independent variables. Chi-Square statistics, using an experiment-wise alpha level of .01, were calculated to identify those test items that best discriminated among the groups. Descriptive statistics and Chi-Square statistics were also used for the Likert-type items on affective responses to the digital captioning. A Pearson Product-Moment Correlation was calculated between test scores and scores on the five-item test identifying the color for different characters (for only those subjects who saw color-coded captions).

Results

Table 1 contains descriptive statistics for test scores (maximum score of 20) for the four conditions. There are significant differences between those who saw color-coded captions and those who saw black-and-white captions, $F(1,68)=12.55$, $p=.001$, $MS_e=14.36$. There are no significant differences for (a) placement, $F(1,68)=.34$, $p=.561$, $MS_e=14.36$, or (b) the color-by-placement interaction, $F(1,68)=1.89$, $p=.174$, $MS_e=27.09$.

Table 1
Mean Test Scores, Standard Deviations, and Number of Subjects for Study Conditions

	Centered	Line21 format	COMBINED MEANS
Color-Coded	16.29 (3.60) (n=17)	17.00 (3.22) (n=18)	16.65
Black-and-White	14.35 (3.98) (n=17)	12.60 (4.22) (n=20)	13.48
COMBINED MEANS	15.32	14.68	14.99

There are significant group differences in the response patterns of subjects for only two of the test questions, #8, $\chi^2(6,N=72)=18.65$, $p=.0048$; #16, $\chi^2(6,N=72)=24.13$, $p=.00049$. Both test items relate to video scenes where the speaker is off-screen and the caption is located near one of the on-screen characters. In the test item, the on-screen character is a distractor (incorrect choice). On item #8, 65% of those who saw black-and-white, Line21 format captions selected that incorrect choice (compared to 41% of those who saw black-and-white, centered captions and only 12% of those who saw color-coded captions). On item #16, 40% and 47% of the subjects who saw black-and-white captions (Line21 format and centered, respectively) selected the on-screen character (compared to 14% of those who saw color-coded captions).

There are no differences among the four groups on their overall affective responses to digital captioning. For this analysis, frequencies for "like and like a lot" are collapsed, as are those for "don't like and don't like a lot." Approximately three-fourths of the subjects like the captioning (78%), font style (78%), font size (74%), and the mixed, upper- and lower-case of the lettering (80%). Of those who saw color-coded captions, 79% like them. For placement of captions, 78% of those who saw centered captions like centering, and 71% of those who saw Line21-format captions like variable placement. The percentage of subjects who had no opinion concerning each feature of the captions ranged from 12% to 21%, and the percentage of subjects who did not like the feature ranged from 3% to 14%.

Approximately two-thirds (67%) of the subjects who saw color-coded captions could name the colors of the captions for all five of the main characters. A moderate, positive correlation exists between test scores and the number of colors correctly identified, $r=.352$, $p=.035$, $n=36$.

Discussion

Results of this study show that color-coding has a significant effect on comprehension, when comprehension is narrowly defined as understanding of speaker identification for potentially-confusing items. These results, however, cannot be interpreted to mean that color-coding should replace the current line21 format that uses placement for speaker identification. First, color-coding is unusable for that portion of the population that is color-blind (thus, redundancy in any coding system that uses color will be necessary). Second, color-coding is unwieldy when the number of main characters is large. Third, complex color-coding systems may not be usable (e.g., The Australian Caption Centre recently reduced the complexity of their color-coding system, based on viewer feedback; C. Grimmer, personal communication, December 15, 1993).

Color-coding for speaker identification, in some combination with placement and inclusion of character names in the captions for off-screen speakers, should be studied further. When there are a limited number of main characters in the story, color-coding, as shown in the present study, appears to increase comprehension. The next study by King and LaSasso (1993-1996) will address how and how well children learn, over time, a color-coding system used for speaker identification.

Summary

The study reported in this paper is the first in a series of studies on captioning format for multimedia and digital television. The results have application for both captioning environments (for caption providers) and screen design (for multimedia developers). Most important, however, is the recognition that--without captioning--deaf and hard-of-hearing people do not have independent access to audio-only information. Tools like *CAP-Media LD* are essential, if this important audience is to have full and equal access to digital media.

References

Armon, C., Glisson, D., & Goldberg, L. (1992). How closed captioning in the U.S. today can become the advanced television captioning system of tomorrow. *SMPTE Journal, 101*, 495-498.

Beldue, I. P. (1983). Fixed versus variable letter width for televised text. *Human Factors, 25*(3), 273-277.

Bess, L. P. (1993, September/October). Closed-Caption TV: A new era. *SHHH Journal, 14*(5), 8-10.

Braverman, B. (1981). Effects of language level and visual display on learning from captioning instruction. *Educational Communication and Technology, 29*(3), 147-154.

DuBow, S. (1991). Television Decoder Circuitry Act: TV for all. *Temple Law Review, 64*, 605-614.

Electronic Industries Association (EIA). (1992). *Draft standard: EIA-608: Line 21 data services for NTSC.* Chicago: Author.

Harkins, J. E. (1993-1994). *Caption features for indicating non-speech information: Research toward standardization.* Grant funded by the US Department of Education, Office of Special Education and Rehabilitative Services.

Hutchins, J. (1993, September/October). The captioning technology of the future is being developed today. *SHHH Journal, 14*(5), 22-24.

Institute for Academic Technology. (1992, Fall). Faculty develop programs for listening comprehension and text synchronization. *IAT Briefings: The Institute for Academic Technology*, pp. 14-15.

Jordan, I. K. (1993). A matter of access: Technology for deaf and hard-of-hearing people. *TechNos, 2*(2), 26-28.

Kaplan, D., & De Witt, J. (1994). *Telecommunications and persons with disabilities: Building the framework. The second report of the Blue Ribbon Panel on National Telecommunications Policy.* Oakland, CA: World Institute on Disability.

King, C. M. (1993). Multimedia captioning: Synchronized text and audio presentations. In H. Maurer (Ed.), *Educational Multimedia and Hypermedia Annual, 1993* (pp. 609-614). Charlottesville, VA: Association for the Advancement of Computing in Education.

King, C. M., & LaSasso, C. J. (1992-1993). *Acceptability, readability, and comprehensibility of caption format features: Perspectives and responses of deaf and hard-of-hearing caption consumers.* Grant funded by the US Department of Education, Office of Special Education and Rehabilitative Services.

King, C. M., & LaSasso, C. J. (1993-1996). *CAP-MEDIA: Captioning for multimedia and digital television.* Grant funded by the US Department of Education, Office of Special Education and Rehabilitative Services.

LaSasso, C. J., & King, C. M. (submitted for publication). *Deaf consumers' views about captioning.*

National Captioning Institute. (1991). *An international bibliography of captioning and subtitling literature.* Falls Church, VA: Author.

Short, D. D. (1992). Instruction in a postliterate culture: Multimedia in higher education [Invited article]. *Journal of Information Systems Education, 4*(3), 11-15.

Short, D. D. [Developer], & King, C. M. [Design Consultant]. (1994). *CAP-Media LD: Synchronized text for laserdiscs* [computer program]. Durham, NC: Institute for Academic Technology.

WGBH Educational Foundation. (1993). *Descriptive video services.* Boston: Author. [125 Western Avenue, Boston, MA 02134]

Acknowledgments

Support for the experimental research described in this paper was provided in part by grants from the U.S. Department of Education, Office of Special Education and Rehabilitative Services: Educational Media Research, Distribution, and Training Program (Grant Nos. H026R10002 and H026R30011), with additional funding from Gallaudet University. Opinions and views expressed in this paper are those of the authors and do not necessarily reflect the positions of any government agency or the university.

The Institute for Academic Technology (IAT) is a partnership between IBM and the University of North Carolina at Chapel Hill on behalf of education. For information about its courses, programs, and services, contact the Institute for Academic Technology, 2525 Meridian Parkway, Suite 400, Durham, NC 27713, (919) 560-5031, e-mail: info.iat@mhs.unc.edu.