CAPTION FEATURES FOR INDICATING NON-SPEECH INFORMATION: RESEARCH TOWARD STANDARDIZATION

Funding

This project received funding from the U.S. Department of Education Office of Special Education Programs (Award #H026R20004-92) and Gallaudet University.

Project Description

This project studied the variations in captioning conventions for representing non-speech information (including sound effects, location of speaker off screen, distortions of speech, speech in a foreign language, music, etc.). The study resulted in recommended guidelines to the captioning industry.

An advisory committee consisting of captioning experts and consumer experts met at the project's outset. This group gave written input into the priority areas within the realm of non-speech information (NSI). In addition, 33 caption writers from 26 companies were interviewed by telephone, to ascertain the principal problem areas for indicating non-speech information and questions the writers would like to see answered through the consumer research. A report on the caption writers' opinions was distributed to the advisory committee.

Working in deaf and hearing teams, staff coded 37 hours of television and motion picture video. Caption features used for NSI were catalogued. More than 1,400 examples of non-speech information portrayed in captions were identified. Data was used as a basis for selection of features and types of NSI to be evaluated.

Instruments and questionnaires were developed for assessing consumer opinions. At the advice of the advisory committee, a test of awareness of caption features in common use was designed. The main portion of the procedure entailed identifying video clips from televised material, such that the targeted types of NSI (e.g., speaker identification, audience reaction, sound effects) could be shown and various features tested. In addition to traditional caption features, new features to be added as a result of the FCC's ruling on decoder displays were also included. These included color, improved placement, and paint-on style. In all, 19 video clips were identified, captions written and encoded on to tape, and editing performed to permit counterbalancing of the choices. Answer forms were also designed, and all materials were pilot tested with consumers before beginning data collection.

A total of 189 deaf and hard of hearing consumers from seven states participated in the study.

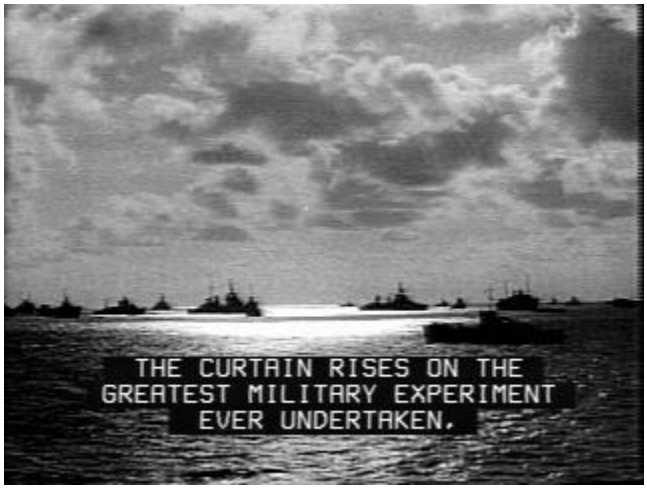
NSI Sample Clip

Nineteen video clips from television were used in the data collection videotape. Each of the 19 clips were captioned two or three different ways. Deaf and hard of hearing consumers watched each clip, and selected the method of captioning NSI that they preferred. They also could note which, if any, captioning methods were unacceptable to them. New uses of captions, incorporating features such as color, were used in the study.

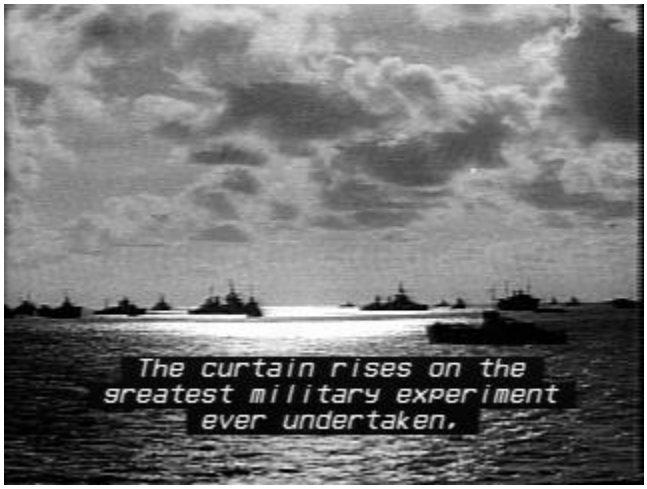
Speaker identification was tested in six clips. In the sample below, speaker identification of an off-screen narrator was tested.

Feature used: speaker identification
(narrator explicitly identified).

Described and Captioned Media Program
VOICE 800-237-6213 • **TTY** 800-237-6819 • **FAX** 800-538-5636 • **E-MAIL** info@dcmp.org • **WEB** http://www.dcmp.org
Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education



Feature used: color (captions for narrator in white; captions for second speaker in yellow).



Feature used: capital letters and italics (capital letters for narrator; slanted letters for second speaker).

Results of Consumer Preference

Speaker identification: 67% (unacceptable: 0%)
Color: 21% (unacceptable: 19%)
Caps and italics: 12% (unacceptable: 5%)

Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education

NSI Executive Summary

(This summary document was presented as a poster session at the National Symposium on Educational Applications of Technology, sponsored by the National Technical Institute for the Deaf in Rochester, New York, July 20-22, 1994.)

The purpose of this study is to improve captioning of "non-speech information" (NSI). NSI is a term to describe aspects of the sound track, other than spoken words, that convey information about plot, humor, mood, or meaning of a spoken passage. Currently, captioning companies vary in how they portray this information. Our study showed that consumers do not recognize features commonly used to represent NSI, when they are presented in isolation. The outcome of the study will be recommendations to the captioning industry to standardize the way it indicates this information.

Examples of NSI

Identification of speaker (off-screen speakers and multiple on-screen speakers)
Sound effects
Music (singing, background music, etc.)
Manner of speaking (whispering, emotion, word emphasis, etc.)
Audience reaction (laughing, groaning, booing, etc.)
Indication of title (book, film, newspaper, play, etc.)
Puns

Examples of Features that Can Now be Used to Identify NSI

- Italics
- Placement
- Double chevrons
- Color
- Music icons
- Paint-on
- Underlining
- Quotation marks
- Explicit description

Methods

An advisory committee composed of consumers and caption industry representatives advised the research staff on all aspects of the study.
Thirty-three caption writers were interviewed for input as to the challenges they face in representing NSI.
Thirty-seven hours of video were analyzed in detail to determine current practice in identifying NSI.

Nineteen video clips from television were selected, giving 19 different examples of NSI. Emphasis was on NSI identified by the advisory committee and industry as being most important, such as speaker identification.

For each of the 19 examples, two or three different ways of captioning to indicate NSI were selected. New uses of captions, incorporating new features in the new caption chips, were included as well as conventional features such as italics. In all, 55 uses of caption features were included.

Described and Captioned Media Program
VOICE 800-237-6213 • **TTY** 800-237-6819 • **FAX** 800-538-5636 • **E-MAIL** info@dcmp.org • **WEB** http://www.dcmp.org
Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education

Deaf (n = 106) and hard of hearing (n = 83) consumers viewed the tape and selected their preference from the choices presented. If any of the choices were unacceptable, consumers were instructed to mark those choices with an X.

Results were analyzed and recommendations to the industry were written. These were circulated for comment to the captioning industry in the summer and early fall of 1994. The final report will be published in early 1996.

Summary of Results

Consumers wish to see more NSI indicated than is now industry practice. They preferred versions that indicated NSI such as background music, audience laughter, accents, and sound effects.

Explicit description was most often preferred as the method of indicating NSI. This was particularly true in speaker identification. The punctuation used--brackets, parenthesis, colon--is relatively unimportant, based on findings of the study, although there was a slight (but non-significant) preference for parenthesis.

Examples:

(audience laughing)
(soft, sad background music)
(southern accent)
(female narrator)

Italics were preferred for indicating when a speaker emphasized a word or phrase.

Example:

DON'T *EVER, EVER* DO THAT AGAIN.

Quotation marks were preferred for indicating titles.

The conventional form of indicating singing, with music icons enclosing the words, was preferred over a paint-on style.

Explanation of a pun was preferred over no explanation.

Color was not a preferred method of indicating NSI, although it was tested five times in this study. This was also found in an earlier study by King and LaSasso (1993).

Flashing captions, paint-on captions, and underlining were not preferred in this study.

The preferences of deaf and hard of hearing consumers followed the same pattern in the vast majority of cases (16 out of 19 clips). The differences were minor and not of consequence for recommendations to the industry.

Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education

NSI Recommendations

Guidelines for Types of Non-Speech Information (NSI)

General Guideline

If a descriptive caption or feature would in any way clarify or enhance the viewer's awareness of the audio, it should be indicated. Consumers prefer that more of such information be included than is often done in current practice.

Background Music

Background music should be indicated, especially if it contributes to the plot or mood of the video. A description of the background music should be given wherever possible.

Sound Effects

Where feasible, a combination of description and onomatopoeia should be used to indicate sound effects. If space or other limitations do not permit the two to be used together, descriptors should be used. Onomatopoeia should not be used alone. A descriptor is particularly important if the source of the sound effect is not obvious from the video.

Singing

Continue the practice of using the musical-note icon surrounding the caption. All-caps and upper/lowercase type are equally acceptable for the caption portion.

Multiple Speakers On Screen

Where multiple speakers appear on the screen, placement should be used to distinguish among them. Explicit identification should be used in combination with placement if dialogue is fast, if faces are obscured, if characters are moving, or if other circumstances could confuse the viewer. If the character cannot be identified by name, then a descriptor should be provided. An acceptable format for explicit identification is the character's name or descriptor in upper/lowercase, surrounded by parentheses, above the caption and left justified with the caption. Other formats are probably uncontroversial.

Narrators

Explicitly identify off-screen narrators, rather than using features, such as italics or color, that require the viewer to interpret the feature/code while reading captions.

Whispered Speech

Whispered lines should be identified as such and combined with upper/lowercase captions.

Emphasis

Indicate with italics the emphasized word(s) within a caption.

Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education

Titles

Use quotation marks when indicating the title of a book, movie, etc.

Audience Reaction

Audience reaction should be captioned. This is particularly important where the reaction itself becomes part of the plot or comedy. Audience laughter should also be described. (It is possible that repeating the descriptor every time the audience laughs, over the length of an entire sitcom episode, would become annoying. This length of exposure was not tested. Therefore, discretion is advised; but audience laughter should be indicated much more often than is now the industry's practice.)

Conveying Emotion

Where strong emotion is conveyed, the emotion should be described with the caption. This feature should be used especially where strong emotion is not entirely obvious in the facial expression and actions of the speaker. Caption writers may be concerned that this feature could be overused. However, based on consumer feedback, caption writers should use this feature more than is current practice.

Accents

Indicate foreign or regional accents with a one-time description at the beginning of the character's lines. (Note: This issue was tested only with a fictional character, and probably should not be generalized to other speakers.)

Puns

Puns should be explained briefly when feasible.

Guidelines for Features

General Guideline

Consumers have indicated a preference for explicit description or identification over features that assume understanding on the part of the viewer. Examples of such features, requiring interpretation by the viewer, include: use of italics for the entire caption, color, and upper and lowercase type without explanation.

Color

Color was not the preferred method of indication in this study, although it was tested in five different circumstances. Color also tested poorly against placement and speaker identification in an earlier study by King and LaSasso (1993). Color is judged unacceptable by more viewers than are many other features. Note that color in real-time captioning (where other options may be problematic) was not tested. (Color in a digital video environment is being studied further by King and LaSasso in 1994-1996.)

Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education

Flashing

Flashing captions were not preferred in the two applications tested in this study, and were unacceptable to an appreciable minority of respondents. Further study may be warranted of whether or how to use this feature.

Paint-On

Paint-on captions were tested in only one context, and they were not preferred. Further study may be warranted of whether or how to use this feature.

Italics

Italics were less desirable than explicit definition in several contexts. Italics are widely used and should be used less frequently, as their intent is frequently lost on viewers.

Underline

Underlining was the last choice of respondents in the two applications tested. Further study may be warranted of whether or how to use this feature.

Quotation Marks

Quotation marks were preferred (contrasted with italics and underlining) for indicating a title.

Project Duration

December 1992 through August 1994

Gallaudet Investigators

Judith E. Harkins
Beth R. Singer
Ellie Korres
Barbara M. Virvan
Yoon K. Lee
(Date of this report: 2/96)

Funding for the Described and Captioned Media Program is provided by the U.S. Department of Education