Computers, Movies, and Nonliteral Language Rehabilitation: An In-depth Look at the Development of an Interactive Video Workstation

Edward Hall, Ph.D.
Department of Communication Disorders
California State University at Northridge

Muriel Goldojarb, M.A.
Audiology and Speech Pathology
Sepulveda Veterans Administration Medical Center
Sepulveda, California

Diana Van Lancker, Ph.D.
Veterans Administration Outpatient Clinic
Los Angeles, California
and
Department of Neurology
University of Southern California School of Medicine
Los Angeles, California

The importance of developing rehabilitative materials for nonliteral language processing, together with the availability of interactive video technology, led to the development of the Nonliteral Language Workstation. The test and training modules present spoken and written speech formulas, idioms, and proverbs paired with scenes on video depicting the meaning of these nonliteral expressions. The design of the test and training materials, the hardware configuration, and the development of software programs are given here in detail. Preliminary results from pilot tests indicate that this workstation can be of value to patients with right and left hemisphere damage as well as to other populations.

An innovative and promising interactive video workstation which integrates new concepts of nonliteral language processing with recent advances in computer technology has been developed for the remediation of communication deficits. The purpose of this workstation is to provide a large repertoire of familiar nonliteral expressions in their social context for testing and training nonliteral language skills in a number of different populations. This review describes the workstation and its deployment with adults who have acquired neurogenic language disorders.

The rationale and efficacy of various types of speech-language therapy, traditional and innovative, have been much debated (Alburt & Helm-Estabrooks, 1988a,

DESCRIPTION

The Nonliteral Language Workstation consists of a computer, a video laser disc player, a computer monitor with Touch Window, a television monitor, and an external cartridge hard disk drive (see Figure 1A and 1B). This is a multi-media system which presents verbal expressions in both voice and print in association with short movie scenes in color and sound. The workstation features both a text and a training module, utilized in the pilot study described here in the following format: pretest, training, posttest. The test module (Figure 1A) consists of short verbal and visual movie scenes which were created by the project team and stored as QuickTime movie files. The training module (Figure 1B) plays scenes from a commercial video laser disc, "Movie, North by Northwest." The computer controls the presentation of the verbal stimuli, the movie scenes, and (in the training module) feedback. It also keeps a data file on each subject's performance in the test phases.

The computer was used to develop the workstation in a Macintosh Quadra 950 with 8 megabytes of RAM and 480 MB of storage on an external hard disk drive. A cartridge-style external drive (Focus Duet Removable Cartridge Drive) which holds two cartridges each having 88 MB of storage is also included (see Figure 1A). Two monitors are used in the workstation. A standard 15-inch RGB monitor is used to present the following visual stimuli: printed verbal expressions, QuickTime movies (see Figure 1A), and audio feedback (training module). The monitor is equipped with a TouchWindow (also shown in Figure 1A) that allows a subject to select an answer by simply touching the screen. This eliminates the intimidating array of the keyboard and the need for good hand-eye motor coordination to control mouse-cursor responses, and is a major advantage for patients who are unfamiliar with mouse operations or have impaired motor control response resulting from their neural pathology. The second screen is a standard color TV monitor (Sharp 21-inch, Model 4662ME) which is connected to the video laser disc player (a smaller screen size would be acceptable). This displays the scenes from the video laser disc module used during the training sessions (see Figure 1B).

The software which interfaces the computer with the video laser disc player is available commercially. However, only certain video laser disc players are supported by the software. This compatibility should be checked before any video laser disc or the software is purchased. The Nonliteral Language Workstation uses a program called Video Disc Tool Kit which supports the following video laser disc players: Pioneer PVR-1000, PVR-500, PVR-4000, PVR-9000, PVR-6010; Sony 5000 and 2000; and Hitachi 9550. A Pioneer 4200 was used for this project. A special cable, also commercially available, is needed to connect the remote control port on the Macintosh computer.

Interactive video requires a special type of laser disc called a CAV (angular velocity) or CLV (linear velocity) disc. A CAV disc is recorded with numeric "addresses" for each picture frame. The computer reads the addresses and advances to the next precise control point on the disc possible. Some interactive video systems use CLV discs that do not have numbered frames. These linear discs can be controlled by a machine with somewhat less precision. Time frames, CLV discs are more expensive than CAV discs, partly because they require more disc space; and the range of available CAV movies is smaller.

Nevertheless, the precision and ease of control provided by the CAV disc makes this format particularly desirable for interactive video. Our workstations' main program record CAV disc format.

Each stimulus (literal and nonliteral expressions) was presented in voice as well as in print because many stroke patients have difficulty with one or the other input mode—auditory/verbal comprehension or reading. The voice recordings were done directly through the computer microphone input using the Sound Edit Pro software by Macromedia. The expressions in the verbal expressions were recorded, edited, compressed, and stored as a resource file in a HyperCard stack, where they were available to be edited by the HyperCard program for the testing and training modules.

For the purpose of program development, the Nonliteral Language Workstation computer has a Video-Spiot Nautilus Digital-Video Frame Grabber card installed along with the Adobe Premiere. Video-Spiot converts the analog signal output from the videocassette recorder to a digital signal which can be used by the computer. This card would not be used in the final version of a workstation not used for program development.

For the Test Module, 120 scenes depicting the meaning of nonliteral speech formulas, 30 literal formulas, and 30 proverbs and (98) literal expressions were recorded on one-half inch video tape using a super-VHS recorder. Speech formulas were selected.
because they are the most preserved in severe aphasia; idioms were included because of their association with right hemisphere processing; proverbs were included because of their role in mental status examinations (Graham, 1986; Strub & Black, 1985) and their association with frontal lobe functioning (Benton, 1988; Stuss & Benson, 1986). Each scene corresponds to four expressions, the target (correct) and three (incorrect) foils (phonologically similar, incorrect nonliteral, and semantically opposite). Original video taped scenes using unknown actors (the research team and their associates) were desirable because they were inexpensive to produce and easy to modify. Environmental sounds were recorded using the built-in microphone of the video recorder. Speech was isolated but care was taken to ensure that the script did not cue the meaning of the test scene. ScreenPlay was used to edit the scenes and to conserve space on the hard drive. On the average, the videotaped scenes were edited to about 15 seconds — sufficient time to dramatize the essential meaning of the target utterance, but still requiring too much storage space. For the final product, there was still a need to reduce the QuickTime movie time and storage space through further editing. This was accomplished using Adobe Premiere. The movies were saved to the external drive cartridges. Storing the 120 QuickTime movies, each around 2.9 MB, required approximately 300 MB of memory. The total number of edited scenes required four cartridges with an average of 31 final scenes being placed on one cartridge. The dual cartridge drive allows for as many as 68 scenes to be viewed at one session without changing cartridges. These files are called up as needed by the main computer workstation program which was written in HyperCard.

HyperCard, an authoring language for Macintosh computers, lends itself to the development of multimedia testing and training programs. It interfaces easily with QuickTime movies, video laser disc players, and audio (voice) resource files. In HyperCard, a series of "cards" (i.e., computer screen presentations) are developed with "buttons" (i.e., active spaces on the screen which can be touched to indicate a response) and "fields" (i.e., areas on the screen with script to provide information and instructions). The cards are presented in sequence on the computer monitor. After each response the computer records and stores the subject's score for later data analysis.

The spatial configuration for each test item appearing on the computer screen is as follows: A series of cards having the test information and response
cues are presented on the computer monitor (see Figure 2A). Each card is displayed for approximately 6 inches by 9 inches and appears in the middle of the computer screen. The background color of the card is gray, although future editions of HyperCard (2.2+4) will support color. In the upper left corner of the card, an icon appears that shows a schematic hand on a button, and directly underneath is written the command "Press to Quit." Selecting this button will allow the patient to quit at any time during the test. At the center of the card, a square box appears with an icon showing a hand and forefinger pointing to a button. Directly below the icon is the command "PLEASE TOUCH TO BEGIN" (see Figure 2A, top). Centered in the middle of a test box (1% X 6") at the bottom of the screen, the viewer sees the written expression for that particular test item.

The expression to be judged is also presented in voice. The audio version is selected by the HyperCard program from preprogrammed utterances stored as resource files in the HyperCard stack designated for this purpose. There follows a pause until the test subject touches the "PLEASE TOUCH TO BEGIN" button (see Figure 2A, top). The appropriate movie file is called up from one of the two 88 MB cartridges and the expression is repeated verbally a second time. The movie scene, which is in color, appears in the middle of the card on the computer screen (Figure 2A, middle). Immediately after the score is completed, two different buttons appear toward the center of the screen, presented side by side 2 inches apart above the box that contains the written form of the expression. The left box contains the word "YES" and the right contains the word "NO." The subject is allowed as much time as is needed to decide whether or not the expression matches the movie scene that was just presented and responds accordingly by touching the "YES" or "NO" response boxes. No feedback regarding correctness of performance is provided in the test module. After each response the correctness of the match is recorded by the computer. After a 2 second delay, the next test item appears.

On completion of the testing, summary scores and a record of the responses to each individual item are available for each individual subject, and for all subjects combined. Tabulations pertaining to each individual subject's performance are presented by the computer in four matrices: three present the scores for each expression (item) in the test, and one (see Figure 3) presents a summary of the scores achieved by an individual subject. Tabulations that pertain to the test hypotheses are also presented in four matrices which represent the accumulated scores of all the subjects. Three matrices present the accumulated scores for every expression, and one presents a summary of the scores for all the subjects tested. With the data recorded by the computer in this manner, it is possible to evaluate performance on individual items and item categories by individual subjects and treatment groups. The effect of training on individual items can be assessed by comparing posttest performance on test items included in the training with test items excluded from training.

The training portion of the Nonliteral Language Workstation is composed of 270 expression/scene pairings, taken from the 1989 movie "North by Northwest." The movie has a total running time of 108 minutes. Music, color, and a host of exciting scenarios provide a realistic social context for the subjects. No recordings or changes were made to the movie proper. Instead, short scenes depicting the meanings of nonliteral expressions are accessed by the computer and presented in randomized order (within the constraints of each side of the video laser disc). This use of commercial movies for therapy/education is referred to as "re-purposing." The selected scenes average 15 seconds in length. These were matched with 90 expressions: 30 speech formulas, 30 idioms, and 30 proverbs. Each expression occurs with three different scenes — two of which are appropriate (correct) and one of which is an incorrect pairing. Each class of 30 expressions (formula, idiom, proverb) includes 10 "target" expressions, 10 "full" expressions, and 10 "new" expressions. The 10 "target" expressions appear as correct matches in
the test module; the 10 “foil” expressions are incorrect/nonsensical expressions in the test module; and the “new” nonsense expressions in the training module do not appear in the test module. The purpose of this design is to explore the generalization factor in recovering nonverbal language processes through training by comparing performance on a posttest protocol with that on the pretest with particular reference to trained and untrained items.

The screen format for the training module is similar to that described for the test module, with the exception that the training module program has a built-in provision for feedback regarding the accuracy of the patient’s response (see Figure 2B). When a response is correct, a verbal and written affirmation “Right!” provided with applause is given by the computer. When the response is incorrect, the computer responds verbally saying, “Good try, but that was not the answer.” The phrase and scene are then repeated until the patient makes the correct choice. A ratio of 2 right to 1 wrong was programmed to enhance learning by offering more correct than incorrect expression/scene pairings. The computer keeps a record of how many times it takes to get the correct response. The data for the viewers’ responses on the training module are recorded and displayed by the computer in eight matrices. The first four matrices reflect the scores of an individual subject, and the last four provide the accumulated scores across subjects. The last matrix of each group is a summary matrix of subject performance. The remaining matrices present the scores for each item. Any and all matrices can be printed out and placed in the subjects’ files.

Our goals in the initial pilot portion of the Nonverbal Language Stimulation project were (1) We wished to determine whether normal and neurological subjects would be able to perform the tasks —to make reasonable choices across the array of stimuli— in both the training and testing; we were especially interested in their ability to identify correct “no” responses. (2) How much could the two groups, in a larger study, be expected to differ in performance on the various stimulus types? (3) In the training module, do subjects make competent selections (i.e., do they correct themselves when given feedback about their response?) (4) In neurological subjects, how much of an effect training could be predicted on trained versus untrained stimuli, and if so, on what kinds of nonverbal expressions? (5) We also were interested in perceived subjective effects of the test and training modules, for example, did subjects catch on to the idea of matching a movie scene with an utterance, giving a “yes” (scene and utterance do match) and “no” (scene and utterance do not match)? Did the materials hold their interest? Could they proceed on their own at their own pace through portions of the testing and training? And, importantly, did they appear to “enjoy” the interactive computer task as designed? We can report preliminary results on all of these points, awaiting data from a larger group of subjects.

The subjects were 10 patients with acquired mild to moderate aphasia following a left hemisphere stroke (see Table 1), ranging in age from 50 to 74 years, matched by age and education to 10 normal/normal (normal control) subjects. Subjects ranged from 1 year, 2 months postonset of injury to 18 years, 18 months postonset, with a mean of 8 years, 1 month. Neurological diagnosis was achieved through neurological examination and computed tomography (CT) scan, where available. Language disorder diagnoses obtained using the Western Aphasia Battery were amnestic (n = 6), Broca’s (n = 4), and Wernicke aphasia (n = 1). Aphasia quotient scores ranged from 15 to 97 with a mean of 70. All 10 subjects completed pretesting and training modules. Nine subjects also completed the posttest. Normal control subjects performed the pretest only.

### PRELIMINARY DATA

Summary data are presented here for patients and normal control subjects on nonverbal utterances (targets—correct “yes” responses and foils—correct “no” responses). Test results, shown in Figure 4, suggest that, overall, the aphasic group performed consistently above chance (60% on the nonverbal expressions, with higher scores on correct “yes” than correct “no” answers, as would be expected. Normal control subjects performed better than test subjects on all tasks, with control group mean scores for correct “yes” responses at 94.5% (range = 88.0-98.5%), compared to patients’ mean correct “yes” responses at 82.4% (range = 70.0-95.5%). On correct “no” responses, control subjects’ mean score for correct “no” responses was 86.8% (range = 71.5-95.8%), compared to patients’ mean correct “no” responses at 93.2% (range = 87.6-98.5%). On the training component, patients performed in a similarly orderly manner. Table 2 shows mean responses on each of three trials for each patient. (Recall that the ratio of correct “yes” to correct “no” responses in the training module is 2:1.) Aphasic subjects (n = 10) answered 85.6% of the 270 stimulus presentations correctly on the first trial, and in nearly all cases, they changed a
feedback on the correctness of their responses. We also found that patients, with few exceptions, remained engaged in interest and attention for up to 1 hour on the workstation. Patients were able to proceed much of the time from item to item on their own, even though they were continuously monitored by a clinician in the event help was needed.

**DISCUSSION**

Following this pilot demonstration, we plan to test the efficacy of the Nonliteral Language Workstation in testing and training patients with unilateral as well as left- and right-sided brain damaged subjects over a 3-year period, comparing performance with matched control subjects. Various modifications in test and training formats have emerged from our early experience. In some instances in the test module, a few patients' performance indicated a tendency to lose focus on the task instructions (i.e., answer "yes" if the scene matches the preceding utterance and "no" if it does not). These patients occasionally reinterpreted the task to mean it was to evaluate the validity of the utterance rather than to compare the statement to the action. For example, shown the proverb "Crime doesn't pay" followed by a mismatching scene, some patients responded "yes"; their comments then indicated that they were expressing an agreement with the moral teaching of the proverb. This occurred less frequently in the training module, where, with the feedback provided, the problem was not as evident. In the revised workstation format, this problem will be addressed by reminding the patient of the instructions for each item by asking "Did you see..." after the movie scene and before the expression is presented for the second time. To provide a better training opportunity for acquiring or relearning nonliteral utterances, the training module has been expanded to include two additional units, for a total of approximately 20–25 hours of training time. For test and training items, patients will have the option of seeing the scene a second time if they wish. In addition, performance on familiar nonliteral utterance recognition will be tested at three points at the beginning and at the end of the study, and at regular intervals during the training period, to compare a proposed trajectory of improvement to a baseline score. Two tests of social efficacy, one using videotaped structured interviews involving use and understanding of nonliteral utterances in "natural" settings before and after nonliteral training, and a second using judgments from family members of communicative functioning before and after training have been devised. Minor modifications to utterance/scene matches in the existing test and training modules have been made following evaluation and analysis of performance by our 10 pilot subjects.

Once results from a controlled study of performance using the Nonliteral Language Workstation have been acquired and analyzed, and modification to the materials are completed, our intention is to make the...
program available at a relatively modest cost, using CD ROM and video laser disc technology. Hardware requirements will be a Macintosh computer with sufficient memory, a CD ROM player, a video monitor. Software requirements will be the workstation HyperCard program, QuickTime and a video disc control program. Users will have the ability to choose among all 3 training modules ("Northwest Yosemite," "Princess Bride," or "Some Like It Hot"). Besides obvious relevance for individuals with acquired communications disorders and nervous system damage, other populations potentially suitable for the workstation training include speakers of English as a second language, developmentally disabled individuals, and persons with hearing impairment.

We hope our presentation of this description will stimulate clinicians and researchers to use and advance computer-based interactive video disc technology.

Acknowledgements  We appreciate the help of Ray Bell, Kristy Kless, Steve Treiman, Evelyn Wallner, and Vitali Sposeti in program development. Actors in the test scenes, in addition to the authors of this article, included Cathy Bowsher, Orian Broderick, John Burnett, Bart Curraway, Dona Czycki, J. Daniel Hines, Wayne Hanson, Joel Hurvitz, Mary Poterwin, Edwina Pierce, and Carol Stackman. Gary Ono assisted in filming scenes for the test modules and Mark and Julianne Hershke assisted with the voice samples for test and training modules. Helen Yule and Philip Murray helped with patient testing. Permission to use the movie "North by Northwest" on "ON" videotape was granted by Voyage Company in Santa Monica, California and New York City. This project was supported by the Department of Veterans Affairs Rehabilitation Research and Development Pilot Proposal No. C91-30747.

Address correspondence to Diana Van Lancker, Ph.D., VA Outpatient Clinic 13(6), 303 E. Temple St., Los Angeles, CA 90012-3328.

REFERENCES

Goldjahr, M. F. (1976, November). The use of video confrontation in teaching Amherst to adult aphasic patients. Presentation at the American Speech-Language-Hearing Association Convention, Houston, TX.
Hall, E., & Goldjahr, M. (1986). Interactive video for the aphasic patient to stimulate speech production. -- Tell me what you have [Software program]. Sepulveda, CA: Veterans Administration, Department of Speech Pathology.