Paper

"What did he say?": Speech Output in Preschool Software

Erik Strommen
Children's Television Workshop
New York

Abstract

Advances in technology have made it possible to include speech in electronic learning materials for small children. Little is known, however, about how young children respond to speech as a component of technology. Based on the results of a series of studies of Sesame Street interactive learning games tested with over 150 children aged 3 to 5, guidelines for evaluating speech as a feature of computer software are presented. Preschoolers responded best to speech segments that were short, contained concrete language, and were not stereotyped or repetitive. Information presented solely via speech was much less effective than when speech segments were combined with supplemental, task-related visual information that reinforced their content. Guidelines derived from these results can be used for judging the quality and effectiveness of speech included in interactive technologies such as computer software and videodiscs.

Recent and rapid changes in technology have made speech a possible feature of interactive electronic learning materials. To date, "talking machines" have most commonly been studied as aids in reading and writing tasks (c.f. Parham, 1988; Reitsma, 1988; Weiner, 1991). However, the potential uses of speech in applications for small children, not as an adjunct to task content but as an effective form of interface, have been given very little study. Because young children cannot benefit from onscreen text messages, spoken requests and feedback can potentially empower young children's learning in complex tasks that currently require verbal guidance and explanation by an adult. Such uses of speech as an interface, however, must be carefully considered. There is abundant evidence that young children have difficulty decoding spoken language in everyday contexts (Beal and Belarad, 1990; Hargrove and Panagos, 1982; Markham, 1979), so issues of how best to integrate speech into interactive multimedia technologies must be carefully considered.

The selection of the nature and content of speech in computer systems is especially important because unlike human speakers, who can tailor their speech to their listeners in an endless variety of ways, the affordable computer of the near future will typically only be able to repeat a limited variety of stock phrases. How long these phrases should be, how they should be worded, and where they should be placed as part of the interactive task, are issues of vital importance if speech is to be used as an effective, engaging component of educational materials for pre-literate children.

The present paper summarizes the results of a series of seven formative studies on computer learning materials for young children that employ speech as an integral component. The goal of this summary is to describe how children reacted to speech as a component of the computer activities, and to make a first attempt at identifying some guidelines for the optimal length, content, and placement of speech in interactive learning materials. These guidelines can serve designers of software using speech, but more importantly they can be used by teachers and others who need to evaluate new products for their suitability with their children.

Method

Participants

One-hundred and fifty children ranging in age from 34 to 62 months, drawn from the greater New York area, served as subjects. Equal numbers of boys and girls were tested.

Materials

Prototypes for seven educational software games for preschoolers, using Sesame Street characters and settings, were used in the present study. These games were designed to teach a variety of familiar preschool skills, such as rhyming, simple arithmetic, etc. Each was identified with a specific Sesame Street character who served as the game's "host" in an animated setting similar to familiar Sesame Street settings (Ernie and Bert's house, Big Bird's nest, etc.). The scripts for the games were written based on conventions from the Sesame Street television series, and were designed to cover the possible alternatives that crop up in interactive tasks: initial instructions, feedback messages, prompts, and end-of-game comments. The spoken portions of the games were recorded by the actual Sesame Street characters, and then incorporated into the games via digitized speech systems available on standard IBM DOS computers.

Six of the games were designed with the same broad format. To play the game, children placed a cursor on the

muppet and pressed a button to start game play. The character then explained the game (the muppet animated during spoken segments), and play commenced. In four of the games, the spoken instructions contained crucial game information (words to find rhymes for, attributes to match with objects, verbal math problems, etc.); in two others, the spoken feedback from responses contained the crucial information. All six of these games also relied almost exclusively on spoken dialogue from the characters to convey incorrectness/correctness of responses, repeat key information, and prompt the child. The response structure was also the same for all six games: The child was required to move the cursor to objects and select them as the basic method of game play. If, for example, Ernie asked the child to "Find things that are blue," the child would move the cursor to an object and click on it; a correct response received a reward animation and Ernie saying "My airplane! It's blue! Good going!" and a randomly varied prompt encouraging the child to find another object that is blue. The child would go on to click on the next object, and so on.

The seventh game tested was slightly different, and involved the child "phoning" different muppets and listening to verbal segments reinforcing basic counting, addition, and initial-sounds concepts.

Procedure

The children, with an adult Researcher as "helper," all played multiple rounds of each game, with each child playing only one game each. The children's interactions with the games were videotaped, and responses to the games, especially to the verbal materials, were scored for the presence of confusion or hesitation, need to repeat spoken segments, requests for assistance from the adult, and affective reaction. The children's spontaneous comments were also recorded.

Results and Discussion

While the results vary somewhat depending on the particular game being described, six consistent findings relating to the use of speech recurred across all seven of the games tested.

1. It appears that instructions lasting 20 seconds or less, and feedback messages and prompts lasting less than ten seconds in length, are ideal. Children tended to ignore, or fail to attend to, important content in a long speech stream. While adults had no trouble picking up the important information, the children often missed it entirely, responded only to the last few words, or reacted only to very simple, concrete language such as "...press the button!". Such performances may reflect a bias by young children toward expecting that communicative messages are meant to prompt an action (Ackerman, 1981; Shatz, 1978).

2. Children tended to respond to the computer as if the character was actually speaking to them, particularly when action words or commands were spoken. If the character queried "Are you ready?" children invariably answered "Yes!" as if they could be heard. Such prompts proved to be very beneficial for cuing children that important verbal information was about to be presented, as did the prompt "Watch this!" for cuing the child to important visual information, and "move your cursor to..." for prompting game action.

3. The children responded to the inflection of the language as much as to the content of what was said. Exuberant language on correct response feedback, such as "That's right!" or "Hooray!" was highly reinforcing for children, and low-key or questioning affect ("My hat?") was effective for feedback to an incorrect response. It was also found that negative feedback, given exuberantly, was often misinterpreted as implying correctness even though the actual content of the speech belied the inflection.

4. Spoken prompts stating the names of objects as they are highlighted can facilitate game play. In several of our games, children were required to select objects based on their name (finding a number six, locating a rhyming object, etc.). Having the object's name spoken if the child stops the cursor on it for a few seconds proved to be an excellent way to support the performance of children who could not identify the objects in question.

5. Children are sensitive to stereotyped or repetitive speech. Children frequently asked "Why does he always say that?" or "I know already!" after a few rounds of game play. They often begin to talk over, or ignore, phrases they had heard repeatedly even if they did not understand the content of the message. Having speech-based systems employ a variety of phrases for the same situation (having four varied error messages, for example) would be beneficial.

6. Visual cues supplementing verbal material, particularly in tasks where the child is required to remember key information, enhanced game appeal and useability. We found that simply presenting problems verbally was not nearly as effective as combining this material with redundant onscreen information. Children drew on both modalities in recalling the information they needed, and used the visual cues as a reminder when they accidently forgot or did not attend to the verbal material.

Conclusions

The use of speech in interactive products constitutes a new area for educational psychologists and teachers. The findings from the present studies suggest several concrete criteria that can be used to determine the effectiveness of speech as a part of preschool software. They are applicable to all interactive media, not just computer software; any multimedia technology (videodiscs, interactive compact discs, etc.) intended for pre-

school use that employs spoken language can be evaluated using the following guidelines.

1. The length of the spoken segments in the program should be brief. Long, expository segments tax young children's listening abilities and attention spans.

2. The language used in the program should be as concrete as possible, and should direct the child to undertake a specific action. For example, if the task is to match animals with their appropriate sound, an ambiguous error phrase such as "Try it again," should be avoided; a phrase like "Pick a new animal," is better.

3. The speech should be as natural as possible, using appropriate inflection wherever possible. Positive affect in messages such as "That's right!" and neutral or questioning affect on wrong answers seem to enhance children's ability to understand the program, as well as help motivate task performance.

4. Programs employing a varied number of messages for the same situation are preferable to those with only one, repeated message. In our studies, children quickly tired of repeated phrases and came to ignore them, defeating their purpose.

5. The language in the program should be supplemented by visual material. The speech segments should have a concrete relationship to objects or events happening on the screen. Such redundancy between spoken and visual elements helps children to "figure out" what to do, if they do not fully grasp the verbal material. In our math game, for example, the verbal problems are actually performed as they are verbally described. This allows children who may initially fail to process the spoken material to nonetheless play along, and it also encourages children to connect the language used to the situation at hand.

References

Ackerman, B. (1981). Performative bias in children's interpretations of ambiguous referential communications. *Child Development, 52*, 1224-1230.

Beal, C. R. & Belarad, S. C. (1990). The development of message evaluation skills in young children. *Child Development, 61*, 705-712.

Hargrove, P. M. & Panagos, J. M. (1982). Role saliency cues and children's sentence comprehension. *Journal of Psycholinguistic Research, 11*, 217-228.

Markman, E.M. (1979). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 986-992.

Parham, C. (1988). Computers that talk. Classroom Computer Learning, 8(6), 26-31.

Reitsma, P. (1988). Effects of guided reading-while-listening and independent reading with computer-based speech feedback. *Reading Research Quarterly, 23*(2), 219-235.

Shatz, M. (1978). On the development of communicative understandings: An early strategy for interpreting and responding to messages. *Cognitive Psychology, 10*, 271-301.

Wiener, R. (1991). Computerized speech: A study of its effect on learning. *T.H.E. Journal, 18*(7), 100-102.