2005

A Dual Modal Presentation of Network Relationships in Texts

Shuang Xu
DePaul University, sxu@cti.depaul.edu

Xiaowen Fang
DePaul University, xfang@cti.depaul.edu

Jacek Brezinski
DePaul University, jbrezinski@cti.depaul.edu

Susy Chan
DePaul University, schan@cti.depaul.edu

A Dual-Modal Presentation of Network Relationships in Texts

Shuang Xu
DePaul University
sxu@cti.depaul.edu

Xiaowen Fang
DePaul University
xfang@cti.depaul.edu

Jacek Brzezinski
DePaul University
jbrezinski@cti.dedpaul.edu

Susy Chan
DePaul University
schan@cti.depaul.edu

ABSTRACT

Based on Baddeley's working memory model, this research proposed a method to convert textual information with network relationships into a "graphics + voice" representation and hypothesized that this dual-modal presentation will result in superior comprehension performance and higher satisfaction than pure textual display.

A simple T-test experiment was used to test the hypothesis. The independent variable was the presentation mode: textual display vs. visual-auditory presentation. The dependent variables were user performance and satisfaction. Thirty subjects participated in this experiment. The results indicate that both user performance and satisfaction improved significantly by using the "graphic + voice" presentation.

Keywords

Multi-modal interfaces, information presentation, human attention, working memory, network relationship.

INTRODUCTION

The benefit of delivering information across different sensory modalities is often justified by the independent nature of multi-modal information processing, which assumes that there will be no interference between tasks and thus no degradation in performance (Cook, Crammer, Finan, Sapeluk & Milton, 1997). However, research in cognitive psychology shows that visual and auditory perceptual processing is closely linked (Eimer, 1999). Problems related to memory and cognitive workload were found in current applications with voice-based interface (Cook et al., 1997). For instance, mental integration of disparate information from different modality channels causes a heavy cognitive memory load. As transient auditory information, speech presentation may impose a greater memory burden. Also, switching attention between modalities may be slow and have a high cost.

The objective of this research is to develop a dual-modal interface that: (1) minimizes the interference in information processing between visual and auditory channels; and (2) improves the effectiveness of mental integration of information from different modalities. This study focuses on the dual-modal presentation of textual information that describes network relationships. Results of this study will help to address the usability problems associated with small-screen computers and the mobile information access via handheld devices.

LITERATURE REVIEW

To develop an effective dual-modal information presentation, we have examined prior research in human attention, working memory, visual and auditory interfaces, and graphical representation of texts.

Human Attention

The interference encountered during multi-modal information perception stems from the allocation of limited attentional resources to concurrent sensory information processing. Researchers have proposed several theoretical attention models to explain the mechanism of resource allocation: bottleneck models, resource pool models, and multiple resource pool models. Bottleneck models (e.g., Broadbent, 1958) specify a particular stage in the information-processing sequence at which the amount of information that humans can attend to is limited. In contrast, resource models (e.g., Kahneman, 1973) view attention as a limited-capacity resource that can be allocated to one or more tasks, rather than as a fixed bottleneck. Among various attention models, multiple-resource models (Navon & Gopher, 1979; Wickens, 1980 & 1984) propose that there are several distinct subsystems, each having their own limited pool of resources. The multiple-resource models assume that two tasks can be efficiently performed together to the extent that they require separate pools of resources.

Allocation of attentional resources during complicated time-sharing tasks across multiple modality channels has long been of interest to cognitive psychology researchers. Research shows that introducing auditory channel into prototypes of civil and military cockpits has resulted in degraded performance (Cook et al., 1997). One explanation is that the total amount of attentional resources is limited. When demanded simultaneously by multi-modal information processing tasks, resources allocated to non-dominant channel decrease, as compared to single-modal information processing. Another explanation is that mental integration of different multi-modal information causes a heavy cognitive load in working memory. Performance will degrade if this integration is critical to understanding information received from different sensory channels.

Although the details of attention allocation mechanism are still under exploration, past research has positively confirmed that human's cognitive resources for attention are relatively limited. Therefore, concurrently performing two or more tasks generally results in a drop of performance for one or all tasks. The amount of shared resources is one of the factors that decide how well people can divide their attention between tasks (Wickens, Gordon, & Liu, 1998). A large body of research work has shown that people are generally better at dividing attention cross-modality, typically on visual and auditory information, than within a single modality channel (Dubois & Vial, 2000; Tremblay & Proteau, 1998; Coull & Tremblay, 2001; Stock, Strapparava, & Zancanaro, 1997; Treviranus & Coombs, 2000). Other researchers (Polson & Friedman, 1988; Wickens & Liu, 1988) show that imagery/spatial and verbal processing demands distinct resources, whether occurring in perception, central processing, or responding. The following literature review in working memory will explain why processing spatial and verbal information concurrently does not cause competition of cognitive resources.

Working Memory

Baddeley (1986) proposes a working memory model that depicts three components: central executive, visuo-spatial sketchpad, and phonological loop (see Figure 1)

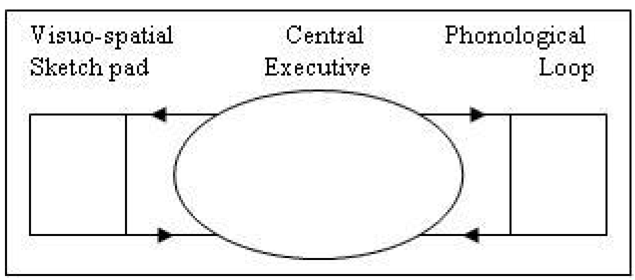

Figure 1. Baddeley's Working Memory Model (1986)

According to this model, human working memory contains two subsystems for storage: phonological loop and visuo-spatial sketchpad. Acoustic or phonological coding is represented by the phonological loop, which plays an important role in reading, vocabulary acquisition, and language comprehension. The visuo-spatial sketchpad is responsible for visual coding and handling spatial imagery information in analog forms. The phonological loop and visuo-spatial sketchpad are able to simultaneously hold verbal and imagery information without interference. Central executive is the control system that supervises and coordinates the information retrieved from the two storage subsystems for further integration. Baddeley's model has been confirmed by many studies. For example, Mousavi, Low, and Sweller (1995) show that students' performance improved significantly when the verbal representation and image representation of a geometry problem were respectively presented in auditory and visual modes. They further suggest that distributing relevant information in visual and auditory modalities might effectively increase working memory.

Visual and Auditory Information Presentation

When comparing visual and auditory information representation, prior research shows that voice is more informal and interactive for handling the complex, equivocal and emotional aspects of collaborative tasks (Chalfonte, Fish, & Kraut, 1991). As Streeter (1998) indicates, universality and mobile accessibility are major advantages of speech-based interface, whereas its disadvantage is the slow delivery rate of voice information. Archer, Head, Wollersheim, and Yuan (1996) compared the user's preferences and the effectiveness of information delivery in visual, auditory, and visual-auditory modes.

They suggest that information should be organized according to its perceived importance to the user, who should also have flexible access to information at different levels of abstraction.

Multi-modal interfaces have been widely used as a support of collaborative work, as well as in teaching systems. Researchers (Nardi et al., 1993) indicate that the integration of video information and other data sources (e.g., aural input, time-based physical data, etc.) help surgeons choose the correct action and interpretation during remote medical operations. Research on interaction between sound, written words, and the image of objects shows that when different sources of information are integrated, a learner's cognitive overload remains light and does not limit learning (Dubois & Vial, 2000). Stock, Strapparava, and Zancanaro (1997) show that hypertext and digital video sequences help users explore information more effectively. By exploring the integration of captioning, video description, and other access tools for interactive learning, Treviranus and Coombs (2000) demonstrated how to make the learning environment more flexible and engaging for students. Dubois and Vial (2000) suggest that several factors affect the effectiveness of integration of multi-modal information. These factors include not only the presentation mode, the construction of co-references that interrelate to the different components of the learning materials, but also the characteristics of the task.

Graphical Representation of Texts

To design an effective dual-modal information presentation based on Baddeley's working memory model, it is important to understand how textual information should be converted to imagery/graphical and verbal representations. Mayer's empirical studies (1989, 1990 & 1991) show that an effective illustration model should use images or diagrams to reorganize and integrate the acquired information. The illustration must be able to guide user's selective attention towards the key items in the presented information. These key items include not only the major entities (such as objects, states, actions, etc.), but also the relationship among them. Dual coding theory (Paivio, 1986 & 1971) further predicts that concrete language should be better comprehended and easily integrated in memory than abstract language because two forms of mental representation, verbal and imagery, are available for processing concrete information.

Schema (or script, frame) has been widely used in knowledge representation (Proctor & Van Zandt, 1994; Johnson-Laird, 1983, 1989). Schemas are frameworks that depict conceptual entities, such as objects, situations, events, actions, and the sequences between them. Schemas not only represent the structure of a person's interest and knowledge, but also enable a person to develop the expectancy about what will occur. Knowledge-based systems were often represented as a network of interrelated units. Such networks are commonly presented to the user as a diagram of nodes connected by lines. These diagrams have provided a powerful visual metaphor for knowledge representation. Travers' study (1989) indicates that when this diagrammatic representation of knowledge structures matches people's virtual metaphor, it improves their information comprehension.

Based on the above discussions, we propose a dual-modal information presentation that presents the internal relationship depicted in texts as network diagrams, and presents the remaining textual information as voice message. The following section discusses this dual-modal presentation in greater details.

PROPOSED DUAL-MODAL INFORMATION PRESENTATION

Based on Baddeley's working memory model, it is assumed that the effectiveness of human information processing can be improved if the verbal representation and the imagery/graphical representation of certain textual information are presented via auditory and visual output, respectively. As shown in Figure 2, if the verbal presentation of the original textual information is presented via auditory channel, the verbal information will be temporarily stored in the auditory sensory register, then sent to and processed in the phonological loop in working memory. Meanwhile, information perceived from the graphical presentation will be stored in the visual sensory register and then transferred to visuo-spatial sketchpad. Verbal and graphical information that are concurrently stored in working memory could be respectively retrieved from the phonological loop and visuo-spatial sketchpad, and then integrated by the central executive for comprehension.

In this proposed dual-modal presentation (see Figure 3), network relationship contained in texts will be extracted and presented in diagrams. The remaining textual information will be delivered through the auditory channel. The following hypothesis is proposed to test the effectiveness of this dual-modal information presentation.

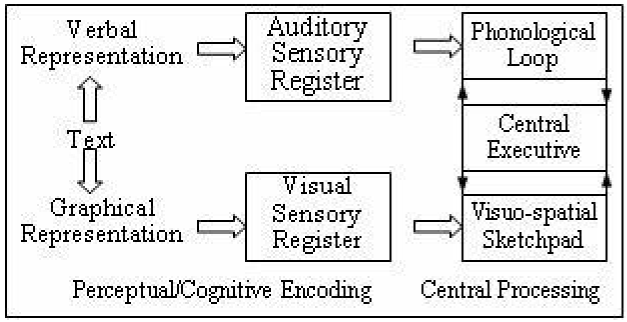

Figure 2. Splitting Textual Information

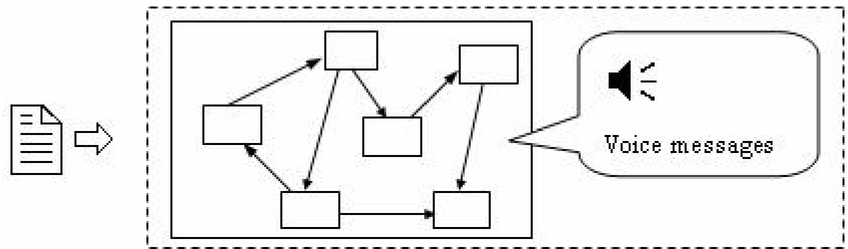

Figure 3. Proposed Dual-modal Presentation

Hypothesis: The dual-modal presentation of network relationship will improve user's comprehension of information and result in higher user satisfaction as compared to pure textual display.

According to Baddeley's model, pure visual display of textual information will be processed entirely in the phonological loop. Non-speech verbal input must go through a sub-vocal rehearsal to be converted to speech input and temporarily saved in the phonological loop of working memory before further processing (see Figure 4).

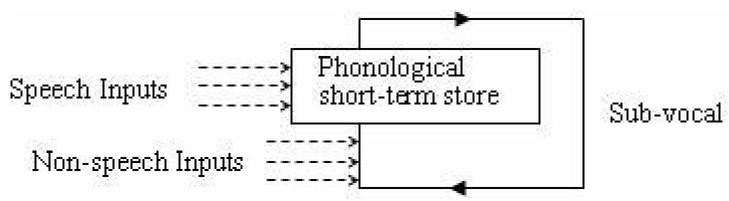

Figure 4. Structure of Phonological Loop

In the proposed dual-modal presentation, the graphical information might be perceived and held in the visuo-spatial sketchpad while the speech input is perceived and directly stored in phonological loop. Therefore, by concurrently utilizing the two subsystems in working memory to process the same amount of information, a reduced cognitive workload is expected during information processing. Research in human attention shows that many voice-based interfaces caused degraded comprehension performance because of the interference between disparate information perceived from visual and auditory channels. In the proposed dual-modal information presentation, the graphic and voice information are derived from the same textual information, and should be highly relevant and complementary to each other. Therefore, the mental integration of the visual and auditory information will be easier during comprehension.

With a reduced cognitive workload and easier mental integration in working memory, the proposed dual-modal information presentation may significantly improve the effectiveness of users' information comprehension.

METHOD

This study used sample analytical tests from the Graduate Record Examination (GRE) for the experiment because these tests are designed to measure subjects' analytical comprehension and reasoning skills without assessing specific content knowledge. An experiment Web site was built to present the GRE analytical tests. The task was to perform GRE analytical tests through the experiment Web site. The following sections discuss the subjects, the tasks and experiment system, independent and dependent variables, and the procedure.

Subjects

Thirty (30) participants were recruited from a Midwest university. The participants included undergraduate students, graduate students, staff, faculty members, and alumni. These participants had a wide range of background in terms of age, ethnicity, computer experience level, and Internet usage. Participants were randomly assigned to one of the two groups: textual display (T-mode) group and "graphics + voice" display (GV-mode) group. As shown in Table 1, participants in both groups shared similar profiles.

Group	Gender		Language		Average Age
	Males	Females	Native English Speakers	Non-Native English Speakers	
GV-mode	6	9	13	2	29.6
T-mode	6	9	13	2	30.9
* Complete summary of subjects' demographic characteristics is available upon request.					

Table 1. Profile of Subjects

Experiment Design and the Tasks

The experiment design was a simple t-test. Subjects performed two GRE analytical tests. The first test served as the pre-test for estimating each individual's analytical comprehension, reasoning, and test-taking skills. The second test was presented in T vs. GV mode for comparing the differences of these two presentation modes.

Recruited participants were evenly and randomly distributed into two treatment groups. Their background information was recorded to ensure a controlled balance in demographic characteristics between groups. Because individual participants' analytical comprehension and reasoning skills may vary greatly and such skills could affect their performance in the experiment, an independent GRE analytical test was conducted as a pre-test to estimate a participant's skills before the actual experiment task was performed. In this 15-minute pre-test, all information was visually presented as texts on a Web page for both groups. The estimate of analytical comprehension and reasoning skills or possibly other test-taking skills from the pre-test was used as a covariate in the analysis of the experiment task performed later.

Independent and Dependent Variables

The only independent variable was information presentation mode. There were two treatments: Text (T) mode and Graphic + Voice (GV) mode. In the T mode, all information was visually presented as texts on a Web page. In the GV mode, the original textual information was split into a network diagram and speech output. Three faculty members with rich teaching experience were asked to manually convert the GRE analytical tests into a graph + voice presentation according to the proposed method (see Figure 3). Only the internal relationship and related entities were converted into graphics. An example of the information presented in the T mode and GV mode is shown in Figure 5.

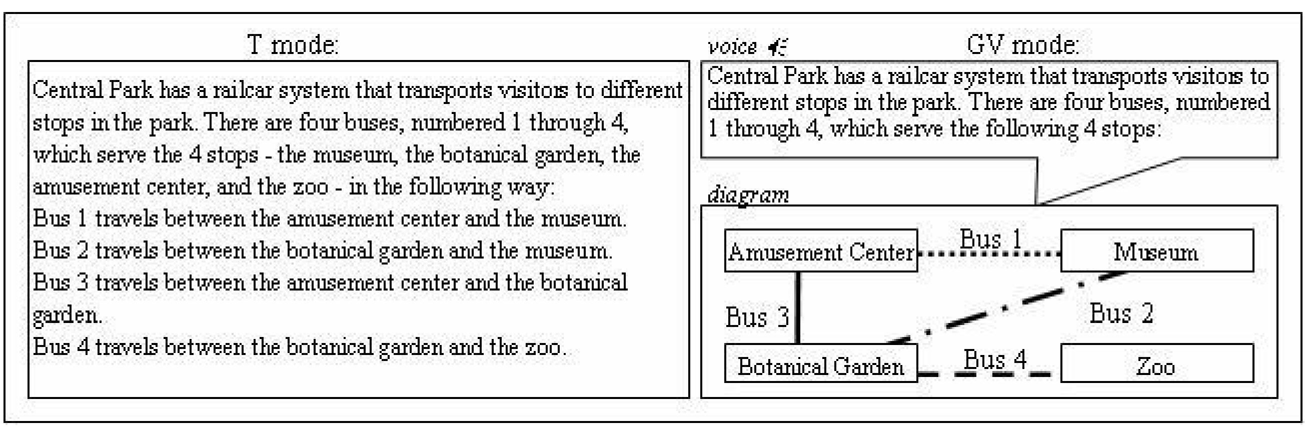

Figure 5. An Example of Display Modes

The two dependent variables were users' performance and satisfaction. Subjects' performance was measured by the number of correctly answered questions within a 30-minute period. The task started when the first analytical problem was presented on the screen, and ended when time was up. User satisfaction was measured by a satisfaction questionnaire using a 7-point Likert scale. Based on the Technology Acceptance Model (TAM) (Davis, 1989; Koufaris, 2002), this satisfaction questionnaire was designed to measure subjects' perceived usefulness and ease of use of the two interfaces. In addition, one question was added to measure user's general satisfaction.

Procedure

Each subject was asked to sign a consent form before participation. During the training session, each subject filled out a background questionnaire and then the experimenter described the tasks included in different groups. A sample problem was used to explain the interface, browsing rules, time limit, graphic notations (for GV-mode group), and voice control (for GV-mode group). Subjects were allowed to ask questions during the training period. They could spend as much time as they needed in the training session. Subjects were encouraged to answer as many questions as they could during the two analytical tests. They were allowed to browse back and forth within each problem to find or correct their answers. Subjects could click the submit button to move on to the next analytical problem after they finished the current one, but they could not go back to the previous problem. For the GV-mode presentation, pre-recorded voice information was automatically played when the Web page was loaded on the screen. Subjects could use controls on the screen to replay voice messages. Subjects were allowed to take breaks before or after the timed tests. Upon completion of these two tests, the subject was asked to fill out a satisfaction questionnaire. There was no time limit for this satisfaction survey. Table 2 presents the experiment procedure.

	Pre-Test (15 min)	**Task (30 min)**	**Satisfaction Survey**
T-mode Group	Solve problems in T-mode presentation	Solve problems in T-mode presentation	Satisfaction questionnaire
GV-mode Group	Solve problems with T-mode presentation	Solve problems with GV-mode presentation	Satisfaction questionnaire

Table 2. Experiment Tasks and Procedure

The following information was collected during the experiment and saved into a database:

- Subjects' background information.
- Subjects' answers to analytical problems in Pre-Test and Task, and time spent on each problem.
- Subjects' response to the satisfaction survey.
- Subjects' online activities (e.g., manipulating voice messages, changing answers, etc.).

RESULTS AND DISCUSSION

The hypothesis postulates that the dual-modal presentation will improve user's comprehension of information and will result in higher user satisfaction as compared to pure textual display. The dependent variables were the number of correctly answered questions, and user satisfaction. A simple T-Test between T and GV groups was used to test the hypothesis.

Table 3 presents the descriptive statistics of performances of the two groups in the pre-test and the task.

	Pre-Test				Task			
	T-Mode		GV-Mode		T-Mode		GV-Mode	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Number of correctly answered questions	4.4	1.35	4.9	2.84	9.5	3.02	23.1	8.55
Accuracy	0.466	0.2140	0.617	0.2294	0.487	0.1663	0.757	0.1794

Table 3. Descriptive Statistics of Performances in Pre-Test and Task

A logarithm transformation was performed to ensure constant variances. The result of the analysis of covariance is shown in Table 4. A significant difference was found between the performance in the T group and GV group ($p<0.0001$). This result suggests that on the average participants in the GV group correctly answered more than twice of the questions that were correctly answered by participants in the T group. Table 4 also lists an additional comparison between the accuracy of participant performance (defined as the number of correctly answered questions divided by the total number of answered questions) in the two groups. Accuracy was not a dependent variable in this experiment. Accuracy was measured for precaution because if a participant guessed a lot during the performance, he/she might be able to correctly answer more questions, with a much lower accuracy. The average probability of hitting the correct answer would be 0.20 (5 choices for each question). The significant difference in accuracy indicates that the greatly improved performance in the GV group was not the result of random guessing. In fact, the graphic + voice presentation increased the GV group's accuracy rate by about 80%, compared to the accuracy on the textual presentation. These results strongly indicate that participants in the GV group did perceive and process information more effectively within the given time.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	4.73041	4.73041	92.50	<0.0001
Pre-Test	1	2.312190	2.312190	45.21	<0.0001

Table 4. Results of Analysis of Covariance

User satisfaction was measured by a questionnaire derived from the technology acceptance model (TAM) (Davis, 1989). According to this model, perceived ease of use and perceived usefulness are predictors of a user's intention to adopt a new technology. If a user feels more satisfied with a technology's ease of use and usefulness, it is more likely that he or she would adopt the technology. Therefore, it is reasonable to use perceived ease of use and perceived usefulness as surrogates of satisfaction. In the satisfaction questionnaire, six questions measure the perceived usefulness, seven questions measure the perceived ease of use, and one question measures the user's general satisfaction. These questions were drawn from Davis's study (1989). The question on user's general satisfaction states: "In general, I am satisfied with visual-auditory (or textual) presentation of the problems." The total scores of the items measuring perceived ease of use and perceived usefulness respectively were calculated and used in the analysis. Table 5 presents the Cronbach's α value, mean, standard deviation, and t-test results of perceived ease of use, perceived usefulness, and general satisfaction. The high Cronbach's α values for perceived ease of use and perceived usefulness suggest that the questionnaire was reliable and valid. Significant differences between the T and GV modes were found in perceived usefulness ($t(28)=12.09$, $p<0.0001$), perceived ease of use ($t(28)=8.40$, $p<0.0001$), and general satisfaction ($t(28)=9.58$, $p<0.0001$), as shown in Table 5. Six items were used in the questionnaire to measure users' perceived usefulness and seven items were used for perceived ease of use. Therefore, the average score of perceived usefulness in the GV group was 39.3/6≈6.6, and 16.8/6=2.8 in the T group. The average score of perceived ease of use in the GV group was 43.4/7=6.2, and 26/7≈3.6 in the T group. In the 7-point Likert scaled questionnaire, the neutral score is 4. These results strongly indicate that participants in the GV group had much higher satisfaction with the information presentation in terms of perceived usefulness and ease of use.

Variables	Cronbach's α	GV group (n = 15)		T group (n = 15)		t	Pr > |t|
		Mean	Std.	Mean	Std.		
Perceived Usefulness	0.95	39.3	3.13	16.8	6.50	12.09	<0.0001
Perceived Ease of Use	0.90	43.4	3.98	26.0	6.97	8.40	<0.0001
General Satisfaction	N/A	6.67	0.49	2.53	1.60	9.58	<0.0001

Table 5. Comparison of User's Satisfaction between GV and T Groups

Therefore, our hypothesis was fully supported by this experiment. The results agree with prior research findings. Based on the multiple-resource human attention model (Wickens, 1980, 1984), two tasks can be performed together more efficiently to the extent that they require separate pools of resources, such as different modalities. According to Baddeley's working memory model (1986), tasks using different subsystems (visuo-spatial sketch pad and phonologic loop) in the working memory should not interfere. While users are visually browsing information, they may be able to receive brief information from the auditory channel. The results of this experiment are also consistent with Mousavi's study (Mousavi et al., 1995), where students' performance was significantly improved when the verbal representation and image representation of a geometry problem were respectively presented in auditory and visual mode. Our findings suggest that when texts describing network relationship are presented in diagram with highly relevant additional voice output, users are able to concurrently perceive the visual and auditory information without interferences or distractions. The proposed dual-modal information presentation significantly improves users' comprehension performance, as well as their overall satisfactions.

CONCLUSIONS

In this study, we aimed to improve the effectiveness of presenting information using multiple modalities with minimum interference. A dual-modal information presentation was proposed and tested through controlled experiments. Findings from this study suggest the following: (1) Users could concurrently integrate perceived visual (diagrams) and auditory (voice messages) input without interference; (2) Highly relevant speech information might facilitate user's understanding of diagrammatic information; (3) The distribution of cognitive workload across modalities might demand less mental effort required in information comprehension, as compared to single-modality presentation; and (4) Users might have higher satisfaction due to the alleviated working memory load during information processing.

These findings have profound implications to future research in multimodal interface design. Multimodal interfaces are specially promising to mobile applications due to the nature of wireless technology. Mobile devices have two main constraints: small screen size and their mobile usage. Compared to desktop or laptop computers, mobile devices typically have a very small screen on which only a very limited amount of information can be presented. When the device is used on the move, it makes reading textual information much more difficult. Multimodal interfaces will be able to address these constraints by delivering information through multiple sensory modalities such as visual and auditory channels.

In this experiment, we focused on the dual-modal presentation of network relationship in texts. This might limit the scope of application of our findings. In the future, we will continue to explore other ways of converting texts to graphics. Currently, the experiment did not examine individual effects of the two components (diagrams and voice) in the proposed dual-modal presentation. In future studies, we will further investigate how the graphics and voice affect user performance and satisfaction respectively.

REFERENCES

1. Archer, N., Head, M., Wollersheim, J., & Yuan, Y. (1996). Investigation of voice and text output modes with abstraction in a computer interface. *Interacting with computers*, 8, 4, 323-245.
2. Baddeley, A. D. (1986). Working memory, Oxford University Press, New York.
3. Broadbent, D. (1958). *Perception and communication*, Pergamon Press, London.
4. Cellario, M. (2001). Human-centered intelligent vehicles: Toward multimodal interface integration. *IEEE intelligent systems*, July, 78-81.
5. Chalfonte, B., Fish, R., & Kraut, R. (1991). Expression richness: a comparison of speech and text as media for revision, in *Proceedings of CHI'91,* April 27-May 2, New Orleans, Louisiana, USA , ACM Press, 21-26.
6. Chan, S., Fang, X., Brzezinski, J., Zhou, Y., Xu, S., & Lam, J. (2002). Usability for mobile commerce across multiple form factors. *Journal of Electronic Commerce Research, 3*, 3, 187-199.
7. Cook, M. J., Cranmer, C., Finan, R., Sapeluk, A., & Milton, C. (1997). Memory load and task interference: Hidden usability issues in speech interfaces. *Engineering psychology and cognitive ergonomic*, 3, 141-150.
8. Coull, J. & Tremblay, L. E. (2001). Examining the specificity of practice hypothesis: Is learning modality specific? *Research quarterly for exercise & sport*, 72, 4, 345-354.
9. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 13, 3, 319-340.
10. Denis, M. (1988). Imagery and prose processing, in M. Denis, J. Engelkamp, & J. T. E. Richardson (Eds.), *Cognitive and neuropsychological approaches to mental imagery,* Martinus Nijhoff Publishers, Dordrecht., 121-132.
11. Dubois, M. & Vial, I. (2000). Multimedia design: the effects of relating multi-modal information. *Journal of comupter assisted learning*, 16, 157-165.
12. Eimer, M. (1999). Can attention be directed to opposite locations in different modalities? An ERP study. *Clinical neurophysiology*, 110, 1252-1259.
13. Faletti, M. V. & Wellens, A. R. (1979). From people to places: Extending a multi-modal information-processing paradigm. *Environmental psychology and nonverbal behavior*, 3, 4, 248-252.
14. Guglielmetti, L. (2003). Standardizing automotive multimedia interfaces. *IEEE multimedia*, 10, 1, 76-78.
15. Johnson-Laird, P. N. (1983). *Mental models*, Harvard University Press, Cambridge, MA.
16. Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.) *Foundations of cognitive science*, MIT Press, Cambridge, MA, 469-499.
17. Kahneman, D. (1973). *Attention and Effort,* Prentice-Hall, Englewood Cliffs, NJ.
18. Koufaris, M. (2002). Applying the technology acceptance model and flow theory to onine consumer behavior. *Information systems research*, 13, 2, 205-233.
19. Mayer, R. E. (1989). Models for understanding. *Review of educational research*, 59, 1, 43-64.

20. Mayer, R. E. & Gallini, J. K. (1990). When is an illustration worth thousand words? *Journal of educational psychology*, 82, 4, 715-726.
21. Mayer, R. E. & Anderson, R. B. (1991). Animations need narrations: an experimental test of the dual-coding hypothesis. *Journal of educational psychology*, 83, 4, 484-490.
22. Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of education psycology*, 87, 2, 319-334.
23. Nardi, B. A., Schwarz, H., Kuchinsky, A., Leichner, R., Whittaker, S., & Sclabassi, R. (1993). Turning away from talking heads: the use of video-as-data in neurosurgery, in *Proceedings of the SIGCHI conference on Human factors in computing systems,* April 24-29, Amsterdam, Netherland, ACM Press, 327-334.
24. Navon, D. & Gopher, D. (1979). On the economy of the human-processing system. *Psychological review*, 86, 3, 214-255.
25. Paivio, A. (1971). Imagery and cognitive processes. Holt, Rinehart & Winston, New York.
26. Paivio, A. (1986). Mental representations: A dual-coding approach. Oxford University Press, New York.
27. Polson, M. C. & Friedman, A. (1988). Task-sharing within and between hemispheres: a multiple-resources approach. *Human factors*, 30, 5, 633-643.
28. Proctor, R. & Van Zandt, T. (1994). Human factors in simple and complex systems. Allyn and Bacon, Needham Heights, MA.
29. Siewiorek, D., Smailagic, A., & Hornyak, M. (2002). Multimodal contextual car-driver interface. In *Proceedings of the fourth IEEE international conference on multimodal interfaces*, October 14-16, Pittsburgh, PA, USA, IEEE, 367-343.
30. Starner, T. E. (2002). Attention, memory, and wearable interfaces. *IEEE pervasive computing,* 1, 4, 88-91.
31. Stock, O., Strapparava, C., & Zancanaro, M. (1997). Multi-modal information exploration. *Journal of educational computing research*, 17, 3, 277-185.
32. Streeter, L. (1998). Applying speech synthesis to user interfaces. In M. Helander (Ed.), *Handbook of human-computer interaction*, Elsevier Science Pub, New York, NY, 312-343.
33. Titsworth, T. (2002). Telematics might steer your car into the future. *IEEE multimedia*, 9,2, 9-10.
34. Travers, M., (1989). A visual representation for knowledge structures. In *Proceedings of the second annual ACM conference on hypertext*, Pittsburgh, PA, USA, ACM Press, 147-158.
35. Tremblay, L. & Proteau, L. (1998). Specificity of practice: The case of powerlifting. *Research quarterly for exercise and sport*, 69, 284-28.
36. Treviranus, J. & Coombs, N. (2000). Bridging the digital divide in higher education. In *Proceedings of the EDUCAUSE 2000 Conference,* Octomber 10-13, Nashville, TN.
37. Wickens, C. (1980). The structure of attentional resource. In R. S. Nickerson (ed.), *Attention and Performance VIII,* Lawrence Erlbaum, Hillsdale, NJ, 239-257.
38. Wickens, C. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (eds), *Varieties of Attention*, Press, New York, NY, 63-102.
39. Wickens, C. D. & Liu, Y. (1988). Code and modalities in multiple resources: A success and a qualification. *Human factors*, 30, 5, 599-616.
40. Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An introduction to human factors engineering.*Addison Wesley Longman, New York, NY.
41. Wickens, C. D. & Ververs, P. M. (1998). Allocation of attention with head-up displays. *Technical report of aviation research lab, Institute of aviation*, DOT/FAA/AM-98/28.