IMPLEMENTING SPEECH AND SIMULATED DATA LINK COMMANDS: THE ROLE OF TASK INTERFERENCE AND MESSAGE LENGTH

Mark W. Scerbo, Matthew R. Risser, Carryl L. Baldwin
Department of Psychology, Old Dominion University
Norfolk, VA

Danielle S. McNamara
Department of Psychology, University of Memphis
Memphis, TN

The present study investigated the ability of individuals to correctly execute different numbers of commands presented in speech and text format (simulating voice and data link communications). These commands were executed in the presence of different sources of task interference drawing upon visual, verbal, and central executive resources. The results showed that performance decreased as the number of commands in the message set increased from two to four, regardless of presentation format. Visual interference was less disruptive than both verbal and central executive interference. These findings have implications for the design of data link systems and suggest that during periods of high workload, communications should include fewer commands within a message set particularly when there is insufficient time to request clarification.

INTRODUCTION

Increased congestion in the national air space (NAS) coupled with technological advances and continual strides toward improving aviation safety and efficiency have led to new methods of digital communication between air traffic controllers and pilots (Wickens, Mavor, Parasuraman, & McGee, 1998). Controller-Pilot Data Link Communication (CPDLC referred to as data link in this paper) allows controllers to communicate with pilots via text messaging and has been in use for over a decade on long haul transoceanic flights where flight paths may be out of range of land-based VHF stations. Data link was implemented in part to reduce reliance on congested radio frequencies, but also to potentially reduce controller workload by giving pilots greater autonomy over flight path selection. An additional objective of data link is to reduce miscommunications between controllers and pilots, a persistent threat to aviation safety.

Researchers have recently begun to study the speech and text communication formats afforded by data link to determine their effects on pilot performance (McGann, Morrow, Rodvold, & Mackintosh, 1998; Risser, McNamara, Baldwin, Scerbo, & Barshi, 2002). One important issue concerns the potential impact of communication format on a pilot's ability to remember and execute commands.

Risser, Scerbo, Baldwin, and McNamara (2003) examined the effects of speech and text formats on the ability to recall and execute procedural commands under different sources of task interference. The foundation for this research draws upon two theoretical models. The first is a model of working memory (WM) proposed by Baddeley and Hitch (1974). According to this model, there are three components of working memory: a central executive (CE) component responsible for controlling attention and decision processes, a visuo-spatial sketchpad for visual and spatial processing, and a phonological loop for verbal processing. One important assumption of the WM model is that verbal and visual-spatial processing draw upon separate resources. The second model is Wickens' (1984) multiple resources theory (MRT) of information processing. According to MRT, the perceptual processing channel (auditory or visual), type of processing code (spatial or verbal), and processing stage (perceptual-central versus response) each draw upon separate resources. Both the WM and MRT theories predict that when two tasks draw on the same pool of resources, there is a greater opportunity for performance decrements than when tasks draw upon separate resources. However, the WM model also makes the unique prediction that utilizing CE resources will limit the processing of multiple sources of information regardless of processing code.

Risser and colleagues (2003) argued that according to the WM model, data link communication should be affected most by visual interference and voice communication should be affected most by verbal interference. On the other hand, because both communication formats utilize an underlying verbal code, both models predict that communication tasks, regardless of format would be affected more by verbal than visual interference. Their results showed that performance was disrupted most by CE interference as the WM model would predict. Further, the effects of speech and text presentation formats were not moderated by visual and verbal interference tasks. Instead, performance suffered more in the context of verbal interference irrespective of presentation format.

It is important to note that the task used by Risser, et al. (2003) required participants to recall and execute four commands on every trial. In operational contexts, ATC communications vary in length and instructions. Research by Barshi (1997) has shown that the ability to correctly execute commands deteriorates with increases in the number of commands contained in a set and that performance bottoms out at three commands. Thus, it is not known whether the effects of visual and verbal task interference on the ability to recall and execute procedural commands depend upon the amount of information that must be maintained in memory.

The goal of the present study, therefore, was to replicate the experiment by Risser and his colleagues (2003) and vary the number of commands in a set from two to four. Overall, it was expected that increases in the number of commands would lead to poorer levels of performance. Expectations for presentation format and visual and verbal interference effects, however, were tied to the number of commands. Regarding interference effects, it was expected that the CE task would be most disruptive because it draws upon resources needed for the storage, rehearsal, and retrieval of information. Expectations for verbal and visual interference were based upon the WM and MRT models. Specifically, verbal interference was expected to be more disruptive than visual interference because both tasks rely on a verbal code irrespective of presentation modality.

Another goal of the present study was to determine whether the effects of task interference would be moderated by command set size. Further, it is also possible that the need to maintain four commands in memory in the presence of task interference might mask potential effects of presentation format that would be observable with small command sets or lower memory load. The final goal was to explore this possibility.

METHOD

Participants

Twenty-four college students participated for credit in a psychology course. All had normal or corrected-to-normal vision and no reported auditory deficits. All participants were native speakers of English.

Stimuli

Commands were presented as speech or text. In the text condition, words were presented one at a time in the center of the screen. An 18-point MS sans serif font was used for the text commands and all text appeared in dark brown (RGB: 128, 64, 64) against a light gray background (RGB: 212, 208, 200). In the speech condition, words were presented one at a time via computer speakers. All words were digitally recorded monophonically at 22 KHz and 16-bit depth in a male voice. The words were presented at a normal conversational level of approximately 60 dBC measured from the position where participants would sit. Visual and verbal interference tasks were presented with the same characteristics as the commands.

Control Panel

The control panel consisted of six controls divided into binary, discrete, and continuous categories with two examples of each category. The binary controls labeled "autopilot" and "lights" could be set on or off. The discrete controls labeled "autobrakes" and "flaps" required the participant to set a pointer to one of four positions, "1, 2, 3, or 4". The two continuous controls labeled "speed" and "heading" consisted of horizontal sliders with a digital readout that could be positioned between 0 and 360 in increments of 10. The controls appeared in the same screen locations across trials. Participants were given an opportunity to familiarize themselves with how to manipulate the controls before beginning the experimental trials.

Procedural Commands

There were a total of 162 procedural commands comprised of three words in a verb-object-indirect object syntax, (e.g., "set heading 160"). Commands were presented in sets of two, three, or four. There were nine trials per command set presented in both speech and text formats. Participants were required to recall and execute as many commands as possible in the correct order. The commands were counterbalanced across control type and command length. All trials for one presentation format (speech or text) were completed before beginning the trials in the other format. The order of speech and text trials alternated for each participant. Each word of each command was presented individually for approximately 500 ms with an interstimulus interval (ISI) of 1000 ms. In the text condition, words were presented on the screen and in the speech condition, words were presented via computer speakers.

Interference Tasks

In the CE task, participants were required to generate and call out letters at random. Participants were instructed not to use any type of strategy (i.e., spelling, grouping letters, etc.) and to maintain a pace of at least one letter per word. This was verified by the experimenter who recorded each letter as it was spoken by the participant. In the visual and verbal tasks, a letter was displayed on the screen after each word of the command. The letter sets were constructed to contain approximately equal numbers of targets and distractors. Letters from within a set were presented in random order with the constraint that no more than two targets were presented sequentially. Participants responded by pressing the space bar whenever they detected a target. For the visual task, targets were letters containing a curved shape. For example, the letters C and D required a response, whereas Z and V did not. For the verbal task, letters were spoken (i.e., presented auditorily through the computer speakers). Targets were letters containing the /ee/ sound. For example, the letters T and D required a response, whereas R and W did not.

Design

A 2x3x2x2 repeated measures design was used in which the independent variables of presentation format (speech, text), command length (2, 3, or 4), interference task (verbal, visual, and CE), and experimental block (first, second) were combined factorially. Two dependent measures included the proportion of commands correctly set in the correct order (CSCO) and responses to controls not mentioned in the command set (false alarms). All participants performed three blocks of three trials each, in which the three interference task conditions were presented in a counterbalanced order. The first block of trials was considered practice. Data from the remaining two blocks were analyzed.

Procedure

Participants were run individually and seated at a computer screen located approximately 3 ft. directly in front of them at eye level. The commands and the interference tasks were presented concurrently. A trial began with the presentation of the first word of a command phrase immediately followed by one of the letters from the interference task. For example, "Set – C – Heading – R – 160 – T". Participants responded to the presentation of the interference task letters as soon as they occurred. This process continued until all command words and interference letters had been presented. Immediately afterward, the participants executed the commands by setting the controls as accurately as possible in the order in which they were presented. Participants completed a total of 54 trials.

RESULTS

The data were analyzed using a repeated-measures ANOVA. A Greenhouse-Geisser correction for repeated measures was used for all statistical analyses using a criterion p value of .05 for statistical significance. Separate analyses were performed for the proportion of CSCO and false alarms to other controls. A Student Newman Keuls post hoc comparison was used to analyze significant main effects and interactions. Experimental block was included in all analyses; however, it resulted in no significant effects across trials. Due to space limitations only the analyses of the dependent measures for the primary control-setting task are presented.

Correctly Set Correct Order

There was a main effect of presentation format, $F(1,23) = 7.20$, (partial Eta-squared = .24), reflecting a greater proportion of CSCO in the speech condition ($M = .43$) than in the text condition ($M = .34$). There was also a significant effect for the number of commands, $F(1.69, 38.91) = 61.21$, (partial Eta-squared = .73). Specifically, performance decreased as a function of command length with reliable differences between the three set sizes (2 commands, $M = .57$; 3 commands, $M = .37$; and 4 commands, $M = .21$). There was also an effect of interference task $F(1.46, 33.67) = 47.40$, (partial Eta-squared = .67). All tasks differed from one another; that is, performance under visual interference was better ($M = .57$) than under verbal interference ($M = .38$), which in turn, was better than CE interference ($M = .20$). Further, there was a significant interaction between command length and interference task, $F(3.18, 73.08) = 3.26$, (partial Eta-squared = .12). This interaction is shown in Figure 1. A post hoc analysis demonstrated that within the 2-command set, CE interference resulted in lower levels of performance than both verbal and visual interference. In the 3-command set, performance was lower for CE and verbal than for visual interference. In the 4-command set, performance under CE was lower than only visual interference. Within interference tasks, both the 3-command and 4-command sets resulted in lower performance than the 2-command set under verbal interference. However, under visual interference all command lengths differed from one another and under CE interference, there were no differences among command set lengths. Further, there were no reliable interactions between presentation format and interference task or command length.

False Alarms to Other Controls

False alarms to other controls represented erroneous responses made to controls that were not in the command set. The analysis showed that more false alarms were made in the text condition (*M* = .16) as compared to the speech condition (*M* = .11), *F*(1, 23) = 6.47, (partial Eta-squared = .22). There was also an effect for the number of commands, *F*(1.95, 44.83) = 15.84, (partial Eta-squared = .41). Specifically, fewer false alarms were made with the 2-command sets (*M* = .07) than with both the 3-command (*M* = .16) and 4-command sets (*M* = .17), which did not differ from one another.

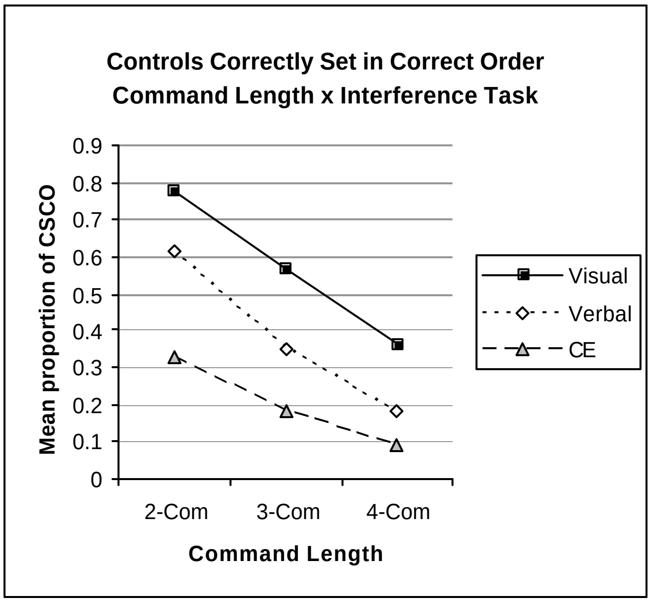

Figure 1. Command length by interference task interaction for proportion of controls correctly set in the correct order.

DISCUSSION

The present study was deigned to examine the ability of individuals to correctly execute different numbers of commands in speech and text formats in the presence of different sources of task interference. A previous study by Barshi (1997) showed that individuals performed more poorly as the number of commands increased. The results from the present study provide partial support for Barshi's findings. On the one hand, as the number of commands in a set increased from 2 to 4, participants set fewer controls correctly and made more false alarms. On the other hand, the level of significance varied as a function of the type of task interference. Under verbal interference, the results were consistent with those of Barshi. Performance was better for two commands than for three or four commands, which did not differ from one another. However, under visual interference there was a significant decline in performance with each additional command in the set. Further, under CE interference performance was quite low and there were no differences among the number of commands in a set. Collectively, these findings show that the ability of individuals to recall and execute commands will decline as the number of commands in a message increases, but that the magnitude of the decline is tied to the source of interference.

Regarding interference effects, it was expected that the CE task would be most disruptive and the results support that prediction. Expectations for presentation format and visual and verbal interference effects, however, were tied to Baddeley and Hitch's (1974) WM and Wickens' (1984) MRT models. The data showed that overall, verbal interference was more disruptive than visual interference and that this effect was not significantly related to presentation format. Moreover, verbal interference was more disruptive than visual interference at the larger command set sizes. Thus, these results are similar to what Risser and his colleagues (2003) observed in their experiment in which command set was held constant at four. More important, however, the results from the present study and from Risser et al. indicate that performance deficits will be greater when two tasks share a common processing code rather than the same perceptual modality.

The data support Baddeley and Hitch's theory of working memory resources and the processing code dichotomy in Wickens' MRT. The data clearly show that it is the verbal code underlying the commands that is susceptible to verbal interference irrespective of presentation format (speech or text). Similarly, according to Baddeley and Hitch's model of working memory both text and speech commands are hypothesized to be maintained in the phonological loop due to their verbal nature. Further, regardless of presentation format the effects of verbal interference were more pronounced than visual interference at the larger command set sizes. The decrease in performance with an increase in command length associated with verbal interference may reflect greater disruption of the ability to rehearse and maintain commands in the phonological loop.

The data also support Wickens' (1991) re-evaluation of the importance of modality in the original MRT model. Wickens suggested that time-sharing efficiency between visual and auditory tasks may stem from structural mechanisms rather than the central resource systems that distinguish the code and stage dichotomies. The current results indicate that processing code impacts verbal processing resources more than presentation format.

Regarding data link, the results from the present study suggest that merely changing the message format from speech to text should not have a detrimental effect on the ability to recall and execute commands. Instead, potential sources of task interference, particularly those that involve verbal processing or the control of attention (i.e., task switching, Baddeley, Chincotta, & Adlam, 2001) and decision-making are a more serious problem.

In the current paradigm, remembering and following procedural ATC commands was the primary task with intermittent interruption stemming from additional tasks. Actual flight conditions would present the reverse scenario in which ongoing flight tasks (of primarily a visual nature) would be disrupted by ATC commands. Initia l evidence suggests that ongoing visual flight tasks may be disrupted more than ongoing auditory flight tasks by intermittent ATC commands (Latorella, 1998). The relevance of this issue to data link warrants further investigation.

Last, it should be noted that in the present study participants had to remember commands that were issued only once while simultaneously performing another task. In this respect, the results suggest how pilots might perform under particularly impoverished conditions. Under normal flight operations, pilots would be able to request clarification or to have instructions repeated.

ACKNOWLEDGEMENTS

This research was supported in part by NASA Ames Research Center, grant NAG2-1481, to Old Dominion University (Danielle S. McNamara, Principal Investigator, Mark W. Scerbo and Carryl L. Baldwin Co-Principal Investigators). We are grateful to Andrea Gajeton and Colleen Moore for help conducting the experiments and to Vasubabu Muppaneni for writing the computer programs.

REFERENCES

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130*, 641-657.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Barshi, I. (1997). *Effects of linguistic properties and message length on misunderstandings in aviation communication.* (Doctoral Dissertation) University of Colorado, Boulder.

Latorella, K.A. (1998). Effects of modality on interrupted flight deck performance: Implications for data link. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting,* 87-91.

McGann, A., Morrow, D., Rodvold, M., & Mackintosh, M. (1998). Mixed-media communication on the flight deck: A comparison of voice, data link, and mixed ATC environments. *The International Journal of Aviation Psychology*, *8*, 137-156.

Risser, M.R., McNamara, D.S., Baldwin, C.L., Scerbo, M.W., and Barshi, I. (2002). Interference Effects on the Recall of Words Heard and Read: Considerations for ATC Communication. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting,* 392-396.

Risser, M.R., Scerbo, M.W., Baldwin, C.L., and McNamara, D.S. (2003). ATC Commands Executed in Speech and Text Formats: Effects of Task Interference. *Proceedings of the 12th Biennial International Symposium on Aviation Psychology, Dayton, OH,* 999-1004.

Wickens, C.D. (1984). Processing resources in attention, In R. Parasuraman & R. Davies (Eds.), *Varieties of Attention,* (pp. 63-101). Orlando, FL: Academic Press.

Wickens, C.D. (1991). Processing resources and attention. In D.L. Damos, (Ed.), *Multiple Task Performance,* (pp. 3-34). London: Taylor & Francis.

Wickens, C.D., Mavor, A.S., Parasuraman, R., & McGee, J. P. (1998). *The future of air traffic control: Human operators and automation*. Washington, DC: National Academy of Press.