Abstract Title: Studying Naturally Occurring Change Points in Navy Radio Communications

Short Title: Managing Multiple Radio Communications Channels

Introduction

It is expected that future Naval forces will be defined by their agility and their capacity for coping with highly dynamic environments [1]. Decision-making in the Combat Information Center (CIC) of the future is exactly such an environment. The individual Naval watchstander might be responsible for the concurrent monitoring of numerous radio communications channels, along with actively monitoring and responding to events on multiple visual displays. Such attentionally demanding environments have motivated various human-computer interaction (HCI) solutions to help warfighters deal with the vast number of information sources needing to be monitored in order to perform their duties successfully.

Monitoring Multiple Communications Channels

Unlike reading, in which the information input rate can be controlled by one's eye movements, comprehension of speech is often dependent on a transient acoustic signal whose input rate is largely controlled by the speaker, not the listener. The information input rate, thus, is determined by the environment, and previous information is often not reviewable. In order to comprehend auditory information effectively, input must be analyzed, segmented, and processed for structure and meaning, all of which must occur even as new auditory information continues to arrive. When auditory input is rapid, listeners have even less time to carry out these integrative processes, and successful comprehension requires greater cognitive effort on the part of the watchstander.

This project addresses a critical consideration when designing HCI solutions in auditory display research: the limitations in watchstanders' abilities to attend to multiple, active communications channels. Instead of presenting messages that must be monitored in parallel, as is the current state of affairs in the CIC, listeners hear messages that are serialized and accelerated in order to facilitate better comprehension. In addition, this project explores the role of natural biases in human attention toward certain categories of information during the listening process. Building an effective system of synthetically accelerated voice communications requires the integration of insights from multiple disciplines including signal processing, cognitive psychology, and linguistics. Each of these disciplines has made and will continue to make contributions to the design and execution of our experiments.

Experimental Design

An auditory test bed was created with communications messages that were synthetically accelerated using a patented NRL speech-rate compression algorithm [2], known as pitch-synchronous segmentation (PSS). PSS retains the fundamental frequency of speech signals and preserves a high degree of intelligibility by representing speech as a combination of individual pitch waveforms that do not destructively interfere with one another. The PSS algorithm was used to construct synthetically accelerated communications messages for our experiments at rates ranging in 15% increments from 50% to 140% faster-than-normal.

The communications messages, themselves, were also structured to enhance comprehension. Based on theories of attention and memory [3], the human attentional capacity is thought to be optimized for detecting relevant changes in the environment, or in this case, the communications message. These changes serve as natural cues to encode and later remember pertinent information. We, therefore, presented auditory information in a way that approximates how listeners more naturally perceive it, that is, by varying the number of natural cues contained in the message. It was expected that comprehension of messages that contained more of these natural cues would be better than comprehension of messages that contained less.

In order to assess the effectiveness of our design, we tested for training effects and practice effects when listening to accelerated speech, and examined whether the number of natural cues included in the messages would influence comprehension.

Results

Comprehension of auditory messages was compared at normal speed and seven accelerated speech rates. Figure 1 shows data from one group assigned to a training condition compared to another group assigned to a condition where accelerated speech messages were presented in a random fashion. As can be seen in Figure 1, overall comprehension for both groups declined as speech rate increased. Training was effective at slower accelerated speech rates (50% and 65% faster-than-normal). The optimum acceleration rate for comprehension, or the fastest rate at which speech can be presented so that performance does not differ from that at a normal rate was 65% faster-than-normal. However, practice did not aid comprehension. That is, there was no systematic improvement shown after listening to multiple auditory messages presented at the same accelerated speech rate. Listeners simply adapted quickly when listening to accelerated speech. Structuring messages to contain more natural cues did enhance listeners' comprehension of information stated in the messages. This relationship is shown in Figure 2.

Summary

This research effort integrates several disciplines, including signal processing, cognitive psychology, and linguistics in order to design a radio communications protocol that is minimally taxing to the watchstander and that will maximize the probability of message retention and understanding.

Acknowledgments

Many individuals have contributed to this effort. Particular thanks is due to Derek Brock, Wende Frost, LT Gregory Gibson, Brian McClimens, and Dennis Perzanowski (NRL). Thanks is also due to Paul Bello and Ray Perez (ONR). This work is supported by ONR, N0001410WX20273.

References

[1] Clark, V. (2002). Sea power 21 series - part I: Projecting decisive joint capabilities. Naval Institute Proceedings Magazine.

[2] Kang, G. S., & Fransen, L. J. (1994). Speech analysis and synthesis based on pitch-synchronous segmentation of the speech waveform (Technical Report 5550-94-9743). Washington, DC: Naval Research Laboratory, Information Technology Division, Transmission Technology Branch.

[3] Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. Psychological Bulletin, 127, 3-21.

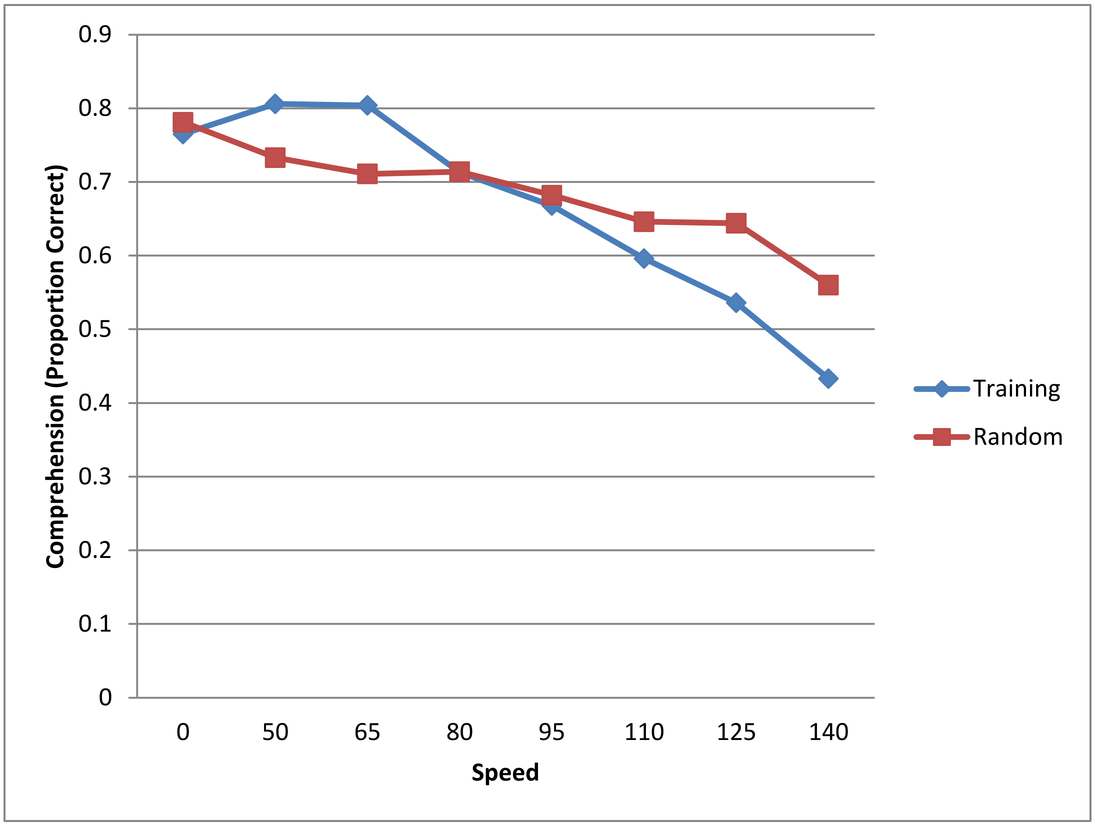

Figure 1. Comprehension (proportion correct) as a function of speed for the training and random presentation conditions. Note that 0 indicates normal speed speech; 50-140 denote % speed faster-than-normal.

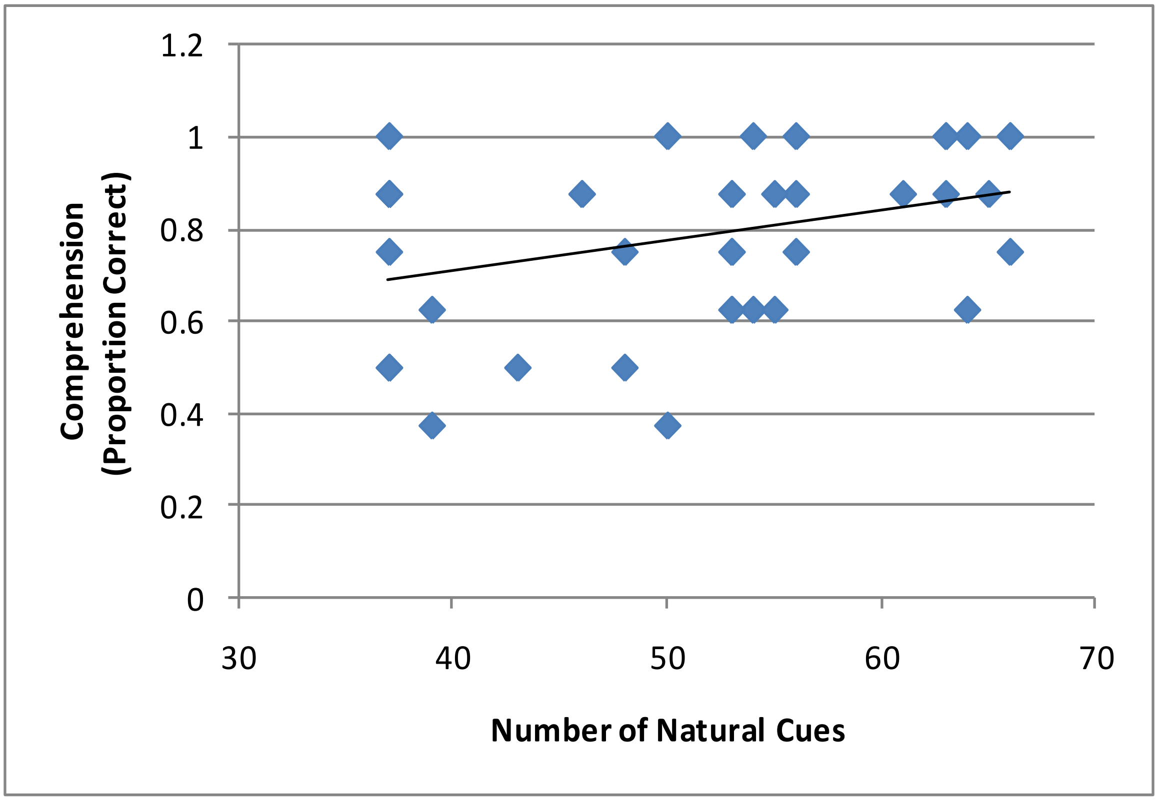

Figure 2. The relationship between comprehension (proportion correct) and the number of natural cues presented in the communications messages.