Modality Specific Attentional Mechanisms Can Govern the Attentional Blink

Kim M. Goddard & Elzbieta B. Slawinski
Department of Psychology, University of Calgary, Calgary, AB., T2N 1N4

INTRODUCTION

The impairment in the ability to detect or identify a secondary target (T2) when it follows within approximately 500 ms of correct detection or identification of a primary target (T1) is a phenomenon known as the attentional blink (1), which may be related to informational masking (2). However, while the Attentional Blink (AB) has been repeatedly demonstrated in the visual domain and more recently, cross-modally (3, 4), pure auditory AB investigations are few. Further, most cross-modal studies have found evidence of auditory ABs with the use of compressed speech (e.g., 4). These types of studies can be problematic for two reasons. First, because it is possible to form visual representations from spoken words or letters, cross-modal studies may make it difficult to unambiguously attribute these effects to the auditory modality. Second, and relatedly, when attentional effects cannot be definitively attributed to a specific modality, it also becomes difficult to determine whether attentional mechanisms operate within or across modalities, or possibly both. To address this issue, this study employed pure tones, for which, no visual representations exist, in a rapid auditory presentation paradigm (11 tones/second). A visual rapid presentation paradigm, consisting of simple visual stimuli (11 lines/second) was also employed to compare visual AB effects with auditory AB effects in a within-subjects design.

METHODS

Participants 20 young adults (Mean = 21.2 yrs) participated in the study for course credit. All participants reported normal hearing and normal or corrected-to-normal vision.

Auditory Stimuli Rapid Auditory Presentation (RAP) stream stimuli consisted of 25 randomly presented tones comprising the range of 1000 Hz to 2490 Hz in 10 Hz multiples. Properties of these tones allowed the processing of the stream as a unit. Tones of 1500 Hz (low), 2000 Hz (medium) and 2500 Hz (high) were not stream items, being reserved for T1 and T2. All tones were equally loud (approximately 50 dB SPL) except for T1 and T2 which were increased in intensity by approximately 10 dB SPL over and above the level of the stream items. All tones, including T1, T2, and stream items were 85 ms in duration, separated by a silent 5 ms Interstimulus Interval (ISI).

Visual Stimuli Rapid Serial Visual Presentation (RSVP) stream stimuli consisted of 25 randomly presented lines in orientations of 30, 60, 120, and 180 degrees. Lines of orientations 45, 90, and 145 degrees were not stream items, being reserved for T1 and T2. All lines were 3 cm long and stream lines were of identical thickness. T1 and T2 were thicker lines, clearly discriminable from stream lines. At a viewing distance of 30 cm, all line stimuli subtended .95 X .76 degrees visual angle, and all were displayed for 15 ms, separated by a blank 75 ms ISI.

Design After receiving training on a frequency identification task and a lines orientation task, all participants performed both the auditory and the visual task. In the experimental condition (Exptl), T1 was presented equally often at positions 5, 9, or 13 in the stream on independent random halves of the trials. T2 was presented on all trials, equally often at all T1-T2 intervals. Participants were asked to report the first loud tone according to pitch (low, medium or high), and then, if a second loud tone was heard, it was also to be reported according to pitch. In the control condition (Ctrl), participants performed only the T2 task, where T2 was present on independent random halves of the trials. The experimenter recorded the number of loud tones reported in both conditions. The control condition was identical to the experimental condition except that in the T1-present trials, T1 was not louder. Procedures were identical for the visual task, where participants reported the thicker lines (targets and probes) according to orientation (45, 90, or 135 degrees). Condition (Exptl, Ctrl) and Task (auditory, visual) was counterbalanced across participants.

RESULTS

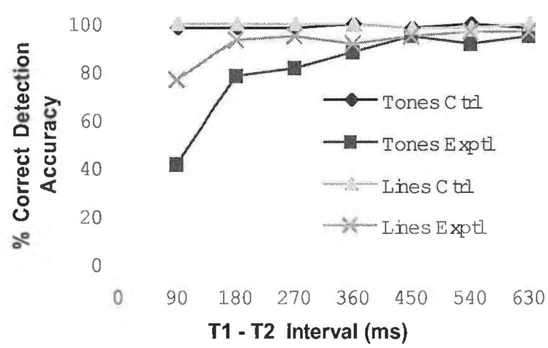

Figure 1. Detection accuracy as a function of T1-T2 Intervals for visual and auditory tasks

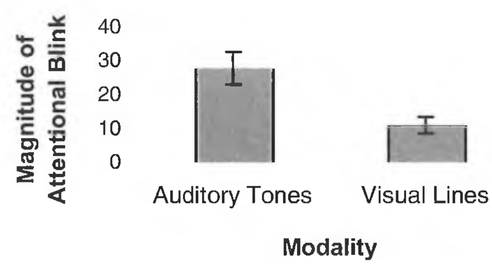

Figure 2. Magnitude of the Attentional Blink as a function of modality.

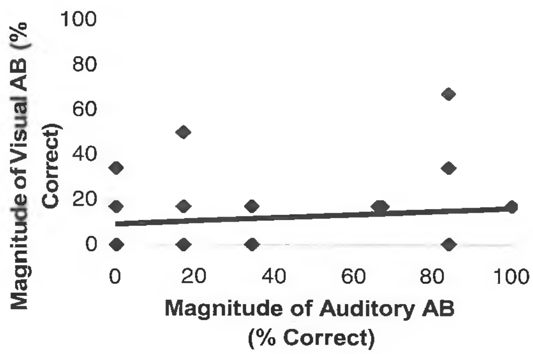

Figure 3. Scatterplot of individual AB magnitudes as a function of modality with trendline (some data points represent more than one individual).

The mean percentage of trials in which T2 was correctly detected is plotted as a function of the T1-T2- interval in the experimental and control conditions for both tasks. These results are presented in Figure 1. In the control condition, participants correctly detected T2 on 97% or better of trials for all T1-T2 intervals in both tasks. However, for the experimental condition, percent correct T2 detection for the range of 90 ms to 360 ms averaged only 72.5% for the auditory task and 89.2% for the visual task. T2 detection accuracy for the interval range of 360 to 630 ms averaged 93.9% for the auditory task and 96.1% for the visual task; T2 detection accuracy at these ranges however, was not significantly different than overall T2 detection accuracy in the control condition for either task. False alarm rates averaged less than 3% across conditions on both tasks.

Multiple paired comparisons revealed that T2 detection accuracy in the experimental condition for both tasks was significantly lower (p's <.05) when T2 was presented at T1-T2 intervals of 90, 180, 270 and 360 ms than the corresponding interval in the control condition; this indicated a significant AB-like T2 detection impairment for those intervals. Furthermore, when collapsed across these intervals, the magnitude of the Attentional Blink for the auditory task was greater than the magnitude of the visual task by a factor of almost 2.5, and this difference was statistically significant (p<.025; see Figure 2). Finally, correlational analyses revealed no significant correlations between individual task performance; that is, the magnitude of the auditory AB did not predict the magnitude of the visual AB. These results are shown in Figure 3.

DISCUSSION

Our results indicate the presence of an Attentional Blink for both the visual and the auditory task. More importantly, because of the nature of the stimuli employed in this study, we may be confident that these ABs are uncontaminated by cross-modal effects. That is, if we assume that no auditory representations exist for visual lines and no visual representations exist for auditory tones, then these ABs would appear to reflect modality-pure attentional effects.

Two interesting findings have emerged from this study. The first is the greater auditory AB magnitude when compared to the visual AB magnitude. This is notable given the less favorable conditions for auditory target detection. That is, it is typically more difficult to detect an auditory target in a stream of auditory stimuli than it is to detect a visual target in a stream of visual stimuli (5), and this was consistent with our participants subjective reports. Nevertheless, Control Condition performance did not differ across tasks, and thus, while differential task difficulty make possibly contribute to magnitude differences between the modalities, it is unlikely to be the major cause.

We have previously argued that the auditory AB reflects an inhibitory mechanism (6), putatively designed to protect the processing of the target (7) during selective attention tasks. Banks, Roberts, and Ciranni (8) note that auditory selective attention is not aided by any structural analogue similar to that of visual fixation that can choose to place important targets within foveal vision and less important targets in peripheral vision. They further suggest that because audition does not have these external capabilities, attentional inhibition must operate solely by internal auditory processes. Consequently, attentional inhibition should be more pronounced in audition than in vision. Our magnitude differences, where magnitude is an indication of the strength and/or extent of attentional inhibition, agree with Bank et al's interpretation.

A second interesting finding is that there was a lack of correlation between individual performance on the auditory task versus the visual task. To the extent that a central, amodal attentional system governs incoming sensory information, then individual performance should be, at least, modestly correlated across visual and auditory tasks. Our lack of a significant correlation, together with pure, modality specific stimuli and magnitude differences between the tasks, suggests that attentional mechanisms can, and do operate in a modality specific fashion.

REFERENCES

1. Raymond, J.E., Shapiro, K.L. & Arnell, K.M. (1992a). Temporary suppression of visual processing in an RSVP task: An attentional blink? Journal of Experimental Psychology: Human Perception and Performance, 18, 849-860.
2. Pollock, I. (1975). Auditory informational masking. Journal of the Acoustical Society of American, 57, S5.
3. Jolicouer, P. (1999). Restricted attentional capacity between sensory modalities. Psychonomic Bulletin & Review, 6, 87-92.
4. Potter, M.C., Chun, M., Banks, B.S. & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: The visual attentional blink and an amodal task-switch. Journal of Experimental Psychology: Learning, Memory & Cognition, 24, 979-992.
5. Leek, M.R., Brown, M.E., & Dorman, M.F. (1991). Informational masking and auditory attention. Perception & Psychophysics, 50, 205-214.
6. Goddard, K.M., Isaak, M.I., & Slawinski, E.B. (1998). The Auditory Attentional Blink in a Congenitally Blind Population. Journal of the Acoustical Society of America, 103, 3018.
7. McDowd, J.M. & Filion, D.L. (1992). Aging, selective attention and inhibitory processes: A psychophysiological approach. Psychology and Aging, 7, 65-71.
8. Banks, W.P., Roberts, D., & Ciranni, M. (1995). Negative priming in auditory attention. Journal of Experimental. Psychology: Human Perception & Performance, 21, 1354-1361.

Subjective Evaluation Of Different Error Correction Schemes For Application With A 900 Mhz Frequency Hopper Communication System

Kimberly Braaten (1), Dean Foster (2), Bruno Korst-Fagundes (2) and Haoye Shen (2)
(1) Faculty of Engineering, University of Regina, Regina, SK, S4S 0A2
(2) Acoustics Group - ENC Nortel, 3705 35th St. N.E., Calgary, AB, T1Y 6C2

1. INTRODUCTION

Wireless communication systems are required to communicate over selected frequency ranges. Each channel has a limited bandwidth or frequency range in which it must operate and the frequency spectrum is becoming increasingly more congested.

Spread spectrum is a modulation scheme that uses the spectrum efficiently and operates with a minimum amount of interference. In a spread spectrum system, the signals are spread over a wide range of frequencies by using a variety of broadband or frequency hopping techniques. Interference is present and subjectively noticeable in some circumstances with the use of frequency hoppers. The effect of having many users utilizing the same frequency bandwidth promotes a special problem since it becomes possible for one user to jam the signal of another. This creates noise or other user perceived anomalies that considerably degrade the audio quality. Errors, caused by jamming and other sources, can be introduced into the signal from anomalies inherent in the transmit and receive modes of a wireless communication unit transporting digital information. These errors are quantified through the bit error rate (BER). An error can occur in transmission from the receiver to transmitter, from transmitter to receiver or from transmitter to transmitter. The bit error rate (BER) is the probability of an error occurring in a bit, or a change in the transmitted information.

Subjective testing was performed on two types of interference associated with such a frequency hopping system. In this article we analyzed two of the simplest techniques used to correct corrupted data. The first correction method studied, called 'repeating', used the previously sent block of data picked up by the receiver and then repeated it. A second correction method used, called 'muting', simply muted any erroneous data that was picked up by the receiver.

2. EXPERIMENT

Digital speech transmission systems can generate degradation's that involve difficulty in the listening path. These degradation's can be perceived to the end user as clicks, pops, distortion, fuzziness, etc. in the receive listening audio path. Since the listening transmission path is involved. we created a test for subjective listener's. Each test person would listen to the same audio file each time creating a consistent test base. The results from this series of tests helped the designer's choose the best error correction scheme that was available to them. To assist the designers in making the correct decision from the results, a method of assessing the subjective listener's opinions on the various audio samples was used. This technique is called the Mean Opinion Score or MOS method [3]. The speech samples used in the listening tests contained audible errors created by software that simulated conditions where jamming and various levels of BER had occurred.

Test 1 determined the type of correction scheme and the threshold of correction for errors preferred by listeners for corrected jammed signals. The threshold determines the level of correction for errors the software is using. Test 2 threshold levels were based on the results from Test 1. For Test 2, since jamming was of more concern for audio quality, the threshold parameters of Test 1 for jamming were incorporated into several selected BER's. Test 3 is based on the chosen threshold and error correction schemes determined from Tests 1 and 2. Test 3 determined when the audio quality would degrade for jamming as the numbers of users increased. It compared two different scenarios that might occur in a jamming situation. The listeners evaluated the audio quality when the jams occurred as users interfered with each other at the same time or when the interference occurred at different times.

3. RESULTS AND DISCUSSIONS

3.1 Test 1

In this project, jamming contributes to the quality of the audio signal to a greater degree than does BER, meaning, if a signal is jammed, it is much more noticeable to a listener than the BER factor. Therefore, Test 1 was performed to find out whether jamming using a correction scheme called muting or using the repeating method of a previous block was preferable. The listeners would find which threshold level was most acceptable using the DCR MOS subjective test method. From Test 1 it appears from Figure 4 (shown on the next page) that Thresholds 1, 2 and 3 have the highest DCR MOS scoring.

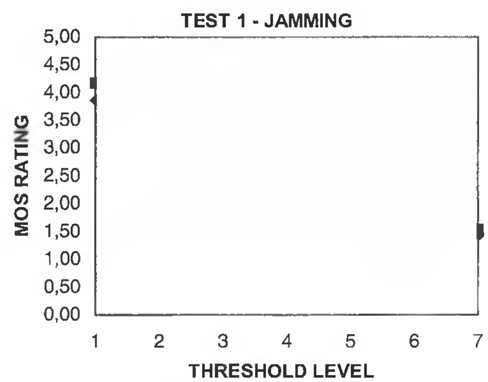

Figure 4 Test 1. Jamming from Threshold's 1 to 7.

3.2 Test 2

Test 2 will use the chosen threshold value and error correction scheme from Test 1 with the selected bit error rates. Since jamming and the BER can only be corrected with one chosen threshold, the need to see how the parameters chosen from Test 1 for jamming compared to the selected BER's became apparent. This became the testing performed for Test 2.

Testing was accomplished by comparing a speech sample that was corrected to the original uncorrected speech sample. All of the samples were corrected using the muting correction method at Threshold's 1, 2 and 3 chosen from Test 1. The threshold test values were so close in Test 1, you cannot really say that a threshold of 2 is completely superior, so 3 threshold's were chosen.

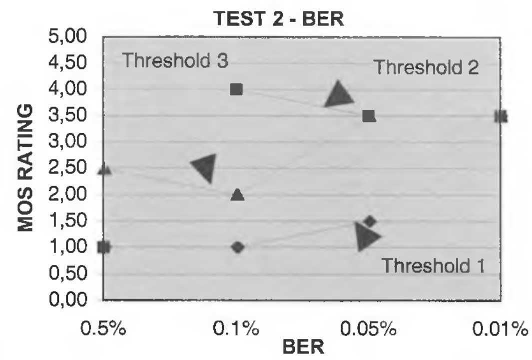

Figure 6 Test 2 for a chosen threshold of 2 with threshold's 1 and 3 using the muting correction scheme and the selected BER's.

3.3 Test 3

Test 3 used an absolute rating schedule was used for each trial based on a single speech sample that was heard one at a time by the listener.

The two graphs shown are for close together jams (to simulate jamming at the same time) and jams far apart to simulate dispersed jams.

From the data in Figure 9, it appears that when the jams are dispersed (highlighted as Far in the figure), most ratings were below fair (MOS < 2.5). The best scenario for dispersed jams is 4 jams since there is a drop off in quality after that. For close together jams, the ratings are fair - up to 9 jams, with an anomaly at 3 jams/s.

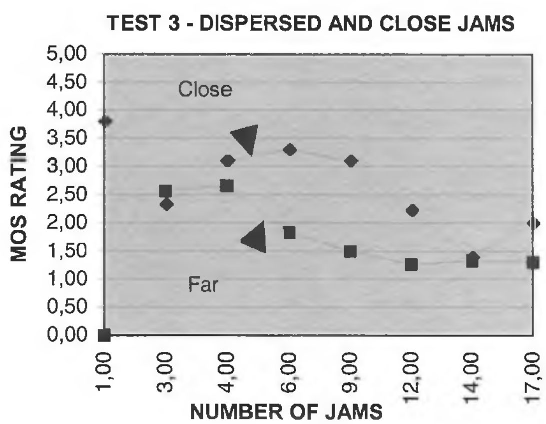

Figure 9 Test 3 for far apart and close Jams from 1 to 17 Jams with a muting error correction scheme at threshold 2.

4. CONCLUSIONS

A 32 kbits/s ADPCM coding scheme gave the best audio quality for the lowest number of bits. For Test 1, the threshold chosen was number 2. This means that the error correction scheme was invoked only after two consecutive errors were detected. From looking at the graph (Figure 4), you can see that these threshold levels received the highest ratings from the listeners. The correction scheme chosen was muting since the speech samples that were corrected using this scheme received higher ratings than the samples corrected with the repeating "previous block" method. See Figure 4 for more information on Test 1. The DCR MOS test rating was used since the basis of the test was to compare a reference sample to a corrected sample.

Test 2 concentrated on finding the BER with the best audio quality for the chosen threshold and error correction scheme from Test 1. To give more variety during testing, the chosen threshold from Test 1 along with the upper and lower threshold were chosen for Test 2, therefore, the threshold's used for Test 2 were 1, 2 and 3. The level chosen from Test 2 that had the best audio quality for BER was threshold 2, which was the same as the threshold chosen for Test 1. Test 2 was conducted for a BER of 0.5%, 0.1%, 0.05% and 0.01%. Any BER meeting or exceeding 0.01% would have an acceptable level of audio quality according to our preliminary tests. Once again, the DCR MOS test for comparison ratings was used. For more information on Test 2, please see Figure 6.

Test 3 incorporated parameters found from Test 1 and Test 2, which are a threshold of 2 with a muting correction scheme, to do a density evaluation for jamming. From reviewing the data, it is evident that using more than 4 dispersed jams does not have an acceptable audio quality. Between 1and occurring at the same time, or close jams, has a fair quality (except at 3 jams), but there is a drop off on either side of these values. Test 3 used the ACR MOS test method that asks for the overall opinion of each sample on its own. For more information on Test 3, please refer to Figures 8 and 9.

5. REFERENCES

[1] Paul G. Flikkema, "Spread-Spectrum Techniques for Wireless Communications", IEEE Signal Processing Magazine, 14:3, May 1997, p.26.

[2] J.D. Laster and J.H. Reed, "Interference Rejection in Digital Wireless Communications", IEEE Signal Processing Magazine, 14:3, May 1997, p. 37.

[3] B.S. Atal, V.Cuperman, and A. Gersho, Speech and Audio Coding for Wireless and Network Applications, (Norwell, MA: Kluwer, 1993), p. 43.

[4] Simon Haykin, Digital Communications (Singapore: John Wiley & Sons, 1988), pg. 210.

[5] Haoye Shen, ENC Nortel, Calgary, Alberta. Memo to Ashfaq A. Choudhury, S.G.S. Thomson Microelectronics, Nepean, Ontario, March 3, 1997.

[6] Leigh L.A. Thorpe, AT30-I BNR, Ottawa, Ontario. Memo to Dean Foster, ENC Nortel, Calgary, Alberta, April 25, 1997.

[7] John Chan, ENC Nortel, Calgary, Alberta. Letter to Kim Braaten, ENC Nortel, Calgary, Alberta.

[8] Leigh L.A. Thorpe, AT30-I BNR, Ottawa, Ontario, in conversation with Kim Braaten, ENC Nortel, Calgary, Alberta, July 16, 1997.

[9] Martin S. Rodin, Analog and Digital Communication Systems, (New Jersey: Prentice-Hall, Inc., 1996), pg.18.