Echoic Memory In Primitive Auditory Selective Attention

Ben Weedon & Zofia Kaminska
Psychology Department, City University, London, EC1V 0HB
email: ben@city.ac.uk

Abstract

This study explored possible confounding effects of echoic memory on estimates of attentional capacity by attempting to estimate capacity while controlling the effect of echoic memory. Detection of a target stimulus and identification of its carrier stream was investigated as a function of variation in number of concurrent non-overlapping auditory streams (1 to 4) and of condition of echoic memory involvement (available or eliminated via articulatory suppression). Error rates were found to increase non-linearly as a function of number of streams, but with a different point of discontinuity – indicative of a processing limitation – under different conditions of memory involvement, a higher capacity (3 as opposed to 2 streams) being achieved with echoic memory contribution. Detection response latencies also increased as a function of number of streams, but the increases were linear. Echoic memory also significantly reduced the response latency. These findings, which implicate echoic memory as a contributory factor in estimates of auditory attention, may help to resolve discrepancies in previous research and have implications for modeling auditory attentional processes.

1 Introduction

Previous research into the processing capacity of selective attention has produced conflicting results. Early findings (e.g., Broadbent, 1958; Treisman, 1964 and Sloboda and Edworthy, 1981) suggested a serial mode of processing in a figure/ground Gestaltian manner, in which a single channel receives full processing, others remaining as background. More recent research (Huron, 1989; Gregory, 1990 and Ludlam, 1993) has discovered what appears to be a processing capacity of 3 auditory streams, suggesting a greater degree of parallel processing.. A possible cause of these discrepant findings may be that the measures of attentional capacity in the latter experiments relied explicitly on memory processes. For example, a typical task used to estimate attentional capacity is to ask listeners to identify which of varying number of simultaneous auditory streams contains a target stimulus. However, listeners may be able to perform this task not because they are able to process all inputs simultaneously, but because they may be able to retrieve traces of streams unattended to at the moment of perception, using their memory. Thus, higher cognitive processes are confounding the measure of attentional limitations.

The aspect of memory most likely to be involved in auditory attention is echoic memory. Cowan (1984), summarising research into auditory memory, noted that there is evidence for two separate systems. Both are contained within the phonological short-term store as proposed by Baddeley (1974). One system deals with elementary processing of auditory stimuli and has a minute temporal span (the short-term auditory process, or ‘sensory register’ in Treisman’s model), whereas the other (the long-term auditory store) has a longer span, a smaller spatial capacity and is synonymous with echoic memory. It is echoic memory which could have been utilised by participants in studies reporting 3-stream capacity (Huron, 1989; Gregory, 1990 and Ludlam, 1993).

To test this hypothesis, the present study explored the effects of blocking participants’ use of echoic memory on a task similar to those used previously to determine attentional capacity (Huron, 1989; Ludlam, 1993). The technique used to preclude the use of echoic memory was articulatory suppression. This has been shown to be a consistent method of occupying the limited space of the phonological loop, thereby preventing subjects from using echoic memory to store new material and resulting in the decay of material already held (Baddeley, 1975; Salamé and Baddeley, 1982; Keller, Cowan and Saults, 1995). It involves participants producing continuous verbal output of minimal cognitive load.

Thus this experiment investigated auditory attention capacity while attempting to control the role of echoic memory.

2 METHOD

2.1 Design

Eighteen participants took part in the experiment. They were randomly assigned to undertake the Echoic or Non-echoic condition first. The experiment took the form of a four by two, completely within-participants, factorial design. The first variable, Streams, was the number of concurrent auditory streams. This had four levels, with the number of simultaneous streams being 1, 2, 3 or 4.

The second variable, Memory, was memory involvement, with two levels, Echoic and Non-echoic, manipulated within –participants. Under the Echoic condition participants performed the target detection task without any experimental blocking of memory involvement. In the Non-echoic condition echoic memory contribution was blocked by requiring the participants to recite aloud randomly generated numbers from 1 to 8 as they appeared on the computer screen during the basic task. This served to suppress any internal auditory imagery activity.

Response times to the target stimulus in each level of the experiment were recorded by the computer, as were the percentage of incorrect responses. The time limit within which participants were to respond was 5 seconds. Incorrect responses were also given a score of 5 seconds. So that participants could get used to the different sounds of the instruments and the demands of the experiment, and also that they might refine their reactions, a practice session was included.

Participants underwent 160 trials in two blocks of 80, one block for each memory condition (Echoic/ Non-echoic). Within each block of 80 there were 20 blocks for each level of Streams, the presentation of which were randomised.

2.2 Stimulus Materials

The stimuli were designed to be 4 individually identifiable auditory streams when presented concurrently, with no stimulus being any more salient than any other when 2 or more were playing simultaneously. Each stream consisted of a ‘melody’ that lasted for a few seconds before a target stimulus was presented as the final note in one of the streams. To achieve the structure of the streams

Bregman’s (1990) guidelines on perceptual streaming and Auditory Scene Analysis were followed.

The auditory stimuli were four individual musical streams varying in timbre by being generated by four different instruments (piano, whistle, harpsichord and saxophone) and also varying in their “brightness”, with the highest pitched instrument having the brightest sound. Each instrument played a chromatic set of notes in different pitch ranges in a symmetrical ‘v’ shape of a 3- semitone spread. The ‘melodies’ of the instruments are shown in Figure 1.

The different streams did not overlap with each other in pitch. When there was more than one instrument playing note onsets were synchronous. The streams were presented in mono at a rate of one note every 200 milliseconds. They were played repeatedly for between 5 and 7 seconds (varying randomly) until the target stimulus was presented. The target stimulus replaced one of the notes in one of the streams, selected at random. It consisted of a burst of white noise. The streams were represented on the computer VDU by their respective names, written in blue, in 4 blue coloured boxes. When a stream was playing its name and box turned red. Each box was ascribed a number, from 1 to 4. Participants were instructed to respond as quickly as possible to the target by pressing the number key that corresponded to the stream in which it was presented.

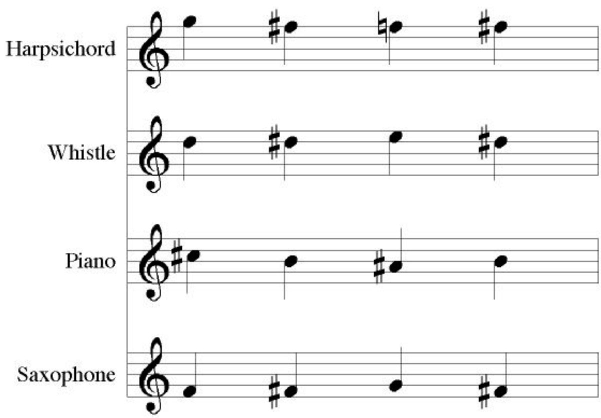

Figure 1 - Melodies of the Stimuli Instruments

2.3 Response Measurements

Participants were randomly assigned to undertake the Echoic or Non-echoic condition first. In the Non-echoic condition subjects were told to respond as quickly as possible to the target stimulus by pressing the key corresponding to the musical instrument in which the stimulus appeared. In addition, they were instructed to repeat aloud digits appearing in a random order in a different portion of the screen. The importance of maintaining accuracy on this latter task was emphasised. In the Echoic condition target detection instructions were the same, but there was no random digit sequence to occupy the subjects. The 80 test trials of each level of Memory followed a practise session for that level. When subjects were confident that they could perform the task, that level of the experiment began. Between each level of Memory there was a 2 minute interval. Participants were fully de-briefed at the end of testing.

RESULTS

Analysis of Errors

A 2-factor repeated measures ANOVA was performed on the error data. There was a significant effect of number of streams presented simultaneously on the subjects' ability to discern in which stream the target stimulus appeared, with error rates increasing as a function of increasing numbers of streams ($F_{36,\ 2} = 196.9$, $p < 0.0005$). Echoic memory involvement also had a significant effect on ability to discern in which stream the target stimulus occurred ($F_{17,\ 1} = 40.83$, $p <0.0005$). Figure 2 shows that echoic memory usage led to an increase in correct responses. There was also a significant interaction between Streams and Memory ($F_{34,\ 2} = 9.65$, $p <0.0005$). This indicates that the pattern of increase in difficulty of discerning the target caused by an increase in number of streams was different for the two memory conditions. This was investigated by more detailed analyses: Bonferroni t-tests were used to determine the points at which there were significant increases in incorrect responses.

In the Echoic condition there was a significant increase in the amount of incorrect responses when the number of simultaneous streams changed from 3 to 4. In the Non-echoic condition there was a significant increase in incorrect responses when the number of simultaneous streams increased from 2 to 3. This indicates that when echoic memory is free to be used, there is a processing capacity of 3 auditory streams, yet when echoic memory usage is suppressed there is a capacity of 2 auditory streams (See Figure 2).

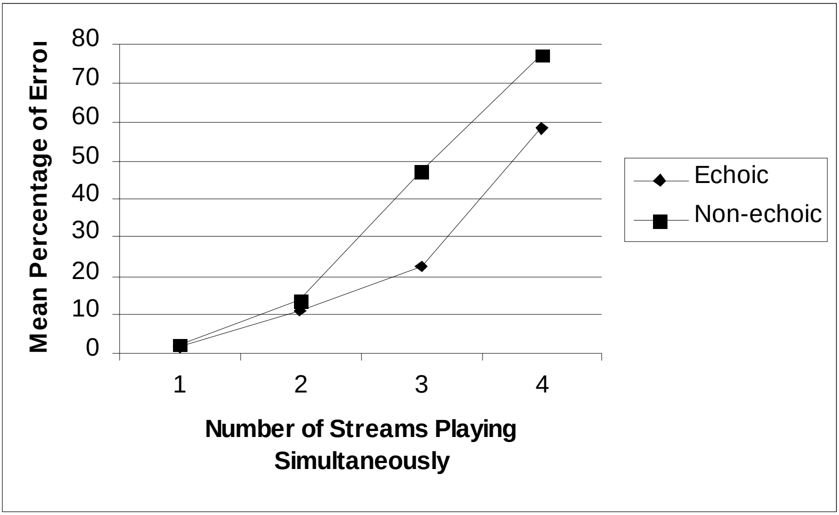

Figure 2 - Percentage of errors as a function of echoic memory use and number of simultaneous auditory streams

2.4 Analysis of Reaction Times

Mean target detection times as a function of echoic memory usage and the number of simultaneous streams are shown in Figure 3.

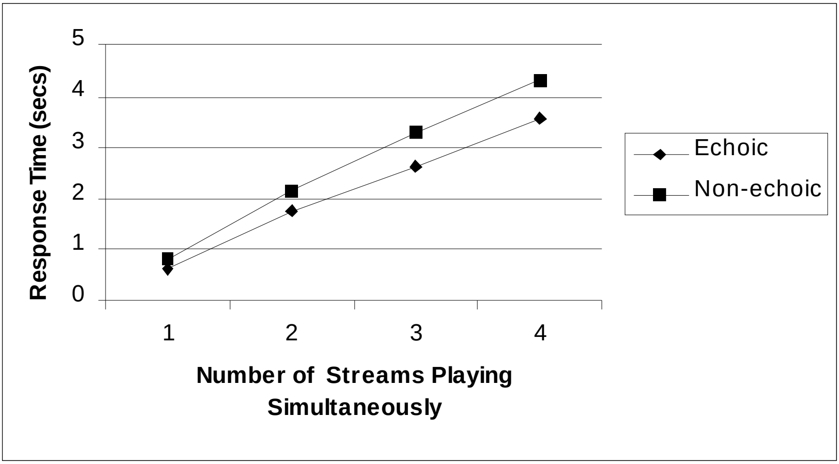

Figure 3 – Mean target detection times (in seconds) as a function of echoic memory usage and the number of simultaneous auditory streams

A 2-way repeated measures ANOVA was performed on these data. Echoic memory led to significantly quicker response rates to the target stimuli ($F_{17, 1} = 41.5$, $p < 0.0005$). The number of streams presented simultaneously had a significant effect upon reaction times ($F_{29, 1} = 387.3$, $p < 0.0005$). As the number of streams played simultaneously increased, the time taken to respond to the target stimulus significantly increased. To check that the increase in reaction times was a linear increase, an unplanned linear contrast was run on the data. In both the echoic and non-echoic conditions there was a significant linear trend. In the echoic condition $F_{7, 18} = 105.285$, $p < 0.00005$. In the non-echoic condition $F_{7, 18} = 62.079$, $p < 0.00005$.

3 Discussion

The findings of this study clearly indicate that echoic memory can play a role in models of auditory attention. Attentional capacity was 3 auditory streams when echoic memory was free to contribute to performance, but when echoic memory input was eliminated by articulatory suppression, the apparent attentional processing capacity dropped from 3 to 2 streams. Both these estimates have parallels in previous work. The present findings may help to explain discrepancies in earlier research, where echoic memory may have contributed to capacity estimates in different degrees, depending on the precise experimental procedures. For example, certain research (Broadbent, 1958; Treisman, 1964) which utilised shadowing techniques between 2 auditory sources, and which as such, did not promote input from echoic memory yielded the Gestaltian figure/ground theories of attention. However, more recent estimates of attentional limitations (Huron, 1989; Gregory, 1990; Ludlam, 1993) whose experimental techniques involved target detection in multiple streams (a procedure in which echoic memory can contribute to the performance) produced the higher attentional limit estimate of 3 streams.

Within the general framework of models of attention, the present findings suggest a more flexible mode and complex model of attentional processing. The finding of 3-stream capacity before a sudden increase in errors indicates a bimodal system, an element of parallel processing giving way under increased pressure to a more serial mode. But the fact that this capacity is lower under the Non-echoic condition indicates that the apparent parallel processing is dependent on support from echoic memory. Furthermore, the fact that reaction times show a steady increase proportional to the number of simultaneous streams implies an underlying serial process, even under conditions of echoic memory contribution. However, under natural listening conditions echoic memory may be viewed as an integral part of the attentional system. For this reason the mechanism of attention might best be depicted as a predominantly parallel and elastic processing system showing evidence of strain (increasing reaction times) but partially absorbing increased demands until its limit is reached. At this point processing changes to a predominantly serial mode (non-linearity of errors), the elastic limit being influenced by specific situational characteristics.

The demonstration of the effect of memory on attention processes shows that the two concepts are virtually in a symbiotic relationship. Previous research has failed to acknowledge this. It appears that the Wundtian quest for the atoms of perception has reached as far as it can into purely auditory attention. It seems impossible to only measure attention without involving memory.

4 References

Broadbent, D. E. (1958). Perception And *Communication. Oxford: Pergamon.*

Treisman, A. M. (1964). Selective Attention In Man. British Medical Bulletin. *20, 12-16.*

Sloboda, J. A. and Edworthy, J. (1981). Attending To Two Melodies At Once: The Effects Of Key Relatedness. Psychology Of Music. *9 (1), 39-43.*

Huron, D. (1989). Voice Denumerability In Polyphonic Music Of Homogenous Timbres. Music Perception. *6 (4), 361-382.*

Gregory, A.H. (1990). Listening To Polyphonic Music. Psychology Of Music. *18 (2), 163-170.*

Ludlam, R. (1993). Undergraduate Dissertation, City University, London.

Cowan, N. (1984). On Short And Long Auditory Stores. Psychological Bulletin. *96 (2), 341-370.*

Baddeley, A. D. and Hitch, G. J. (1974). Working Memory. In G.H. Bower (Ed.) The Psychology Of Learning And Motivation. *(Vol. 8). London: Academic Press.*

Baddeley, A. D., Grant, S., Wight, E. and Thompson, N. (1975). Imagery and Visual Working Memory. In Rabbitt, P. M. A. and Dornic, S. (Eds.) Attention And Performance *(Vol.5). London: Academic Press.*

Salamé, P. and Baddeley, A. (1982). Disruption Of The Short-Term Memory By Unattended Speech. Implications For The Structure Of Working Memory. Journal Of Verbal And Learning Behaviour. *21 (2), 150-164.*

Keller, T. A., Cowan, N., and Saults, J. S. (1995). Can Memory For Tone Pitch Be Rehearsed? Journal Of Experimental Psychology: Learning, Memory and Cognition. *21 (3), 635-645.*

Bregman, A. S., (1990). Auditory Scene Analysis: The Perceptual Organisation Of Sound. *MIT Press. Cambridge.*