Scandinavian Audiology

ISSN: 0105-0397 (Print) (Online) Journal homepage: http://www.tandfonline.com/loi/iaud20

Aspects of Short-term Auditory Memory as Revealed by a Recognition Task on Multi-tone Sequences

Silvano Prosser

To cite this article: Silvano Prosser (1995) Aspects of Short-term Auditory Memory as Revealed by a Recognition Task on Multi-tone Sequences, Scandinavian Audiology, 24:4, 247-253, DOI: 10.3109/01050399509047544

To link to this article: http://dx.doi.org/10.3109/01050399509047544

Published online: 12 Oct 2009.

Aspects of Short-term Auditory Memory as Revealed by a Recognition Task on Multi-tone Sequences

Silvano Prosser

Audiology Service of ENT Clinic, University of Ferrara, Italy

Prosser S. Aspects of short-term auditory memory as revealed by a recognition task on multi-tone sequences. *Scand Audiol* 1995;24:247–53.

Retention capacity and temporal aspects of auditory short-term memory have been investigated through a multi-tone sequence paradigm and recognition task. There were 37 normally hearing subjects, with (16) and without (21) musical education. They were required to judge whether or not a tonal sequence comprised a tone probe, which was presented following the sequence. The subjects' performance appears to be mainly dependent on the number of the sequence components ($n = 2, 4, 6$), while sequence-probe interval (1, 3, 7 sec) represents a factor of minor weight. In addition, a strong recency effect has been shown for the last sequence component, also extending to the antecedent ones when sequence-probe interval is short. In contrast to studies on verbal short-term memory, no primacy effect has been demonstrated. Musically experienced subjects performed significantly better than naive counterparts. These results have proved to be repeatable in separate groups of subjects, and sensitive to auditory skills associated with musical practice. It is likely that tests of auditory memory based on tonal sequences could be useful in the clinical assessment of subjects with suspected central auditory dysfunction, or subjects who have had a cochlear implant after variable periods of auditory deprivation.

Key words: Auditory memory, auditory tonal memory, short-term memory

Received: July 12, 1994/Accepted January 31, 1995

Address for offprints: *Silvano Prosser, Servizio di Audiologia-Clinica O.R.L., Università di Ferrara, 44100-Ferrara, Italy*

Introduction

Short-term auditory memory (STAM) has mostly been investigated through experimental designs based on speech material. STAM is thought to play an important role in the stages of speech processing. The ability to compensate for the fragmentation of the speech signal and functions involving serial analysis on speech units are examples where STAM could be active, allowing portions of the acoustic signal to be temporarily stored before being definitively processed. On the other hand, measurements of STAM based on speech stimuli have been criticized, since they could reflect more general aspects of memory, and be influenced by mechanisms specific to speech processing, or, in short, by linguistic competence. The hypothesis that acoustic stimuli lacking semantic content should be more effective in directly assessing STAM functions was put forward by Pollack (1972). Recently, McFarland & Cacace (1992) and Cacace & McFarland (1992), using a paradigm based on a binary tone pattern and three alternative forced choice (AFC) responses, were able to demonstrate the typical phenomena of STAM, i.e. performances decreasing with forgetting time and pattern complexity, primacy and recency effects.

The present work aimed to find out whether such phenomena were also evident with an item recognition task based on multi-tone sequence and two AFC responses. In addition to investigating the consequences of changing some stimulus parameters (sequence length and forgetting time), we were also interested in the STAM performances in two selected groups, teenage students and musicians.

Material and Methods

Subjects

All were volunteers, right-handed, normally hearing, without histories of neurological diseases. 1. Pilot group (PIL): 5 subjects, age 27–32 years, high educational level, without musical education. 2. Student group (STU): 16 subjects, age 16–17 years, frequenting a technical school, without musical education. 3. Musician group (MUS): 16 subjects, age 18–38 years, belonging to a concert group; musical education ranging between 4 and 12 years.

Procedure

A PC-386 computer was programmed for stimuli generation, experimental condition control, and response data processing. A set of 14 pure-tone stimuli were used ranging between 504 and 1557 Hz (i.e.: 504, 549, 599, 653, 712, 777, 847, 924, 1008, 1100, 1199, 1308, 1427, 1557 Hz) and separated from each other by intervals progressively rising from 43 to 73 mels. These intervals were chosen to avoid a close similarity to the musical scale: they correspond to a frequency ratio between contiguous frequencies of 1.09 vs 1.03 of the tempered scale. Sequences were constructed consisting of 2, 4, 6 tones, followed at an interval of 1, 3, 7 s, by a tone probe, for a total of nine experimental conditions. The probe could be equal to one of the sequence tones, or different. When different, it was randomly chosen from the remaining tones of the original set. The sequence-probe complex constituted a trial. All the pure-tone stimuli had a duration of 500 ms ('on'), with a rise/decay of 25 ms; they were delivered with an inter-stimulus pause of 500 ms ('off'). The stimuli constituting a sequence differed in frequency from each other, being quasi-randomly chosen from the original set of 14. This process occurred every time a sequence was delivered, such that each sequence was different from the preceding one. The subject's task was to decide whether the tone probe was present or absent within the sequence. The presence/absence ratio was set at 10:7 (1:1 in the PIL group), and the order of presence-absence conditions was random. In the trials with the probe present, the ordinal position of the corresponding target (the stimulus equal to the probe) was random. At least five trials of probe-target equality (15 in PIL group) were repeated for each ordinal position. Therefore the number of trials for each of the nine experimental conditions (blocks) was dependent on the length of the sequence. For example, a block for sequences of 4 stimuli consisted in total of 34 trials: 20 of them had the probe present and 14 (10:7) had the probe absent. Out of the 20, 5 had the target corresponding to the first stimulus of the sequence, 5 to the second and so on.

The blocks were consecutively presented to the subjects, their order of presentation was random. The acoustic stimuli were transduced through a loudspeaker frontally placed at 50 cm from the subject's head, and delivered at 65 dB SPL, in a quiet, but not sound-treated, room. At each trial the response was manually given by pressing one of two adjacent keys of the computer keyboard, and according to the signal detection theory, stored as hit (HI), miss (MI), correct rejection (CR) or false alarm (FA). Two seconds after the response another trial was presented, and so on, until the completion of the block. The presentation of the blocks was subdivided into sessions of 20 min and a rest pause of 10 min was given after each session. Subjects underwent no more than 4–5 sessions daily. The subjects were left free in response criterion, and no specific instructions were given as regards the ratio of the presence/absence of the probe within the sequence. PIL group underwent all nine experimental conditions. STU group underwent two conditions: 4-tone sequence with 1 and 7 sec sequence-probe intervals. MUS group underwent four conditions: 4- and 6-tone sequences, with 1 and 7 sec sequence-probe intervals.

Analysis of Data

The subject's performance was defined for each experimental condition by a score defined HICO. It was calculated by assuming the HI and FA response rates to be independent, according to the formula: HICO = (HI-FA/100-FA)*100 (Green & Swets, 1966). The aim of this transformation is to compensate for the different criteria the individuals may adopt in the discrimination task, or, in other words to 'clean up' the hit rate from effects related to casual guessing. Statistical analysis on percentage data was performed after arcsin transformation.

Results

Fig. 1 shows individual and mean data as obtained by the PIL group. To give an impression of inter-subject variability, HI and FA response rates are plotted separately as a function of the three sequence-probe intervals (1, 3, 7 sec) for the 2-, 4-, 6-tone sequences. For each sequence the HI response rate clearly decreases as the sequence-probe interval increases from 1 to 7 s, while the FA rate appears to be relatively invariant. In addition, the HI rate also decreases and the FA rate increases with sequence length. Fig. 2 shows mean HICO data as a function of the sequence-probe interval, with sequence length as parameter.

HICO data were analysed by an analysis of variance (factorial design 5 subjects × 3 sequences × 3 intervals). HICO values are significantly influenced by sequence length ($F(8,2) = 17.0$, $p < 0.000$) and less by intervals ($F(8,2) = 5.6$, $p = 0.004$), while the contribution of inter-subject difference is insignificant ($F(8,4) = 2.3$, $p = 0.06$). These factors account for 57% of the total variance, and their effects are additive (2-way and 3-way interactions are insignificant).

Fig. 3 shows HICO values as observed in the PIL group, for 2-, 4-, 6-tone sequences and 1, 3, 7 sec sequence-probe intervals, as a function of the ordinal position occupied by the target within the sequence. In the 2-tone sequence a recency effect, i.e. the higher probability of the last element of the sequence to be correctly matched to the probe, is clearly evident with the longest sequence-probe interval. In the 4- and 6-tone sequences the recency effect is even more marked: it appears to be maximal with 6-tone sequences, being the mean scores enhanced from percentages relatively close to chance level (corresponding to 0% for HICO) to percentages approximating 100%. In addition, with sequence-probe intervals of 1 sec a recency also seems evident, albeit of a minor degree, on the penultimate (in both 4- and 6-tone sequence) and ante-penultimate (in the 4-tone sequence) elements. HICO data relative to ordinal positions of 4- and 6-tone sequences were analysed through an analysis of variance (factorial design: 3 intervals × 3 positions (the last 3 positions of each

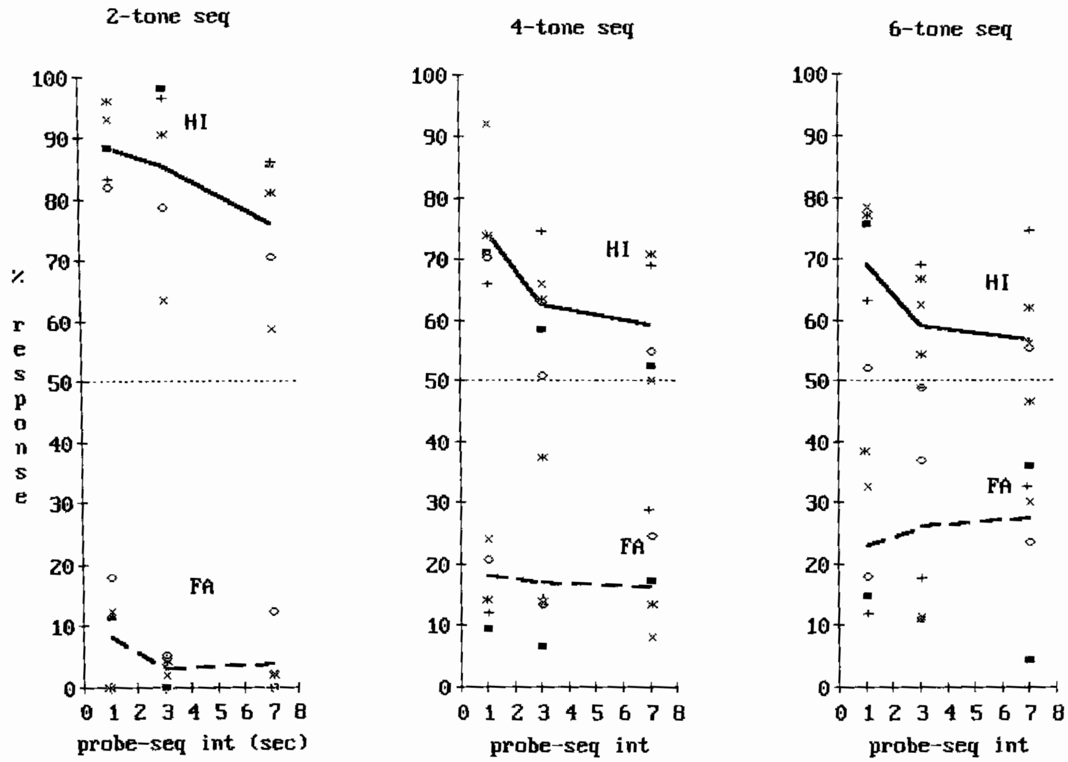

Fig. 1. Data from pilot group (PIL, $n = 5$). Individual hit (HI) and false alarm (FA) response rate as a function of probe-sequence interval, for sequences of 2, 4, 6 tones. Mean data are linked by continuous and interrupted lines. Pointed line: level of chance.

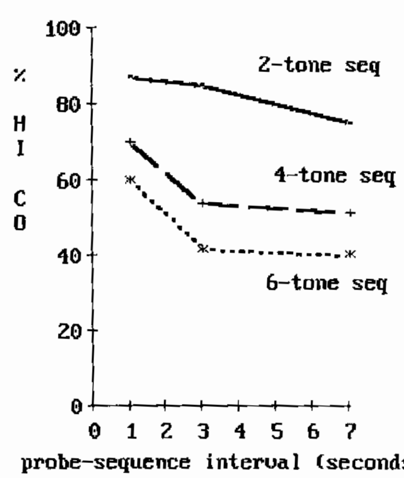

Fig. 2. PIL group. Percentage mean values of corrected hit (HICO) response as a function of probe-sequence interval. For HICO, level of chance = 0%.

sequence) × 2 sequences (4-, 6-tone)). Altogether, these factors account for 55% of the total variance. The scores are significantly influenced mainly by the target position (74% of the explained variance. $F(5,2) = 38.8$, $p < 0.000$) and less by sequence-probe intervals (17% of the explained variance, $F(5,2) = 9.2$, $p < 0.000$). The contribution of sequence length is not significant, indicating that the observed recency effect does not depend on the number of sequence elements here considered. Effects of the three independent variables are additive, since 2- and 3-way interactions are not significant.

Fig. 4 shows HICO mean rates for the experimental conditions common to the PIL, STU and MUS groups. As the sample size of the PIL group ($n = 5$) was smaller than that of STU ($n = 16$) and MUS groups ($n = 16$), a t-test for independent measures was used to test group differences, independently of sequence and interval, with PIL group data taken as reference. Statistical results indicate that performance of the STU group is similar to that of PIL group,

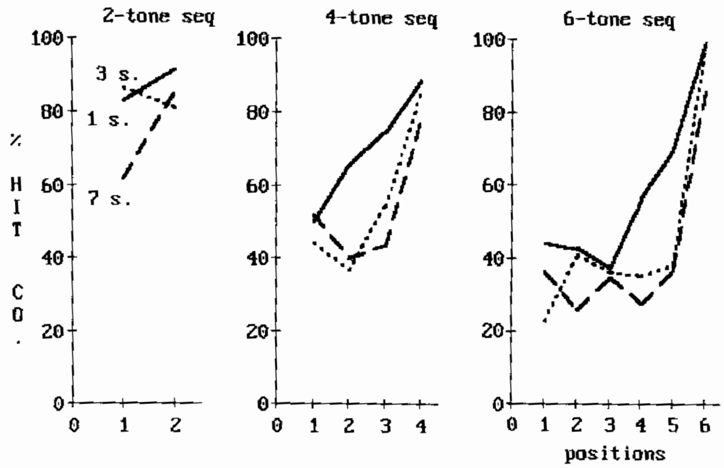

Fig. 3. PIL group. Percentage mean values of HICO as a function of the ordinal position occupied by the target within 2-, 4-, 6-tone sequence. Data are shown separately for sequence-probe intervals of 1, 3, 7 sec.

while performance of the MUS group is significantly higher ($t(82) = 2.02$, $p < 0.05$). In addition, the relative reduction of HICO mean values with probe-sequence intervals, appears to be similar in the three groups. This indicates that the effect of probe-sequence interval is relatively constant, in spite of group differences in the HICO absolute values.

Fig. 5. shows the HICO mean rates as a function of the ordinal position occupied by the target within the sequences for the experimental conditions common to the three groups. A recency effect is confirmed by both the STU and MUS groups for all the conditions. The HICO rate of the MUS group appears to be higher than that of the PIL and STU groups across the experimental conditions. With 6-tone sequences the better performance of MUS group compared to the PIL group appears to be caused by a more consistent effect of recency, which manifests itself on the 2 or 3 tones before the last one.

The arrows in the graphs indicate significant differences ($p < 0.05$) between the STU and MUS groups and the PIL group, as obtained by a t-test for independent measures on data relative to the ordinal positions within each sequence and interval condition.

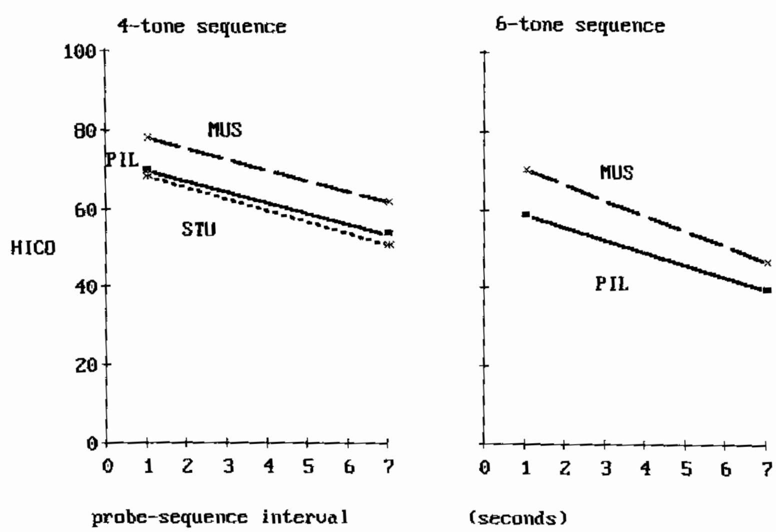

Fig. 4. Percentage mean values of HICO for PIL ($n = 5$), student (STU, $n = 16$), and musician (MUS, $n = 16$) groups.

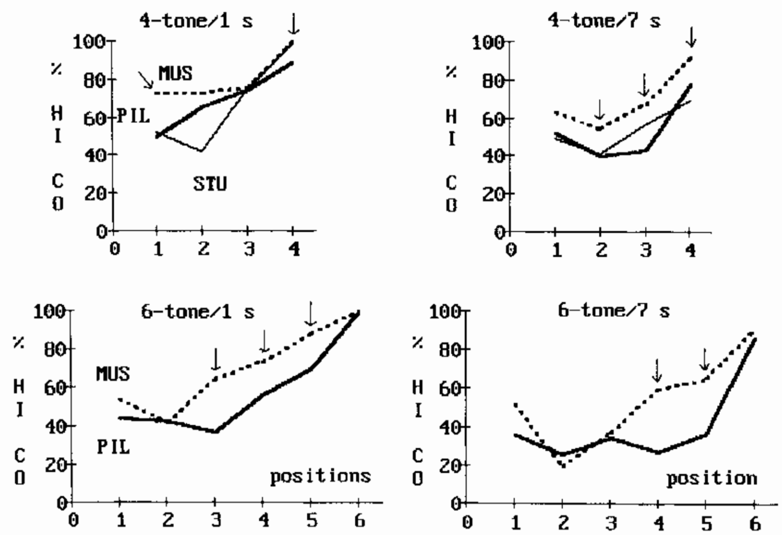

Fig. 5. Percentage mean values of HICO as a function of the target ordinal position, for the three experimental groups. Arrows indicate significant difference from PIL data.

Discussion

These results indicate that, in general, the correct recognition of a single tone within a multi-tone sequence depends on the number of elements constituting the sequence, and to a lesser extent on the sequence-probe interval. However, by analysing the effect of the ordinal position occupied by the target within the sequence it appears that, at least with 4- and 6-tone sequences and intervals of 3 and 7 sec, the recognition rate is mainly determined by a recency effect. With 2-tone sequences the recognition rate is well above the chance level, relatively independent of sequence-probe intervals, being a recency effect clear enough with the 7 sec interval only. The classical model of auditory memory comprising an echoic stage and a short-term stage may provide a first framework for explaining our results. The echoic stage should be characterized by a quick decay (within 1 sec) of acoustic information (Crowder & Morton, 1969); the short memory stage is classically described as limited in capacity (about seven items according to Miller, 1956) with a decay time estimated in the order of 10 sec. It is accepted that active processes possibly involving an articulatory loop take place at this latter stage (Atkinson & Shiffrin, 1971; Baddeley, 1992). The decay of memory for sequentially presented items has been explained by the interference each newly presented element should exert on the preceding ones, or by a time-dependent cancellation of the mnesic traces, or by the concomitant effects of these two factors (Lindsay & Norman, 1973).

Because of the kind of stimuli we applied, assumed as lacking in semantic content, at least for the musically naive subjects, and the temporal parameters of the presentation, it is likely that the results reflect functions mostly related to echoic memory, and to a lesser extent to short-term memory. Tonal sequences should indeed be difficult to rehearse, since they presumably lack corresponding perceptual categories which, instead, facilitate the retention of speech elements through silent repetition. The timing of sequence-probe presentation could also contribute to make rehearsal less effective: in fact, our experimental design includes some conditions in which the sequence-probe interval duration (i.e. 1 sec) is too short for a silent repetition of long sequences to be completed (i.e. 4 or 6 tones delivered at 1/sec). However, in these conditions, the recognition rates of the penultimate and ante-penultimate elements are significantly higher than those obtained at sequence-probe intervals of 3 and 7 sec, which theoretically should facilitate an element-by-element rehearsal. While it is intuitive that echoic memory facilitate the recognition of the last sequence element, it is likely that its effect may extend also on the antecedent elements, provided the sequence-probe interval is short. As this interval increases to 3 and 7 sec there is a reduction in the recognition rate of the targets in intermediate positions. However, the performance remains above the level of chance, probably reflecting the effect of an active process of rehearsal which should contrast the mechanisms responsible for temporal memory decay. Independent of the number of elements, we observed that increasing the sequence-probe interval from 1 to 7 sec causes a decrease in HI rate of 12–15% (17–21% for HICO rate). Comparable results have been obtained by other authors (Eriksen & Johnson, 1964; Cacace & McFarland, 1992), albeit with different experimental designs. Cowan et al. (1990), using speech items found a decrease of 21% and 26% in recognition rate of vowels and consonants respectively, when the interval changed from 1 to 10 sec.

A primacy effect, i.e. the advantage of the first items in a sequence in being correctly recognized, has been generally found consistent with speech material, and interpreted as reflecting functions related to long-term memory (Murdock, 1962). Our results failed to show a clear primacy effect: an interpretation could be that mnesic traces do not exist in long-term memory specific enough to allow for a categorical recognition and late retrieval of tonal stimuli.

Our results show that the decrease in performance due to an increase of the sequence-probe interval is relatively small compared to the more robust effect due to the number of tones within the sequence. Watson et al. (1990) observed a comparable performance decrease in a series of experiments designed to investigate the detection of frequency changes occurring in single elements of multi-tone patterns. Although these authors used tonal patterns of shorter duration than ours, the similarity of results substantiates the hypothesis that also in a recognition task the number of acoustic components that can be held in immediate memory is the main limiting factor. In fact, Kidd & Watson (1992) demonstrated that, in addition to the number of sequence elements, the proportion of time occupied by the target within the sequence also substantially contributes in changing performance. It is possible that in our results the effect of the number of elements may have been confused with the target duration proportionality, since our sequences consisting of isochronous elements had a duration depending on the number of elements. This leads to a different proportionality of the target duration, relative to the sequence duration: 25%, 12.5%, 8.3% respectively for the 2-, 4-, 6-tone sequences.

Another confounding factor which may increase the difficulty of the recognition task is related to the distance in frequency between the tonal elements of the sequence-probe complex. In our experiments each sequence consisted of stimuli differing in frequency from each other, randomly chosen from the original set of 14 tones; however, there was a probability that two or more elements were contiguous in frequency. Although the maximal contiguity in frequency corresponded to 43 mel, well above the threshold of frequency discrimination, it is likely that within some sequences the pitch contiguity between elements contributed to the difficulty in perceptually retaining a sufficient separation between the corresponding mnesic traces. A similar factor, defined as acoustic similarity, has been put forward by Conrad (1964) to explain the difficulty in recalling speech elements scarcely differentiated in their phonetic characteristics.

The data of the STU group confirmed, at least for the common experimental conditions, the mean values of correct response rates and their dependency on sequence-probe intervals previously measured in the PIL group. The data of the MUS group, significantly higher than those of the PIL group, may be interpreted as reflecting different auditory skills. It is likely that musicians are able to perceive tonal differences in a more categorical way than naive subjects, so that the mnesic trace corresponding to sequences of tonal elements could be robust enough to contrast the adverse effects of interference and temporal memory decay. Consequently, even the silent repetition of the tonal sequence should be more effective in enhancing the probability of a correct probe-target match. This ability is probably responsible for the higher recognition rate the musicians demonstrated for the two or three intermediate sequence elements.

In conclusion, the above results indicate that an experimental paradigm based on multi-tone sequences and recognition task may provide some measurements which reflect the typical mechanisms of STAM. These measurements have been shown to be repeatable within musically naive groups (PIL and STU); the results from the MUS group indicate that the recognition task we employed is sensitive enough to reveal the auditory skills associated with musical education. In our opinion this kind of paradigm could be useful in assessing STAM functions in subjects with a suspected central auditory dysfunction, or subjects who have had a cochlear implant after variable periods of auditory deprivation.

References

Atkinson RC, Shiffrin RM. The control of short-term memory. Sci Am 1971; 225: 82–90.

Baddeley A. Working memory. Science; 1992; 255: 556–8.

Cacace AT, McFarland DJ. Acoustic pattern recognition and short term memory in normal adults and young children. Audiology 1992; 31: 334–41.

Conrad R. Acoustic confusion in immediate memory. Br J Psychol 1964; 55: 75–84.

Cowan N, Lichty W, Grove TR. Properties of memory for unattended spoken syllables. J Exp Psychol 1990; 2: 258–69.

Crowder RG, Morton J. Perception acoustic storage. Percept Psychoph 1969; 5: 365–73.

Eriksen CW, Johnson HJ. Storage and decay characteristics of unattended auditory stimuli. J Exp Psychol 1964; 68: 28–36.

Green DM, Swets JA. Sensory processes in detection. In: Green DM, Swets JA, eds. Signal detection theory and psychophysics. New York, Sidney: Wiley, 1966: 149–208.

Kidd GR, Watson CS. The 'proportion-of-duration rule' for the discrimination of auditory patterns. J Acoust Soc Am 1992; 92: 3109–18.

Lindsay PH, Norman D. Transient memories. In: Lindsay PH, Norman D, eds. Human information processing. New York, London: Academic Press, 1973; 328–72.

McFarland DJ, Cacace AT. Aspects of short-term acoustic recognition memory: modality and serial position effects. Audiology 1992; 31: 342–52.

Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956; 63: 81–97.

Murdock BB Jr. The serial effect of free recall. J Exp Psychol 1962; 64: 482–8.

Pollack I. Memory for auditory waveform. J Acoust Soc Am 1972; 52: 1209–15.

Watson CS, Foyle DS, Kidd G. Limits of auditory pattern discrimination for patterns with various durations and number of components. J Acoust Soc Am 1990; 88: 2631–8.