Perception, 2003, volume 32, pages 1393 – 1402

DOI:10.1068/p5035

Cross-modal and non-sensory influences on auditory streaming

Robert P Carlyon, Christopher J Plack¶, Deborah A Fantini¶, Rhodri Cusack
Medical Research Council, Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 3EF, UK; e-mail: bob.carlyon@mrc-cbu.cam.ac.uk; ¶ Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
Received 20 January 2003, in revised form 6 June 2003

Abstract. Carlyon et al (2001 *Journal of Experimental Psychology: Human Perception and Performance* **27** 115 – 127) have reported that the buildup of auditory streaming is reduced when attention is diverted to a competing auditory stimulus. Here, we demonstrate that a reduction in streaming can also be obtained by attention to a visual task or by the requirement to count backwards in threes. In all conditions participants heard a 13 s sequence of tones, and, during the first 10 s saw a sequence of visual stimuli containing three, four, or five targets. The tone sequence consisted of twenty repeating triplets in an ABA – ABA ... order, where A and B represent tones of two different frequencies. In each sequence, three, four, or five tones were amplitude modulated. During the first 10 s of the sequence, participants either counted the number of visual targets, counted the number of (modulated) auditory targets, or counted backwards in threes from a specified number. They then made an auditory-streaming judgment about the last 3 s of the tone sequence: whether one or two streams were heard. The results showed more streaming when participants counted the auditory targets (and hence were attending to the tones throughout) than in either the 'visual' or 'counting-backwards' conditions.

1 Introduction

The extent to which perceptual organisation is influenced by attention has recently been the subject of research in both the visual (eg Joseph et al 1997; Mack et al 1992; Rock et al 1992) and the auditory (Bregman and Campbell 1971; Carlyon et al 2001a; Cusack et al, in press; Macken et al 2003; Sussman et al 1998, 1999) modalities. One reason for this interest is that the answer may shed light on the stages of neural processing involved in forming perceptual organisation, and, at the same time, on models of sensory processes. For example, the phenomenon of 'auditory streaming' (eg Anstis and Saida 1985; Bregman and Campbell 1971; van Noorden 1975) is sometimes modeled as an automatic process, in which the role of attention is restricted to selecting from competing representations at the output of low-level neural mechanisms (Beauvois and Meddis 1991, 1996). If, however, attention is shown to affect the streaming process itself, then these models will have to be recast, either in terms of cortical processes, or with more peripheral structures receiving input from those neural pathways involved in attention (eg McCabe and Denham 1997).

In audition, research on this question has, with few exceptions (eg Alain et al 2001), focused on the sequential organisation of sounds. The stimulus features that govern this organisation, or auditory streaming, can be illustrated with reference to the tone sequence introduced by van Noorden (1975), in which a pair of tones with frequencies A and B is presented in the sequence ABA – ABA ... (figure 1). When the repetition rate is slow, or when the frequencies of A and B are close, participants hear a galloping rhythm corresponding to the repeating triplets (figure 1a). At faster rates and wider separations, the sequence splits into two streams and the galloping rhythm is lost (eg figure 1b). The percept 'flips' pseudorandomly between these two possibilities throughout the sequence, but there is a strong tendency for stream segregation to build up over several seconds;

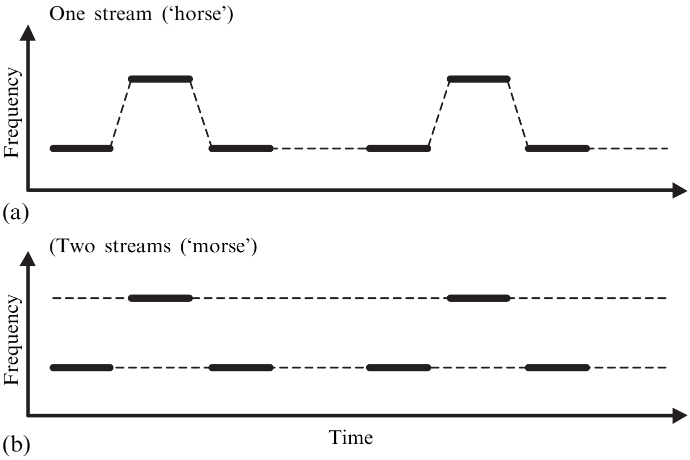

Figure 1. Schematic representation of van Noorden's 'galloping rhythm' stimulus, used in the instructions for the experiment. The lines in the top (a) and bottom (b) panels of the figure indicate the perceptual organisation into one stream ('horse') and two streams ('morse'), respectively.

hence participants often hear the beginning of a long sequence as a single stream, but hear later portions as two streams.

A challenge facing researchers in all modalities is measuring the perceptual organisation of a stimulus to which the participant is not fully attending. The challenge arises because, to test the effects of attending to the stimulus, one must divert attention away from it in some conditions. Consequently, reports of the perceived organisation may either be unreliable (for example not all attributes of the stimulus may be encoded in short-term memory), or reorient attention to the stimulus. In audition, three different approaches have been used to determine the extent to which auditory streaming is affected by attention.

One way of avoiding the requirement to make a judgment about unattended sounds is to measure the auditory evoked potential. The mismatch negativity (MMN) component of this potential is a negative wave that occurs in response to a rare 'deviant' sound presented in a sequence of otherwise homogeneous sounds. It can be measured even when participants are not required to attend to the sounds (Cheour-Luhtanen et al 1996; Kane et al 1996; Naatanen et al 1993). Sussmann and her colleagues (Sussman et al 1998, 1999) played participants a tone sequence that alternated regularly between a high-frequency and a low-frequency range, with the tones in each range playing a simple melody. On a minority of trials, they altered the order of tones in the low range, and argued that this should elicit an MMN only when the interfering high-frequency tones were pulled into a separate auditory stream. They found that, when the tones alternated at a fast rate, an MMN was observed even when participants were told to ignore the tones and to read a book. They concluded that, under these conditions, attention was not necessary for streaming to occur. However, at slower alternation rates, where streaming is typically weaker, they found that an MMN was obtained only when participants were told to attend to the tones.

An alternative method, adopted by Jones and his colleagues, is to measure the effects of streaming on a secondary task (Jones et al 1995, 1999; Macken et al 2003). The irrelevant sound effect (ISE) refers to the disruptive effect of auditory stimuli on the serial recall of a sequence of visually presented letters. Repetition of a single sound does not appear to disrupt performance, and Jones and colleagues have argued that "changes in state" from one sound to the next are crucial for the effect; so that, for example, two alternating tones produce more disruption than a single repeated tone.

Importantly, they have shown that when a sequence splits into two streams, such that each stream consists of a repetition of a single sound, then disruption is also reduced—*re* the case where participants hear a single stream of alternating tones (Jones et al 1995; Jones and Macken 1995; Macken et al 2003). They have argued that, because participants are instructed to ignore the tones, the streaming must have occurred in the absence of attention.

A somewhat different method was used by Carlyon et al (2001a), who measured the effects of diverting attention away from the *first* half of a 20 s tone sequence on the streaming observed during the *second* half. This method relies on the fact that streaming builds up over time. Carlyon et al presented 20 s sequences to the left ear of each participant, and, for the first 10 s, presented a sequence of noise bursts to his/her right ear. When participants ignored the noise bursts, and made streaming judgments about the tones throughout, then this streaming built up as expected. However, when they attended to the noise bursts for the first 10 s, and only then started making streaming judgments about the tones, the amount of streaming was greatly reduced compared to the first condition. Carlyon et al concluded that attending to a sequence of tones is important for the streaming of those tones to build up. A second experiment showed that although it was necessary for participants to attend to the tones, it was not necessary for them to continue to make streaming judgments about them (see also Cusack et al, in press). In that experiment, the tones were amplitude modulated (AM) at either a fast or a slow rate, with the rate of the AM switching on average every four tone bursts. The amount of streaming observed after 10 s was as great when participants judged the AM rate for the first half of the sequence and then switched to making streaming judgments, as when they made streaming judgments throughout. This controlled for an explanation in terms of response bias, according to which participants might default to making 'one stream' responses whenever they started to make streaming judgments. This finding is also important for the design of the experiment described here.

The results presented by Carlyon et al (2001a) concerned only the effects of a competing *auditory* task. There is some debate as to the extent to which the auditory and visual components of attention are independent (Arnell and Jolicoeur 1999; Duncan et al 1997). It is therefore not clear a priori whether the buildup of auditory streaming can be affected by attention to a visual stimulus. If so, then we can conclude that the neural locus of auditory streaming is quite central, and can be mediated by multimodal pathways. Conversely, if only auditory attention affects the buildup of streaming, then diverting attention to a visual task during early parts of a sequence should not affect the streaming reported at the end. This would help explain why experiments with the ISE and MMN often reveal some evidence of streaming when participants are reading a book or memorising visual sequences. With the present experiment we also investigated whether streaming could be disrupted by a cognitive task that required no sensory input.

2 Method

Before the start of the experiment, participants were played examples of the ABA sequences shown in figure 1, with the A-tone frequency equal to 1000 Hz and the B-tone frequency either one semitone or one octave higher. These stimuli were chosen because they produce unambiguous percepts of one and two streams respectively (eg van Noorden 1975). They were in a different frequency region to the tones in the main experiment, to convey the general concept of streaming without teaching participants to respond to a particular physical stimulus. The experimenter explained that the one-semitone separation gave rise to the percept of a galloping rhythm, and that the one-octave separation did not. As this latter stimulus sounds a little like Morse code, the abbreviations 'horse' and 'morse' were introduced to describe the two percepts.

The main part of the experiment consisted of three conditions, in which the auditory and visual stimuli were identical, and which differed only in the instructions to the participants. Demonstrations of the example stimuli and of the three conditions used in the main experiment are available at: http://www.perceptionweb.com/misc/p5035/.

2.1 *Stimuli*
The auditory stimulus was always a 13 s sequence of tones in the ABA sequence shown in figure 1, with the duration of each tone equal to 125 ms. The frequency of the A tone was always 400 Hz, and that of the B tones was, in different sequences, either four or six semitones higher. Most of the tones were steady, but, during the first 10 s of each sequence, either three, four, or five tones (the 'targets') were sinusoidally amplitude modulated (AM) at a rate of 16 Hz and a depth of 100%. Averaged across trials, equal numbers of A and B tones were modulated and, for a given trial, the difference between the number of A and B targets was always 0 or 1. This was done to encourage participants to attend to both the A and the B tones. All tones were diotically presented at a level of 65 dB SPL over Sennheiser HD580 headphones. The participant was seated inside a double-walled sound-attenuating booth.

During the first 10 s of the tone sequence, the visual stimulus consisted of a series of cross-hatched white ellipses presented on a black background. The orientation of the ellipse alternated every 125 ms between 0° and 90° (figures 2a and 2b, respectively). During each sequence, the diagonal lines making up the crosshatching in either three, four, or five target ellipses were replaced with broken lines (figures 2c and 2d). After 10 s the ellipses were replaced by the written message "horse or morse?", which prompted participants to start making streaming judgments (see section 2.2). The dimensions of

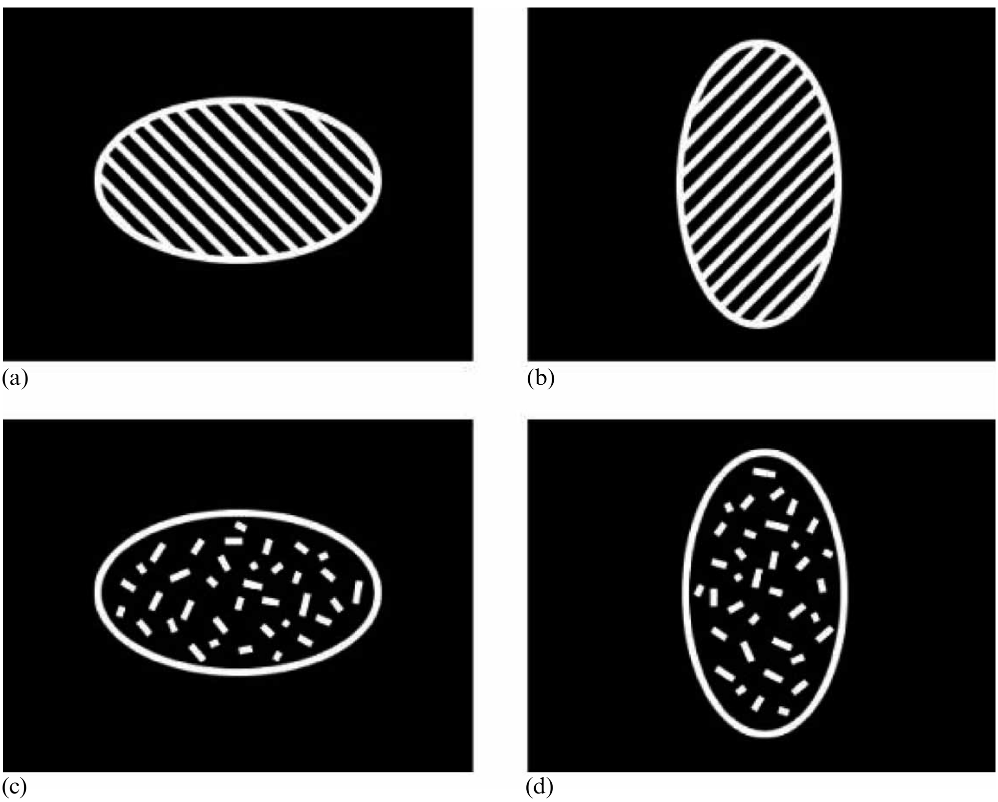

Figure 2. Snapshots of different frames of the movies used in the experiment. Panels (c) and (d) were targets in the 'visual' condition.

the video (including background) were $\sim$6.75 cm $\times$ 8.25 cm. The videos were displayed on a computer monitor located outside the booth, and viewed from a distance of $\sim$50 cm through a window.

2.2 *Procedure*

Eighteen participants, all of whom reported normal hearing, performed three conditions. Each of the six possible orderings of these conditions was assigned to three participants, and so the design was fully counterbalanced. In the 'auditory' condition, participants were instructed to pay attention to the tones, and to count the number of AM tones (the AM was described as a 'warble'). They were to ignore the ellipses, but to keep looking at the screen, and, when the message "horse or morse?" appeared, to judge whether the end of the auditory sequence was heard as one or two streams. This response was entered by clicking on one of two virtual buttons on the computer screen (labeled 'horse' and 'morse'), and they then entered the number of auditory targets heard.

In the 'visual' condition, participants were told to count the visual targets and to ignore the tones until they saw the prompt "horse or morse?", at which point they were to switch their attention to the tones and to judge whether the end of the auditory sequence was heard as one or two streams. They then entered their responses (horse versus morse, and number of visual targets) in the same way as for the auditory condition.

In the 'count-backwards' condition, participants were presented with a randomly generated number between 258 and 390 before each sequence. They were instructed to ignore the tones and ellipses, but to keep looking at the screen and to (silently) count backwards in threes until they saw the "horse or morse?" prompt. They then entered their streaming judgment and the last number reached by the counting procedure, as in the other two conditions. A response on the competing task (counting backwards) was defined as correct if the difference between the (given) starting number and the (reported) final number was an integer multiple of three.

2.3 *Predictions*

The main predictions were that (i) if diverting attention to a visual task can reduce auditory streaming, then there should be fewer two-stream (morse) judgments in the visual condition than in the auditory condition, and (ii) if diverting attention by means of counting backwards in threes can reduce auditory streaming, then there should be fewer two-stream judgments in the 'count-backwards' condition than in the auditory condition. Both of these predictions are derived from the finding that performing a competing task on the tones does not reduce the buildup of streaming, but that performing a task on another stimulus does (Carlyon et al 2001a). An additional prediction is that, in all conditions, more stream segregation should be observed at the six-semitone than at the four-semitone frequency separation.

It should be noted that, although we measured performance on the competing tasks in order to check that participants were indeed attending to them, it is quite possible that the cognitive load may have differed between the tasks. Our rationale is that, if we find that either the visual task or the counting-backwards task do produce fewer morse judgments than when participants attend to the tones throughout, then we will have demonstrated an effect of non-auditory attention on streaming. What we can not conclude is whether one or other of these effects would have been larger if the competing task had been made more demanding.

3 Results

The proportion of two-stream judgments is shown for each condition in figure 3. It can be seen that, consistent with previous findings (Anstis and Saida 1985; van Noorden 1975), more stream segregation is observed at the wider frequency separation. More importantly, the proportion of two-steam judgments is greater for the auditory condition

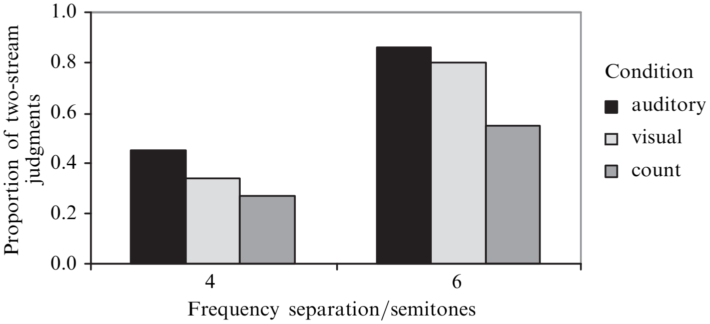

Figure 3. Proportion of two-stream (morse) judgments made in the three conditions of the experiment, averaged across participants. Data obtained with a frequency separation of four semitones are shown on the left; those with a separation of six semitones are on the right.

than for the other two conditions. To assess the statistical significance of these trends, the data were entered into a three-way ANOVA, with within-participants factors of competing task and frequency separation, and a between-participants factor of the order in which the conditions were performed. There were main effects of task ($F_{2,24} = 6.4$, $p = 0.006$) and frequency separation ($F_{1,12} = 49.4$, $p < 0.001$), but no effect of order, and no interactions. Pairwise comparisons revealed that the auditory condition produced significantly more two-stream judgments than either the visual ($p = 0.004$) or the counting-backwards ($p = 0.014$) conditions. The difference between the visual and the counting conditions, although numerically quite large, was not significant ($p = 0.074$). This was because it varied substantially across participants.

The proportion of trials in which participants correctly performed the competing task was lower in the auditory condition (60%) than in either the visual (94%) or the counting-backwards (82%) conditions. However, this was largely due to the results of five participants, who appeared to have misunderstood the instructions for the auditory task. They reported hearing about twenty targets on some trials, consistent with them either counting the number of B tones, or the total number of triplets. To determine whether the effects observed were strongly influenced by these participants, the data were re-analysed with their results excluded. Because there were now unequal numbers of participants for the six possible orderings of condition, the data were re-analysed with a two-way (frequency separation $\times$ task) ANOVA, without order being entered as a between-participants factor. This was considered justified as there was no effect of order in the main analysis. The new analysis produced results that were very similar to the original. Again, there were main effects of frequency separation ($F_{1,12} = 47.8$, $p < 0.001$) and of task ($F_{2,24} = 6.0$, $p = 0.008$), and no interaction. Pairwise comparisons also yielded similar results as before, with the auditory condition remaining significantly different from the visual ($p = 0.007$) and the counting-backwards ($p = 0.011$) conditions. Finally, we should mention that this finding is sufficiently robust that we have obtained similar results on two other occasions, when the experiment was run in an undergraduate laboratory class (Carlyon et al 2001b).

4 Discussion

4.1 *Effects of attention on auditory streaming*

There is now converging evidence from several different paradigms in favour of an important role for attention in the time course of auditory streaming. Carlyon et al's (2001a) original experiments with healthy volunteers were accompanied by evidence that stroke patients with a deficit in attending to visual objects in the left side of space also showed reduced streaming of sounds presented to their left ears. This reduction, obtained with reference both to stimuli presented to the left ears of controls and to the right ears of the same patients, complements our conclusion that cross-modal or supra-modal attentional mechanisms affect auditory streaming. More recently, Cusack et al (in press) have shown that, once streaming has been allowed to build up, diverting attention for a few seconds to a competing task causes that buildup to disappear. Conversely, simply suspending streaming judgments for a few seconds while continuing to attend to the tones did not affect the buildup, thereby providing further evidence against the idea that Carlyon et al's (2001a) results were due to participants defaulting to making "one stream" responses whenever they were required to start making streaming judgments about the tones. Additional evidence against an interpretation in terms of response bias comes from the present study, in which participants made only a single judgment about the streaming of each sequence.

Cusack et al (in press) proposed an interesting variation of Carlyon et al's (2001a) interpretation. They suggested that hearing a single stream at the beginning of a sequence may reflect an ecologically desirable form of processing. It is most preferable that when something new is to be attended, the initial perceptual organisation is of higher-level, less fragmented groups. On entering a room containing someone talking and some music, the objects available for attention should be either the speech or the music—not some fragment of one of them. Cusack et al proposed hierarchical decomposition, with some initial grouping, and then only attended streams in the auditory scene fragmenting further. The current study extends this idea to cross-modal switches of attention. It may be that, when a switch is made to a set of auditory stimuli following attention to another modality or task, a few large-scale perceptual groups are available for selection. Whichever of these is attended to then fragments further over time.

Cusack et al's (submitted) results are also consistent with an alternative explanation, according to which the default representation is for an auditory scene to fragment, only to be reset to a single stream whenever participants pay attention to it. According to this interpretation, streaming may be reset whenever attention is drawn to a sequence, either because participants have been instructed accordingly or simply because the sequence has just started. Hence the buildup observed in many previous studies (eg Anstis and Saida 1985; van Noorden 1975) may be more properly thought of as a "recovery from resetting". The available evidence, including that from the present study, does not distinguish between this explanation and one in which the default is for a single stream to be heard, with an active buildup being dependent on the extent to which the sequence is attended. What does seem clear, though, is that supra-modal attentional mechanisms mediate auditory streaming.

4.2 *Counting-backwards task*

The number of "two stream" responses was lower when participants were instructed to count backwards in threes during the first 10 s of the tone sequence than in the auditory condition, where they were required to attend to the tones throughout. Because they were instructed to count silently, this finding provides evidence that diverting attention away from a sequence can reduce the buildup of streaming, even when the auditory input is held constant across conditions. An alternative interpretation is that the phonological representation of the numbers and of the tones interfered in some form of common memory store, such as the phonological loop of the working memory model (Baddeley and Hitch 1974). Of course, some form of this interference may also have occurred in the visual and the auditory conditions, as a result of the need to count the targets. However, there are two reasons to think that, if this interference did take place, then it would have been greater in the counting-backwards condition. First, whereas the average number of targets reported (and hence the 'amount of counting') was very similar in the visual (3.9) and the auditory (4.0) conditions, participants counted backwards an average of 6.3 times.[1] Second, the numbers used in the counting-backwards condition contained three digits (eg three hundred and sixty-nine), and would each have involved more phonological activity than the single-digit numbers of auditory and visual targets. Hence, it is possible that the reduction in streaming produced in the counting-backwards condition, while being attributable to some form of cognitive activity, can be explained without recourse to attention per se. What we can conclude, however, is that a purely cognitive task performed during the first part of a tone sequence can affect the streaming perceived at the end—a finding inconsistent with purely stimulus-driven accounts of streaming (eg Beauvois and Meddis 1991, 1996).

4.3 *Visual competing tasks*

The difference in two-stream responses between the auditory and the visual conditions, although robust, was fairly small: 45% versus 34% at the four-semitone separation, and 86% versus 80% at the six-semitone separation. One reason for this may be that, in our concern to avoid such a difference arising from factors unrelated to the effects of attention on streaming, we designed features of the experiment to act against the predictions of the experiment. For example, in the auditory condition, we took care to ensure that the A and B tones were equally likely to be targets, because this would more strongly encourage participants to integrate the percepts of the A and B tones than if, for example, only the B tones were modulated (which might encourage subjects to 'hear out' a separate stream consisting of the B tones).

Nevertheless, it is clear that, despite the requirements to perform a visual task, and the fact that average performance on this task reached 90%, significant amounts of auditory streaming were still observed at the end of the sequence. This may help explain why effects of auditory streaming can be observed in the MMN and ISE paradigms, even when attention is diverted to a visual task. For example, Macken et al (2003) have interpreted the reduction in ISE caused by increasing the frequency separation between tones (Jones et al 1999) in a sequence as evidence that auditory streaming occurs outside the focus of attention. Similarly, the fact that the MMN to a frequency or duration difference in a sequence of tones in one frequency is reduced when 'interfering' tones are presented in the same but not a different frequency region has been interpreted as evidence for pre-attentive auditory streaming (eg Shinozaki et al 2000). We think it likely that, if we could have replaced our participants' direct report of the number of streams heard at the end of the sequence with an indirect measure such as the ISE or the MMN, then we would still have observed an effect of frequency separation on the results. However, this would not mean that streaming is independent of attention. Rather, the fact that a very large reduction in streaming can be obtained with a different type of competing task suggests the opposite: when participants spend the first 10 s of a tone sequence judging whether each of a sequence of noise bursts presented to the contralateral ear has a rising or falling amplitude contour, the percentage of two-steam judgments at a six-semitone separation observed after 10–13 s drops from about 50% to less than 10% (Carlyon et al 2001a).

Overall, we believe that the results of this and of previous studies lead to the conclusions that (i) the buildup of streaming can be inhibited by diverting attention to a competing task in either the visual or the auditory modality, and (ii) this inhibition can be very large but depends on the precise nature of the competing task. What we do not yet know is whether the smaller reduction observed here, compared to that reported with the noise-burst task, is due to the difference in stimulus modality or to other features of the competing task.

[1] These calculations excluded those five participants who reported hearing very large numbers of targets in the auditory condition (see section 3). The number of times that a participant counted backwards in threes was defined as one third of the average difference between the (given) starting number and the (reported) final number.

Acknowledgments. We thank Brian Cox for his help in designing and producing the movie sequences used in the experiment, and Dennis Norris for helpful discussions.

References

Alain C, Arnott S R, Picton T W, 2001 "Bottom – up and top – down influences on auditory scene analysis: evidence from event-related brain potentials" *Journal of Experimental Psychology: Human Perception and Performance* **27** 1072 – 1089

Anstis S, Saida S, 1985 "Adaptation to auditory streaming of frequency-modulated tones" *Journal of Experimental Psychology: Human Perception and Performance* **11** 257 – 272

Arnell K M, Jolicoeur P, 1999 "The attentional blink across stimulus modalities: Evidence for central processing limitations" *Journal of Experimental Psychology: Human Perception and Performance* **25** 1 – 19

Baddeley A D, Hitch G J, 1974 "Working memory", in *The Psychology of Learning and Motivation* Ed. G H Bower (New York: Academic Press) pp 47 – 89

Beauvois M W, Meddis R, 1991 "A computer model of auditory stream segregation" *Quarterly Journal of Experimental Psychology A* **43** 517 – 542

Beauvois M W, Meddis R, 1996 "Computer simulation of auditory stream segregation in alternating-tone sequences" *Journal of the Acoustical Society of America* **99** 2270 – 2280

Bregman A S, Campbell J, 1971 "Primary auditory stream segregation and perception of order in rapid sequences of tones" *Journal of Experimental Psychology* **89** 244 – 249

Carlyon R P, Cusack R, Foxton J M, Robertson I H, 2001a "Effects of attention and unilateral neglect on auditory stream segregation" *Journal of Experimental Psychology: Human Perception and Performance* **27** 115 – 127

Carlyon R P, Plack C J, Cusack R, 2001b "Cross-modal and cognitive influences on the build-up of auditory streaming" *British Journal of Audiology* **35** 139 – 140

Cheour-Luhtanen M, Alho K, Sainio K, Rinne T, Reinikainen K, Pohjavuori M, Renlund M, Aaltonen O, Eerolo O, Näätänen R, 1996 "The ontogenetically earliest discriminative response of the human brain" *Psychophysiology* **33** 478 – 481

Cusack R, Deeks J, Aikman G, Carlyon R P, in press "Effects of location, frequency region, and time course of selective attention on auditory scene analysis" *Journal of Experimental Psychology: Human Perception and Performance*

Duncan J, Martens S, Ward R, 1997 "Restricted attentional capacity within but not between sensory modalities" *Nature* **387** 808 – 810

Jones D M, Alford D, Bridges A, Tremblay S, Macken W J, 1995 "Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect" *Journal of Experimental Psychology: Learning, Memory and Cognition* **25** 464 – 473

Jones D M, Alford D, Bridges A, Tremblay S, Macken W J, 1999 "Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect" *Journal of Experimental Psychology: Learning, Memory and Cognition* **25** 464 – 473

Jones D M, Macken W J, 1995 "Organizational factors in the effect of irrelevant speech: The role of spatial location and timing" *Memory & Cognition* **23** 192 – 200

Joseph J S, Chun M M, Nakayama K, 1997 "Attentional requirements in a 'preattentive' feature search task" *Nature* **387** 805 – 807

Kane N M, Curry S H, Rowlands C A, Manara A R, Lewis T, Moss T, Cummins B H, Butler S R, 1996 "Event related potentials—neuropsychological tools for predicting emergence and early outcome from traumatic coma" *Intensive Care Medicine* **22** 39 – 46

McCabe S L, Denham M J, 1997 "A model of auditory streaming" *Journal of the Acoustical Society of America* **101** 1611 – 1621

Mack A, Tang B, Tuma R, Kahn S, 1992 "Perceptual organization and attention" *Cognitive Psychology* **24** 475 – 501

Macken W J, Tremblay S, Houghton R J, Nicholls A P, Jones D M, 2003 "Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory" *Journal of Experimental Psychology: Human Perception and Performance* **29** 43 – 51

Naatanen R, Schroger E, Karakas S, Tervaniemi M, Paavilainen P, 1993 "Development of a memory trace for a complex sound in the human brain" *Neuroreport* **4** 503 – 506

Noorden L P A S van, 1975 *Temporal Coherence in the Perception of Tone Sequences* PhD, Eindhoven University of Technology

Rock I, Linnett C M, Grant P, 1992 "Perception without attention: results of a new method" *Cognitive Psychology* **24** 502 – 534

Shinozaki N, Yabe H, Sato Y, Sutoh T, Hiruma T, Nashida T, Kaneko S, 2000 "Mismatch negativity (MMN) reveals sound grouping in the human brain" *Neuroreport* **11** 1597 – 1601
Sussman E, Ritter W, Vaughan H G Jr, 1998 "Attention affects the organization of auditory input associated with the mismatch negativity system" *Brain Research* **789** 130 – 138
Sussman E, Ritter W, Vaughan H G Jr, 1999 "An investigation of the auditory streaming effect using event-related potentials" *Psychophysiology* **36** 22 – 34

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 32 2003

www.perceptionweb.com