CHAPTER

52

Effects of Attention on Auditory Perceptual Organization

Robert P. Carlyon and Rhodri Cusack

ABSTRACT

This chapter summarizes the evidence on the effects of attention on auditory perceptual organization, with particular reference to the phenomenon of auditory streaming. Four approaches are described, and the limitations of each discussed. It is concluded that, although some streaming can take place without subjects devoting all their attention to the sounds, attention can have a profound effect on auditory perceptual organization.

I. INTRODUCTION

The extent to which perceptual organization is influenced by attention has recently been the subject of research in both the visual and auditory modalities
1 (Visual: see Kovacs, this volume. Auditory: Bregman and Campbell, 1971; Sussman et al., 1998; Sussman et al., 1999; Carlyon et al., 2001a; Macken et al., 2003; Cusack et al., in press). The answer may shed light on the stages of neural processing involved in forming perceptual organization, and, at the same time, inform models of sensory processes. For example, many aspects of "auditory streaming" (e.g., Bregman and Campbell, 1971; van Noorden, 1975; Anstis and Saida, 1985) can effectively be modeled by automatic processes, 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). However, evidence that attention can affect the streaming process itself (Carlyon et al., 2001a; Carlyon et al., 2001b; Cusack et al., in press) suggests that 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 (e.g., McCabe and Denham, 1997).

A challenge facing researchers in all modalities is that of measuring the perceptual organization of a stimulus to which the participant is not fully attending. The challenge arises because, in order to test the effects of attending to the stimulus, one must divert attention away from it in some conditions. Consequently, subjective reports of the perceived organization may either be unreliable (for example, not all attributes of the stimulus may be encoded in short-term memory) or cause attention to be reoriented to the stimulus. In audition, research into the effects of attention on perceptual organization has focused not on the grouping of simultaneous frequency components into the same or different objects, but on the sequential organization of successive sounds ("auditory streaming"). Here we briefly describe auditory streaming and then review the different approaches that have been adopted when studying its modulation by attention.

II. AUDITORY STREAMING

The ability to organize sounds into separate streams is important in everyday situations when a listener needs to track the voice of one speaker in the presence of an interfering talker, and has also been widely used by composers of polyphonic music. It is commonly studied using simplified sequences of pure tones, like the one shown in Fig. 52.1 (van Noorden, 1975). This stimulus consists of a sequence of tones of frequencies A and B, in a series of repeating triplets "ABA-ABA-ABA . . ." (the dashes represent silent intervals). When the frequencies of A and B are similar, or the presentation rate is slow, subjects typically report hearing all

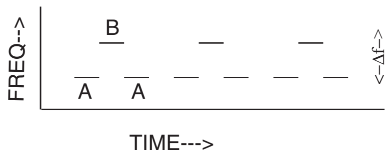


FIGURE 52.1 Example of a tone sequence used to study auditory streaming. The sequence, invented by van Noorden (1975) consists of repeating triplets of tone. The frequency of the "B" tones differs from that of the "A" tones by an amount Δf.

the tones in a single auditory stream, with a galloping rhythm. In contrast, when the A and B frequencies are far apart, and/or the presentation rate is fast, the "A" and "B" tones stream apart, and the percept of the galloping rhythm is lost. Furthermore, when subjects attend to a long sequence and monitor what they hear throughout, two further interesting findings emerge. First, for a wide range of intermediate frequency separations and rates, the percept flips between one and two streams. Second, superimposed on this flipping, the tendency to report hearing two streams increases markedly over the first 5 to 10 seconds. As we shall see in section V, this buildup of streaming has proved extremely useful in the study of the effects of attention on auditory streaming

III. ELECTROPHYSIOLOGICAL MEASURES

A conceptually straightforward way of avoiding the problems associated with subjective reports, described in section I, is to perform an EEG experiment. One measure, the "mismatch negativity (MMN)," lends itself to this approach because it can be measured when subjects are not attending to the stimulus, and are instead reading a book or watching a silent movie. It is measured in a paradigm where subjects receive repeated presentation of a common standard sound, mixed with a small proportion (usually <20%) of "deviants." (For example, the deviants might be quieter than the standards or have a different frequency.) The MMN takes the form of a negative wave, with a latency of 150 to 200 ms, observed when the response to the standards is subtracted from that to the deviants. Although the MMN is thought to have multiple generators, its major source has been shown to be in auditory areas along the supratemporal plane.

A. Streaming Outside the Focus of Attention

Some MMN experiments have led to the conclusion that streaming can occur when participants are not listening to the to-be-streamed sounds. Sussman et al. (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 observed an MMN when the stimulus onset asynchrony (SOA) between successive tones was 100 ms, at which behavioral experiments suggest that streaming should occur. In contrast, when the SOA was 750 ms, for which streaming should be much weaker, no MMN was observed. They attributed the difference in MMN between the two conditions to differences in streaming, which, as subjects were reading a book, they described as occurring outside the focus of attention. However, it should be noted that even when the low tones were omitted, the MMN to the high tone deviants was very weak at the 750-ms SOA, indicating that the observed effects of repetition rate may not have been entirely due to its effects on streaming.

Yabe et al. (2001) used MEG to measure the magnetic MMN ("MMNm") produced by the omission of an occasional high (3000 Hz, "H") tone in a sequence of alternating high and low ("L") tones (i.e., HLHLHL_LHL). Previous work had suggested that stimulus omissions only elicit an MMN when the SOA between successive tones is less than 160 to 170 ms. Using an SOA of 125 ms, they observed an MMN when the lower tone frequency was 2800 Hz, but not when it was 500 Hz. They attributed the lack of an MMN in the latter condition to the high and low tones being pulled into separate auditory streams, so that the SOA in the high stream was now too long (250 ms) to elicit the MMN. Their finding provides a clear demonstration that the omission MMN operates on a frequency selective basis, and, as described in section II, so does auditory streaming. However, a demonstration that the MMN difference is *due* to streaming would, we believe, require one to show that it varies with other parameters known to affect streaming. Indeed, the results of Shinozaki et al. (2000) show that size of the frequency separation between pairs of tone sequences can affect the MMN, even when participants hear all pairs of sequences as a single stream.

B. Manipulating Attention

One general limitation of the approach described above is that, although subjects were instructed to ignore the sounds, one can never be sure that they did not "sneak a little listen". For this reason, researchers are careful to conclude that some streaming occurs outside the *focus* of attention. An alternative approach is to manipulate the amount of attention paid to the sounds and see whether this affects the MMN response. This was adopted by Sussmann et al., using a paradigm similar to that in their 1989 experiment described above. They used an SOA of 500 ms and observed an MMN only when subjects were attending to the tones; they concluded that, for this stimulus, attention did indeed modulate streaming.

IV. INDIRECT EFFECTS OF STREAMING ON COMPETING TASKS

Another way of measuring the perceptual organization of unattended stimuli has been developed by Jones and his colleagues, who measured the effects of streaming on a secondary task—namely, serial recall of visually presented letters (Jones et al., 1995; Jones et al., 1999; Macken et al., 2003). Performance on this task can be disrupted by sequences of sounds that subjects are instructed to ignore, provided that the sounds change over time. For example, two tones that alternate in frequency disrupt performance, whereas a single repeated tone does not. Importantly, this "irrelevant sound effect (ISE)" appears to be sensitive to streaming, so that when a sequence splits into two streams, each of which is heard as a repetition of a single sound, its deleterious effect on visual recall is reduced. Furthermore, this effect has been observed when streaming is produced by a number of different manipulations, including differences in frequency separation, in perceived location, and in repetition rate (for a summary see Macken et al., 2003). For example, Macken et al. (2003) visually presented letters accompanied by a sequence of tones that alternated at low, medium, or high rate. When the alternation rate increased from low to medium, the number of errors increased, consistent with the increased number of changes in the (ignored) auditory sequence. However, when the rate was increased further, the percentage of errors dropped, consistent with the two tones now falling into separate auditory streams, each containing a repeating single tone.

A strong advantage of using the ISE to measure auditory streaming is that it provides a *performance* measure of the effect. Hence, any effects that one observes cannot be attributed to nonperceptual factors such as response biases. However, it too suffers from some limitations when it comes to studying the effects of attention on auditory streaming. The main effect is that it is hard to be sure that the task of remembering visually presented letters is sufficiently demanding to prevent subjects from attending to the sounds. Furthermore, because attention is not explicitly manipulated, one does not know whether any streaming observed would have been greater if subjects had been paying attention to the sounds. Essentially, these experiments lead to the important conclusion that *some* streaming can occur without *full* attention, but do not show whether, or by how much, streaming can be affected by attention.

V. MANIPULATING ATTENTION DURING THE BUILDUP OF AUDITORY STREAMING

A third approach, developed by Carlyon et al. (2001a), exploited the fact that the initial percept of a sequence of tones as a single stream can turn into a two-stream percept after several seconds. In their first experiment they presented a 20-sec sequence of repeating ABA-triplets (see section II and Figure 52.1) to subjects' left ears, and manipulated attention during the first 10 sec of the sequence. In a baseline condition, they simply asked subjects to monitor how many streams they heard throughout the sequence by switching between two response buttons marked "1 stream" and "2 streams." In the remaining two conditions, a series of noise bursts was presented to the opposite ear during the first 10 sec of the sequence. For one of these, subjects were instructed to perform a demanding task on the noise bursts for the first 10 sec, and then switch their attention to the tones and start making streaming judgements. The question was, when they did this, whether the streaming would be the same as if they had been attending to the tones all along, or more like the tones at the very start of an attended sequence. The results, shown in Fig. 52.2*a*, reveal that the latter prediction was correct; when attention was diverted to the tones (circles), the buildup of streaming was much less than in the baseline condition (triangles). A third condition (squares) showed that the noise bursts had no effect when subjects ignored them and attended to the tones throughout.

An important feature of Carlyon et al.'s paradigm, which differs from the MMN and ISE experiments, is that it requires subjects to make direct reports of what they hear. It overcomes the problems of such reports themselves affecting attention by manipulating attention during the *first* half of a sequence but measuring streaming during the *second* half. However, the use of a subjective measure does make it possible that, in Carlyon et al.'s first experiment, a form of response bias was operating. Specifically, subjects may have decided to always respond "1 stream" whenever they

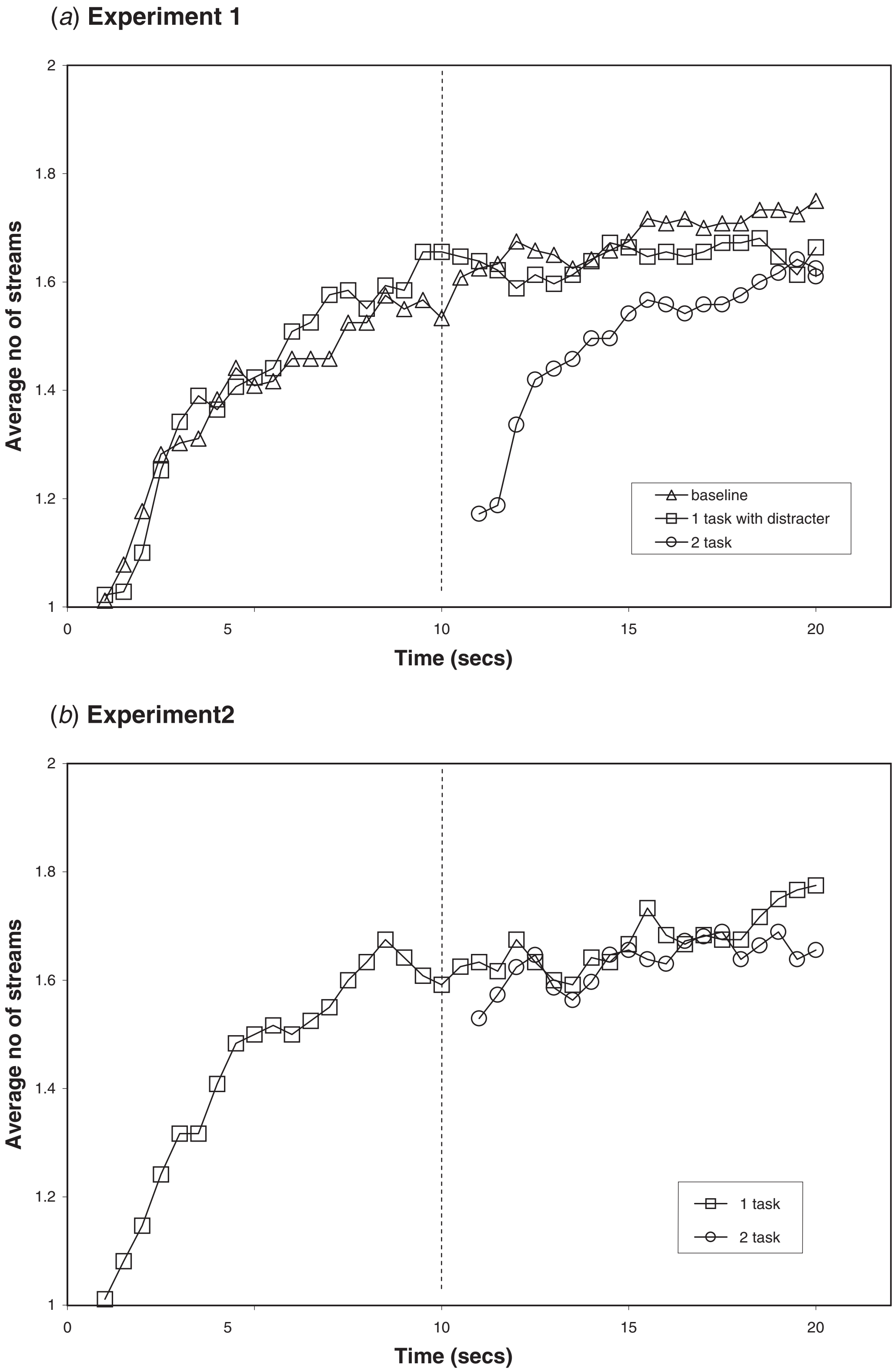


FIGURE 52.2 Average number of streams, plotted as a function of time, for (a) experiment 1 and (*b*) experiment 2 of the study by Carlyon et al. (2001a).

started to make streaming judgments about a sequence. To control for this, they performed a second experiment with no noise bursts and in which the tones were amplitude modulated (AM), with the rate of modulation switching every few seconds. In one condition, subjects were told to detect the switches in AM rate for the first 10 seconds, and then start making the streaming judgments. These judgments (Fig. 52.2*b*, circles) resembled those after 10 sec of a baseline condition where subjects made streaming judgments throughout (Fig. 52.2*b*, squares). Carlyon et al. noted that this result was inconsistent with a response bias and concluded that streaming does build up provided one is attending to the tones, even if one is performing a different type of task on them. Further experiments, using slightly different paradigms, have lent support to this conclusion (Carlyon et al., 2001b; Cusack et al., in press). These experiments also showed that the buildup of streaming can be reduced by attending to sounds in the same ear, or even to visual stimuli.

Although the experiments by Carlyon and colleagues indicate a clear influence of attention on streaming, they do leave one important question unanswered. This is that we only know what subjects have reported after they have switched attention back to the tones. It may be that instead of diverting attention really preventing the buildup of streaming, the act of switching attention back to a sequence "resets" streaming. Indeed, Cusack et al. have suggested that the default organization might always be for a sequence to be heard as two streams, except when attention is directed toward it—either by instructions to the listener or simply because it has just been turned on. According to this explanation, the "buildup" of streaming observed in numerous experiments should instead be thought of as recovery from an active attentional influence. What does seem clear is that attention can have a profound effect on auditory perceptual organization, either by preventing a gradual buildup or by an active resetting mechanism.

VI. THE HIERARCHICAL DECOMPOSITION MODEL

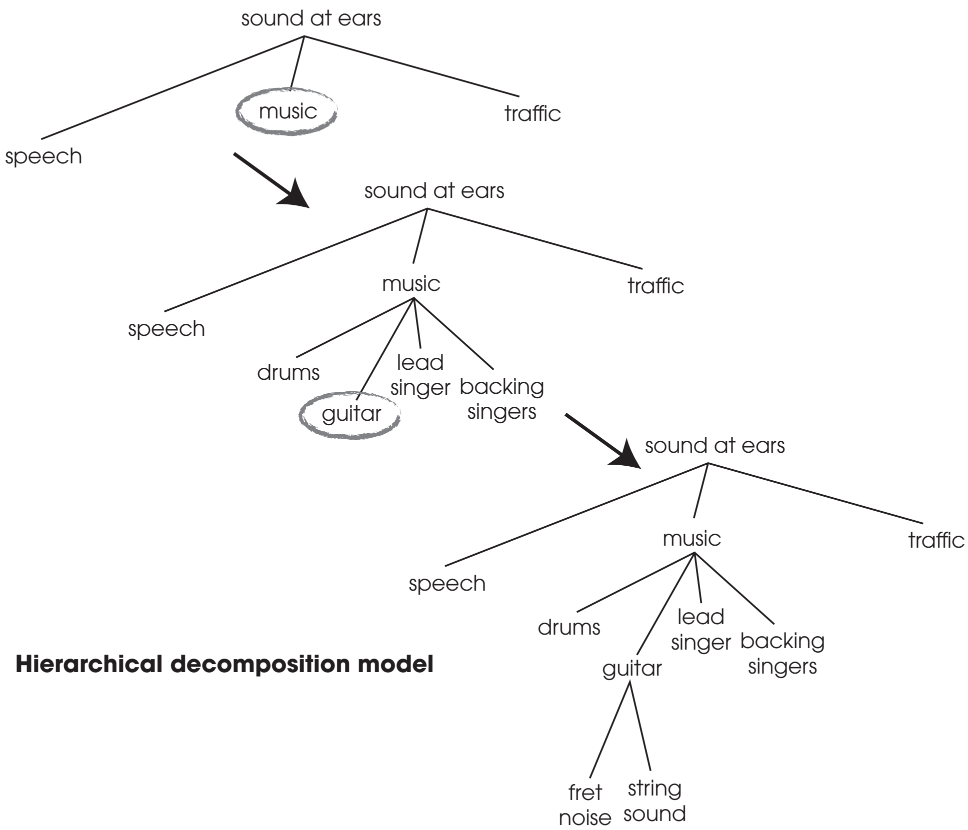
Cusack et al. (in press) further investigated the effect of attention on the buildup of auditory streaming by examining the domain over which auditory streaming was prevented when attention was directed elsewhere. They found that, rather than just those sounds in unattended locations or frequency regions failing to show a buildup of streaming, *any* stream that was outside the focus of attention failed to show further fragmentation. This result was interpreted in terms of a *hierarchical decomposition model*, illustrated in Fig. 52.3. In this example of a real-world auditory scene, initial (perhaps automatic) grouping allows the listener to focus attention on (e.g.) the music, at the expense of the speech and traffic noise. This attentional focusing then causes the internal representation of the music to fragment, and several further streams become available for attention. If one of these is then attended, further fragmentation takes place and so on. As shown by Cusack et al. in a further experiment, if attention is withdrawn from a sequence, its grouping is then reset: in our example, if attention is switched to the traffic, the fragmentation of the music hierarchy will collapse. The advantages of such hierarchical decomposition are that on arrival in the auditory scene the listener is not bombarded with tiny components of all the sounds, and that any limited capacity processes involved in grouping are not wasted by the organization of unattended parts of the auditory scene.

VII. NEUROPSYCHOLOGICAL APPROACH

Only one study has adopted a neuropsychological approach to the effects of attention on auditory streaming. Carlyon et al. (2001a) studied patients with unilateral neglect, who had been diagnosed on the basis of having attentional deficits toward the left side of visual space. They asked patients to report how many streams they heard when 6-second sequences of alternating tones (Fig. 52.1) were presented to their left or right ears. The amount of stream segregation for sequences presented to their left ears was reduced compared both to the same sequences presented to their right ears and to sequences presented to either ear of healthy or brain-damaged controls. Carlyon et al. suggested that the reduced streaming in the left ears of neglect patients was caused by damage to cortical attentional processes. However, their patients had a variety of lesions, and so it is not possible to specify which areas of cortex are important for the effect.

VIII. CONCLUSIONS

Measuring the perceptual organization of an unattended stimulus, without requiring the participant to attend to it, stretches the imagination of the cognitive neuroscientist to the limit. We have described four approaches to this problem, each of which has its own limitations. Nevertheless, two broad principles emerge. First, the ISE experiments of Jones and colleagues allow us to conclude that some streaming can



2

FIGURE 52.3 An example of the interaction between selective attention and grouping as proposed by the hierarchical decomposition model (Cusack et al., in press).

take place without full attention. Second, Sussman et al.'s (1998) MMN experiment, combined with Carlyon et al.'s studies of the buildup of streaming, clearly indicate that attention can affect streaming. Crucially, those results show that that attention can have a profound effect on the streaming process per se, rather than simply allowing subjects to choose from the outputs of a purely automatic streaming mechanism (van Noorden, 1975).

References

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

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

Bregman, A. S., and 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., and 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., and Cusack, R. (2001b). Cross-modal and cognitive influences on the build-up of auditory streaming. *Brit. J. Audiol.* **35**, 139–140.

Cusack, R., Deeks, J., Aikman, G., and 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*.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., and Macken, W. J. (1995). Organizational factors 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., and 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.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., and 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.

McCabe, S. L., and Denham, M. J. (1997). A model of auditory streaming. *J. Acoust. Soc. Am.* **101**, 1611–1621.

Shinozaki, N., Yabe, H., Sato, Y., Sutoh, T., Hiruma, T., Nashida, T., and Kaneko, S. (2000). Mismatch negativity (MMN) reveals sound grouping in the human brain. *Neuroreport* **11**, 1597–1601.

Sussman, E., Ritter, W., and Jr., H. G. V. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research* **789**, 130–138.

Sussman, E., Ritter, W., and Vaughan, H. G., Jr. (1999). An investigation of the auditory streaming effect using event-related potentials. *Psychophysiology* **36**, 22–34.

van Noorden, L. P. A. S. (1975). Temporal coherence in the perception of tone sequences. Ph.D. Eindhoven University of Technology.

Yabe, H., Winkler, I., Czigler, I., Koyama, S., Kakigi, R., Sutoh, T., Hiruma, T., and Kaneko, S. (2001). Organizing sound sequences in the human brain: the interplay of auditory streaming and temporal integration. *Brain Research* **897**, 222–227.

AUTHOR QUERY FORM

Dear Author,
During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.
Many thanks for your assistance.

Query References	Query	Remarks
1	Au: Please provide chapter number for Kovacs.	
2	Au: Update.	