Attentional and perceptual sources of the auditory attentional blink

SÉBASTIEN TREMBLAY and FRANÇOIS VACHON
Université Laval, Quebec City, Quebec, Canada

and

DYLAN M. JONES
Cardiff University, Cardiff, Wales

When a rapid succession of auditory stimuli is listened to, processing of the second of two successive targets among fillers is often impaired, a phenomenon known as the *attentional blink* (AB). Three experiments were conducted to examine the role of filler items in modulating the size of the auditory AB, using a two-alternative forced choice discrimination paradigm. In the first experiment, dual-stream presentations in which low- and high-pitch items were separated by six semitones were tested. A transient deficit in reporting the probe was observed in the presence of fillers that was greater when fillers were in the same stream as the probe. In the absence of a filler, there was a residual deficit, but this was not related to the time lag between the target and the probe. In the second and third experiments, in which single-stream presentations were used, a typical AB was found in the presence of homogeneous fillers, but heterogeneous fillers tended to produce a greater deficit. In the absence of a filler, there was little or no evidence of a blink. The pattern of results suggests that other attentional and perceptual factors contribute to the blink.

A notable example of a limitation to attentional selectivity is manifest in the attentional blink (AB) phenomenon. When two successive targets among nontargets (or fillers) in a rapidly presented sequence are to be reported, the second target (usually called the *probe*) is often missed if it is presented within approximately 500 msec of the first target (e.g., Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992). Whether the AB reflects a fundamental limitation shared among some or all modalities is not certain (for a discussion, see, e.g., Arnell & Larson, 2002). In the present study, we sought to contribute toward characterizing the generality of the AB phenomenon by examining the degree to which the well-established features of the AB in the visual modality extend also to the auditory modality.

Although there is a substantial body of work on visual AB (see McLaughlin, Shore, & Klein, 2001, for a review), a thorough understanding of its auditory counterpart has been hampered by a relative dearth of studies. Certainly, the existence of auditory AB would constitute a challenge to theories that explain the AB purely in terms of visual mechanisms (e.g., Shapiro, Raymond, & Arnell, 1994). To date, however, the few results available point to a relatively more inconsistent effect in the auditory modality than in the visual modality. Some have gone so far as to claim that there is simply no AB for auditory sequences (Potter, Chun, Banks, & Muckenhoupt, 1998; see also Chun & Potter, 2001); others posit that the effect is attenuated in audition, relative to vision (e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Soto-Faraco & Spence, 2002), whereas still others have reported a marked auditory AB (e.g., Duncan, Martens, & Ward, 1997; Goddard & Slawinski, 1999). Although on balance the evidence points to the existence of auditory AB phenomena—and hence, to the likelihood that the AB reflects a general limitation on cognition—it is certainly the case that a full functional characterization of the auditory AB is some way off (see Mondor, 1998, for a discussion; see also Arnell, 2001).

In the present article, interest centers on the role played by the context in which the probe and the target appear—that is, the role played by the presence and nature of the filler items in determining the magnitude of the blink. Such interest arises from two sources: (1) studies of the visual AB, in which fillers have been implicated as a key variable modulating the magnitude of the AB (e.g., Kawa-

hara, 2003; Ross & Jolicœur, 1999), and (2) the generalizations suggested by a body of work on auditory streaming that suggests that fillers play a significant role in isolating or incorporating single events within streams (see Bregman, 1990, for an overview). This body of work speaks to the organization of auditory events into streams of temporally extended auditory *objects*. Evidence is emerging that the principles of auditory perceptual organization play a key role in a range of cognitive phenomena, particularly those related to the isolation of events within sequences. For example, auditory streaming has been shown to modulate cognitive effects, such as memory for pitch (Jones, Macken, & Harries, 1997) and the auditory suffix effect (Nicholls & Jones, 2002). In each of these cases, the perception of particular items can be altered by modifying the extent to which these events are embedded within streams. Given that a key characteristic of the AB paradigm is the isolation of two events within an auditory sequence, it seems not too fanciful to suggest that streaming constructs could be applicable here also.

Auditory and visual modalities certainly differ in the temporal coherence of the stimulus sequence. The powerful organizational processes manifest in auditory perception can be construed as a means of deriving stability from an evanescent stimulus. The binding and embedding of events into a temporally extended perceptual object is a key characteristic of audition. This may have particular relevance in the AB paradigm, inasmuch as it seems reasonable to expect that as a result of this prominent characteristic, the effect of the context in which the target and the probe appear will be particularly influential. Evidence for context effects in the perception of the order of simple auditory sequences is compelling. Take the perception of order for two immediately successive tones, differing only slightly in pitch. When these are presented in isolation, there may be little difficulty in discerning the order. However, if two other tones flank the pitches of interest, the order now becomes much more difficult to discern. Adding yet more flankers, so that they now form a stream, distinct from the tone pair, significantly improves the perception of their order. The flanking stimuli may either capture or partition the tone pair, depending on the relation in time and pitch to the flankers (see, e.g., Bregman & Rudnicky, 1975).

The impact of context on the auditory AB is tested in the present study through the manipulation of fillers that surround a target and a probe. In the first experiment, the effect of the grouping of targets and fillers within different streams was investigated. In this experiment, the grouping of target items and fillers was studied in streams defined by pitch. On each trial, low- and high-pitch items were presented. These could be just a target and a probe of different pitch, presented in isolation, or a target and a probe surrounded by fillers of the same pitch. Such manipulations might lead the target or the probe to pop out of the sequence, thus making that item easier to distinguish and select. In Experiments 2 and 3, pitch was not used as a means of perceptual organization; rather,

grouping was promoted by manipulating the heterogeneity among fillers. Here, only a single stream was formed, and we manipulated the coherence of that stream by changing its composition. Homogeneous fillers (repeated items) tend to be more coherent than heterogeneous fillers (changing items). The homogeneous fillers are therefore perceived as a whole and tend to stream apart from the target and probe. Again, the perceptual organization of the target and the probe in relation to the fillers should affect the magnitude of the AB.

One of the key issues within the visual blink literature is the contribution of perceptual interference by masking. The impact of masking in modulating the size of AB has been tested extensively in the visual modality (see Enns, Visser, Kawahara, & Di Lollo, 2001, for a review). There is ample evidence that the item following the target (T+1 item)[1] serves to mask the target and, hence, is a key determinant of the magnitude of the AB (e.g., Chun & Potter, 1995; Raymond, Shapiro, & Arnell, 1995; Seiffert & Di Lollo, 1997). The action of the item following the probe (the P+1 item) also appears to be a critical aspect of the AB (e.g., Giesbrecht & Di Lollo, 1998; Jolicœur, 1999a; see also Kawahara, Zuvic, Enns, & Di Lollo, 2003). There is a broad consensus with regard to the effects of masking in the visual AB (see, e.g., Enns et al., 2001), but evidence of interference effects by masking is not yet available in the auditory AB (see Mondor, 1998). In vision, the absence of target masking reduces the size of the blink, but it does not abolish the effect, whereas the absence of probe masking does appear to eliminate the AB (e.g., Giesbrecht & Di Lollo, 1998; Jolicœur, 1999a). In the first experiment of the series that follows, rather than exploring specifically the role of T+1 and P+1 items, we compared the effect of the presence and absence of the whole filler set. Therefore, the emphasis in the present study was on investigating the role of the context, rather than the role of a single posttarget mask. The impact of masking is difficult to dissociate from that of streaming; both point to the contribution of perceptual sources to the AB (see Arnell & Larson, 2002, for a related discussion).

That perceptual organization plays a major role in modulating the AB may be taken to imply that the AB will be subject to the structural constraints of the particular modality and, hence, will take different forms in the visual and the auditory modalities. Typically, it is claimed that central, rather than modality-specific, mechanisms are responsible for the AB (e.g., Jolicœur & Dell'Acqua, 1998; see also Ruthruff & Pashler, 2001, for a discussion). For example, the AB has been interpreted as a process related to the transfer of events from a sensory or perceptual encoding stage into short-term storage, and interference is assumed to result either from confusion in a short-term memory (STM) overloaded with stimuli (e.g., *retrieval competition theory*; Shapiro et al., 1994) or from a bottleneck arising from the serial nature of the transfer process, leading to a failure of consolidation in STM (e.g., Chun & Potter, 1995; see also Jolicœur, 1999a, 1999b). As we already mentioned, some approaches

suggest that the AB is restricted to vision, taking place in visual STM (VSTM), where filler items compete with the probe for retrieval, thus inducing the ensuing AB (Shapiro et al., 1994). However, it is plausible that the auditory modality is endowed with a mechanism that is equivalent to VSTM—namely, the echoic short-term store (e.g., Crowder & Morton, 1969)—within which confusion may occur between the auditory stimuli competing for retrieval (Duncan et al., 1997; for a contrary view, see Mondor & Terrio, 1998, and Nicholls & Jones, 2002). These notions of similar but separate modality-based processes are consistent with a central process (Jolicœur, 1999b). The concept of central limitations to STM consolidation does not necessarily imply that the character of the AB will be universal and independent of modality; indeed, modality-specific factors can modulate the limits on central processes (see, e.g., Arnell & Larson, 2002; Shih, 2000).

The present experimental series began with a partial replication and extension of a frequently cited study, that of Duncan et al. (1997), in which auditory AB was found. In Experiment 1, we extended the range of variables used by Duncan et al. by exploring whether the presence or absence of fillers plays a role in determining the magnitude of the auditory AB.

EXPERIMENT 1

Duncan et al. (1997) demonstrated a compelling auditory AB. In addition to replicating that original finding, in Experiment 1 we sought to show the extent to which acoustic context is important to the detection of a probe. The same general procedure and auditory stimuli as those used by Duncan et al. were employed here. In Duncan et al.'s study, targets and fillers were presented in two concurrent auditory sequences, each of a different pitch. One set of targets (*nab* or *nap*) was presented at a low pitch embedded in a sequence of filler items at the same pitch as the target, where fillers consisted of repetitions of the item *guh*. Another set of targets (*cod* or *cot*) was presented at high pitch, again with repetitions of the filler item *guh*, but this time at the same high pitch of the target. Coupled to a fast rate of presentation, the pitch difference was of sufficient magnitude that the sequences were likely to be perceived as two separate streams.[2] The experiment was designed to distinguish the influence of the presence of the fillers independently for the target and the probe. That is, fillers could be presented either with both sequences, be absent from the sequence containing the probe (but present with the target), absent from the sequence containing the target (but present with the probe), or absent from both sequences, leaving the target and the probe in isolation (see Figure 1).

These manipulations allowed us to examine whether auditory AB-like effects can be observed even in the absence of any filler. According to the existing literature, AB deficits should be eliminated when masks are removed, but there might be some cost due to the pitch distance between the target and the probe, given the re-

quired reallocation of attention from, for example, a low-pitch target to a high-pitch probe (see, e.g., Mondor & Bregman, 1994). At the same time, the design allows some separate estimation of the extent to which the respective filler contexts of the target and the probe contribute to the magnitude of the blink. These comparisons may also be informative about the relative roles of target and probe masking in the auditory AB.

Method

Participants. Twenty-four volunteers, each of whom reported normal hearing, were recruited from Cardiff University students. They received course credit for their participation.

Materials. There were two concurrent rapid auditory presentations (RAPs), one high pitch and one low pitch. The low-pitch sequence contained a single instance of either *nab* or *nap*, surrounded by fillers (*guh*) or no filler. The high-pitch sequence contained a single target, either *cod* or *cot*, with or without fillers. First, all the items (*nab*, *nap*, *cod*, *cot*, and *guh*) were digitally recorded in a male voice. Great care was taken to produce the vowels at an even pitch (using a reference pure tone of 103.8 Hz, available to the male speaker at recording) and level. All the samples were digitally edited to 16-bit resolution at a sampling rate of 48 kHz and were compressed in order to last exactly 150 msec. Each filler was a digital copy of an original. The high-pitch sequence was created by a digital-processing method (using Sound Designer II software) that shifted the items (the targets *cod* and *cot* and the filler *guh*) up in pitch by six semitones. Importantly, these digital transformations did not decrease the intelligibility of the individual items.

Syllables lasted for 150 msec each and were separated by silent gaps of 100 msec. The sequences were presented in such a way that the items in the high- and the low-pitch sequences were slightly out of phase with one another: One of the two concurrent sequences, chosen at random on each trial, began 125 msec before the other. The first sequence to begin also contained the target, presented after five filler items or, in the case in which no filler was presented, after the corresponding silent interval. The probe was presented following delays of 125, 625, or 1,375 msec, measured from target onset to probe onset (stimulus onset asynchrony, or SOA). Each sequence contained 13 items per stream and lasted 3,275 msec. Within each sequence, the rate of presentation was 250 msec per item (150 msec on, 100 msec off).

There were four experimental conditions in Experiment 1, and they are referred to here by a notation that signifies the relation of the target (T) and the probe (P) to the presence of fillers (F) in the sequence (see Figure 1). In the TF + PF sequence, both of the items for report were embedded in a sequence—one of high pitch, the other of low pitch—of fillers. In one condition, fillers surrounded only the target, with no filler around the probe (TF + P), and in another, fillers surrounded only the probe, with no filler around the target (T + PF). In a fourth condition (T + P), no filler was present.

Experimental design. The three independent variables were treated as repeated measure variables: task (single vs. dual), target–probe SOA (125, 625, or 1,375 msec), and filler distribution (TF + PF, T + PF, TF + P, or T + P). Each participant took part in two single-task conditions (one with instructions to focus their attention on the low-pitch stimuli and one with instructions to focus on the high-pitch stimuli) and one dual-task condition in which the participants were told to report both targets. These conditions were conducted in blocks, and there was one experimental block per task condition, the order of the three blocks being counterbalanced across participants, using a Latin square procedure. Within each block, target–probe SOAs and filler distribution variables were randomized. There were 96 trials in each block, preceded by 24 practice trials. It is important to note that the items to be detected in the single-task condition corresponded to the probe items from the
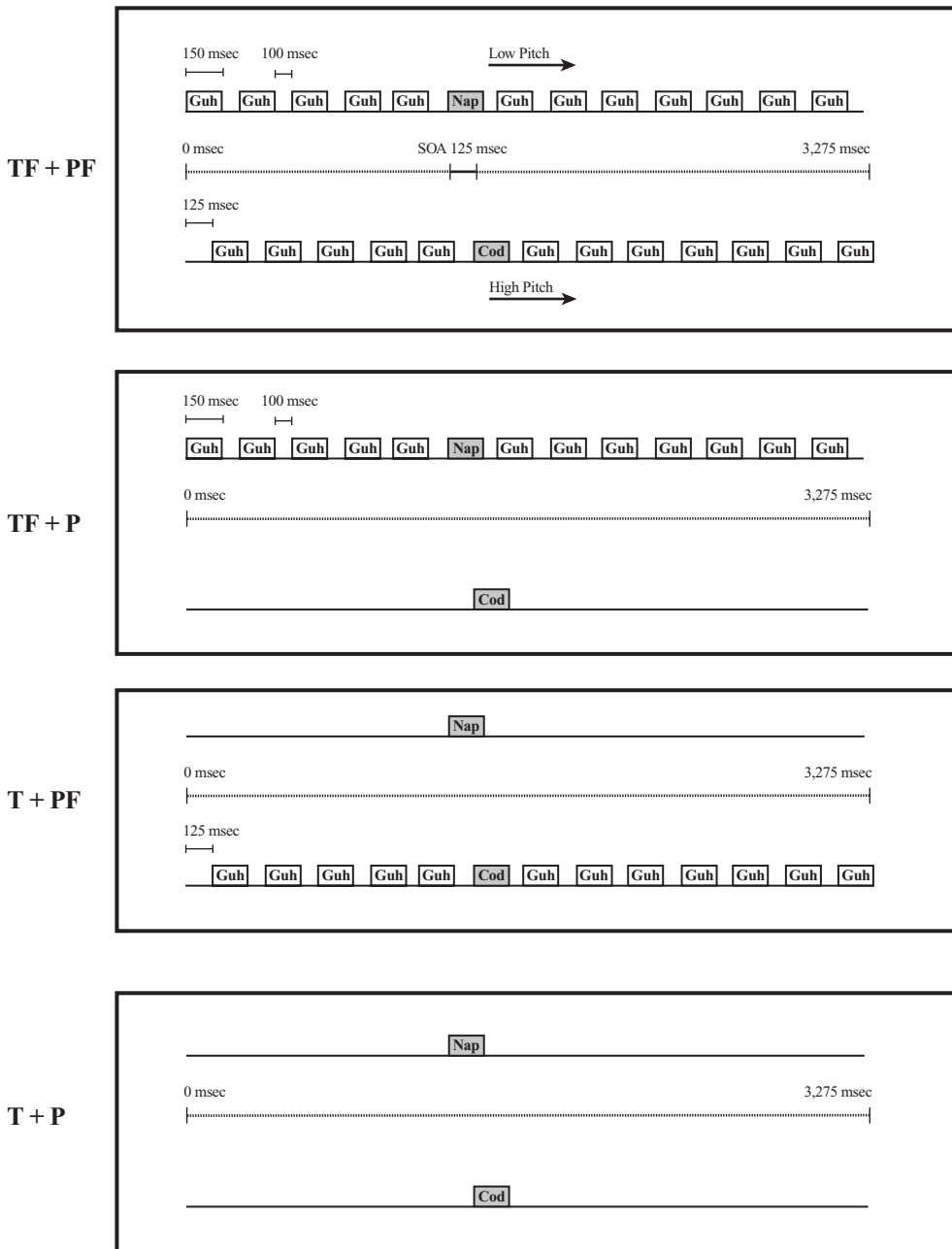
Figure 1. Schematic diagram illustrating the four dual-stream presentations employed in Experiment 1: TF + PF, TF + P, T + PF, and T + P (examples at an SOA of 125 msec). T, target; P, probe.

dual-task condition, thus making the two conditions comparable. There was always one *na* word (*nab* or *nap*) and one *co* word (*cod* or *cot*) item on each trial, regardless of the task condition.

Procedure. The participants had to perform a two-alternative forced choice discrimination; they were told to discriminate between *nab* and *nap* for the *na* words and between *cod* and *cot* for the *co* words. The participants initiated a trial by a mouse click on the start button displayed on the computer screen. Presentation of the auditory stimuli began after a fixed delay of 250 msec. The stimuli were presented via headphones at approximately 65 dB(A), with a PC computer running a Visual Basic 5.0 program. The participants' responses were typed in following the presentation of each set of stim-

uli, using keys labeled appropriately for the targets. Under the single-task conditions, there was a single response identifying the attended target. The participants were instructed to focus on either the *na* or the *co* targets. Under dual-task conditions, two responses were typed in, in either order. The participants were encouraged to take time with their responses, taking care that typing errors were not made.

Results

Target discrimination. In the dual-task condition, the mean target accuracy was 89.9%. A repeated measures analysis of variance (ANOVA) was performed on

these data, with SOA (three levels) and filler distribution (four levels) as factors. There were significant effects for both SOA [$F(2,46) = 10.39$, $MS_e = 0.01$, $p < .01$] and filler distribution [$F(3,69) = 4.95$, $MS_e = 0.011$, $p < .01$], but the interaction between the two factors was not significant ($p = .71$). Target discrimination was at its lowest level with TF + P and TF + PF when the target was embedded with fillers (see Table 1).

Probe discrimination. The data in the single-task conditions were pooled over targets (*co* and *na* as single targets) and were analyzed only for trials in which the target was in the second stream to begin with (the same stream as that for the probe in the dual-task condition; this precaution was taken in order to make single- and dual-task conditions comparable). Probe accuracy in the dual-task condition was computed only for trials on which a correct discrimination was made for the target. The probability of identifying the probe, given correct identification of the target, is presented in Figure 2 as a function of SOA, task, and filler distribution. A repeated measures ANOVA was carried out on the data, with task, SOA, and filler distribution conditions. The main effects of task [$F(1,23) = 83.17$, $MS_e = 0.038$, $p < .01$] and SOA [$F(2,46) = 15.37$, $MS_e = 0.019$, $p < .01$] were significant, but that of filler distribution did not reach significance [$F(3,69) = 2.23$, $MS_e = 0.015$, $p = .09$]. The interaction between SOA and filler distribution was significant [$F(6,138) = 5.88$, $MS_e = 0.017$, $p < .01$], but there was no significant interaction between task and filler distribution ($p = .55$).

The interaction between SOA and task, which is the empirical signature of the AB, was significant [$F(2,46) = 8.12$, $MS_e = 0.019$, $p < .01$]. Of particular importance for the purposes of the present experiment is the significant three-way interaction between task, SOA, and filler distribution [$F(6,138) = 6.51$, $MS_e = 0.015$, $p < .01$]. This significant three-way interaction reinforces the impression given by a visual inspection of Figure 2 that the critical relation between SOA and task differs according to filler distribution. Here, and applicable elsewhere in this study, a stricter alpha level of .0125 was used in order to compensate for the family-wise error. Further analysis of the three-way interaction revealed that the interaction between task and SOA was significant with TF + PF [$F(2,46) = 5.80$, $p < .01$], at T + PF [$F(2,46) = 12.86$, $p < .01$], did not reach significance with TF + P [$F(2,46) = 3.99$, $p = .03$], but was far from significant at T + F [$F(2,46) = 0.16$, $p = .79$]. In the absence of fillers, the difference between single and dual tasks was significant ($p < .01$), but it did not interact with SOA.

Context effects. The critical interaction between SOA and task that reflects the time-related nature of AB-like deficits was not significant for filler distributions in which there was no filler in the probe stream (T + P and TF + P) but was highly significant for distributions in which fillers surrounded the probe (T + PF and TF + PF). In order to test the relative impact of the context on the dual-task cost in a target stream (absence or presence of fillers) and a probe stream (absence or presence of fillers), the conditions were contrasted with respect to SOA, using the difference in performance between single (control) and dual tasks as the dependent variable (see Figure 3). A repeated measures ANOVA with target stream (two levels), probe stream (two levels), and SOA (three levels) as factors was carried out on the data. The main effect of target stream was not significant [$F(1,23) = 0.29$, $MS_e = 0.048$, $p = .60$]. The effect of SOA was significant [$F(2,46) = 17.39$, $MS_e = 0.029$, $p < .01$], but more important, the effect of probe stream was also significant [$F(1,23) = 14.96$, $MS_e = 0.026$, $p < .01$]. There was no significant interaction between target stream and probe stream [$F(1,23) = 1.66$, $MS_e = 0.038$, $p = .21$]. The interaction of target stream and SOA also was not significant [$F(2,46) = 0.34$, $MS_e = 0.033$, $p = .71$], nor was the three-way interaction significant [$F(2,46) = 0.61$, $MS_e = 0.026$, $p = .55$]. However, the interaction between probe stream and SOA reached significance [$F(2,46) = 5.03$, $MS_e = 0.034$, $p = .01$]. Simple main effects were performed on the data in order to pinpoint the source of the interaction between SOA and probe stream. The difference between absence and presence of fillers in the probe stream was significant at both 125 msec [$t(23) = 3.27$, $p < .01$] and 625 msec [$t(23) = 2.88$, $p < .01$] but was not significant at the longest SOA [$t(23) = -0.71$, $p = .48$]. The effect of SOA was significant when there were fillers in the probe stream [$F(2,46) = 15.33$, $p < .01$] but was not significant when no filler surrounded the probe [$F(2,46) = 2.37$, $p = .12$].

Discussion

The results in Duncan et al. (1997) in relation to the auditory AB were replicated with the demonstration of an SOA-related deficit in identifying the probe (given correct identification of the target) within a dual-stream presentation. In relation to the manipulation of context, the results of Experiment 1 may be summarized as follows: The context in which targets appear is important. Indeed, when there is no filler, there is no blink (although there is a penalty to be paid for switching between stimuli differing in pitch that is roughly the same at all SOAs). In some sense, therefore, context determines the blink. However, the effect appears to be asym-

Table 1
Probability of Target Report ($+SE$) Under a Dual Task as a Function of Stimulus Onset Asynchrony (SOA) (125, 625, and 1,375 msec) and Filler Distribution (TF + PF, TF + P, T + PF, and T + P)

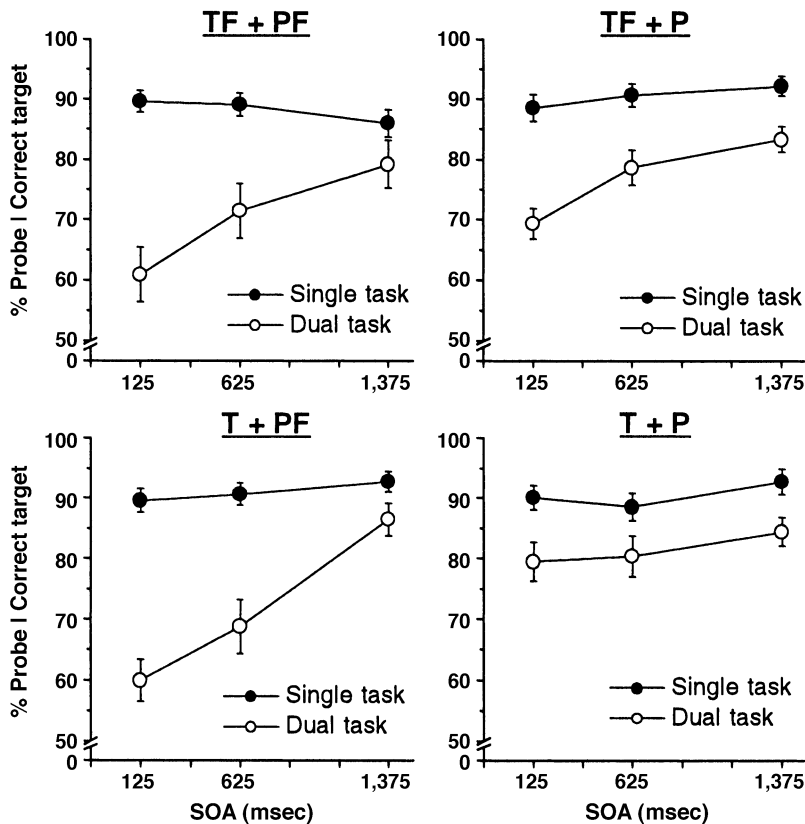| | SOA | | | | | |
| | 125 msec | | 625 msec | | 1,375 msec | |
Filler Distribution	p	SE	p	SE	p	SE
TF + PF	.906	.019	.891	.022	.813	.025
TF + P	.912	.018	.912	.018	.849	.024
T + PF	.912	.021	.922	.018	.875	.020
T + P	.953	.015	.943	.017	.906	.019

Note—T, target; P, probe; F, filler.

Figure 2. Results from Experiment 1: probability of probe report, given correct report of the target, as a function of task, SOA, and filler distribution (TF + PF, TF + P, T + PF, and T + P). Error bars represent standard errors.

metric; the context provided for the probe plays a more powerful role than does the context provided for the target (see Figure 3).

The results of Experiment 1 suggest that masking of the probe is more important than masking of the target, at least in the auditory AB. A greater blink was observed when the probe was surrounded by fillers of the same pitch (T + PF) than when the target was surrounded by fillers (TF + P). Also in line with this finding is the demonstration by Mondor (1998) that the auditory AB is substantial even when the stimulus immediately following the target is replaced with a silent gap. Mondor suggests that the survival of the blink in the absence of a T+1 item reflects a bottleneck limitation at the late stage of response selection (see Jolicœur, 1998, 1999a, 1999b). In support of the latter view, Vachon and Tremblay (in press) showed, in a follow-up to Mondor, that the auditory AB is eliminated when the stimuli following a probe are replaced with a silent gap. The question remains as to whether the observed time-related deficit reflects a true AB or some other dual-task cost.

Some authors claim that the character of a true AB is revealed in its relationship with SOA. In contrast to the present results, in many studies there has been little impairment in reporting the probe when it was temporally adjacent to the target (an effect often referred to as *lag 1 sparing*), as compared with when there was one or more filler items between the target and the probe. However, this U-shaped appearance of the relationship between AB and SOA is not universal: In some studies, this relationship has been linear, as in the present case, with a significant deficit in probe detection even when the target and the probe were adjacent (e.g., Arnell & Jolicœur, 1999). In the visual case, lag 1 sparing is more likely to occur when the target and the probe are presented in the same spatial location, whereas failure to report the lag 1 probe seems to occur when there is a location switch between the target and the probe (e.g., Visser, Bischof, & Di Lollo, 1999; but see Shih, 2000). The same reasoning may apply in audition to pitch distance in the case of dual-stream presentations such as the one used here (see Mondor & Bregman, 1994). Thus, one possibility is that the poorer performance with a probe at T+1 (i.e., absence of lag 1 sparing) is related to the cost of having to reallocate attention to a different frequency region.

The same reallocation cost of having to switch attention from the frequency region of the target to that of the probe could also be responsible for the residual time-independent deficit observed in T + P presentations. If we construe reallocation of attention not merely in terms
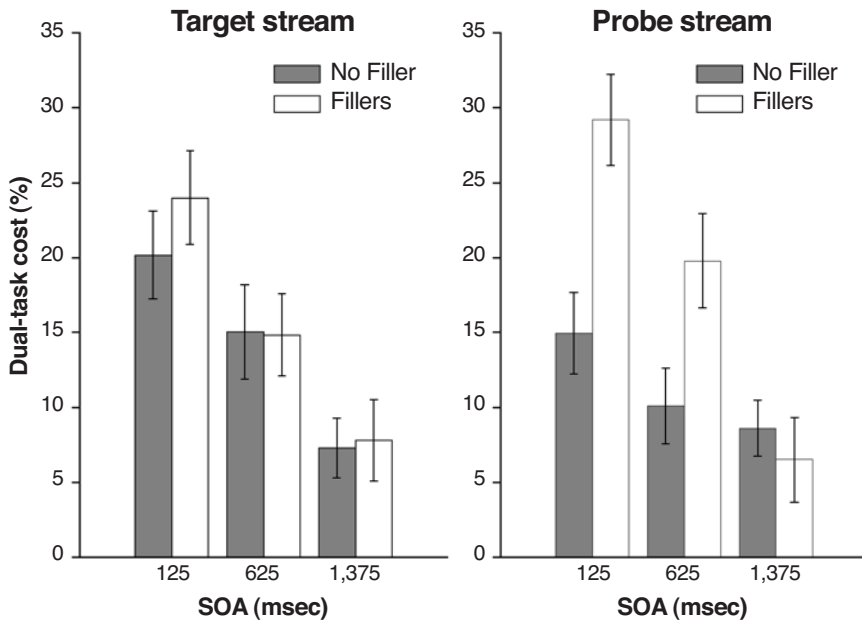
Figure 3. Results from Experiment 1: dual-task cost as a function of SOA, target stream, and probe stream. This cost corresponds to the difference in performance between single and dual tasks. Error bars represent standard errors.

of task requirements (e.g., Potter et al., 1998), but also in terms of a locus of operation (as when the target and the probe are presented in different spatial locations; see Visser et al., 1999), we gain some renewed insight into the effects of context. Studies with single targets reveal that detection is impaired (mostly, slowed reaction time) if attention is cued away from the target. Such a cost in reallocating attention to the appropriate locus has been observed with visual and auditory stimuli cuing spatial location (e.g., Spence & Driver, 1994), as well as with auditory stimuli cuing frequency (e.g., Mondor & Bregman, 1994; see also Woods, Alain, Diaz, Rhodes, & Ogawa, 2001). The penalty of this form of switching is reflected in the small deficit not related to time observed in the T + P condition in Experiment 1.

It is plausible that the pitch-switching[3] cost is additional to the AB, as is suggested by the greater deficit observed in the presence of fillers (see Chun & Potter, 2001, for a discussion on the issue of additivity). That pitch switching may have been partly responsible for the AB found in Experiment 1 does not rule out the possibility of finding an effect in a single stream. There is some evidence, although limited, of auditory AB within a single-stream presentation (e.g., Arnell & Jolicœur, 1999, Experiment 4; Arnell & Larson, 2002; Soto-Faraco & Spence, 2002), and in the next experiment of the present series, we attempted to provide further evidence that an auditory AB can be observed—this time, using single-stream presentations. A key objective of the following experiments was to further characterize the effect of context by exploring the impact of the nature of the fillers on the auditory AB.

EXPERIMENT 2

In Experiment 2, we followed a procedure similar to that employed in Experiment 1 and used the same discrimination task with the same target items. However, instead of dual-stream presentations, a single-stream presentation was used in which there was no overlap in time between successive stimuli. Here, the effect of context was manipulated by the homogeneity of the filler sequence, the items being either repeated or varying. In Experiment 2, the SOA was 150 msec, and there were lag 1, lag 2, and lag 4 conditions, so that, potentially, lag 1 sparing could be observed. Also, the role of pitch switching was minimized by the use of a single stream, rather than a dual stream.

One of the key manipulations in Experiment 2 was of the composition of the filler sequences; these could be a heterogeneous set (*guh*, *gah*, *gih*) or a homogeneous set, consisting of repetition of the same filler (*guh*). This was done with the aim of trying to understand the role of perceptual organization and, in particular, the grouping of sounds into streams. Experiment 1 pointed to the role of the context in which a probe appeared. One interpretation would be that the role of fillers is to mask the probe item and make its recognition more difficult—hence, the lack of a blink when the fillers are removed (for example, in the T + P condition in Experiment 1).

From the standpoint of perceptual organization (especially in the auditory domain), the similarity of the filler items determines the likelihood of their grouping with each other. There is more than ample evidence that streaming depends on the coherence of a sequence, and we may

expect that the nature of grouping depends on the degree of change in a sequence. It is expected that repeated fillers will tend to be grouped together and make the target and the probe stream apart from that sequence. A possible consequence of the latter streaming outcome is that target discrimination will be facilitated. On the other hand, heterogeneity among fillers should lower the probability that filler items cohere and stream out of target items. Hence, the magnitude of the auditory AB might be larger in the presence of changing fillers than in the presence of repeated fillers. However, such an increased AB may also be the result of more items (changing fillers) competing for limited memory or processing resources or causing more confusion in some auditory STM.

Just as in Experiment 1, the design in Experiment 2 incorporated a control condition in which there was no filler. Without any filler, a single-stream presentation in which all items are played at the same pitch (thereby circumventing a possible pitch-switching cost) should not produce any AB-like deficit.

Method

Participants. Twenty-four students volunteered to take part in this experiment. Each reported normal hearing. They received either a small honorarium or course credit for their participation.

Materials. The stimuli were the same as those employed in Experiment 1, with the following exceptions: All the items were presented at the same pitch (low voice at 103.8 Hz), the items were within a single sequence, and in the changing-filler condition, there were three different filler items presented in a fixed order (*guh*, *gah*, and *gih*). There was no gap between successive items. The probe was presented following one of four delays of 150 msec (adjacent, or lag 1), 300 msec (lag 2), 600 msec (lag 4), or 1,350 msec (lag 9). These were measured from target onset to probe onset. Sequences with fillers were all made up of 17 items in total; there were six items preceding the target item, and that item was then followed by 10 items (including one probe and nine fillers). There was either no filler (lag 1) or one (lag 2), three (lag 4), or eight (lag 9) fillers between the target and the probe.

Experimental design. There were three repeated measures variables: task (single vs. dual), target–probe SOA (150, 300, 600, or 1,350 msec), and homogeneity of fillers (no filler, repeated filler, or changing fillers). Each participant took part in two single-task conditions (one with instructions to attend to the *na* word and the other to attend to the *co* word) and the dual-task condition (with instructions to attend to both targets). There was one experimental block per task condition, and the order of the three blocks was counterbalanced across participants, using a Latin square design. Within each block, SOA and homogeneity conditions were randomized from trial to trial. There were 96 trials in each block, preceded by 24 practice trials.

Procedure. The procedure was the same as that in Experiment 1.

Results and Discussion

Target discrimination. On average, the target was reported correctly on 96.1% of the trials in the dual-task condition. An ANOVA with SOA (four levels) and homogeneity (three levels) as repeated measures factors was carried out on the data. None of the effects was significant [SOA, $F(3,69) = 1.49$, $MS_e = 0.014$, $p = .24$; homogeneity, $F(2,46) = 0.91$, $MS_e = 0.016$, $p = .41$; and SOA $\times$ homogeneity, $F(6,138) = 1.99$, $MS_e = 0.011$, $p = .07$].

Probe discrimination. The probability of reporting the probe as a function of task and SOA is displayed in Figure 4 for the no-filler, repeated-filler, and changing-filler conditions. The means were calculated using only those trials in which report of the target was correct. The data were submitted to a repeated measures ANOVA with task (two levels), SOA (four levels), and homogeneity (three levels) as independent variables. The analysis revealed a main effect of SOA [$F(3,69) = 7.97$, $MS_e = 0.01$, $p < .01$] and a main effect of homogeneity [$F(2,46) = 9.52$, $MS_e = 0.014$, $p < .01$], but the main effect of task was not significant [$F(1,23) = 3.09$, $MS_e = 0.033$, $p = .09$]. The interaction of task and homogeneity was significant [$F(2,46) = 4.10$, $MS_e = 0.008$, $p = .02$], whereas the interaction of SOA and homogeneity was not significant [$F(6,138) = 1.50$, $MS_e = 0.012$, $p = .21$]. The critical interaction between SOA and task [$F(3,69) = 10.65$, $MS_e = 0.01$, $p < .01$] and the three-way interaction [$F(6,138) = 2.33$, $MS_e = 0.008$, $p = .04$] were significant.
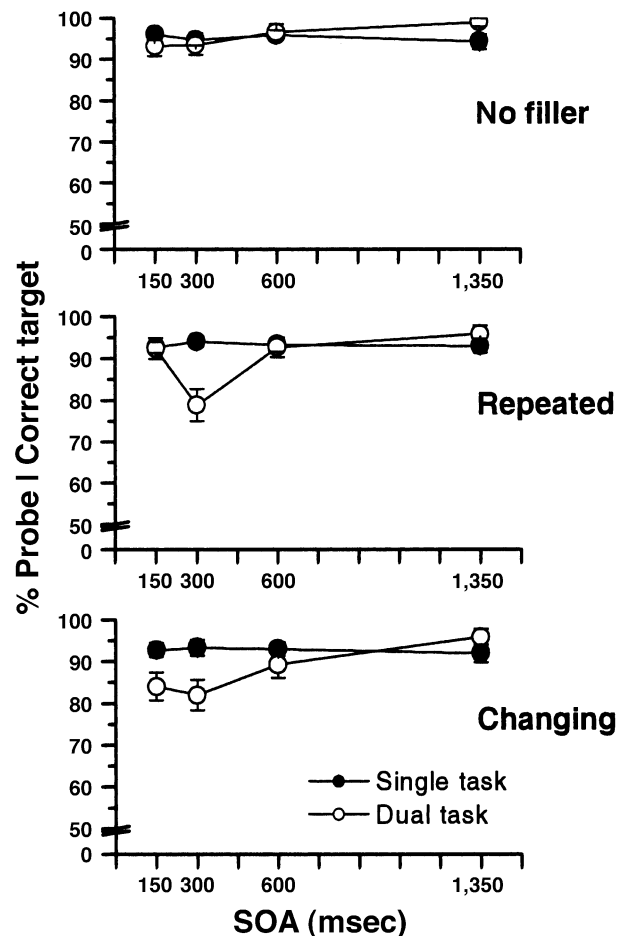

Figure 4. Results from Experiment 2: probability of probe report, given correct report of the target, as a function of task, SOA, and homogeneity of fillers (no filler, repeated, and changing). Error bars represent standard errors.

The homogeneity of fillers seems to modulate the interaction of task and SOA. A closer look at Figure 4 suggests the occurrence of a typical AB in the presence of fillers, and the lag 1 sparing appears to be slightly greater in the presence of repeated fillers. Tests of simple effects were performed on the data in order to decompose the three-way interaction. In the absence of fillers, the two-way interaction of task and SOA did not reach significance [$F(3,69) = 2.40, p = .08$]. In the presence of repeated fillers, task interacted with SOA [$F(3,69) = 8.20, p < .01$]. Further analysis revealed the following pattern: The effect of task was not significant at the shortest SOA or lag 1 [$t(23) = 0.14, p = .89$] but was significant at the 300-msec SOA [$t(23) = 3.56, p < .01$] and then did not reach significance at 600 msec [$t(23) = 0.23, p = .82$] and 1,350 msec [$t(23) = -1.06, p = .30$]. In the case of changing fillers, again the interaction of task and SOA was significant [$F(3,69) = 4.82, p < .01$], and the pattern of results was similar to the latter: The effect of task was marginally significant at 150 msec [$t(23) = 2.49, p = .02$], significant at the following SOA of 300 msec [$t(23) = 2.95, p < .01$], but not significant thereafter ($ps > .20$).

Homogeneity of fillers. The relative impact of heterogeneous and homogeneous fillers on the dual-task cost was tested using the difference in performance between single and dual tasks as the dependent variable for the repeated- and changing-filler conditions at the SOAs that revealed an AB-like deficit (i.e., 150 and 300 msec). A repeated measures ANOVA with homogeneity (two levels: repeated vs. changing fillers) and SOA (two levels: 150 and 300 msec) as factors was carried out on the data (see Figure 5). The main effect of homogeneity was not significant [$F(1,23) = 0.41, MS_e = 0.029, p = .53$]. The effect of SOA was significant [$F(1,23) = 5.43, MS_e =$

$0.034, p = .03$] and so was the interaction of homogeneity and SOA [$F(1,23) = 4.64, MS_e = 0.019, p = .04$]. Decomposition of the interaction revealed the following pattern: The difference between repeated and changing fillers was marginally significant at 150 msec [$t(23) = 1.93, p = .07$] but was far from significant at 300 msec [$t(23) = -0.85, p = .41$]. As for the effect of SOA, the difference between 150 and 300 msec was significant in the presence of repeated fillers [$t(23) = 3.00, p < .01$] but was not significant with changing fillers [$t(23) = -0.59, p = .56$]. The latter result reflects the observed lag 1 sparing.

The results of Experiment 2 provide further evidence that AB-like deficits within the auditory modality are not restricted to dual-stream presentations (see also, e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002). By presenting the various filler conditions (no filler, repeated filler, and changing fillers) randomly from trial to trial, the possibility that the AB effects observed here (and also in Experiment 3) were the result of preparatory strategy differences is ruled out. The presence of fillers seems to be a necessary condition for the blink to occur, and the perceptual organization of the fillers appears to have a differential effect on the blink response. There was no residual dual-task cost in the absence of filler items, this time within a single-stream presentation. Without a change of pitch between the target and the probe, there was very little difference in discrimination performance between single- and dual-task conditions. Both repeated- and changing-filler presentations seem to show the U-shaped function of task and SOA, the so-called signature of the true AB (although the sparing of lag 1 appears to have been larger for the repeated-filler presentations). The magnitude of the blink was reduced, albeit only slightly, when the auditory sequence was homogeneous. Fillers in the heterogeneous condition were phonologically similar at their onset, which may have diminished the degree of streaming between homogeneous and heterogeneous filler streams. In Experiment 3, an attempt was made to enhance the effect of heterogeneity on the size of the AB by increasing the degree of change between fillers in the changing-filler presentations.

EXPERIMENT 3

The design and procedure in Experiment 3 were the same as those employed in Experiment 2, except for the stimuli used as fillers. In Experiment 2, given that the three syllables were presented in fixed order and shared a similar onset (*guh*, *gah*, and *gih*), the degree of change among them can arguably be regarded as moderate and, hence, insufficiently heterogeneous to break up the coherence of the auditory stream. In the present experiment, we used a set of six items phonologically dissimilar at both onset and offset, and the presentation order of these items within filler sequences was randomized, rather than arranged in the same fixed and predictable order, thereby diminishing the coherence of the filler sequences still further.
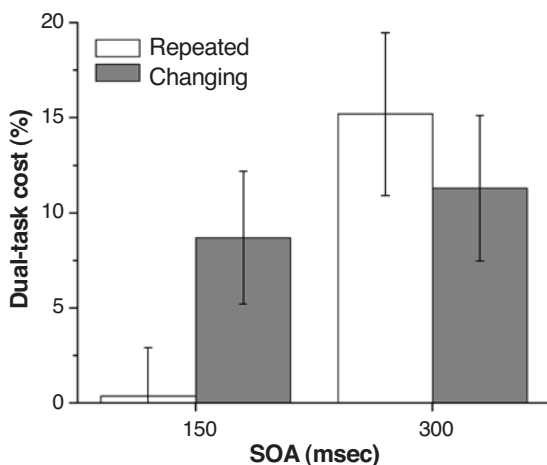

Figure 5. Results from Experiment 2: dual-task cost as a function of SOA (150 and 300 msec) and homogeneity (repeated vs. changing fillers). This cost corresponds to the difference in performance between single and dual tasks. Error bars represent standard errors.

Method

Participants. Twenty-four students volunteered to participate in this experiment in exchange for a small honorarium. Each reported normal hearing.

Materials. As in Experiment 2, all the items were presented at the same pitch within a single sequence. In the changing-filler condition, there were six different fillers (*guh*, *kev*, *del*, *rum*, *tic*, and *bow*) presented in a random order. In the repeated-filler condition, one of the six items was presented repeatedly. The filler to be repeated was chosen randomly, with the restriction that all six were presented at least twice in the experimental session. The construction of the individual items and the item sequences were the same as those in Experiment 2 in terms of both procedure and parameters. As in the previous experiment, the probe was presented following one of four delays of 150 msec, corresponding to lag 1, 2, 4, or 9. The experimental design and the procedure were the same as those described in the previous experiments.

Results and Discussion

Target discrimination. On average, performance at target discrimination in the dual-task condition was 84.2%. A repeated measures ANOVA with SOA (four levels) and homogeneity (three levels) as factors showed a significant effect of homogeneity [$F(2,46) = 6.07$, $MS_e = 0.024$, $p < .01$], but neither the effect of SOA [$F(3,69) = 2.45$, $MS_e = 0.013$, $p = .07$] nor the interaction between homogeneity and SOA [$F(6,138) = 0.38$, $MS_e = 0.02$, $p = .82$] was significant. Post hoc comparisons (Bonferroni t tests; alpha = .01) were carried out on the main effect of homogeneity. Performance in the presence of changing fillers (80.0%) was significantly worse than performance in the presence of repeated fillers (84.6%) and in the absence of fillers (87.8%).

Probe discrimination. The probability of reporting the probe (given a correct report of the target) as a function of task (two levels), SOA (four levels), and homogeneity (three levels) was analyzed with a three-way repeated measures ANOVA. The data are displayed in Figure 6. All three main effects were significant [task, $F(1,23) = 12.05$, $MS_e = 0.051$, $p < .01$; SOA, $F(3,69) = 11.43$, $MS_e = 0.018$, $p < .01$; homogeneity, $F(2,46) = 41.09$, $MS_e = 0.019$, $p < .01$]. The interaction between task and homogeneity was significant [$F(2,46) = 4.95$, $MS_e = 0.016$, $p < .05$], as was that between SOA and homogeneity [$F(6,138) = 6.16$, $MS_e = 0.012$, $p < .01$]. Importantly, the interaction between SOA and task was significant [$F(3,69) = 12.40$, $MS_e = 0.011$, $p < .01$].

The three-way interaction of task, SOA, and homogeneity of fillers was also significant [$F(6,138) = 3.30$, $MS_e = 0.012$, $p < .01$]. The composition of the three-way interaction was revealed by further analysis on the data (alpha = .01). Again, the two-way interaction of task and SOA was not significant in the absence of fillers [$F(3,69) = 0.19$, $p = .90$]. The same interaction was significant in the repeated-filler condition [$F(3,69) = 2.80$, $p < .05$] and also in the changing-filler condition [$F(3,69) = 13.32$, $p < .01$]. The pattern of significance in relation to the effect of task at each SOA was different whether the fillers were repeated or changing (Bonferroni t tests; alpha = .01): With changing fillers, there was a significant deficit at the first two SOAs (150 and 300 msec),
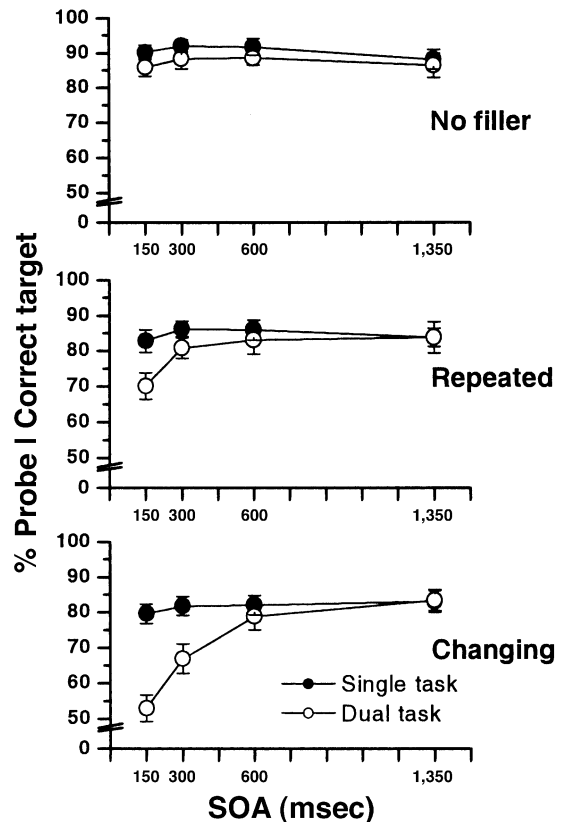

Figure 6. Results from Experiment 3: probability of probe report, given correct report of the target, as a function of task, SOA, and homogeneity of fillers (no filler, repeated, and changing). Error bars represent standard errors.

but this was far from significant thereafter, whereas a significant difference was observed only at 150 msec in the repeated condition.

Homogeneity of fillers. As in Experiment 2, a repeated measures ANOVA with homogeneity (two levels: repeated vs. changing fillers) and SOA (two levels: 150 or 300 msec) as factors was performed on the dual-task cost in order to test the impact of the heterogeneity of the fillers on the blink. A visual inspection of the data suggests that heterogeneous fillers were more disruptive than homogeneous fillers at both SOAs (see Figure 7). Both main effects were significant [homogeneity, $F(1,23) = 12.44$, $MS_e = 0.027$, $p < .01$; SOA, $F(1,23) = 7.40$, $MS_e = 0.030$, $p = .01$]. The interaction of homogeneity and SOA was not significant [$F(1,3) = 0.38$, $MS_e = 0.033$, $p = .54$].

As was the case in Experiment 2, there was a marked difference between single and dual tasks in terms of probe report, and importantly, that difference was related to SOA. The latter finding provides strong evidence for the existence of the AB in the auditory modality. Our attempt to increase the difference between repeated- and changing-filler conditions in relation to the size of the blink was successful. In Experiment 2, a small set of changing fillers with a low degree of heterogeneity yielded
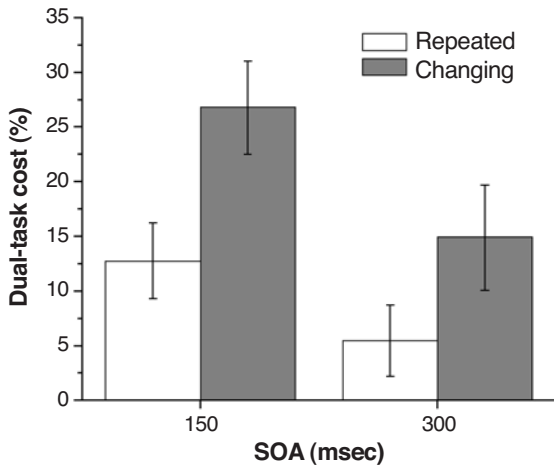
Figure 7. Results from Experiment 3: dual-task cost as a function of SOA (150 and 300 msec) and homogeneity (repeated vs. changing fillers). This cost corresponds to the difference in performance between single and dual tasks. Error bars represent standard errors.

a slightly greater blink than repeated fillers did, whereas in the present experiment, heterogeneous fillers produced a deficit of significantly greater magnitude to that produced by repeated fillers. Overall accuracy in Experiment 3 was lower than that observed in Experiment 2. One possible reason for the latter result is that using a design in which filler conditions are mixed led to carry-over effects of heterogeneous trials onto repeated trials. A rather puzzling result is that we did not replicate the sparing effect at the shorter SOA observed in Experiment 2. These results require further exploration.

GENERAL DISCUSSION

A substantial deficit in reporting the second of two auditory targets was observed in both dual-stream and single-stream auditory sequences. In all three experiments, it was demonstrated that the magnitude of these effects depends critically on the context provided by the items that surround the stimuli to be reported. Experiment 1 contributed to an understanding of those contextual factors in a setting that involved switching attention between streams. Presenting the probe embedded with fillers of the same pitch increased the AB appreciably. Experiments 2 and 3 served to clarify the role played by the nature of fillers within a single-stream presentation. Whether the fillers were homogeneous or heterogeneous had a significant influence on the magnitude of the blink. Generally, lag 1 and lag 2 probes are susceptible to the effects of the coherence of the filler sequence. When the fillers were identical (homogeneous conditions in Experiments 2 and 3) or phonologically similar and predictable (heterogeneous condition in Experiment 2), the deficit was small, but when the fillers were phonologically heterogeneous and random (heterogeneous condition in Experiment 3), the deficit was large. These effects of con-

text suggest that perceptual organization contributes to the overall auditory AB. Certainly, there are similarities with the visual modality, but it is not clear whether the same factors modulate the effect or, indeed, whether there is a general sensitivity to context.

Implications for Theoretical Accounts of the AB

Most recent models of the AB cleave along the distinction between retrieval competition (e.g., Raymond et al., 1995; Shapiro et al., 1994) and *bottleneck* type of processing (e.g., Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998). In short, the bottleneck interpretation assumes that target stimuli are processed in a serial fashion.[4] While resources are occupied with the target, there is a period during which no or fewer attentional resources are available for processing those stimuli that closely follow the target. According to the other type of explanation, the retrieval competition model, each item is processed to some degree, but only a few items are transferred to STM. Items may be admitted into STM if they match a preset template of the target or the probe and if they are temporally contiguous to the target or the probe. This type of interpretation is similar to theoretical accounts in the STM literature that assume that interference is caused by a similarity of content between target stimuli (e.g., to-be-remembered items) and nontargets (e.g., irrelevant sound; see Jones & Tremblay, 2000, for a review).

Bottleneck models are in agreement that the AB arises from attentional limitations but differ with respect to the stage at which to place the bottleneck. Jolicœur's (1998; see also Jolicœur & Dell'Acqua, 1998) central interference theory places the bottleneck at a stage, that of short-term consolidation (STC), later than that at which it is located in the two-stage model (Chun & Potter, 1995). A key assumption specific to the central interference theory is that STC requires central processing. Therefore, while the target stimulus goes through STC, any other operations (such as probe STC or task switch) that require central processing have to wait (see Jolicœur, 1998). Another possible source of the AB deficit is that task or location switching makes use of central resources, thus postponing the central processing required by STC of the probe (see McLaughlin et al., 2001). With a similar argument, the central interference theory can account for the observed cost of pitch switching.

Processes other than retrieval competition and consolidation may also produce AB-like effects. Some versions of the blink paradigm may impose different processing requirements on the target and the probe. The blink may therefore be a product of switching from one type of activity to another, and this switch may be wholly or partly responsible for the loss of efficiency in reporting the probe. In all three experiments of the present study (as in Duncan et al., 1997), there was no task switch between the target and the probe; both the target and the probe required a qualitatively similar two-alternative forced choice discrimination. The same task and same target set size were used for the target and the probe. Therefore, the auditory AB observed here, as well as that

found by Duncan et al., is unlikely to be caused by an amodal task switch (as claimed by Potter et al., 1998). However, the possibility remains that there was some cost attributable to a switch of target set (e.g., *co* for the target followed by *na* for the probe; see Arnell & Larson, 2002).

In the present study, we explored the impact of the context in which a target and a probe are presented on the auditory AB effects. Context effects were tested through the manipulation of coherence across streams (Experiment 1: pitch segregation) and within a stream (Experiments 2 and 3: heterogeneity of fillers). Segregation by pitch (e.g., in T + PF and TF + P) leads to inserting the probe and the target in different but coherent groups, and heterogeneity is concerned with placing them into one group that is more or less coherent. Whether it is through stream segregation or within-stream coherence, our results revealed that perceptual organization plays a major role in modulating the auditory AB.

Pitch segregation: Experiment 1. A logic of perceptual organization and grouping can be applied to account for the results of Experiment 1 on the basis of the principle of similarity by frequency between target and filler items (as opposed to similarity by repetition or phonology among fillers). Indeed, one explanation is that auditory sequences in which the probe was surrounded by fillers of the same pitch (TF + PF and T + PF) yielded a greater AB by making the probe harder to distinguish from the other items in that sequence. On the contrary, sequences in which the probe differed from the other items in pitch (TF + P and T + P) induced the probe to pop out of the sequence and, hence, substantially reduced the size of the AB.

Bottleneck models can account for the results of Experiment 1 with the assumption that by popping out, the probe suffers less perceptual interference while waiting for STC. The retrieval competition model can also explain the same pattern of results by assuming that the allocation of attentional resources is based on the similarity of targets with preset templates. As the similarity between filler and target items increases, the amount of resources, devoted to fillers temporally contiguous with the targets, also increases. In Experiment 1, the probe and its surrounding fillers were presented at the same pitch; that increased similarity between the probe and the filler items resulted in more competition for retrieval among the items admitted in STM—hence, the greater AB observed in the TF + PF and T + PF conditions. According to the retrieval competition model, it is not clear why the probe–filler similarity produced greater AB effects than the target–filler similarity did.

Heterogeneity of fillers: Experiments 2 and 3. There is very little research on the impact of placing targets among heterogeneous fillers on the AB. In demonstrating that the processing of color information is susceptible to the AB, Ross and Jolicœur (1999) provided data that speak to the issue of heterogeneity. Within a rapid serial visual presentation of color stimuli, a probe that had to be detected was immune from the AB when the fillers were homogeneous and of a different color (e.g., performance at detecting a red probe among blue fillers was very high at any lag). However, when the probe was embedded in fillers of heterogeneous colors, a significant AB was observed. The authors concluded that effectiveness of masking is reduced when targets pop out of the filler stream (see also Arnell & Jolicœur, 1999).

The effect of heterogeneity on target detection has been demonstrated on many occasions in visual search studies (see, e.g., Bauer, Jolicœur, & Cowan, 1996; Duncan, 1980; Hoffman, 1978; Pashler, 1987; Treisman & Gelade, 1980). Homogeneity among fillers usually makes them easier to suppress when a target is searched for. In other words, heterogeneous fillers make target detection less likely. The results from a variety of experiments using auditory analogues of the visual search paradigm provide evidence that similar context effects are observed in auditory selection (e.g., Botte, Drake, Brochard, & McAdams, 1997; Mondor, Zatorre, & Terrio, 1998; see also Cusack & Carlyon, 2003). Mondor and his colleagues (Mondor & Terrio, 1998; Mondor et al., 1998) proposed a model of auditory selection attention based on the visual selection theory of Duncan and Humphreys (1989). Stimuli that match a preset template are more likely to be selected as targets, and other stimuli are rejected. However, selection is constrained by the outcome of the perceptual organization of stimuli into streams, at some preattentive stage. There is ample evidence that selection of information is preceded by preattentive perceptual organization.

By enhancing the heterogeneity among the changing fillers (Experiment 3), we obtained a greater effect of heterogeneity than that observed in Experiment 2. The phonological similarity among heterogeneous fillers used in Experiment 2 may have promoted grouping, although to a lesser extent than repeated fillers did. That heterogeneous fillers provoke a greater auditory AB can be ascribed to the effects of grouping. Both the retrieval competition theory and the bottleneck models can accommodate the construct of grouping. The development of the retrieval competition theory (e.g., Shapiro et al., 1994) has its roots in Duncan and Humphreys's (1989) theory of visual selection. It is assumed that competition between target and filler items is influenced by principles of grouping. Homogeneous fillers will likely tend to be grouped together and rejected as a whole (which includes items following targets acting as masks). One can make the assumption that, as a consequence of the easier rejection of fillers as a group, there is less competition for retrieval and the AB is, therefore, greatly reduced. Bottleneck models can be taken to posit that homogeneity among fillers makes the target and the probe pop out of the sequence. Hence, the target is processed more efficiently, and the delay before the processing of the probe is shorter. In addition, if the probe suffers less perceptual interference by masking, the probability of overwriting during the delay is reduced.

Another possibility that could account for the effect of heterogeneity observed in Experiments 2 and 3 is the

concept of stimulus uncertainty. Research on the impact of auditory masking upon target identification has shown that a mask is generally more disruptive with a high degree of stimulus uncertainty (see, e.g., Durlach et al., 2003). Predictability of the frequency or location of the item acting as a mask has been shown to influence the degree of masking efficiency (see Hawkins & Presson, 1986). Therefore, the effect of heterogeneity could also be explained by the fact that the identity of the masks was unpredictable in the heterogeneous streams (especially in Experiment 3).

Conclusion

In conclusion, there is mounting evidence that more than one factor is responsible for AB interference. Organization of stimuli in terms of spatial location (e.g., Spence, Ranson, & Driver, 2000), pitch (results of Experiment 1), and heterogeneity (Experiments 2 and 3) has been shown to affect the AB. The present study provides evidence that there is a so-called conventional AB in the auditory modality and that its effect can be additive to other effects, such as pitch switching. A promising line of research might be to apply the principles and theories of auditory selection (e.g., Mondor & Terrio, 1998) to the study of the auditory AB.

REFERENCES

ARNELL, K. M. (2001). Cross-modal interactions in dual-task paradigms. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 141-177). Oxford: Oxford University Press.

ARNELL, K. M., & JOLICŒUR, P. (1999). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 630-648.

ARNELL, K. M., & LARSON, J. M. (2002). Cross-modality attentional blinks without preparatory task-set switching. *Psychonomic Bulletin & Review*, **9**, 497-506.

BAUER, B., JOLICŒUR, P., & COWAN, W. B. (1996). Distractor heterogeneity versus linear separability in colour visual search. *Perception*, **25**, 1281-1294.

BOTTE, M.-C., DRAKE, C., BROCHARD, R., & MCADAMS, S. (1997). Perceptual attenuation of nonfocused auditory streams. *Perception & Psychophysics*, **59**, 419-425.

BREGMAN, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 380-387.

BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

BREGMAN, A. S., & RUDNICKY, A. I. (1975). Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 263-267.

BROADBENT, D. E., & BROADBENT, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics*, **42**, 105-113.

CHUN, M. M., & POTTER, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 109-127.

CHUN, M. M., & POTTER, M. C. (2001). The attentional blink and task switching within and across modalities. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 20-35). Oxford: Oxford University Press.

CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, **5**, 365-373.

CUSACK, R., & CARLYON, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 713-725.

DUNCAN, J. (1980). The locus of interference in perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.

DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.

DUNCAN, J., MARTENS, S., & WARD, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, **387**, 808-810.

DURLACH, N. I., MASON, C. R., KIDD, G., JR., ARBOGAST, T. L., COLBURN, H. S., & SHINN-CUNNINGHAM, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, **113**, 2984-2987.

ENNS, J. T., VISSER, T. A. W., KAWAHARA, J., & DI LOLLO, V. (2001). Visual masking and task switching in the attentional blink. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 65-81). Oxford: Oxford University Press.

GIESBRECHT, B., & DI LOLLO, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1454-1466.

GODDARD, K. M., & SLAWINSKI, E. B. (1999). Modality specific attentional mechanisms can govern the attentional blink. *Canadian Acoustics*, **27**, 98-99.

HAWKINS, H., & PRESSON, J. (1986). Auditory information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. II. Cognitive processes and performance* (pp. 1-64). Oxford: Wiley.

HOFFMAN, J. E. (1978). Search through a sequentially presented visual display. *Perception & Psychophysics*, **23**, 1-11.

JOLICŒUR, P. (1998). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded $Task_1$ decisions. *Memory & Cognition*, **26**, 1014-1032.

JOLICŒUR, P. (1999a). Concurrent response-selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1097-1113.

JOLICŒUR, P. (1999b). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin & Review*, **6**, 87-92.

JOLICŒUR, P., & DELL'ACQUA, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, **36**, 138-202.

JOLICŒUR, P., DELL'ACQUA, R., & CREBOLDER, J. M. (2001). The attentional blink bottleneck. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 82-100). Oxford: Oxford University Press.

JONES, D. M., MACKEN, W. J., & HARRIES, C. (1997). Disruption of short-term recognition memory for tones: Streaming or interference? *Quarterly Journal of Experimental Psychology*, **50A**, 337-357.

JONES, D. M., & TREMBLAY, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review*, **7**, 550-558.

KAWAHARA, J.-I. (2003). Mere presence of distractors: Another determining factor for the attentional blink. *Japanese Psychological Research*, **3**, 150-151.

KAWAHARA, J.-I., ZUVIC, S. M., ENNS, J. T., & DI LOLLO, V. (2003). Task switching mediates the attentional blink even without backward masking. *Perception & Psychophysics*, **65**, 339-351.

MCLAUGHLIN, E. N., SHORE, D. I., & KLEIN, R. M. (2001). The attentional blink is immune to masking-induced data limits. *Quarterly Journal of Experimental Psychology*, **54A**, 169-196.

MONDOR, T. A. (1998). A transient processing deficit following selection of an auditory target. *Psychonomic Bulletin & Review*, **5**, 305-311.

MONDOR, T. A., & BREGMAN, A. S. (1994). Allocating attention to frequency regions. *Perception & Psychophysics*, **56**, 268-276.

MONDOR, T. A., & TERRIO, N. A. (1998). Mechanisms of perceptual organization and auditory selective attention: The role of pattern structure. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1628-1641.

MONDOR, T. A., ZATORRE, R. J., & TERRIO, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 66-79.

NICHOLLS, A. P., & JONES, D. M. (2002). Capturing the suffix: Cogni-

tive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 12-28.

PASHLER, H. (1987). Target–distractor discriminability in visual search. *Perception & Psychophysics*, **41**, 285-292.

POTTER, M. C., CHUN, M. M., BANKS, B. S., & MUCKENHOUPT, M. (1998). Two attentional deficits in serial target search: The visual attentional blink and an amodal task-switch deficit. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 979-992.

RAYMOND, J. E., SHAPIRO, K. L., & ARNELL, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 849-860.

RAYMOND, J. E., SHAPIRO, K. L., & ARNELL, K. M. (1995). Similarity determines the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 653-662.

ROGERS, W. L., & BREGMAN, A. S. (1998). Cumulation of the tendency to segregate auditory streams: Resetting by changes in location and loudness. *Perception & Psychophysics*, **60**, 1216-1227.

ROSS, N. E., & JOLICŒUR, P. (1999). Attentional blink for color. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1483-1494.

RUTHRUFF, E., & PASHLER, H. E. (2001). Perceptual and central interference in dual-task performance. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 100-123). Oxford: Oxford University Press.

SEIFFERT, A. E., & DI LOLLO, V. (1997). Low-level masking in the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1061-1073.

SHAPIRO, K. L., RAYMOND, J. E., & ARNELL, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 357-371.

SHIH, S.-I (2000). Recall of two visual targets embedded in RSVP streams of distractors depends on their temporal and spatial relationship. *Perception & Psychophysics*, **62**, 1348-1355.

SOTO-FARACO, S., & SPENCE, C. J. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology*, **55A**, 23-40.

SPENCE, C. J., & DRIVER J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 555-574.

SPENCE, C. [J.], RANSON, J., & DRIVER, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Perception & Psychophysics*, **62**, 410-424.

TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

VACHON, F., & TREMBLAY, S. (in press). Auditory attentional blink: Masking the second target is necessary, delayed masking is sufficient. *Canadian Journal of Experimental Psychology*.

VISSER, T. A. W., BISCHOF, W. F., & DI LOLLO, V. (1999). Attentional switching in spatial and non-spatial domains: Evidence from the attentional blink. *Psychological Bulletin*, **125**, 458-469.

WOODS, D. L., ALAIN, C., DIAZ, R., RHODES, D., & OGAWA, K. H. (2001). Location and frequency cues in auditory selection attention. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 65-74.

NOTES

1. The notation used here and elsewhere in this article indicates the relation of the target (T) and the probe (P) to the serial position of their mask in the sequence. For example, the notation P+1 signifies that there is an item, acting as a mask, immediately following the probe.

2. In the AB literature, dual-stream presentation does not necessarily refer to the percept of two separate streams, since it would prejudge the nature of the perceptual organization. Also, given the short duration of RAP, streams may not have enough time to build up fully (see Bregman, 1978; Rogers & Bregman, 1998).

3. Here and elsewhere in this article, we refer to such terms as *pitch switching* or *switch pitch* to indicate that attention has to be reallocated from the frequency region of the target to that of the probe.

4. The bottleneck model proposed by Jolicœur and Dell'Acqua (1998; see also Jolicœur, Dell'Acqua, & Crebolder, 2001) to account for AB phenomena assumes limited capacity and serial processing, but the authors do not exclude parallel processing as a possible explanation of the effects.