

Modality-specific and amodal sources of interference in the attentional blink

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

When two masked targets (T1 and T2) are visually or auditorily presented in rapid succession, processing of T1 produces an attentional blink (AB)—that is, a transient impairment of T2 identification. The present study was conducted to compare the relative impact of masking T1 and T2 between vision and audition. Within a rapidly presented sequence, each of the two verbal targets, discriminated by their offset (Experiment 1) or their onset (Experiment 2), could be followed by either a single item, acting as a mask, or a blank gap. Masking of T2 appeared to be necessary for the occurrence of the AB for both the visual and the auditory modality. However, whereas masking of T1 affected the expression of the visual AB in both experiments, the same effect was observed in the auditory modality only when the targets varied at the onset. These results provide further evidence that processing auditory and visual information is restricted by similar attentional limitations but also suggest that these limits are constrained by properties specific to each sensory system.

When interacting with its environment, an organism is constantly confronted with an overwhelming amount of perceptual input from all the senses, continuously changing over time and space. The efficient management of this flow of information is under the control of selective attention, and this process can be viewed as the cognitive capacity of preferentially processing stimulations that are the most relevant as guides to action, while concurrently disregarding unwanted stimuli (see, e.g., Pashler, 1998). Many studies have shown that attentional mechanisms are limited as to the number of stimuli they can handle simultaneously, as reflected, for example, by breakdowns in an individual's ability to efficiently process multiple stimuli presented closely in time (see, e.g., Shapiro, 2001). Temporal-processing deficits have been observed over a wide range of paradigms within different sensory modalities or even across modalities. In the present study, we examined whether temporal constraints on the processing of auditory information are functionally similar to those observed in the visual domain. We did so by using the attentional blink (AB) phenomenon, which is a manifestation of the temporal limitations on human information processing, to contrast audition and vision.

The AB occurs when two to-be-processed masked targets are presented in rapid succession (Raymond, Shapiro, & Arnell, 1992). This is true whether the two targets are embedded within a rapid stream of distractors (a procedure known in the visual domain as rapid serial visual presentation; see, e.g., Broadbent & Broadbent, 1987; Weichselgartner & Sperling, 1987) or followed only by a single posttarget stimulus (a procedure usually referred as the two-target paradigm; see, e.g., Duncan, Ward, & Sha-

piro, 1994; Ward, Duncan, & Shapiro, 1997). The typical expression of the phenomenon is a transient deficit in reporting the second (T2) of two targets when it follows the first target (T1) by less than 500 msec. A well-accepted explanation of the AB suggests that processing of T2 is impaired while limited attentional resources are committed to the processing of T1 (see Shapiro, Arnell, & Raymond, 1997). Such a constraint in the temporal distribution of attention has been extensively investigated in vision, but far fewer studies have been carried out in the auditory domain. One key feature of the visual AB phenomenon is the determinant role of target masking (see Enns, Visser, Kawahara, & Di Lollo, 2001). When applied to the auditory AB, the concept of masking also appears to contribute to the phenomenon (e.g., Vachon & Tremblay, 2005); however, its role seems to be restricted, as compared with the role of masking in the visual AB (e.g., Shen & Mondor, 2006). Such findings suggest that the function of masking in the AB would differ according to the modality, which has important implications for understanding the nature of the AB phenomenon and the limitations of temporal processing in both audition and vision.

The investigation of the AB in various sensory modalities is of great importance, since theoretical models diverge with regard to the origin of the attentional limitations underlying the phenomenon. Indeed, some models assume that the AB reflects attentional constraints specific to each modality (e.g., Duncan, Martens, & Ward, 1997; Soto-Faraco & Spence, 2002) or even that the effect is restricted to vision (e.g., Potter, Chun, Banks, & Muckenhaupt, 1998; Shapiro, Raymond, & Arnell, 1994), whereas others suggest that the AB relies on central and

F. Vachon, f.vachon@umontreal.ca

amodal processing limitations (e.g., Arnell & Jenkins, 2004; Jolicœur & Dell'Acqua, 1998). However, investigations of the auditory AB do not allow unequivocal conclusions about the nature of the phenomenon. Although a number of studies have provided direct AB comparisons between the visual and the auditory modalities (e.g., Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Duncan et al., 1997; Soto-Faraco & Spence, 2002), in none of these has the role of masking been examined. Since the literature on the auditory AB has revealed conflicting evidence regarding the role of masking and its correspondence with the visual modality, in the present study we sought to uncover the source of such discrepancies by comparing the impact of target masking on the AB between the auditory and the visual modalities.

Masking in the Visual AB

Despite the well-established postperceptual nature of the AB (e.g., Martens, Wolters, & van Raamsdonk, 2002; Shapiro, Caldwell, & Sorensen, 1997; Vachon, Tremblay, & Jones, 2007; Vogel, Luck, & Shapiro, 1998), the contribution of perceptual interference to the phenomenon is essential. Within the visual AB literature, most of the studies have focused on the influence of perceptual interference provided by the presence of a single posttarget item, denoted as a mask. The masking of a target, either T1 or T2, in a visual sequence usually yields detrimental effects on the probability of reporting T2 correctly. However, there is ample evidence that the role of masking T1 in the visual AB differs from that of masking T2 (see Enns et al., 2001).

T1 masking was recognized early as an important factor in the visual AB phenomenon (see Raymond et al., 1992). Replacing the item that immediately followed T1 in the visual sequence (T1 + 1) by a blank usually leads to the attenuation of the visual AB (e.g., Chun & Potter, 1995; Grandison, Ghirardelli, & Egeth, 1997; Moore, Egeth, Berglan, & Luck, 1996; Seiffert & Di Lollo, 1997). Backward interference suffered by T1 from the subsequent item is not the only form of masking that affects the visual AB. Indeed, even when T1 + 1 is omitted, visual AB effects can be observed if an item is displayed simultaneously with T1, either spatially superimposed (integration masking) or in close spatial proximity (metacontrast masking; see, e.g., Grandison et al., 1997; Seiffert & Di Lollo, 1997). Nevertheless, the presence of perceptual interference upon T1 does not appear to be necessary to observe the visual AB, since the effect is reduced but not abolished in the absence of T1 masking (see Visser, 2007, for a discussion).

On the other hand, the contribution of T2 masking seems to be essential for the visual AB to occur. In fact, there is no AB deficit when T2 terminates the sequence (e.g., Giesbrecht & Di Lollo, 1998; Jolicœur, 1999; Vogel & Luck, 2002). Whereas the type of masking is rather unimportant for T1, the precise form of masking is critical in the case of T2. If T2 is masked by integration (i.e., T2 is overlaid by a nontarget item), its identification is compromised, but the deficit is not related to the temporal distance between the two targets (Brehaut, Enns, & Di Lollo, 1999; Giesbrecht & Di Lollo, 1998). Thus, the presence of at least one item after the presentation of T2, even delayed, is a prerequi-

site for observing reliable visual AB effects (e.g., Brehaut et al., 1999; Giesbrecht & Di Lollo, 1998).

Among the theoretical accounts of the AB making explicit predictions about the influence of masking, one class of theories refers to a processing bottleneck at the stage of transferring information from a perceptual stage to short-term memory (STM). Bottleneck models, such as the two-stage model (Chun & Potter, 1995) and the central interference theory (Jolicœur, 1998, 1999; Jolicœur & Dell'Acqua, 1998), posit that the AB arises because the consolidation of T2 into STM—a process assumed to be capacity limited—is postponed while T1 is being consolidated. During this delay, the representation of T2 is vulnerable to interference from subsequent stimuli. According to bottleneck models, the role of the T2 mask is to prevent T2 report by corrupting its representation while awaiting consolidation (Giesbrecht & Di Lollo, 1998). On the other hand, the role of masking T1 is to degrade the target, thereby slowing down its processing. Such an increase in the time needed to process T1 extends the delay for consolidating T2.

Although the bottleneck models agree that a postperceptual bottleneck is responsible for the AB, they support different views as to whether the phenomenon will extend outside the visual modality. Indeed, whereas the two-stage model confines the bottleneck to visual information processing (see Chun & Potter, 2001), the central interference theory explicitly predicts the existence of the AB in audition by locating that bottleneck after the convergence of the information coming from the different sensory systems (e.g., Arnell & Jolicœur, 1999).

Masking in the Auditory AB

In the auditory modality, little is known about the role of target masking, since in only a few studies has masking within the auditory AB paradigm been directly manipulated (Mondor, 1998; Shen & Mondor, 2006; Vachon & Tremblay, 2005, 2006). Some authors have claimed that auditory information is more resilient with regard to interference from subsequent events than is its visual counterpart, given the larger capacity and longer duration of the echoic buffer, relative to iconic memory (Chun & Potter, 2001; Potter et al., 1998). According to this claim, perceptual interference provided by masking would not be an important factor in the auditory AB. Conversely, Vachon and Tremblay (2005) showed that when T2 terminated the auditory sequence, the AB deficit was abolished, suggesting that backward masking was necessary for obtaining auditory AB effects, as is the case in the visual modality (see also Shen & Mondor, 2006). Moreover, no AB occurs in the auditory domain if T2 is masked by integration or if the T2 mask is delayed long enough to appear after T2 has acceded to the so-called short-term consolidation stage of processing (Vachon & Tremblay, 2006), paralleling findings from the visual AB literature (Brehaut et al., 1999; Giesbrecht & Di Lollo, 1998). In the case of the impact of masking T1 in the auditory AB, Mondor (1998; Shen & Mondor, 2006) showed that the removal of T1 + 1 from the auditory sequence had no effect on the size of the AB, rather than markedly reducing the size of the deficit, as is

observed in vision (e.g., Chun & Potter, 1995; Seiffert & Di Lollo, 1997). Such a divergence raises some questions about the equivalence of the AB phenomenon between modalities and, thus, deserves more attention.

Mondor's (1998; Shen & Mondor, 2006) findings may be taken to support the hypothesis that auditory stimuli are somewhat insensitive to backward interference (e.g., Chun & Potter, 2001). However, the results from Vachon and Tremblay (2005, 2006) strongly suggest that auditory information is rather sensitive to interference from subsequent stimuli. The failure of Mondor to find an effect of removing T1+1 from the auditory sequence may be explained if the special role played by perceptual organization in audition is taken into account. The selection of target stimuli is assumed to be constrained, in part, by an early automatic perceptual organization stage during which the auditory system tends to group stimuli on the basis of their relative similarity (e.g., Bregman, 1990; Cusack & Carlyon, 2003; Mondor & Terrio, 1998; Mondor, Zatorre, & Terrio, 1998). According to Mondor and Terrio, the "perceptual organization process acts dynamically over time to determine whether each successive tone is a member of a larger sequence" (p. 1634). Since tones grouped at a perceptual level tend to be selected or rejected as a group, the selection of a target will be more difficult if it shares perceptual attributes analogous to those of the other stimuli in the sequence. There is evidence that an increase of target-distractor similarity within an auditory sequence reduces accuracy (e.g., Tysiaczny & Mondor, 2005) and lengthens detection time for a target tone (e.g., Mondor et al., 1998). Given the high similarity between T1 and the distractors in Mondor's (1998) study (T1 was a pure tone of 4000 Hz embedded within pure tones whose frequencies ranged from 452 to 3462 Hz), the interference provided by the whole sequence of distractors may have exceeded that of the T1 mask, rendering the presence of the latter obsolete.

The Present Study

The following experiments were tailor-designed to equate testing conditions for T1 and T2 masking and to compare the AB phenomenon between audition and vision. In order to isolate the influence of the masks from that of the whole sequence of distractors, we adopted a simplified version of the AB paradigm, in which each of the two targets was followed by a single item (e.g., Brehaut et al., 1999; Duncan et al., 1994; Ward et al., 1997). Under visual conditions in which all four items were displayed at the same spatial location, the results showed a performance comparable to that with the traditional stream-like presentation (e.g., McLaughlin, Shore, & Klein, 2001; Ward et al., 1997). However, to our knowledge, such a paradigm has never been used in audition. Hitherto, the examination of the impact of masking in the auditory AB has been based solely on studies employing pure and complex tones (Mondor, 1998; Shen & Mondor, 2006; Vachon & Tremblay, 2005, 2006). Because these stimuli have no equivalent in the visual domain, the use of non-verbal auditory stimuli restricts the comparison with the visual AB. Thus, the target and nontarget items employed

here were verbal stimuli. In either the visual or the auditory modality, the participants' task was to report the identity of two target syllables. Each of the two target stimuli was discriminated by either their offset (Experiment 1) or their onset (Experiment 2). The use of the same task requirements in both modalities in a within-subjects design enabled a straightforward comparison of AB patterns between the auditory and the visual conditions.

The manipulation of masking in the present study consisted of a systematic variation of the presence of a mask after each target. So, T1 and T2 could be independently followed by either a mask item or a blank interval of the same duration. In this way, the contribution of T1 masking to visual and auditory AB effects could be examined separately from that of T2 masking. Following the literature, visual and auditory AB deficits should occur only when T2 is masked. If the removal of the whole sequence of distractors is effective in reducing the influence of perceptual organization in audition, we should observe an enlargement of the AB with the addition of a mask after T1 not only in the visual condition, but also in the auditory condition.

EXPERIMENT 1

The verbal stimuli employed in Experiment 1 were similar to those utilized by Duncan et al. (1997). In either the visual or the auditory condition, the participants had to report the identity of two target syllables: One was a *na'* syllable (either *nab* or *nap*), and the other was a *co'* syllable (either *cod* or *cot*). Therefore, the discriminating information was located at each target syllable's offset. Given that the participants in the present study were all native French speakers, it is noteworthy that the syllables employed here have no particular meaning in French. Auditory targets could be masked by the syllable *guh*, whereas an XXX string mask could follow visual targets.

Method

Participants. Thirty French-speaking adults reporting normal or corrected-to-normal vision and no hearing problems were recruited on the campus of Université Laval. They received a small honorarium for their participation in the experiment.

Apparatus. A Pentium PC computer, with a 15-in. (38.1-cm) VGA monitor and a 16-bit soundcard, was used for presenting stimulus sequences and recording responses. The same computer ran a Visual Basic 6.0 program for the auditory condition¹ and an E-Prime 1.1 program for the visual condition.

Stimuli. In both the visual and the auditory conditions, a stimulus sequence always contained two target items. One target item was a *na'* syllable (*nab* or *nap*), and the other was a *co'* syllable (*cod* or *cot*). Each target could be immediately followed by a single nontarget item, denoted as a mask. Figure 1 presents a schematic illustration of the four types of sequence employed in Experiment 1: (1) T1M+T2M, in which both targets were masked; (2) T1+T2M, in which a mask followed T2 only; (3) T1M+T2, in which only T1 was masked; and (4) T1+T2, in which the two targets were presented with no other item.

Each visual target lasted 80 msec. The visual mask consisted of the XXX string presented for 150 msec. When absent, the mask was replaced by a blank frame of 150 msec. All of the letters forming the visual stimuli were capitalized (NAB, NAP, COD, COT, and XXX) and were presented in 25-point Arial font, subtending approximately 1.26° of visual angle in height on the computer screen.

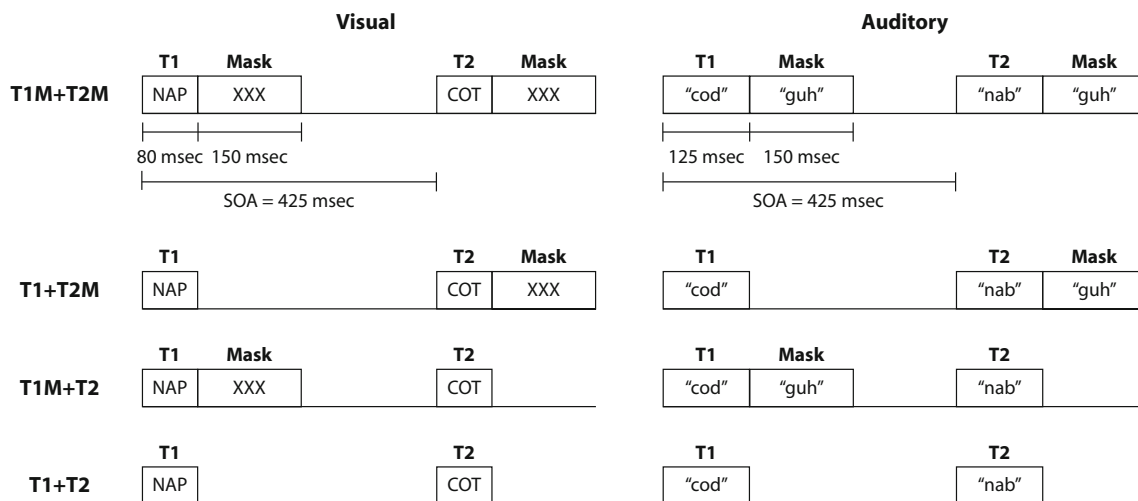


Figure 1. Schematic diagram illustrating the four types of stimulus sequence employed in visual and auditory modalities in Experiment 1: T1M+T2M, T1+T2M, T1M+T2, and T1+T2. Visual targets were presented for 80 msec and auditory targets for 125 msec. Mask items lasted 150 msec in both modalities. T2 followed T1 by one of three delays of 275, 425, or 1,025 msec (all examples show a T1–T2 stimulus onset asynchrony [SOA] of 425 msec).

All the visual stimuli were white and appeared in the center of a black background.

Sounds were digitally edited to a 16-bit resolution at a sampling rate of 48 kHz, using Sound Forge 5.0, and were presented binaurally via headphones at approximately 65 dB(A). All the auditory items were digitally recorded in a male voice. Great care was taken to produce the vowels at an even pitch and level. Each auditory target was compressed to a duration of 125 msec. The auditory mask item was the syllable *guh* compressed to last 150 msec. These digital compressions did not decrease the intelligibility of the individual items. In the absence of the auditory mask, a 150-msec silent gap followed the target.

Both the visual and the auditory sequences began with the presentation of T1, chosen at random on each trial between the *na'* and the *co'* syllables. T2 was presented following delays of 275, 425, or 1,025 msec measured from T1 onset to T2 onset (stimulus onset asynchrony, or SOA). So, even when a T1 mask was absent, T2 never occurred immediately after T1—that is, at a lag of 1. The identity of T2 depended on that of T1: If T1 was a *co'* syllable, T2 was a *na'* syllable, and vice versa. When presented, the mask followed the target immediately, with no interstimulus interval.

Design. A within-subjects design was employed with four factors: modality (visual or auditory), task (single or dual), T1–T2 SOA (275, 425, or 1,025 msec), and type of sequence (T1M+T2M, T1+T2M, T1M+T2, or T1+T2). All the participants took part in two experimental sessions, one for each modality condition; the order of these sessions was counterbalanced across participants. Within each session, the participants performed two single-task conditions (one with the instructions to identify the *na'* syllable and one the *co'* syllable) and one dual-task condition in which both targets had to be reported. These conditions were blocked so that there was one experimental block per task condition. The order of these three blocks was counterbalanced across participants. T1–T2 SOAs and types of sequence were randomized from trial to trial within each block. There were 96 experimental trials per block, preceded by 24 practice trials.

Procedure. The participants had to perform a two-alternative forced choice discrimination; they were told to discriminate between *nab* and *nap* for the *na'* syllable and between *cod* and *cot* for the *co'* syllable. When performing the auditory session, the participants were familiarized with the sounds before attempting any trials. The

participants initiated a trial by a mouse click on a “start” button displayed on the screen in the auditory session and by pressing the space bar on the keyboard in the visual session. Each trial began with the presentation of a fixation cross in the center of the screen for 500 msec, followed by a delay of 100 msec before the start of the stimulus sequence. Each trial was 2,100 msec in duration, measured from the onset of the fixation to the onset of the response prompt. The participants’ responses were typed in following the presentation of each set of stimuli, using keys labeled appropriately for the targets, without time pressure. Under the single-task conditions, there was a single response identifying the attended target. Under dual-task conditions, two responses were typed in, in either order.

Results

Data in the single-task conditions were pooled over target syllables (*na'* and *co'* as single targets) for both T1 and T2. Within the present study, the mean accuracy for T1 and T2 were calculated regardless of whether or not the other target was correctly discriminated. However, the same pattern of data was also found when T2 performance was conditional on T1’s being correctly reported. The mean accuracy for T1 (negative SOAs) and T2 (positive SOAs) is plotted as a function of task, SOA, and type of sequence in Figure 2 for the auditory condition and in Figure 3 for the visual condition.

T1 performance. The mean percentage of correct reports of T1 was submitted to a repeated measures ANOVA with modality (2 levels), task (2 levels), SOA (3 levels), and type of sequence (4 levels) as within-subjects factors. Here and elsewhere in the present study, the Greenhouse–Geisser procedure was applied on every effect for which the sphericity assumption was violated. The analysis showed significant effects of modality [$F(1,29) = 72.22, p < .001, d = 3.16$], task [$F(1,29) = 27.83, p < .001, d = 1.96$], and SOA [$F(2,58) = 7.99, p < .001, d = 1.05$]. There was also a significant effect of type of sequence [$F(3,87) = 79.19, p < .001, d = 3.31$], which indicates

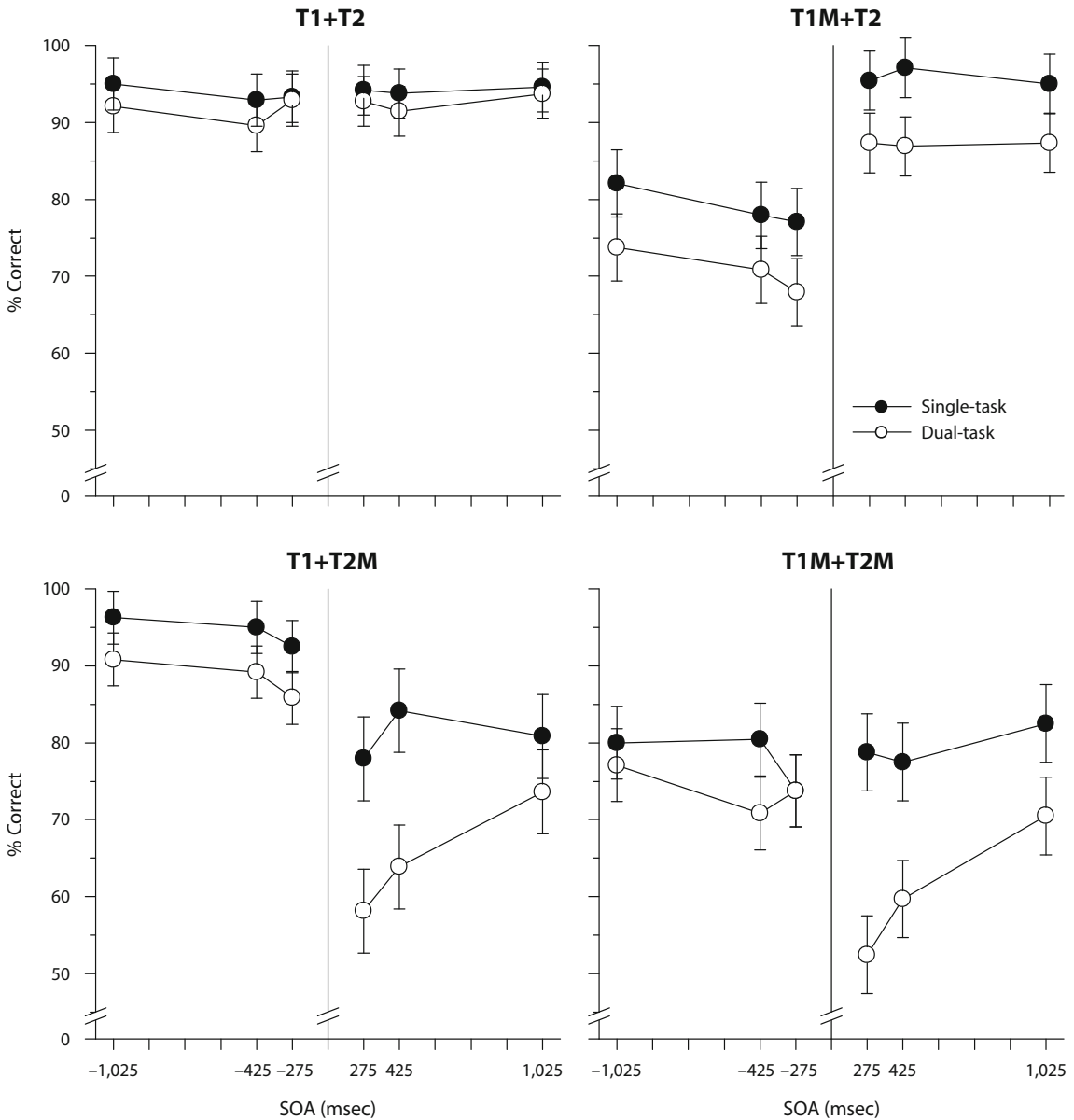


Figure 2. Results from the auditory condition in Experiment 1: Mean percentages of correct target responses as a function of task (single vs. dual), T1–T2 stimulus onset asynchrony (SOA), and type of sequence (T1M+T2M, T1+T2M, T1M+T2, and T1+T2). Positive SOAs refer to T2, and negative SOAs refer to T1. Error bars represent 95% within-subjects confidence intervals.

that T1 performance was at its lowest level when T1 was followed by a mask. The significant modality \times type of sequence interaction [$F(3,87) = 34.67, p < .001, d = 2.19$] provided evidence that this detrimental effect of T1 mask on T1 performance was larger in the auditory condition. The interaction between modality and SOA was significant [$F(2,58) = 3.25, p = .03, d = 0.67$], as was that between modality, task, and type of sequence [$F(3,87) = 3.58, p = .01, d = 0.70$]. All the remaining effects were nonsignificant ($ps > .07$ and $ds < 0.47$).

T2 performance. A $2 \times 2 \times 3 \times 4$ repeated measures ANOVA was performed on T2 accuracy, with modality, task, SOA, and type of sequence as within-subjects factors. All the main effects were significant [modality, $F(1,29) = 112.81, p < .001, d = 3.95$; task, $F(1,29) = 104.83, p < .001, d = 3.80$; SOA, $F(2,58) = 31.21, p < .001, d = 2.08$; and type of sequence, $F(3,87) = 125.68, p < .001, d = 4.16$]. The latter effect meant that T2 was reported less accurately when it was followed by a mask. The interaction of modality and type of sequence was significant [$F(3,87) =$

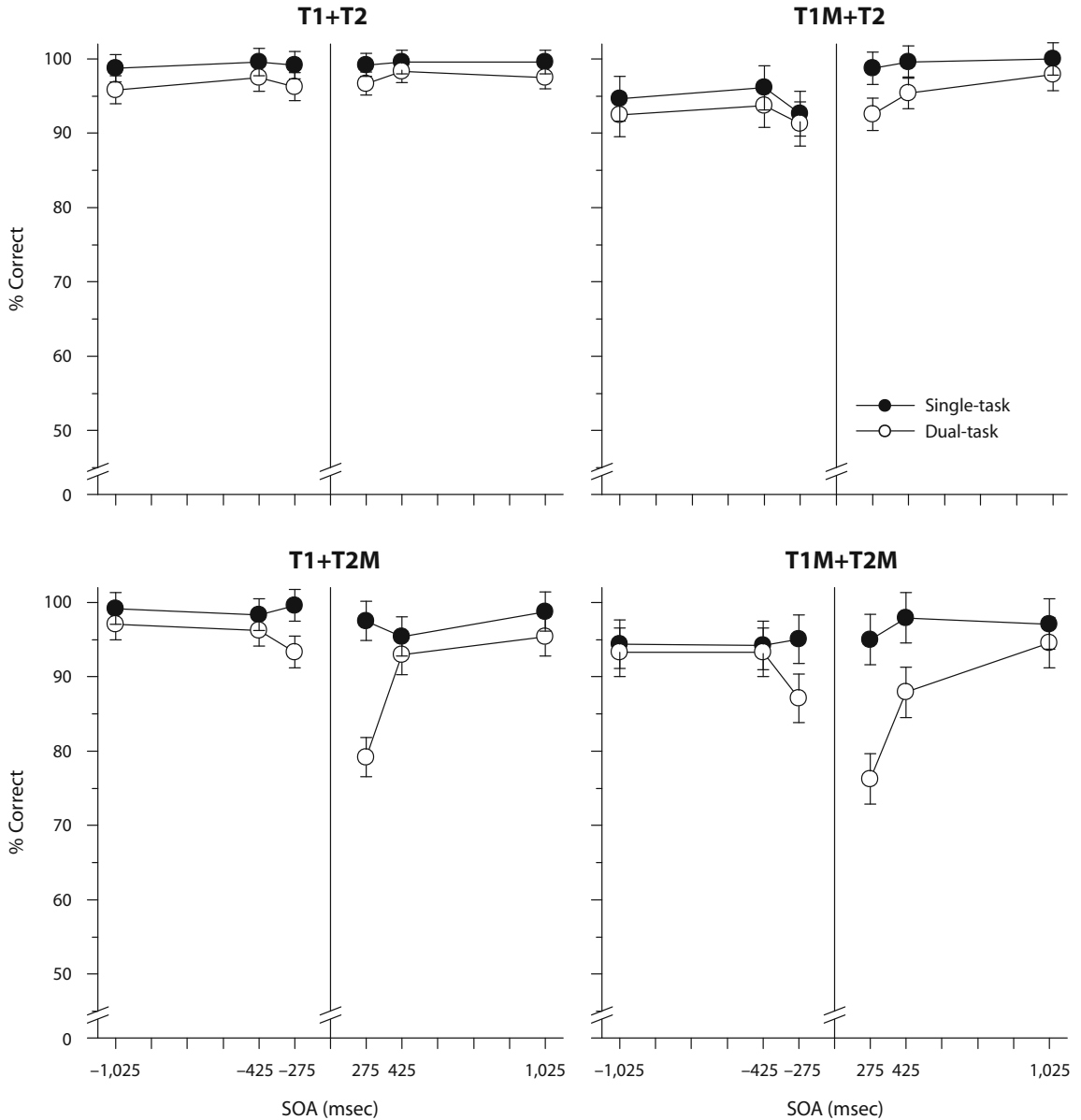


Figure 3. Results from the visual condition in Experiment 1: Mean percentages of correct target responses as a function of task (single vs. dual), T1–T2 stimulus onset asynchrony (SOA), and type of sequence (T1M+T2M, T1+T2M, T1M+T2, and T1+T2). Positive SOAs refer to T2, and negative SOAs refer to T1. Error bars are 95% within-subjects confidence intervals.

42.94, $p < .001$, $d = 2.43$], indicating that the disruptive effect of the T2 mask was larger in the auditory condition. The significant task \times type of sequence interaction [$F(3,87) = 25.89$, $p < .001$, $d = 1.89$] revealed lower T2 accuracy with dual tasks when T2 was masked. The analysis showed reliable interactions between task and modality [$F(1,29) = 11.36$, $p = .001$, $d = 1.25$]; between SOA and type of sequence [$F(6,174) = 8.61$, $p < .001$, $d = 1.09$]; and between modality, task, and type of sequence [$F(3,87) = 4.01$, $p < .01$, $d = 0.74$]. However, there was no significant interaction

between modality and SOA [$F(2,58) < 1$, $d = 0.36$] or between modality, SOA, and type of sequence [$F(6,174) < 1$, $d = 0.36$]. The interaction of task and SOA, which is the empirical signature of the AB, was significant [$F(2,58) = 15.94$, $p < .001$, $d = 1.48$]. Of particular importance for the purpose of the present experiment is the significant interaction between task, SOA, and type of sequence [$F(6,174) = 5.00$, $p < .001$, $d = 0.83$]. This reliable three-way interaction indicates that the relation between task and SOA differs according to the type of sequence. In the same way, the

significant modality \times task \times SOA interaction [$F(2,58) = 3.04, p = .03, d = 0.65$] points to variations in the expression of the visual and the auditory AB. However, the four-way interaction did not reach significance [$F(6,174) = 1.37, p = .12, d = 0.44$].

Since AB deficits appear to vary according to the type of sequence, the relation between task and SOA was analyzed for each of the four types of sequence separately for each modality. A stricter alpha level of .01 was used in order to compensate for the increase in family-wise error rate. In the auditory condition, the interaction of task and SOA was significant at T1M+T2M [$F(2,58) = 4.00, p = .01, d = 0.74$] and at T1+T2M ($F = 3.68, p = .01, d = 0.71$) but was far from significant at T1M+T2 and T1+T2 ($F_s < 1, d_s < 0.19$). These results reinforce the impression given by a visual inspection of Figure 2 that auditory AB-like effects were observed when T2 was masked. In the visual condition, analysis showed a significant task \times SOA interaction at T1M+T2M ($F = 12.02, p < .001, d = 1.29$) and at T1+T2M ($F = 25.87, p < .001, d = 1.89$). However, this interaction was not significant at T1M+T2 ($F = 2.19, p = .07, d = 0.55$), or at T1+T2 ($F < 1, d = 0.22$). It appears that, as is suggested by a visual inspection of Figure 3, there were reliable visual AB-like deficits when a mask followed T2.

Dual-task cost. To evaluate the contribution of T1 masking on the magnitude of the AB in both the auditory and the visual conditions, the dual-task cost—the difference in T2 accuracy between single and dual tasks—was contrasted between conditions in which an AB was obtained—that is, at T1+T2M (T1 mask absent) and T1M+T2M (T1 mask present), with respect to SOA. These data were submitted to a 2 (T1 mask) \times 3 (SOA) repeated measures ANOVA separately for each modality. In the auditory condition, the main effect of T1 mask did not reach significance [$F(1,29) = 1.13, p = .15, d = 0.40$], and the T1 mask \times SOA interaction was far from significant [$F(2,58) < 1, d = 0.33$]. Conversely, the main effect of T1 mask [$F(1,29) = 2.87, p = .05, d = 0.63$] and the T1 mask \times SOA interaction [$F(2,58) = 2.98, p = .03, d = 0.64$] were significant in the visual condition. When the dual-task cost was compared between T1+T2M and T1M+T2M at each SOA, a significantly larger dual-task cost for T1M+T2M was obtained at 425 msec [$t(29) = 3.28, p < .01, d = 1.22$], revealing a protracted visual AB effect when T1 was followed by a mask.

Discussion

As has been demonstrated before in vision (e.g., Giesbrecht & Di Lollo, 1998), as well as in audition (e.g., Vachon & Tremblay, 2005), the present experiment showed that the presence of a mask after T2 is a prerequisite for revealing the AB phenomenon. Moreover, the presence of visual and auditory time-locked deficits in the T1+T2M condition provides evidence that the sole presence of a T2 mask with no other distractor is sufficient to produce reliable visual (see Moore et al., 1996) and auditory AB effects. The similarities between the impact of visual and auditory T2 masking suggest that the AB is caused by attentional processes that are independent of modality.

However, the present results also reveal discrepancies between the visual and the auditory AB: T1 masking seemed to modulate T2 deficits only in the visual condition. Masking of T1, although not essential, produces larger and protracted visual AB effects (see Visser, 2007). Although a similar trend was observed in the auditory condition (see Figure 2), the presence of a mask after T1 did not significantly modify the expression of the auditory AB. This failure to find a modulating effect of T1 masking on the AB in the auditory condition cannot be attributable to the inefficient perceptual interference provided by the mask or the relative insensitivity of auditory items to masking (see, e.g., Chun & Potter, 2001). In fact, the presence of a mask after either T1 or T2 in the auditory condition produced a major drop in overall performance with both single and dual tasks. Such important masking effects may have contributed to the failure to show an effect of masking T1 on the auditory AB. Indeed, T2 accuracy for T1+T2M and T1M+T2M was near chance level at short SOAs, which might partly have prevented seeing any subtle changes in T2 performance caused by the presentation of T1 mask.

With regard to overall T1 and T2 performance, the interference produced by the mask appeared to be much more detrimental in the auditory than in the visual condition. The disparity highlighted in the present experiment between audition and vision in relation to the influence of masking may ensue from a difference in the temporal nature of the auditory and visual stimuli, which could have modulated the interference provided by the mask. The information required to discriminate each pair of target syllables used in Experiment 1 (i.e., *nab-nap* and *cod-cot*) was located at the end of the stimuli. In the visual condition, this crucial information was displayed for the whole stimulus duration, since the three letters forming the target syllables were presented simultaneously on the screen. In the auditory condition, however, the discriminating information was available solely at the end of the stimulation, given the successive presentation of phonemes in audition. Since the offset of the target coincided with the onset of the mask (i.e., no target-mask interval), the time the discriminating information was accessible within the sensory buffer tended to be much shorter in the auditory than in the visual condition. Besides, some studies have shown that listeners generally require more information and more time to identify a syllable when the discriminating information is located at the end, rather than at the beginning, of the syllable (e.g., Salasoo & Pisoni, 1985; Wingfield, Goodglass, & Lindfield, 1997).

The effects of masking in the auditory condition can be accounted for by considering the processes implicated in interference by masking—that is, interruption and integration processes (Massaro, 1975; see also Bazana & Stelmack, 2002). The successive presentation of the mask and the target yielded the premature interruption of the processing of the latter. When applied on T2, this form of masking led to the appearance of the AB. However, the very close temporal proximity of the mask and the target syllable in the auditory condition may also have promoted the integration of the last part of the target, which allows

discrimination, with the onset of the mask (e.g., Massaro, 1973, 1975; see also Vachon & Tremblay, 2006). The addition of sensory noise to the target-relevant information by integration masking may then have strongly reduced discriminability (e.g., Kallman & Massaro, 1979; Massaro, 1975; see also Mattys, 1997). This hypothesis can account for the notable impairment of the overall performance for masked targets in the auditory modality. The next experiment was designed to verify whether the results from Experiment 1 are attributable to the fact that target syllables were discriminated by their offset. If shifting the relevant information away from the offset produced analogous AB effects across modalities, it would strongly suggest that perceptual affordances are responsible, at least to some extent, for setting up attentional effects.

EXPERIMENT 2

The paradigm and design employed in Experiment 2 were identical to those used in Experiment 1. However, target syllables were discriminated henceforth by a variation located at their onset, rather than at their offset. A number of studies showed that discrimination between two consonants is easier when it takes place at the onset of a consonant–vowel–consonant stimulus (e.g., Content, Kearns, & Frauenfelder, 2001; Sidwell & Summerfield, 1986). In Experiment 2, the participants performed the same discrimination task on the two targets; to do so, they had to determine whether the target syllable began with the letter *m* or the letter *n*. As in Experiment 1, two target syllables were employed: one syllable ending in *'ab* (either *mab* or *nab*) and another ending in *'ot* (either *mot* or *not*). Shifting the discriminating information from the offset to the onset of the syllables should reduce the low-level interference provided by the masks in the auditory modality and, consequently, should facilitate identification of auditory targets. Hence, the experimental setting used in Experiment 2 should promote the appearance of a T1 masking effect on the auditory AB. Given that the letters forming the visual target syllables appeared simultaneously on the screen, this manipulation was not expected to influence the pattern of results in the visual condition.

Method

The method was identical to that employed in Experiment 1, except as noted below.

Participants. Twenty-nine French-speaking adults who reported having normal or corrected-to-normal vision and normal hearing received a monetary compensation for their participation. None of these participants had taken part in Experiment 1.

Apparatus and Stimuli. E-Prime 1.1 was used to run both the auditory condition and the visual condition. A new set of target syllables was employed: One target stimulus was an *'ab* syllable (*mab* or *nab*), and the other was an *'ot* syllable (*mot* or *not*). The task performed on each target item was identical, given that the participants had always to discriminate between a syllable beginning with the letter *m* and another beginning with the letter *n*.

Results

As in Experiment 1, data in the single-task conditions were pooled over target syllables (*'ab* and *'ot* as single tar-

gets) for both T1 and T2. The mean accuracy for T1 (negative SOAs) and T2 (positive SOAs) is presented in Figure 4 for the auditory condition and in Figure 5 for the visual condition as a function of task, SOA, and type of sequence.

T1 performance. A $2 \times 2 \times 3 \times 4$ repeated measures ANOVA was carried out on mean percentages of correct discrimination of T1, with modality, task, SOA, and type of sequence as within-subjects factors. The analysis showed significant main effects of modality [$F(1,28) = 10.04, p < .01, d = 1.20$], task [$F(1,28) = 4.11, p = .03, d = 0.77$], and SOA [$F(2,56) = 6.70, p = .001, d = 0.98$]. The significant main effect of type of sequence [$F(3,84) = 49.27, p < .001, d = 2.65$] points to a lower T1 accuracy when T1 was masked. This time, the magnitude of the negative impact of masking T1 was similar across modalities [modality \times type of sequence, $F(3,84) < 1, d = 0.19$]. There was a significant interaction between SOA and type of sequence [$F(6,168) = 3.34, p < .01, d = 0.69$], as well as between modality, SOA, and type of sequence [$F(6,168) = 2.33, p = .04, d = 0.58$]. The latter interaction seems attributable to an effect of SOA on T1 accuracy restricted to the auditory T1M+T2M condition. All the remaining effects were not significant ($ps > .14$ and $ds < 0.43$).

T2 performance. T2 accuracy was submitted to a 2 (modality) \times 2 (task) \times 3 (SOA) \times 4 (type of sequence) repeated measures ANOVA. The main effects of modality [$F(1,28) = 8.91, p < .01, d = 1.13$], task [$F(1,28) = 192.17, p < .001, d = 5.24$], and SOA [$F(2,56) = 63.81, p < .001, d = 3.02$] were significant. The significant main effect of type of sequence [$F(3,84) = 76.01, p < .001, d = 3.30$] indicates that T2 performance was impaired when T2 was followed by a mask. This detrimental effect of T2 masking was larger in the visual condition [modality \times type of sequence, $F(3,84) = 11.83, p < .001, d = 1.30$] and with a dual task [task \times type of sequence, $F(3,84) = 32.97, p < .001, d = 2.17$]. Except for the interaction between modality and SOA [$F(2,56) = 1.65, p = .10, d = 0.49$], all other interactions reached significance [modality \times task, $F(1,28) = 5.31, p = .02, d = 0.87$; SOA \times type of sequence, $F(6,168) = 5.67, p < .001, d = 0.90$; modality \times task \times type of sequence, $F(3,84) = 4.98, p < .01, d = 0.84$; and modality \times SOA \times type of sequence, $F(6,168) = 2.29, p = .02, d = 0.57$]. The interaction between task and SOA, which indicates the presence of an AB effect, was significant [$F(2,56) = 32.16, p < .001, d = 2.14$]. The critical relation between task and SOA seems to be influenced by both modality and type of sequence, as is suggested by the significant interactions between modality, task, and SOA [$F(2,56) = 2.86, p = .04, d = 0.64$]; between task, SOA, and type of sequence [$F(6,168) = 2.23, p = .02, d = 0.56$]; and between the four factors [$F(6,168) = 1.91, p = .04, d = 0.52$].

To test the combined impact of modality and type of sequence on the AB, the interaction between task and SOA was analyzed for each of the four types of sequences separately for each modality, using an alpha level of .01. In the auditory condition, the task \times SOA interaction was significant at T1M+T2M [$F(2,56) = 5.30, p < .01, d = 0.87$], at T1+T2M ($F = 8.83, p < .001, d = 1.12$), and even at T1M+T2 ($F = 5.37, p < .01, d = 0.88$), but not

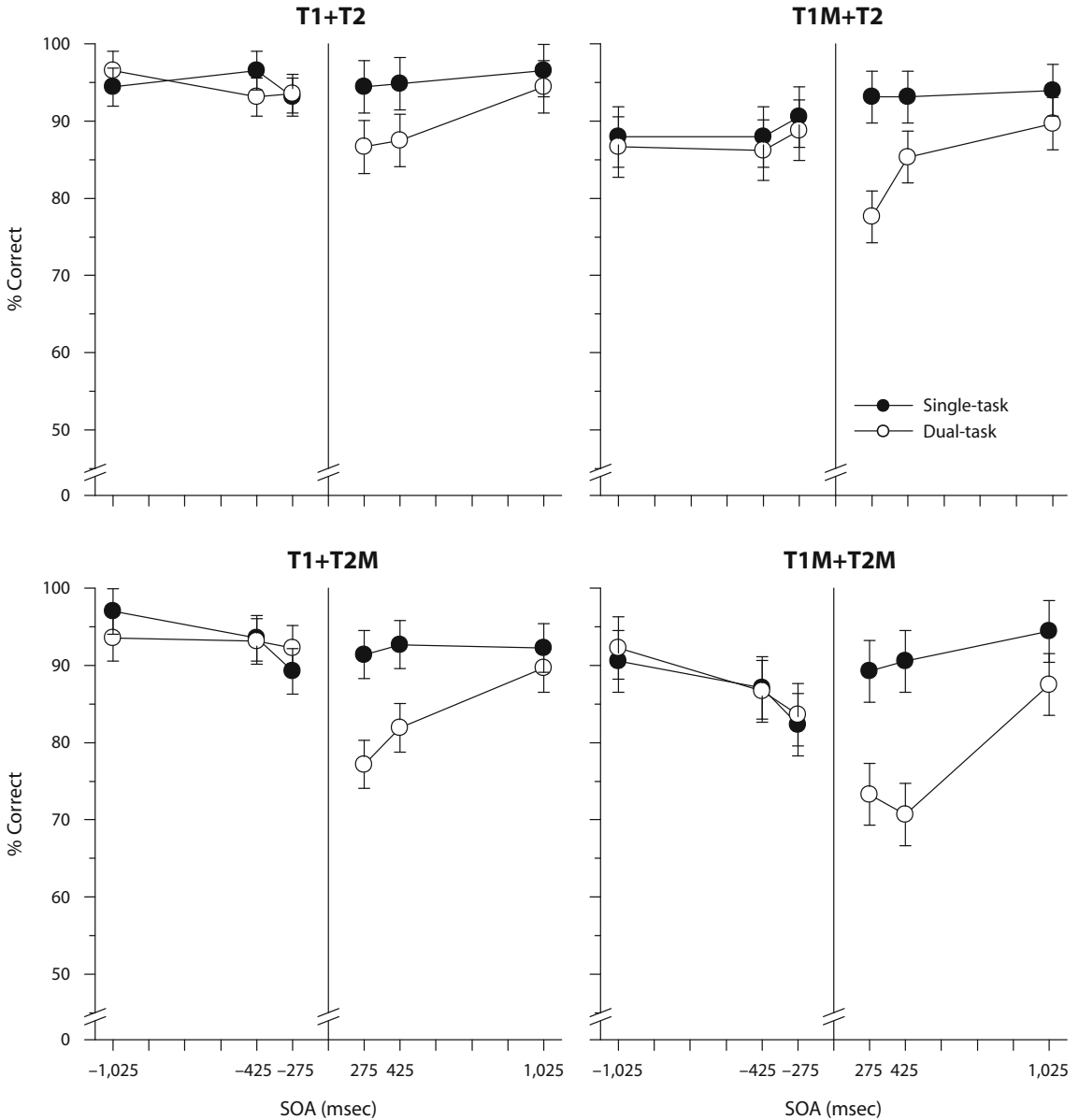


Figure 4. Results from the auditory condition in Experiment 2: Mean percentages of correct target responses as a function of task (single vs. dual), T1–T2 stimulus onset asynchrony (SOA), and type of sequence (T1M+T2M, T1+T2M, T1M+T2, and T1+T2). Positive SOAs refer to T2, and negative SOAs refer to T1. Error bars represent 95% within-subjects confidence intervals.

at T1+T2 ($F = 1.31, p = .14, d = 0.43$). These results reinforce the impression given by a visual inspection of Figure 4 that an AB effect was observed in every auditory condition except when both targets were unmasked. In the visual condition, analyses revealed that the task \times SOA interaction was significant at T1M+T2M ($F = 10.91, p < .001, d = 1.25$), approached significance at T1+T2M ($F = 3.54, p = .018, d = 0.71$), but was far from significant at T1M+T2 and T1+T2 ($F_s < 1$ and $d_s < 0.26$). These results are in agreement with a visual

inspection of Figure 5, showing AB effects only when T2 was masked.

Dual-task cost. In order to examine the impact of T1 masking on the AB in both the visual and the auditory conditions, the dual-task cost was compared with respect to SOA between an AB-inducing condition in which the T1 mask was absent (T1+T2M) and another in which the T1 mask was present (T1M+T2M). To do so, a 2×3 repeated measures ANOVA with T1 mask and SOA as factors was carried out on the dual-task cost separately for

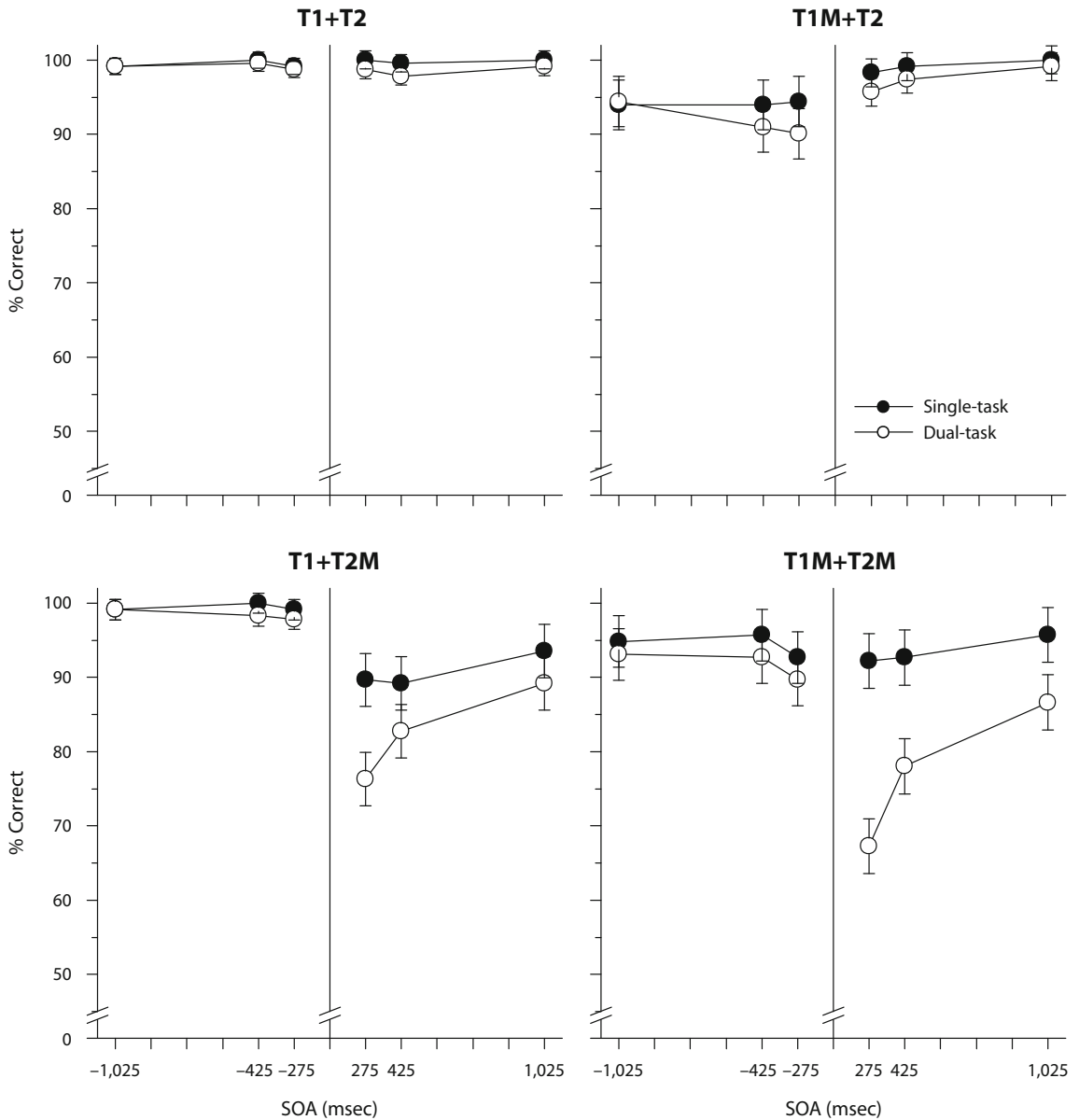


Figure 5. Results from the visual condition in Experiment 2: Mean percentages of correct target responses as a function of task (single vs. dual), T1–T2 stimulus onset asynchrony (SOA), and type of sequence (T1M+T2M, T1+T2M, T1M+T2, and T1+T2). Positive SOAs refer to T2, and negative SOAs refer to T1. Error bars are 95% within-subjects confidence intervals.

each modality. In the auditory condition, the critical effect of T1 mask was significant [$F(1,28) = 10.68, p < .01, d = 1.24$], but the two-way interaction was not significant [$F(2,56) < 1, d = 0.36$]. When contrasted at each SOA, the dual-task cost was significantly larger when T1 was masked at 425 msec [$t(28) = 2.11, p = .02, d = 0.55$]. In the visual condition, the ANOVA also showed a significant effect of T1 mask [$F(1,28) = 28.79, p < .001, d = 2.03$], with no significant two-way interaction [$F(2,56) < 1, d = 0.35$]. Analyses performed at each SOA revealed a

significantly larger dual-task cost in the presence of a T1 mask at 275 msec [$t(28) = 3.55, p < .001, d = 1.34$] and 425 msec [$t(28) = 2.26, p = .02, d = 0.85$].

Discussion

As was expected, moving the discriminating information from the offset to the onset of the target syllables did not affect the pattern of results obtained in the visual condition of the previous experiment: The visual AB emerged exclusively when T2 was masked and was enlarged by the

presentation of a masking item following T1. In the auditory condition, introducing a mask after T2 also guaranteed that the AB would take place, as in Experiment 1. The transformation of the target syllables induced some changes in the results observed in the auditory condition. First, backward masking did not lead to an important overall impairment in target identification, which diminishes the likelihood of observing floor effects, and performance was much less variable (see Figure 4). In fact, the detrimental effect of target masking on overall performance was similar between the auditory and the visual conditions. Such results were possibly the consequence of the reduction of the low-level interference provided by auditory integration masking and also the relative ease of discriminating an auditory syllable by its onset, rather than by its offset (e.g., Content et al., 2001; Sidwell & Summerfield, 1986; Wingfield et al., 1997).

Experiment 2 also revealed a different pattern of results concerning the impact of masking T1 on the auditory AB. The magnitude of the T2 deficit in the auditory condition was larger when T1 was followed by a mask (i.e., for T1M+T2M) than when the target was presented alone (i.e., for T1+T2M). This result is central because it demonstrated for the first time that the influence of masking T1 on the expression of the AB is not restricted to the visual modality but extends also to the auditory modality, providing new evidence of functional similarities between the visual and the auditory AB. Therefore, it seems that in the auditory as well as in the visual domain, T1 processing can be impaired by the subsequent presentation of a mask, which lengthens the postponement of T2 transfer into STM (e.g., Chun & Potter, 1995; Jolicœur, 1998). The finding that manipulating the presence of the T1 mask can affect the expression of the auditory AB when the two targets were presented with no other stimuli than their respective masks strongly suggests that the absence of such an effect in Mondor's studies (Mondor, 1998; Shen & Mondor, 2006) was due to the overriding influence of the perceptual organization of the whole sequence of distractors over that of the T1 mask.

The impact of the auditory masking of T1 within the present experiment was not restricted to conditions in which T2 was masked: A significant auditory AB effect was found in the T1M+T2 condition—that is, when only T1 was followed by a mask. Although the size of the AB deficit was much smaller than that observed in the presence of a T2 mask, such a result may be taken to go against the conclusion common to Experiment 1 and the study of Vachon and Tremblay (2005), according to which T2 must be followed by a nontarget item for the auditory AB to occur. The fact remains, nonetheless, that even in the absence of a T2 mask, T2 ended with “nontarget” information, since the discriminating information of T2 was located at the onset of the target item. Hence, the auditory AB obtained in the T1M+T2 condition could have been the consequence of backward interference exerted by the offset of T2 on its onset. The occurrence of such a time-locked deficit in the absence of an item after T2 was restricted to the condition in which T1 was masked. There is evidence suggesting that onset processing is particularly

sensitive to attentional manipulations (e.g., Content et al., 2001; Mattys, 1997). Given that masking the T1 is said to increase the attentional demand required to process T1 effectively (e.g., Chun & Potter, 1995; Seiffert & Di Lollo, 1997), it could be argued that the onset phoneme is particularly vulnerable to the interference from the remainder of the stimulus when little or no attentional resources are available. The presence of a larger AB deficit when T2 was masked (i.e., in T1M+T2M) suggests that the deterioration of T2 onset would increase when a distinct auditory event (i.e., a mask) follows the target within the sequence (see also Vachon & Tremblay, 2006).

GENERAL DISCUSSION

In the present study, the contribution of target masking to the AB phenomenon was compared between vision and audition, using analogous verbal stimuli. The role of masking was investigated by systematically manipulating the presence of a trailing item after T1 and T2. In order to isolate the impact of the masks from the interference ensuing from the perceptual organization of a whole stream of distractors, the stimulus sequences employed here consisted of presenting the two targets and their respective masks with no other distractor. Two experiments replicated a key aspect of the AB phenomenon—namely, that the occurrence of a masking item after T2 guarantees the appearance of an AB effect in both vision (e.g., Giesbrecht & Di Lollo, 1998) and audition (e.g., Vachon & Tremblay, 2005). Moreover, the present study provides evidence that masking T1 modulates the expression of the AB not only in the visual domain (e.g., Chun & Potter, 1995), but also in the auditory domain. However, this effect of T1 masking on the auditory AB was noticeable when target syllables varied at their onset (Experiment 2), but not when they were discriminated by their offset (Experiment 1). Although the failure to observe the impact of T1 masking in the auditory condition in Experiment 1 may have ensued, at least in part, from floor effects, the fact remains that, taken together, the results from the auditory conditions in the two experiments highlighted the interaction between masking interference and the temporal arrangement of relevant information in audition.

Similarities and Discrepancies Between the Visual and the Auditory AB

By comparing visual and auditory AB effects via a within-subjects procedure similar in both modalities, the present study clearly established the existence of functional similarities between vision and audition with regard to the role of masking interference in modulating the limitations of the deployment of attention in time. Indeed, the necessity of masking T2 to observe visual and auditory AB effects indicates that for both vision and audition, the role of T2 masking consists of deteriorating the perceptual representation of the target awaiting consolidation, which thus prevents its conscious report (e.g., Giesbrecht & Di Lollo, 1998; Jolicœur, 1999). Furthermore, by establishing masking of T1 as a way of influencing the expression of the auditory AB, our study suggests that, as in the

visual modality (e.g., Seiffert & Di Lollo, 1997), auditory masking may slow down the processing of an auditory target, so that it yields a longer delay in the processing of a subsequent target stimulus. The present findings provide further evidence that, in contrast with previous claims (see, e.g., Chun & Potter, 2001), auditory information is as sensitive as its visual counterpart to the interference from subsequent stimulations (Shen & Mondor, 2006; Vachon & Tremblay, 2005, 2006).

The comparable role played by target masking in the visual and the auditory AB highlighted in the present study suggests that the temporal deployment of attention is controlled by mechanisms common to every sensory modality that are probably central in nature (e.g., Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Jolicœur & Dell'Acqua, 1998).² There are also a number of studies in the AB literature that provide evidence of similarities between visual AB and auditory AB phenomena. For instance, heterogeneous distractors tend to produce a greater AB than do homogeneous distractors, whether in visual (Ross & Jolicœur, 1996, reported in Arnell & Jolicœur, 1999) or in auditory (Tremblay, Vachon, & Jones, 2005) sequences. Moreover, Arnell and Jenkins extended to the auditory domain the modulating effect that similarity between T2 and its surrounding distractors exerted on the visual AB (e.g., Maki, Bussard, Lopez, & Digby, 2003). This accumulating evidence of functional equivalence between the visual and the auditory AB leads us to conclude that the phenomenon reflects a fundamental limit of human cognition. Such a conclusion is in line with the central interference theory (e.g., Arnell & Jolicœur, 1999; Jolicœur, 1998, 1999; Jolicœur & Dell'Acqua, 1998), according to which the AB originates from a central bottleneck that is amodal in nature. Given that this theory was developed within the bottleneck framework, which generally provides the best account for the effects of masking on the AB (see Kawahara, Enns, & Di Lollo, 2006), it appears to be particularly suitable for the present findings.

The present study revealed not only similarities but also discrepancies between the visual and the auditory AB. Probably the most striking divergence between the two phenomena was related to the sensitivity to the temporal arrangement of relevant information. Indeed, the results from the two present experiments showed that whether the discriminating information is located at the beginning or at the end of a target syllable can modulate the expression of the AB in audition, whereas it has no particular effect in vision. This difference between the two modalities appears to ensue from a fundamental distinction in perceptual processing between vision and audition. According to several authors (e.g., Cusack & Carlyon, 2003; Kubovy & Van Valkenburg, 2001; Mondor & Terrio, 1998; Näätänen & Winkler, 1999), the formation of visual objects is based mainly on the integration of features in space (see also Treisman & Gelade, 1980), whereas in audition, the mediums of object formation are time and frequency. So, a change in the temporal nature of stimuli should affect only processes involved in the formation of auditory representations. The present findings indicate that the way perceptual objects are formed may modulate

temporal processing limitations (e.g., Kellie & Shapiro, 2004; Raymond, 2003).

At first blush, the influence of T1 masking on the AB deficit appears to vary according to the modality of presentation: Whereas such an effect is robust in the visual domain (see Visser, 2007), it is not systematically observed in the auditory domain. In effect, the impact of masking T1 on the auditory AB seems rather sensitive to the experimental context, since the removal of the T1 mask had no effect when spoken syllables were discriminated by their offset (Experiment 1 of the present study) or when targets were embedded in a sequence of distractor tones (Mondor, 1998; Shen & Mondor, 2006). Although floor effects may have obscured the potential effects of T1 masking in Experiment 1, such a factor cannot account for Mondor's results.

One way to reconcile Mondor's (1998; Shen & Mondor, 2006) findings with the results in the visual AB literature is to refer to the special role played by perceptual organization in rapid auditory sequences. In both vision and audition, a perceptual process acts preattentively to integrate and group incoming stimuli on the basis of their environmental origin. Besides, the dynamics of perceptual grouping of visual and auditory information are based on the same Gestalt principles of similarity and proximity (see Aksentijević, Elliott, & Barber, 2001, for a discussion). Despite these similarities in perceptual organization between the two domains, information grouping operates mainly from the spatial dimension in vision and from the temporal and the spectral dimensions in audition (e.g., Aksentijević et al., 2001; Bregman, 1990; Koffka, 1935). This divergence across modalities is particularly consequential when one considers the paradigm to be employed to study the AB. Indeed, within a visual sequence, sensory information typically varies in time but not in space, which reduces the likelihood that visual stimuli will be grouped. Although the temporal dimension can contribute to perceptual organization in vision (e.g., Bregman & Achim, 1973), it plays a more important role in audition because of the sequential nature inherent in an auditory stimulation. Within the experimental context of the auditory AB, stimuli change *rapidly* not only in time, but also in frequency, promoting the action of perceptual organization processes (e.g., Beauvois, 1998; Bregman, Ahad, Crum, & O'Reilly, 2000; Macken, Tremblay, Houghton, Nicholls, & Jones, 2003). Furthermore, there is evidence that perception of an auditory stimulus within a sequence depends to a great extent on the context in which it is embedded (e.g., Brochard, Drake, Botte, & McAdams, 1999; Johnston & Jones, 2006; Nicholls & Jones, 2002), illustrating the superiority of the sequence over the item in audition. Therefore, given the higher propensity of auditory stimuli, as compared with visual stimuli, to be perceptually grouped in rapid sequences, it is possible that in the auditory sequences used by Mondor (1998; Shen & Mondor, 2006), distractors, including the T1 mask, were designated members of the same perceptual stream by virtue of their similarity—they were all pure tones—and were then inhibited as a group, rendering the interference provided by T1 mask ineffective in

affecting the auditory AB. When the influence of perceptual organization is minimized—for example, by removing all distractors (except the masks) from the sequence, as we did in the present research—the interference coming from the T1 mask can have an effect on the magnitude of the auditory AB (see Experiment 2).

The hypothesis according to which central and amodal attentional limitations are responsible for the AB does not seem to chime harmoniously with the differences highlighted in the present study between the visual and the auditory AB. Other variations in the manifestation of visual and auditory AB patterns have been reported in the literature. For example, the auditory AB rarely appeared U-shaped (Soto-Faraco & Spence, 2002; Tremblay et al., 2005), as opposed to the visual AB (see Visser, Bischof, & Di Lollo, 1999). One way to account at once for the similarities and discrepancies existing between the visual and the auditory AB is to posit that the phenomena ensue from mechanisms specific to each modality (e.g., Hein, Parr, & Duncan, 2006; Soto-Faraco & Spence, 2002), but working in a similar fashion. However, the idea that distinct modules are performing the same operations in the same way is not the most parsimonious. Moreover, this approach has no provision that would account for the existence of cross-modal AB effects—that is, T2 deficits occurring when the two targets are presented in different modalities (e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Soto-Faraco et al., 2002). Nevertheless, the notion that constraints specific to each sensory system are responsible for the AB is not necessarily incompatible with the concept of central bottleneck processing limitations.

Multiple Sources of Interference in Human Information Processing

One promising way to consider the AB as a phenomenon sensitive to amodal, as well as to modality-specific, factors is to apply the multiple-sources-of-dual-task-cost approach put forward by Arnell and Duncan (2002; see also Arnell, 2001; Arnell & Jenkins, 2004). According to these authors, temporal processing in dual tasking is limited by both central and within-modality constraints. Such an account not only provides an answer for the existence of auditory and cross-modal ABs, but also explains why AB deficits are larger and more robust within than between modalities (e.g., Arnell & Jenkins, 2004; Arnell & Larson, 2002). Indeed, it is suggested that when the two targets are presented in the same modality, two sources of interference contribute to the T2 deficit: one source that is specific to the modality of presentation and another that is central and amodal. On the other hand, cross-modal AB effects are exclusively the consequence of central limitations. A similar view was proposed by Chun and Potter (2001; Potter et al., 1998), in which the visual AB reflects limits of visual-processing capacity accentuated by a central bottleneck, whereas the origin of T2 deficits in audition is restricted to the latter source of interference. However, the results from the present study showed that factors affecting only auditory information processing can modulate the auditory AB, suggesting that the phenomenon is also sensitive to modality-specific constraints, as in vision.

In an attempt to refine the mechanisms responsible for modality-specific and amodal limitations in temporal information processing, we set the multiple-sources-of-interference approach of Arnell and Duncan (2002) within the bottleneck framework of the central interference theory (Jolicœur, 1998; Jolicœur & Dell'Acqua, 1998). According to this theory, information presented within a stimulus sequence goes through two stages of processing: (1) a first stage of formation and selection of perceptual representations and (2) a second stage during which selected representations are consolidated into STM. We suppose that within-modality constraints restrict processing at the first stage, whereas, as is assumed by the central interference theory, a central bottleneck is responsible for limitations at the second stage of processing. The theoretical approach described here is illustrated in Figure 6.

During the first stage, information from the stimulus sequence is extracted and then integrated so as to create perceptual representations that must be selected in order to be consolidated in STM (e.g., Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998). There is recent evidence suggesting that constraints in the selection of information play an important role in the AB (e.g., Nieuwenstein, 2006; Nieuwenstein, Chun, van der Lubbe, & Hooge, 2005; Olivers & Watson, 2006). As initially put forward by Potter, Staub, and O'Connor (2002), we assume that stimuli compete for attention in order to be identified and then are selected for further processing. The level of competition for selection is, at least in part, determined by the presence of perceptual interference from irrelevant stimuli, especially that ensuing from similarity with the target stimuli, and by target structural descriptions (e.g., Duncan & Humphreys, 1989; Mondor & Terrio, 1998). As a consequence, the result of the competition relies on the product of object formation and perceptual organization, which, as was mentioned earlier, appear to differ at some point according to the modality. Thus, differences between the visual and the auditory AB are likely to come up from the first stage of processing. Unlike Arnell and Duncan (2002), who suggested that modality-specific limitations arise because “to some extent at least, different perceptual resources are required for processing in visual and auditory modalities” (p. 142), we ascribe such limits instead to the selection process, which is constrained by the way representation formation and organization processes operate.

Because representations at the first stage are sensitive to perceptual interference and cannot serve as a basis for response, they must be transformed into a more durable form for a subsequent report (e.g., Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998). Selection of a potential target initiates a second stage of processing in which the representation is consolidated into STM. Short-term consolidation is assumed to require central bottleneck processing (e.g., Jolicœur & Dell'Acqua, 1998). So, during consolidation of T1, operations necessitating the participation of central mechanisms, such as short-term consolidation of T2 or response selection (e.g., Arnell & Duncan, 2002; Jolicœur, 1999), must wait. When T2 is selected before the end of T1 consolidation, its transfer in STM is temporarily suspended and its representation is de-

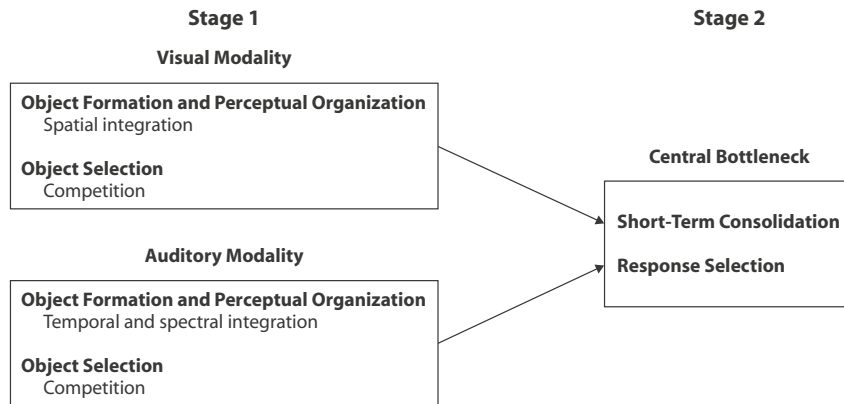


Figure 6. Illustration of the multiple-sources-of-interference approach applied to the central interference theory (e.g., Jolicoeur, 1998; Jolicoeur & Dell'Acqua, 1998). Temporal information processing is first limited by the selection process (Stage 1), which is delineated by the properties of object formation and perceptual organization that are specific to each sensory modality, and then by an amodal bottleneck for operations requiring central processing (Stage 2), such as short-term consolidation and response selection.

layed in the first stage. During this delay, if T2 terminates the sequence, its representation persists until the second stage becomes available (e.g., Jolicoeur, 1999). However, when T2 is followed by distractors, its representation in the first stage is deteriorated by the subsequent item (i.e., a T2 mask), leading to the AB deficit (Giesbrecht & Di Lollo, 1998; Vachon & Tremblay, 2005). Because it is centrally located, this bottleneck produces interference that is amodal in nature and, thus, common to all modalities. Besides, electrophysiological data have suggested that the consolidation process of T2 is delayed during the AB for both visual and auditory targets (Arnell, 2006).

CONCLUDING REMARKS

A direct comparison of the visual and the auditory AB with regard to the effects of masking showed functional similarities and discrepancies between the two phenomena. Although backward masking can be explained in a similar fashion in vision and audition (see Massaro, 1975; Massaro & Loftus, 1996), the present experiments revealed that visual and auditory masking effects can be differentially affected by the temporal structure of the stimuli. Given that the close examination of the specific processes involved in visual and auditory masking is beyond the scope of the present study, further research is needed to fully understand the differences between the mechanisms by which stimulus perceptual and temporal parameters interact to produce interference in the visual and auditory domains. One conclusion that can be drawn from the present findings is that temporal processing of visual and auditory information is restricted by similar attentional limitations but that these limits are constrained by properties specific to each sensory system. In fact, by manipulating the perceptual characteristics of targets and their sensitivity to masking, we demonstrate that auditory and visual rapid presentations can lead to identical

patterns of AB when perceptual parameters are tuned to equate affordances across modalities.

AUTHOR NOTE

This study was presented at the 12th Meeting of the Cognitive Science Association for Interdisciplinary Learning in August 2006, Hood River, Oregon. This work was supported by a grant from the Fonds québécois pour la recherche sur la nature et les technologies to S.T. F.V. received support from the Natural Sciences and Engineering Research Council of Canada in the form of a doctoral scholarship. We are grateful to Nadine Hamel and Sonia Packwood for running the experiments and to Richard Lapointe-Goupil for his technical assistance. Correspondence concerning this article should be addressed to F. Vachon, Département de psychologie, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montreal, QC, H3C 3J7 Canada (e-mail: f.vachon@umontreal.ca).

REFERENCES

- AKSENTIJEVIĆ, A., ELLIOTT, M. A., & BARBER, P. J. (2001). Dynamics of perceptual grouping: Similarities in the organization of visual and auditory groups. *Visual Cognition*, *8*, 349-358.
- 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. (2006). Visual, auditory, and cross-modality dual-task costs: Electrophysiological evidence for an amodal bottleneck on working memory consolidation. *Perception & Psychophysics*, *68*, 447-457.
- ARNELL, K. M., & DUNCAN, J. (2002). Separate and shared sources of dual-task cost in stimulus identification and response selection. *Cognitive Psychology*, *44*, 105-147.
- ARNELL, K. M., & JENKINS, R. (2004). Revisiting within-modality and cross-modality attentional blinks: Effects of target-distractor similarity. *Perception & Psychophysics*, *66*, 1147-1161.
- ARNELL, K. M., & JOLICOEUR, 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.
- BAZANA, P. G., & STELMACK, R. M. (2002). Intelligence and information processing during an auditory discrimination task with backward

- masking: An event-related potential analysis. *Journal of Personality & Social Psychology*, **83**, 998-1008.
- BEAUVOIS, M. W. (1998). The effect of tone duration on auditory stream formation. *Perception & Psychophysics*, **60**, 852-861.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREGMAN, A. S., & ACHIM, A. (1973). Visual stream segregation. *Perception & Psychophysics*, **13**, 451-454.
- BREGMAN, A. S., AHAD, P. A., CRUM, P. A. C., & O'REILLY, J. (2000). Effects of time intervals and tone durations on auditory stream segregation. *Perception & Psychophysics*, **62**, 626-636.
- BREHAUT, J. C., ENNS, J. T., & DI LOLLO, V. (1999). Visual masking plays two roles in the attentional blink. *Perception & Psychophysics*, **61**, 1436-1448.
- 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.
- BROCHARD, R., DRAKE, C., BOTTE, M.-C., & MCADAMS, S. (1999). Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1742-1759.
- 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.
- CONTENT, A., KEARNS, R. K., & FRAUENFELDER, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory & Language*, **45**, 177-199.
- CUSACK, R., & CARLYON, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 713-725.
- 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.
- DUNCAN, J., WARD, R., & SHAPIRO, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, **369**, 313-315.
- ENNS, J. T., VISSER, T. A. W., KAWAHARA, J.-I., & 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.
- GRANDISON, T. D., GHIRARDELLI, T. G., & EGETH, H. E. (1997). Beyond similarity: Masking of the target is sufficient to cause the attentional blink. *Perception & Psychophysics*, **59**, 266-274.
- HEIN, G., PARR, A., & DUNCAN, J. (2006). Within-modality and cross-modality attentional blinks in a simple discrimination task. *Perception & Psychophysics*, **68**, 54-61.
- JOHNSTON, H. M., & JONES, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 2-17.
- JOLICŒUR, P. (1998). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded Task₁ decisions. *Memory & Cognition*, **26**, 1014-1032.
- JOLICŒUR, P. (1999). Dual-task interference and visual encoding. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 596-616.
- JOLICŒUR, P., & DELL'ACQUA, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, **36**, 138-202.
- KALLMAN, H. J., & MASSARO, D. W. (1979). Similarity effects in backward recognition masking. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 110-128.
- KAWAHARA, J.-I., ENNS, J. T., & DI LOLLO, V. (2006). The attentional blink is not a unitary phenomenon. *Psychological Research*, **70**, 405-413.
- KELLIE, F. J., & SHAPIRO, K. L. (2004). Object file continuity predicts attentional blink magnitude. *Perception & Psychophysics*, **66**, 692-712.
- KOFFKA, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.
- KUBOVY, M., & VAN VALKENBURG, D. (2001). Auditory and visual objects. *Cognition*, **80**, 97-126.
- MACKEN, W. J., TREMBLAY, S., HOUGHTON, R. J., NICHOLLS, A. P., & JONES, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 43-51.
- MAKI, W. S., BUSSARD, G., LOPEZ, K., & DIGBY, B. (2003). Sources of interference in the attentional blink: Target-distractor similarity revisited. *Perception & Psychophysics*, **65**, 188-201.
- MARTENS, S., WOLTERS, G., & VAN RAAMSDONK, M. (2002). Blinks of the mind: Memory effects of attentional processes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1275-1287.
- MASSARO, D. W. (1973). A comparison of forward versus backward recognition masking. *Journal of Experimental Psychology*, **100**, 434-436.
- MASSARO, D. W. (1975). *Experimental psychology and information processing*. Chicago: Rand-McNally.
- MASSARO, D. W., & LOFTUS, G. R. (1996). Sensory and perceptual storage: Data and theory. In E. C. Carterette & M. P. Friedman (Series Eds.), E. L. Bjork & R. A. Bjork (Vol. Eds.), *Handbook of perception and cognition: Memory* (2nd ed., pp. 67-99). San Diego: Academic Press.
- MATTYS, S. L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin & Review*, **4**, 310-329.
- 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., & 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.
- MOORE, C. M., EGETH, H., BERGLAN, L. R., & LUCK, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, **3**, 360-365.
- NÄÄTÄNEN, R., & WINKLER, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**, 826-859.
- NICHOLLS, A. P., & JONES, D. M. (2002). Capturing the suffix: Cognitive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 12-28.
- NIEUWENSTEIN, M. R. (2006). Top-down controlled, delayed selection in the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 973-985.
- NIEUWENSTEIN, M. R., CHUN, M. M., VAN DER LUBBE, R. H. J., & HOOGHE, I. T. C. (2005). Delayed attentional engagement in the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 1463-1475.
- OLIVERS, C. N. L., & WATSON, D. G. (2006). Input control processes in rapid serial visual presentations: Target selection and distractor inhibition. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1083-1092.
- PASHLER, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- POTTER, M. C., CHUN, M. M., BANKS, B. S., & MUCKENHOUT, 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.
- POTTER, M. C., STAUB, A., & O'CONNOR, D. H. (2002). The time course

- of competition for attention: Attention is initially labile. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1149-1162.
- RAYMOND, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science*, **14**, 54-59.
- 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.
- ROSS, N., & JOLICŒUR, P. (1996, August). *An attentional blink for colour signals*. Poster presented at the annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Montreal.
- SALASOO, A., & PISONI, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory & Language*, **24**, 210-231.
- 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.] (ED.) (2001). *The limits of attention: Temporal constraints in human information processing*. Oxford: Oxford University Press.
- SHAPIRO, K. L., ARNELL, K. M., & RAYMOND, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, **1**, 291-296.
- SHAPIRO, K. L., CALDWELL, J., & SORENSEN, R. E. (1997). Personal names and the attentional blink: A visual "cocktail party" effect. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 504-514.
- 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.
- SHAPIRO, K. [L.], & TERRY, K. (1998). The attentional blink: The eyes have it (but so does the brain). In R. D. Wright (Ed.), *Visual attention* (pp. 306-329). Oxford: Oxford University Press.
- SHEN, D., & MONDOR, T. A. (2006). Effect of distractor sounds on the auditory attentional blink. *Perception & Psychophysics*, **68**, 228-243.
- SIDWELL, A., & SUMMERFIELD, Q. (1986). The auditory representation of symmetrical CVC syllables. *Speech Communication*, **5**, 283-297.
- SOTO-FARACO, S., & SPENCE, C. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology*, **55A**, 23-40.
- SOTO-FARACO, S., SPENCE, C., FAIRBANK, K., KINGSTONE, A., HILLSTROM, A. P., & SHAPIRO, K. [L.] (2002). A crossmodal attentional blink between vision and touch. *Psychonomic Bulletin & Review*, **9**, 731-738.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREMBLAY, S., VACHON, F., & JONES, D. M. (2005). Attentional and perceptual sources of the auditory attentional blink. *Perception & Psychophysics*, **67**, 195-208.
- TYSIACZNY, C. E., & MONDOR, T. A. (2005, November). *The effects of stimulus similarity on auditory search*. Poster presented at the 4th Annual Auditory, Perception, Cognition and Action Meeting, Toronto.
- VACHON, F., & TREMBLAY, S. (2005). Auditory attentional blink: Masking the second target is necessary, delayed masking is sufficient. *Canadian Journal of Experimental Psychology*, **59**, 279-286.
- VACHON, F., & TREMBLAY, S. (2006). Delayed masking and the auditory attentional blink: A test for retrieval competition and bottleneck models. *Experimental Psychology*, **53**, 182-190.
- VACHON, F., TREMBLAY, S., & JONES, D. M. (2007). Task-set reconfiguration suspends perceptual processing: Evidence from semantic priming during the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 330-347.
- VISSER, T. A. W. (2007). Masking T1 difficulty: Processing time and the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 285-297.
- VISSER, T. A. W., BISCHOF, W. F., & DI LOLLO, V. (1999). Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink. *Psychological Bulletin*, **125**, 458-469.
- VOGEL, E. K., & LUCK, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, **9**, 739-743.
- VOGEL, E. K., LUCK, S. J., & SHAPIRO, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1656-1674.
- WARD, R., DUNCAN, J., & SHAPIRO, K. [L.] (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics*, **59**, 593-600.
- WEICHELGARTNER, E., & SPERLING, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, **238**, 778-780.
- WINGFIELD, A., GOODGLASS, H., & LINDFIELD, K. C. (1997). Word recognition from acoustic onsets and acoustic offsets: Effects of cohort size and syllabic stress. *Applied Psycholinguistics*, **18**, 85-100.

NOTES

1. The potential timing issues of running a Visual Basic program under Windows were avoided here by using prerecorded auditory sequences. More specifically, instead of choosing each auditory stimulus individually to build a sequence on each trial, the Visual Basic program randomly selected on each trial a prerecorded sound file of the sequence to play. As a consequence, the timing of the sequences did not rely on Windows's internal clock.

2. One may argue that the resemblance of the auditory and visual AB patterns observed in the present study arose because verbal information, regardless of the modality of presentation, was converted into a verbal representation that underwent visual processing limitations (e.g., Shapiro & Terry, 1998). However, the evidence reported in many laboratories of auditory AB effects using pure tones (e.g., Arnell & Jolicoeur, 1999; Goddard & Slawinski, 1999; Shen & Mondor, 2006; Vachon & Tremblay, 2006), which have no corresponding visual representations, suggests that this was not the case (see also Arnell, 2001). Moreover, the offset/onset manipulation of discriminating information affected only auditory conditions. If auditory (as well as visual) targets had been transformed into visual verbal code, we would have expected the impact (if any) of such a manipulation of phonological properties to be analogous across modalities.

(Manuscript received October 27, 2006;
revision accepted for publication March 3, 2008.)