Sound source location modulates the irrelevant-sound effect

AXEL BUCHNER AND RAOUL BELL
Heinrich-Heine-Universität, Düsseldorf, Germany

KLAUS ROTHERMUND
Friedrich-Schiller-Universität, Jena, Germany

AND

DIRK WENTURA
Universität des Saarlandes, Saarbrücken, Germany

Participants memorized lists of visually presented digits in silence or while ignoring distractor sounds that either came from the front and thus from the direction in which participants' attention was oriented, or from behind. Distractor sounds impaired recall performance, but the largest impairment was observed when the sound source was directionally close to the frontal visual target display. The results are consistent with the assumption of cross-modal attentional links in models of attention, and they are problematic for explanations of the irrelevant-sound effect within working memory models that do not specify an explicit role of attention in the maintenance of information for immediate serial recall.

In a series of experiments on cross-modal spatial attention (see Driver & Spence, 2004, for a review), Spence, Ranson, and Driver (2000) asked participants to shadow an auditory message from behind while attending a visual stream of events in one of the frontal hemifields. In addition to the to-be-shadowed auditory message from behind, an irrelevant stream of spoken words was displayed in front of the participants. Shadowing performance was worse when the irrelevant stream of spoken words originated near the visual stream than when the irrelevant spoken words came from the opposite hemifield. This same-hemifield performance decrement was larger when participants were actively attending the stream of visual events (in order to detect and report visual targets) than when they simply had to fixate the visual stream. These data were interpreted as evidence in favor of models that posit cross-modal links in spatial attention in that apparently it was more difficult to ignore auditory distractors at visually attended locations than at some other point in space.

Spence et al. (2000) suggested that cross-modal links in spatial attention should also become apparent in other phenomena, most notably the so-called *irrelevant-sound effect*—that is, an impairment in immediate serial recall of short lists of visually presented items when irrelevant auditory stimuli are presented either during encoding or during retention (Banbury, Tremblay, Macken, & Jones, 2001). Spence et al. lamented a lack of studies investigating whether the size of the irrelevant-sound effect is

affected by the spatial proximity of the distractor sound source and the target display location "despite the fundamental importance of this issue for models of cross-modal attention" (p. 411). The situation does not seem to have changed much in the past few years. To our knowledge, a study by Spence and Driver (1999, cited in Spence et al., 2000) is still the only one that speaks to the issue just mentioned, but this study is not readily available for evaluation. In essence, irrelevant spoken digits impaired the serial recall of visually presented digits, but this irrelevant-sound effect was not modulated by the proximity of the sound-source location and the visual-target display. Spence et al. speculated that this null result may have been due to the insensitivity of the "retrospective" serial recall measure as an indicator of on-line distractor rejection. One major goal of the research reported here was to reexamine the possible spatial modulation of the irrelevant-sound effect.

Interestingly—and apart from the question of the sensitivity of the serial recall measure—it is not at all clear that attention is involved in the processes leading to the irrelevant-sound effect in the first place (Jones, Macken, & Mosdell, 1997; Tremblay & Jones, 1998). Immediate serial recall is a typical working memory task in that information has to be maintained for a short amount of time until it can be recalled. The detrimental effects of irrelevant auditory distractors are therefore to be located in working memory as well. Elliott (2002) pointed out that theories of human working memory fall into one of two categories—theories

A. Buchner, axel.buchner@uni-duesseldorf.de

that explicitly specify a role for attention in the maintenance of information and theories that do not. To explain the disruption of serial recall by auditory distractors, in the latter theories, it is assumed that irrelevant sounds have automatic access to the representational structure that is also used for the primary task of maintaining a suitable representation of the to-be-recalled words.

Both the *modular working memory model* (Baddeley, 1986, 1996; Baddeley & Logie, 1999) and the *object-oriented episodic record model* (Jones, 1993; Jones & Macken, 1993; Jones & Tremblay, 2000) fall into the latter category. An explanation of the irrelevant-sound effect within the modular working memory model begins by assuming that the preferred strategy for the immediate serial recall of short visually presented word lists is to convert the words into a phonological representational format for maintenance rehearsal in the limited-capacity articulatory loop module. Irrelevant sounds may gain automatic access to this representational structure where they compete with the target representations, thereby impairing later recall (Salamé & Baddeley, 1982). A strong assumption in this model is that the component identified with attentional function—the so-called *central executive*—is "not involved in temporary storage" (Baddeley & Logie, 1999, p. 28). Thus, the temporary maintenance of items must not be affected by manipulations of attention if this assumption is to be maintained. The situation is similar for the object-oriented episodic record model's explanation of the irrelevant-sound effect (Jones, 1993). The crucial process in immediate serial recall is that of seriation of the to-be-recalled objects, which are "linked" by a series of production rules. Temporary representations of to-be-recalled visual items and their interconnecting links are constructed by means of articulation. Distractor sounds form an auditory stream, which is parsed into objects by preattentive segmentation processes. The number of changing states in the auditory signal (roughly defined as rapid changes in frequency and amplitude) determines the number of objects that will be formed (up to a limit). Additional links are established automatically among auditory distractor objects. The process of seriation along these additional links competes with the seriation of the mutually linked visual targets. This competition may cause the loss of the link to a target, which would then be unavailable for recall. Note that in both of these models, nonacoustical distractor sound properties—and variables that affect attention in particular—must not affect the size of the irrelevant-sound effect.

This prediction has already been shown to be problematic. Elliott (2002) showed that the magnitude of the irrelevant-sound effect decreases from childhood to adulthood. Assuming that attentional control improves during this time period, Elliott suggested that this was evidence of an involvement of attention in the genesis of the irrelevant-sound effect. Neely and LeCompte (1999) showed that a strong semantic association between visual targets and auditory distractors may play a role for serial recall performance (but see Buchner, Irmen, & Erdfelder, 1996; LeCompte & Shaibe, 1997, for conflicting results). Buchner and Erdfelder (2005) showed that the frequency of auditory distractor words may affect the serial recall of visually presented words in that rare words cause more disruption than frequent words. Buchner, Rothermund, Wentura, and Mehl (2004) showed that both positively and negatively valent distractor words may cause more disruption of serial recall performance than would neutral distractors, and negative distractors may cause more disruption than would positive distractors. Buchner, Mehl, Rothermund, and Wentura (2006) were able to replicate the basic pattern of this finding using nonwords that were or were not artificially associated with negative valence. Negative distractor nonwords caused more disruption than did neutral distractors.

Such findings pose no problems for the *embedded-processes* model of working memory (Cowan, 1995, 1999) and the *feature* model (Nairne, 1990; Neath, 2000). Within the embedded-process model, "working memory" is defined by the set of cognitive processes needed to retain memory representations in a highly accessible state. The most highly activated working memory elements represent the focus of attention. In immediate recall tasks, the focus of attention comprises the to-be-recalled targets. The activation of a target's representation may be reduced if task-irrelevant stimuli attract attention. Loss of activation reduces the probability of successful recall of the target. Therefore, nonphonological distractor properties may influence the capacity of the irrelevant material to distract attention and, hence, may modulate the disruption of serial recall. The feature model (Nairne, 1990; Neath, 2000) allows the derivation of similar predictions, because it includes an attentional parameter reflecting the amount of processing resources available for the primary memorization tasks. If competing processes—such as the orienting toward and the suppression of auditory distractors—required attentional resources to a greater or lesser degree, then short-term memory may also be disrupted to a greater or lesser degree.

Thus, investigating whether possible cross-modal links in spatial attention may be reflected in the irrelevant-sound effect has implications for models of working memory—in addition to the importance of such findings for models of cross-modal attention, as was pointed out by Spence et al. (2000)—because attentional effects as such are informative about the validity of working memory models. In the present experiments, participants memorized sequences of digits while ignoring environmental sounds (Experiments 1 and 2) or spoken words (Experiment 3). The digits appeared on a display in front of the participants. The sounds either came from a location in front of the participants—that is, from the direction in which participants' attention was oriented—or from behind. If cross-modal attentional processes play a role in the maintenance of information in working memory, then serial recall performance should have been worse when distractor sounds came from the front—and thus from the same direction as the targets—than when distractor sounds and visual targets came from different directions.

As an additional manipulation in Experiments 1 and 2, sound sources could be located at the same distance from the participant as the visual target display, or further away. Because of differences between close and distant sound sources in sound level, in the relative proportion of high-

to low-frequency components, and in the relative proportion of direct to indirect sound, identifying differences in sagittal distance was straightforward (this was confirmed by pilot testing). The distance manipulation—in addition to the direction manipulation—was implemented in order to test whether a possible cross-modal attentional modulation of the irrelevant-sound effect was constrained only by the directional proximity of target and distractor or by their relative locations in space in analogy to effects that were reported for visual spatial attention (Downing & Pinker, 1985) as well.

EXPERIMENT 1

Method

Participants. Participants were 67 students (47 women) who were paid for their participation. Their ages ranged from 19 to 48 years old ($M = 26$). All participants were tested individually.

Materials. For the digit span test, which preceded the experiment proper, 60 sequences of digits were created that varied in lengths from 4 to 18 digits. Four sequences of each length were created. For each sequence, the digits were sampled randomly from the set $\{1, 2, \ldots, 9\}$, with the restriction that two adjacent digits were not to be identical.

For the experiment proper, the to-be-remembered lists consisted of digits sampled in the same manner as that during the digit span assessment. A total of 80 such lists were generated for each participant, with 16 lists in each of the five different conditions (see the Design section below). Five lists—one of each condition—served as practice lists. The items were presented at the center of a 14-in. TFT screen controlled by an Apple iMac computer via digital video interface. The numbers were written in black upright Helvetica font on a white background. Viewing distance was approximately 70 cm, although head position was not constrained. At this distance, each target digit subtended about 1.4° horizontally and 2.1° vertically.

Participants were seated at the center of a room that was completely dark (the entire interior of the room was painted in black) except for the light emitted by the TFT screen at which participants had to look. Due to the adaptation of the participants' eyes to the white light emitted by the screen, they were essentially blind to other objects in the dark environment around them.

Participants were surrounded by eight loudspeakers that were positioned as follows. Two small, black loudspeakers were placed immediately to the left and to the right of the display (but slightly behind its surface plane) and thus about 70 cm in front of the participant's head. The loudspeakers were not placed directly behind the computer screen, because there had to be an unobstructed path for sound waves from the loudspeakers to the participants' ears. In this way, the sound sources were easy to localize by the direct sound components. Two additional loudspeakers were placed at a distance of about 210 cm from the participant so that from his or her perspective, the distant loudspeakers were behind the close loudspeakers in the same line of sight. Four more loudspeakers were placed behind the participants at the same distances and eccentricities as those of the four loudspeakers in front (see Figure 1).

The digitized environmental sounds that were used in the irrelevant-sound trials consisted of animal sounds (such as those of a duck or eagle), sounds emitted by the operation of various tools (such as a hammer or a power drill), human-produced sounds (such as footsteps or a cry of pain), and various other environmental sounds that are less straightforward to classify (such as a splash of water, the tearing apart of paper, various squeaking sounds, etc.). Sounds were played sequentially—that is, only one sound was played at a time. Sound duration varied between 500 and 5,000 msec, depending on the particular kind of sound. The average (IEC 60804) sound level at the participant head position was 72 and 65 dB(A) for the close and distant loudspeakers, respectively.
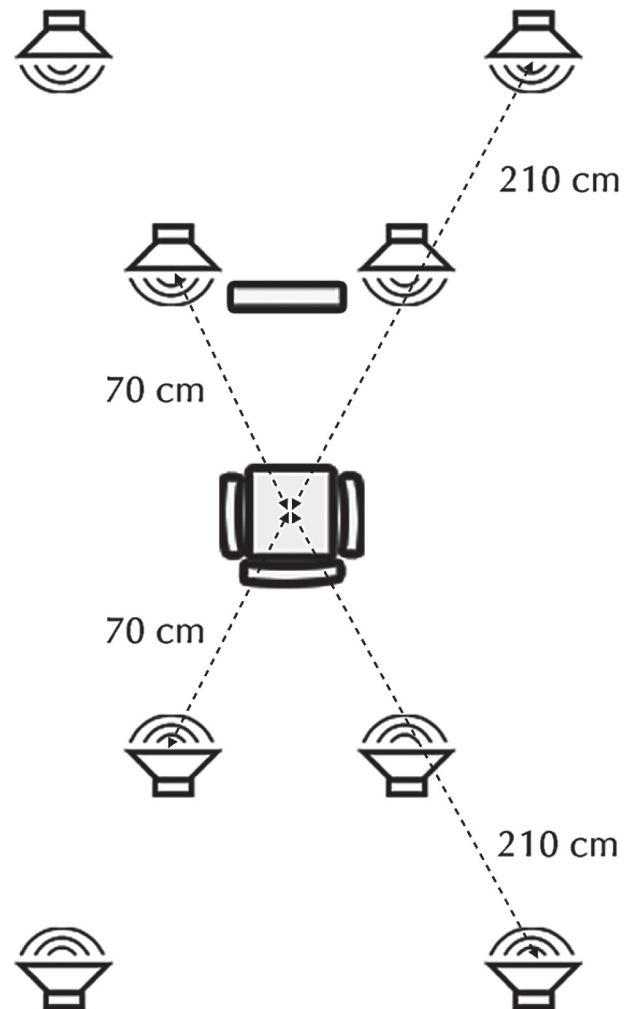
Figure 1. Schematic bird's-eye view of the experimental setup, showing the position of the loudspeakers that could be used to present the distractor sounds.

Procedure. Prior to the experiment proper, simple digit span of the participants was assessed. Except for the length of the sequences, the procedure of the digit span test was identical to that used in the silent condition of the experiment. Testing began with sequences of four digits. The length of the sequences was increased by one out of every four trials. Prior to each block of sequences of the same length, the participants were informed about the number of digits that they had to recall in the following set of four sequences. Testing ended when the participants failed to report correctly two or more of the four same-length sequences. Recall of a sequence was considered correct only when all digits were recalled at the serial position at which they were presented. Digit span was defined as the longest sequence correctly recalled in more than 50% of the trials. To anticipate, the average digit span in this experiment was 6.4 digits (from 5 to 9 digits). In the experiment proper, the length of the sequences that the participants had to recall corresponded to their digit spans.

After the digit spans were determined, participants received the instructions for the experiment proper. They were informed that sounds would be played during the presentation of many of the to-be-remembered digit sequences and that the sounds should be ignored. Participants were instructed to keep their hands near the keyboard and not to speak out loud the to-be-remembered items during the presentation phase or the retention interval. An experimenter was always present in

the experimental room—seated behind a screen placed approximately 1.5 m to the left of the participant—to make sure that the instructions were followed and to interrupt the experiment should a participant decide to quit (which did not occur). Five practice trials—one from each condition—were used to illustrate the task. The test phase consisted of 75 trials with 15 repetitions of each of the five conditions. The order of the trials during the test phase was determined randomly.

Each trial began with a visual warning signal centered on the computer screen. Following the offset of that signal, the screen went blank for 1,200 msec. Next, the sequence of to-be-recalled digits was presented. Each digit was presented for 800 msec, followed by a 400-msec blank interstimulus interval. Presenting the to-be-remembered digits took about 7 sec (from 6 to 11 sec, depending on the digit span as determined in the first phase of the experiment). The retention interval after each list was 5 sec long. Participants were to watch the computer screen during the retention interval in order not to miss a series of question marks—one for each of the serial positions of the digits—which were the signal for the participants to immediately commence recalling the digits in the order in which they had been presented. The digits were entered via the number keys of the computer keyboard. Typing the first digit replaced the first question mark with that digit; typing the second digit replaced the second question mark, and so on. Participants pressed a button labeled *don't know* (the "0" key on the number keypad) for each digit they could not recall. The participants were allowed to correct their responses. The arrow keys of the computer keyboard could be used to move the current selection to another position at which any prior entry could be replaced. After replacing all of the question marks by numbers or *don't know* responses, the participants were asked to initiate the next trial by pressing the spacebar. If the spacebar was pressed before all question marks were replaced, then a 1,500-msec visual warning was shown. No time pressure was imposed.

In trials with distractor sounds, the participants heard random sequences of environmental sounds during both the list presentation and the retention interval, but not during the recall phase. Depending on the condition, the sequence of sounds was played by a loudspeaker that was either close or far away and either in front of or behind the participant. It was randomly determined for each trial whether the sounds were played by the left or the right loudspeaker in each of these four categories.

After a block of five trials, the participants received a summary feedback about the number of items correctly recalled and were encouraged to rest briefly. On average, the experiment lasted about 50 min, after which participants were offered an explanation of the purpose of the experiment.

Design. The basic within-subjects independent variable was distractor type (silence vs. four different types of distractors). Within the distractor conditions, a 2×2 subdesign was formed by the independent variables sound direction (from the front vs. from behind) and distance (close [70 cm] vs. far [210 cm]). The comparisons within the 2×2 subdesign are most important in the present context. In order to detect "medium" effects ($f = .25$) as defined by Cohen (1977), given the assumption that the average population correlation between two levels of the repeated measures variable is $\rho = .5$ (estimated from pilot data) with a probability of $1-\beta = .95$ at $\alpha = .05$, a total sample size of $N = 54$ was needed (the power calculations were conducted using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). We were able to recruit 67 participants, so the power was actually as large as $1-\beta = .98$.

In all experiments reported in this article, a multivariate approach was used for all within-subjects comparisons. In our applications, all multivariate test criteria correspond to the same (exact) F statistic, which is reported. The level of α was set to .05 for all analyses.

Results

Figure 2 illustrates the serial recall performance in all five experimental conditions. A repeated measures MANOVA showed that performance varied systematically as a func-

tion of the distractor type variable [$F(4,63) = 4.04, p < .01, \eta_p^2 = .20$]. The contrast comparing the silent condition with the distractor conditions combined was significant [$F(1,66) = 10.44, p < .01, \eta_p^2 = .14$], confirming the typical irrelevant-sound effect for the present experiment.

More interestingly, a 2×2 MANOVA with sound direction and distance as independent variables showed a significant effect of sound direction [$F(1,66) = 8.01, p < .01, \eta_p^2 = .11$], but neither the main effect of distance [$F(1,66) = 0.40, p = .53, \eta_p^2 < .01$] nor the interaction between the two variables was significant [$F(1,66) = 0.27, p = .61, \eta_p^2 < .01$].

Discussion

Serial recall varied systematically as a function of whether distractor-sound and visual-attention direction concurred or not. Serial recall was worse when distractor sounds came from a source in front of the participant and thus from the direction in which visual attention was oriented. This pattern is consistent with the assumption that ignoring distractors is easier when the region from which they are heard is different from the region of visual attention.

The irrelevant-sound effect observed in Experiment 1 was relatively small when compared with what is typically observed in these kinds of experiments, almost certainly due to the lower-than-usual number of changing states in the mostly nonverbal auditory distractor material. We will come back to this issue in the introduction to Experiment 3. The sound direction effect also was rather small. However, in terms of raw performance scores, the size of the sound direction effect was about 46% of the size of the average irrelevant-sound effect.

The distance of the sound source was irrelevant for the size of the irrelevant-sound effect, as was the sound level that was confounded with sound–source distance (72 and 65 dB[A] for close and distant loudspeakers, respectively). Taken in isolation, the latter aspect is consistent with earlier findings showing that the irrelevant-sound effect is not sensitive to the sound level of the auditory distractors when distractors are displayed via headphones at 40 dB and above (see, e.g., Colle, 1980; Ellermeier & Hellbrück, 1998; Jones, Miles, & Page, 1990). Sound level is one important cue to the sagittal distance of a sound source. Additional distance cues that were present in Experiment 1 (but not in experiments in which headphones were used) include the relative proportion of both high- to low-frequency components and direct to indirect sound. The present experiment shows that even with these distance cues present, sagittal distance does not affect the irrelevant-sound effect.

EXPERIMENT 2

One problem with Experiment 1 is that a certain proportion of the distractor sounds were played while the sequence of to-be-memorized digits was presented. As was mentioned in the Materials section of Experiment 1, participants could not see objects such as the black front loudspeakers in the dark environment around them because their eyes were adapted to the light emitted by the computer screen. Also note that the loudspeakers were positioned to
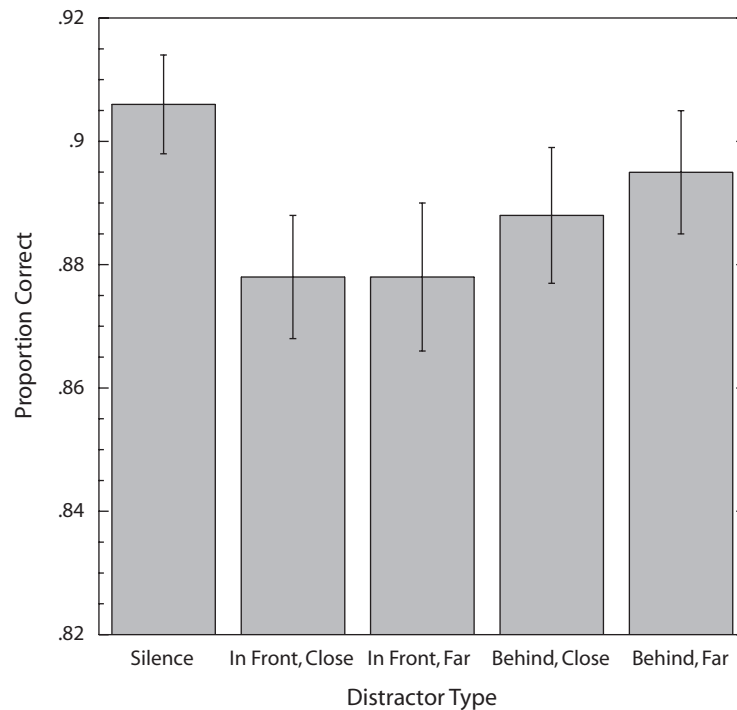
Figure 2. Average proportion correct as a function of the distractor sound source (Experiment 1). The error bars represent the standard errors of the means.

the left and right slightly behind the computer screen. Nevertheless, sudden onsets of distractor sounds could have resulted in occasional gaze shifts toward the slightly off-target location of the sound source; in this case, a visual digit could have been missed. As a result of missing a target every once in a while, serial recall performance can be expected to suffer. To the degree to which gaze shifts occurred selectively for conditions with frontal sound sources, serial recall performance in these conditions could have been reduced relative to conditions with distractor sound sources behind the participants in which gaze shifts would not be expected. Thus, occasional gaze shifts may also explain the sound-direction effect on serial recall performance.

Experiment 2 was designed to test whether the sound-direction effect disappears when the presence of auditory distractors is restricted to the retention interval. In this way, target perception cannot be impaired by gaze shifts because the sudden onsets that could possibly induce such shifts occur only after all targets have been perceived—namely during the interval in which participants attended the computer screen only to wait for the visual signal for commencing serial recall.

Method

Participants were 59 students (43 women) who were paid for their participation. Their ages ranged from 19 to 42 years old ($M = 25$). All participants were tested individually.

The materials, procedure, and design were identical to those of Experiment 1, with the following exceptions. First, no auditory distractors were played during the target presentation. Second, the retention interval (during which distractor sounds were played in all conditions except the silent control condition) was increased to

13 sec. On average, the experiment lasted about 60 min. We were able to recruit 59 participants, so the power was $1 - \beta = .96$.
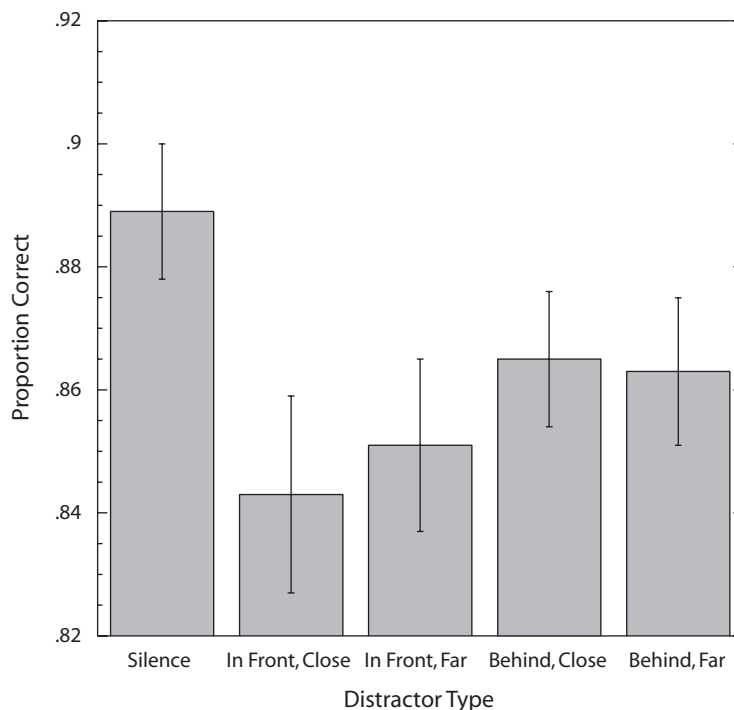
Results

Figure 3 illustrates the serial recall performance in all five experimental conditions. A repeated measures MANOVA showed that performance varied systematically as a function of the distractor type variable [$F(4,55) = 6.23, p < .01, \eta_p^2 = .31$]. The contrast comparing the silent condition with the distractor conditions combined was significant [$F(1,58) = 22.61, p < .01, \eta_p^2 = .28$], confirming the typical irrelevant-sound effect for the present experiment.

A 2×2 MANOVA with sound direction and distance as independent variables showed a significant effect of sound direction [$F(1,58) = 4.97, p < .05, \eta_p^2 = .08$], but neither the main effect of distance [$F(1,58) = 0.23, p = .64, \eta_p^2 < .01$] nor the interaction between the two variables was significant [$F(1,58) = 0.59, p = .45, \eta_p^2 = .01$].

Discussion

The results of Experiment 2 conceptually replicate those of Experiment 1 in showing that serial recall was worse when distractor sounds originated from the front—and thus from the direction that was attended—than when they originated from a direction that was not attended. Sagittal distance—and, hence, location in (3-D) space—was again irrelevant for the size of the irrelevant-sound effect. As in Experiment 1, the irrelevant-sound effect was relatively small, and so was the sound direction effect—the size of which was about 52% of the size of the irrelevant-sound effect in terms of raw performance scores.

Figure 3. Average proportion correct as a function of the distractor sound source (Experiment 2). The error bars represent the standard errors of the means.

The results of Experiment 2 thus complement those of Experiment 1 in showing that the basic directional effect is still observed when distractor presentation is restricted to the retention phase. The alternative explanation that acoustically induced gaze shifts may have caused the distractor direction effect by disrupting the encoding of the visual targets in Experiment 1 is inconsistent with the results of Experiment 2 and needs to be rejected.

EXPERIMENT 3

A problem with both Experiments 1 and 2 is that the direction from which the sound was played also affected the sound signal at the ear entrance. For instance, directional (or head-related) transfer functions[1] that determine the spectral differences between sounds from the front and sounds from behind may be such that the components most relevant for the changing-state effect (Jones, 1993; Jones, Beaman, & Macken, 1996; Jones & Tremblay, 2000) are less pronounced in the signal reaching the ear from behind a participant's head than from the front. A hint that this may indeed be the case comes from research showing that the front–back spectral cues seem to be located mainly in the 8–16 kHz range (Langendijk & Bronkhorst, 2002). Considering that low-pass filtering of auditory distractors reduces the irrelevant sound effect (Jones, Alford, Macken, Banbury, & Tremblay, 2000)—presumably because this type of filtering affects the changing-state components of the auditory distractor signal—one might expect that front–back differences in the auditory signal perhaps play

a role for the changing-state effect. Note, however, that the filter used by Jones et al. (2000) changed the sound quality of the auditory distractors rather drastically, whereas the differences between sounds played from the front versus sounds played from behind are comparatively subtle.

Nevertheless, it seemed at least possible that the sound direction may influence the amount of changing states within the irrelevant sound stream arriving at the ear. Ideally, in order to determine whether sound direction modulates the irrelevant speech effect via changing states, we would need to manipulate independently the sound direction (from the front, from the rear) and the spectral composition of the sound signal. This ideal manipulation is not available, because appropriate manipulations of the spectral composition of the sound signal would also manipulate the direction from which it is perceived (and vice versa, as in Experiments 1 and 2). As an approximation, we decided to use sounds that were recorded by in-ear microphones when played through a front as opposed to a rear speaker. These sounds were then used as distractors played through either a front or a rear speaker. Thus, front-recorded and rear-recorded sounds were played through both front and rear speakers. In this way, we could try to replicate the basic front–back sound direction effect observed in Experiments 1 and 2 while at the same time testing for possible differences as a function of spectral composition of the sound signal (front recorded vs. rear recorded).

An additional consideration concerned the finding that in sound localization studies in which sounds are typically played through headphones, the use of nonindividual ("av-

erage") directional transfer functions degrades localization performance and increases the frequency of front–back confusions (Wenzel, Arruda, Kistler, & Wightman, 1993). This finding indicates that the spectral properties determining front–back sound localizations are quite specific for each individual. We therefore decided to record the distractor sounds separately for each individual using in-ear microphones so that the differences between front-recorded and rear-recorded sounds would reflect individual transfer functions. We did this immediately before the experiment proper. These recordings thus had spectral properties that approximated very closely what each listener was used to.

Note that it was not expected that the recording location (front recorded vs. rear recorded) would actually lead to a substantial number of illusory sound locations, because we expected the sound-direction cues caused by the actual sound direction to dominate. Nevertheless, the front-recorded sounds differ considerably from the rear-recorded sounds in amplitude and spectral composition (sound files with representative examples of front-recorded and rear-recorded words may be obtained from the first author). Thus, if these amplitude and spectral differences were responsible for the sound-direction effect in the previous experiments, then we would expect an effect of recording location. However, if the sound-direction effect of the previous experiments was actually an effect of the perceived sound direction (and not due to an alteration of their changing–state properties), then we would expect recording location to have no effect on serial recall performance, whereas sound direction at presentation should have an effect like that in Experiments 1 and 2.

Another change from Experiments 1 and 2 was that the environmental sounds were replaced by spoken words. This was done in an attempt to increase the overall irrelevant speech effect—that is, the difference between the silent and the distractor conditions. As was mentioned previously, with effect sizes of $\eta_p^2 = .14$ and $\eta_p^2 = .28$, this effect was substantial, but may still be considered not too impressive in comparison with the size of the effect in studies using more typical verbal distractors that can sometimes be as large as $\eta_p^2 = .68$ (Buchner et al., 2004, Experiment 1). This difference appears even larger in terms of raw performance scores: The decrement in the noise conditions in comparison with that of the silent conditions was 2% and 4% in the present Experiments 1 and 2, which looks rather small in comparison with the remarkable 34% decrement in Buchner et al. (2004, Experiment 1). Thus, one incentive for our attempts to increase the overall irrelevant speech effect was to arrive at a data pattern that can be considered more typical of that type of research. However, the more important point is that our attempt to increase the overall irrelevant-sound effect served a theoretical purpose. We wanted to test whether the irrelevant-sound effect and the sound-direction effect are functionally independent—in the sense that the irrelevant speech effect increases as a function of the increase in changing states in the distractor materials, whereas the sound-direction effect remains unaffected—or whether both effects change as a function of the increase in changing states in the auditory distractors. The result of this test

determines how models of working memory account for the sound-direction effect.

An informal inspection of the data from Experiments 1 and 2 suggested that participants who were attested low digit spans in the initial assessment later received digit sequences that were so short that they produced ceiling effects: Their serial recall performance in the experiment proper was almost perfect in all conditions. Thus, in order to maximize the irrelevant-sound effect, all participants received to-be-remembered sequences of nine visually presented digits.

Method

Participants. Participants were 67 students (46 women) who were paid for their participation. Their ages ranged from 19 to 41 years old ($M = 26$). All participants were tested individually.

Materials. The materials were identical to those of Experiment 1, with the following exceptions. First, for all participants, the to-be-remembered lists consisted of nine digits sampled randomly from the set $\{1, 2, . . . , 9\}$, with the restriction that two adjacent digits were not to be identical. A total of 80 such lists were generated for each participant, with 16 lists in each of the five different conditions (see the Design section below). Second, there were only four loudspeakers: Two small loudspeakers were placed immediately to the left and right of the display, and two more loudspeakers were placed behind the participants at the same distances and eccentricities as the two loudspeakers in front. Third, distractor sounds were 20 two-syllable German nouns with a frequency of 1/1,000,000 according to the German language corpus available in the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). These words were spoken by a male voice and were digitally recorded at 16-bit 44.1 kHz. All word recordings were edited to be 700 msec long. Fourth, before the experiment proper, each distractor sound was played twice, once through one randomly selected front loudspeaker and once through one randomly selected rear speaker. These sounds were digitally recorded at 16-bit 44.1 kHz using a set of two small electret-condenser in-ear microphones (Soundman OKM-II Professional), placed in the participant's ears. When these distractors were played during the experiment proper, the average (IEC 60804) sound level at the participant head position was approximately 64 and 58 dB(A) for the front-recorded and the rear-recorded sounds, respectively.

Procedure. The procedure was identical to that of Experiment 1, with the following exceptions. First, the individual digit-span assessment was dropped. All to-be-remembered lists consisted of nine digits (see Materials section). Second, presenting the to-be-remembered digits on each trial took 10.6 sec and was followed by a retention interval of 2.5 sec. Third, in trials with distractor sounds, the participants heard random sequences of the previously recorded distractor words beginning 700 msec before the list presentation, during the list presentation, and during the retention interval. For each trial, the distractor sounds were exclusively (1) front-recorded sounds displayed from the front, (2) rear-recorded sounds displayed from the front, (3) front-recorded sounds displayed from behind, or (4) rear-recorded sounds displayed from behind, depending on the condition of the trial.

Design. The basic within-subjects independent variable was distractor type (silence vs. four different types of distractors). Within the distractor conditions, a 2×2 subdesign was formed by the independent variables sound direction (from the front vs. from behind) and spectral property (spectral composition of front-recorded vs. rear-recorded sounds). The power to detect "medium" effects ($f = .25$) (as defined by Cohen, 1977) given the assumptions spelled out in Experiment 1 and $N = 67$ was $1-\beta = .98$.

Results

Figure 4 illustrates the serial recall performance in all five experimental conditions. A repeated measures MANOVA showed that performance varied systematically

as a function of the distractor type variable [$F(4,63) = 15.86$, $p < .01$, $\eta_{\mathrm{p}}^2 = .50$]. The contrast comparing the silent condition with the distractor conditions combined was significant [$F(1,66) = 64.83$, $p < .01$, $\eta_{\mathrm{p}}^2 = .49$] and thus much larger than the irrelevant speech effect in Experiments 1 and 2 ($\eta_{\mathrm{p}}^2 = .14$ and $\eta_{\mathrm{p}}^2 = .28$, respectively) and closer to the $\eta_{\mathrm{p}}^2 = .68$ observed for this contrast in Buchner et al. (2004, Experiment 1). The same holds for the decrement—in raw performance scores—in the noise conditions in comparison with the silent condition, which was 17% and thus much larger than the 2% and 4% decrements seen in Experiments 1 and 2.

A 2 × 2 MANOVA with sound direction and distance as independent variables showed a significant effect of sound direction [$F(1,66) = 4.83$, $p < .05$, $\eta_{\mathrm{p}}^2 = .07$], but neither the main effect of spectral property [$F(1,66) = 1.58$, $p = .21$, $\eta_{\mathrm{p}}^2 = .02$] nor the interaction between the two variables was significant [$F(1,66) = 1.89$, $p = .17$, $\eta_{\mathrm{p}}^2 = .03$].

Discussion

The results of Experiment 3 conceptually replicate those of Experiments 1 and 2 in showing that serial recall was worse when distractor sounds originated from the front and thus from the direction that was attended than when they originated from behind and thus from the direction that was not attended. Differences in spectral composition of front-recorded and rear-recoded sounds did not affect performance. The alternative explanation

that an elimination or reduction of the spectral components relevant for producing the changing-state effect in sounds played from behind may have caused the distractor direction effect in Experiments 1and 2 thus needs to be rejected.

As expected, it was possible to substantially increase the irrelevant-sound effect by using speech distractors. Nevertheless, the sound-direction effect was in the same order of magnitude as that in Experiments 1 and 2 and thus was not amplified by the change of the changing-state properties in the auditory distractors.

GENERAL DISCUSSION

Auditory distractors are more difficult to ignore when they emanate from the same direction as that of the to-be-attended visual stimuli (Spence et al., 2000). The experiments reported in the present article extend this finding to the irrelevant-sound effect. The typical decrement in serial recall of visually presented digits was worse when the distractor sound source was located in the direction in which visual attention was oriented. In that sense, the data reported here favor models of attention that posit cross-modal attentional links.

Why were cross-modal attentional effects on the irrelevant-sound effect observed in the present experiments, but not in those of Spence and Driver (1999, cited in Spence et al., 2000)? This is of course difficult to tell,
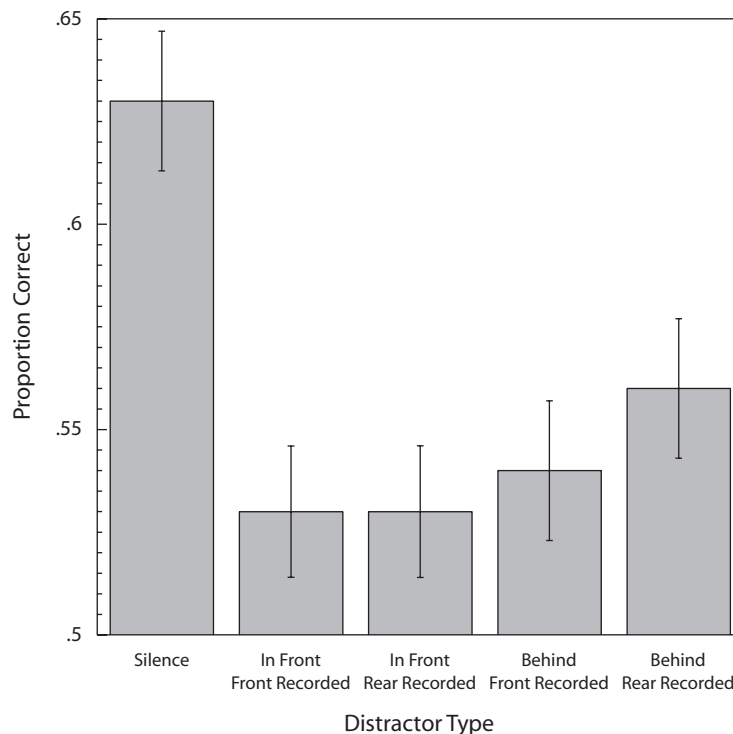

Figure 4. Average proportion correct as a function of the distractor sound source (Experiment 3). The error bars represent the standard errors of the means.

but one possibly relevant variable comes readily to mind. The locus-of-distractor-sound effects found by Spence et al. were descriptively much larger than the effects in the present experiments. The sample effect sizes were $\eta_p^2 = .82$, $\eta_p^2 = .74$, and $\eta_p^2 = .84$ in Spence et al.'s Experiments 1 ($N = 16$), 2 ($N = 10$), and 3 ($N = 12$). These huge effect sizes may have encouraged Spence and Driver (1999, cited in Spence et al., 2000) to regard sample sizes as sufficient that were in the order of magnitude of those used by Spence et al. If this was so, then it should not come as a surprise that the much smaller modulations of the irrelevant-sound effect by spatial attention were not detected in their sample data. To illustrate, the probability of detecting an effect as large as $\eta_p^2 = .74$ in a sample of $N = 10$ participants (as in Spence et al.'s Experiment 2) is larger than .99. In contrast, the probability of detecting an effect of $\eta_p^2 = .11$ (as in the present Experiment 1) in such a sample would be as small as .17. Thus, Spence and Driver (1999) may have simply missed an effect that was there but that was much smaller than they had expected.

Why is the effect so small in comparison with the huge effects observed by Spence et al. (2000)? First, due to its retrospective, memory-based nature, serial recall performance may reflect effects of spatial attention to a much smaller degree (in comparison with other processes relevant for this performance measure) than would shadowing performance that is measured simultaneously with the distraction and presumably disrupts perceptual processes already. Second, in Spence et al.'s experiments, relevant and irrelevant stimuli (1) were of the same modality and (2) were only defined by their spatial locations, whereas in the irrelevant speech paradigm, modality alone is already sufficient for distinguishing between relevant and irrelevant streams so that ignoring the irrelevant stream in the latter case should be both easier and much less dependent on spatial attentional processes. Both of these arguments fit with the fact that even the overall irrelevant speech effect observed in the present experiments is much smaller than the huge locus-of-distractor-sound effects observed by Spence et al. (see the previous paragraph).

Also note that Spence and Driver (1999, cited in Spence et al., 2000) presented relevant and irrelevant information to the left and right of the participants. It seems that this is a weaker manipulation than the front–back manipulation used in the present experiments, which is why the sound-direction effect in their experiment may have been even smaller than the size of the sound-direction effect observed in the present experiments. In essence, then, the present results fill a gap in the literature on cross-modal attentional effects.

The present findings are consistent with results on cross-modal attention effects that were obtained using other experimental paradigms. For instance, Spence and Read (2003) asked participants to shadow an auditory message while being seated in the driver's seat in a driving simulator. Shadowing performance was better when the relevant auditory message came from the front—that is, from the direction in which participants were looking—

than when the message came from the side. What is more, the advantage of the frontal over the lateral sound source was much larger when participants actually performed a simulated driving task than when they simply looked ahead. Thus, when visual attention was intensely drawn to the front during the driving task, auditory attention seems to have been directed even more to the front than was the case in the stationary look-ahead condition. Similarly, Hublet, Morais, and Bertelson (1976, 1977) reported that identifying synthetic speech syllables was easier when those syllables come from a front rather than from a rear or lateral loudspeaker. The front loudspeaker was also the location at which participants were to maintain the narrow beam of a headlight they were wearing—that is, the location that they were attending while trying to identify the syllables. Finally, when visual information directs spatial attention to a certain location, reactions to auditory targets occurring at that location are faster than reactions to targets occurring at other locations (Ward, McDonald, & Lin, 2000). In sum, these findings converge with the evidence reported in the present article that the orienting of spatial attention in one modality (the visual modality in the present case) affects attention in another modality (the auditory modality in the present case). That is, the processing of sounds that come from the direction in which visual attention is oriented is in some sense more efficient—or, as in the present case, more likely to be inevitable—than the processing of sounds that originate elsewhere. Note that in the present experiments, the focus was on the effect of the orienting of visual attention on the processing of auditory stimuli, but there is evidence suggesting that symmetrical links exist between vision and audition in transient spatial attention (see, e.g., Eimer & Schröger, 1998; Spence & Driver, 1996).

The present results also add to the growing body of evidence of an attentional involvement in the irrelevant-sound effect (Buchner & Erdfelder, 2005; Buchner et al., 2006; Buchner et al., 2004; Elliott, 2002; Lange, 2005; Neely & LeCompte, 1999). Like those results, the finding that the irrelevant-sound effect is modulated by the proximity of the auditory distractors to the direction of attention can be readily explained within the embedded-processes model of working memory (Cowan, 1995, 1999) and the feature model (Nairne, 1990; Neath, 2000), because both models comprise an attentional component. In the former model, rehearsed target elements represent the focus of attention, that is—the elements of memory kept at the highest activation levels. Distractor events imply a diversion of resources away from the currently attended object representations, which reduces their activation levels and, hence, the probability of successful recall. As we have seen in the previous paragraph, one may assume that distractor sounds attract more processing resources and are more highly activated when visual attention is oriented toward the distractor sound source. Therefore, in the present experiments, the effect of irrelevant sound distractors should be largest when the distractor sound source is located in front of the participant, which is what was ob-

served. This data pattern can also be accounted for by the feature model (Nairne, 1990; Neath, 2000), within which an attentional parameter reflects the amount of processing resources available for the primary memorization task. One could assume that the ignoring of auditory distractors that come from the same direction as the target is more difficult and that it requires, on average, more attentional resources than does the ignoring of distractors that come from a different direction. This assumption implies that the memorization task should be disrupted to a greater degree by auditory distractors that come from the same direction as the target than by distractors that come from a different direction. An interesting aspect of the feature model is that it does not rely on the concept of attention to explain the standard irrelevant speech effect. Rather, the feature model contains the assumption that the features of the irrelevant sounds overwrite a certain number of elements of the feature vectors of the targets, which degrades the target representations in working memory. With fewer intact features, the probability of successfully matching a degraded target-item representation to the representations in long-term memory is reduced. Recall depends on a successful match, which is why irrelevant speech reduces the probability of successful recall from working memory according to this model. Note that when environmental sounds were used as distractors in Experiments 1 and 2, the size of the standard irrelevant speech effect was $\eta_p^2 = .14$ and $\eta_p^2 = .28$ and was thus rather small (the decrement in the noise conditions in comparison with the silent condition was just 2% and 4% in Experiments 1 and 2). In contrast, in Experiment 3, the standard irrelevant speech effect with words as distractors was much more substantial in that the decrement in the noise condition in comparison with the silent condition was as large as 17% ($\eta_p^2 = .49$). From the point of view of the feature model, this is to be expected, because verbal target and distractor information (as that in Experiment 3) share more features so that overwriting may become more likely when compared with a condition with mostly nonverbal environmental sounds as distractors (as in Experiments 1 and 2). Interestingly, the size of the effect of sound direction was $\eta_p^2 = .11$, $\eta_p^2 = .08$, and $\eta_p^2 = .07$ in Experiments 1, 2, and 3, and was thus approximately constant in all experiments. The assumption that the standard irrelevant speech effect and the effect of sound direction are caused by different processes (feature overwriting as opposed to attentional distraction) and thus may be functionally independent seems broadly consistent with this data pattern.

In contrast, the modular working memory model (Baddeley & Logie, 1999) does not specify a role for attention in the maintenance of information for immediate serial recall. In this model, visually presented words are converted into an articulatory representational format so that they can be maintained in the limited-capacity articulatory loop module for subsequent serial recall. Although the modular working memory model does specify attentional functions, these are identified with the so-called central executive that "is not involved in temporary storage" (Baddeley & Logie, 1999, p. 28). This is why the modular model—at least in its present form—cannot account for differences in serial recall as a function of whether distracting sound comes from an attended or an unattended direction.

Meiser and Klauer (1999) suggested extending the modular working memory model by adding the auxiliary assumption that the central executive may contribute to performance in short-term retention tasks through the involvement of coordinative and supervisory functions. This extension is appropriate for making the modular model compatible with Meiser and Klauer's finding that secondary tasks with high demands on central-executive processes interfered more with serial recall performance than did tasks with lower central executive demands. Is it not clear that this extension would also be appropriate for making the modular model compatible with the present findings. The reason for this skeptical assessment is that when cross-modal attentional effects are analyzed by looking at ERPs, Eimer and Schröger (1998) found that the relevant processes occurred very early (160–280 msec poststimulus), which is why they concluded that these processes must be independent of the timing of executive processes (see also McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003). This conclusion fits with the finding that an increase in the disruption of serial recall due to greater variability in the irrelevant auditory stream covaries (although imperfectly) with an increase in the auditory N1 (Campbell, 2005; Campbell, Winkler, Kujala, & Näätänen, 2003), an ERP component that has been associated with the switching of attention to a significant discontinuity (onset, energy changes, or transitions) in a sound sequence (Näätänen, 1990).

The situation is different for the object-oriented episodic record model (Jones, 1993), which may be extended so that the effects of sound direction on serial recall could be explained within this model. For instance, one could assume the seriation of objects to require attention, or—as was suggested recently by Hughes, Vachon, and Jones (2005)—to remain open to interruption by previously unattended but potentially important information. Serial recall would then suffer not only from competing seriation processes within the representational structure used to temporarily maintain information, but also from attention distraction to task-irrelevant events. If one assumed that distractors coming from the front have a higher potential to capture attention in this sense than do distractors coming from behind, then the present results can be integrated into this model. What is more, just like the feature model, the object-oriented episodic record model in conjunction with the attentional capture concept offers a nice explanation of the relatively constant effect of sound direction, despite large differences in the standard irrelevant speech effect observed in the present series of experiments. The fact that the irrelevant-sound effect in Experiment 3 was larger than the irrelevant-sound effect in Experiments 1 and 2 may readily be explained by the very plausible assumption that the verbal distractors in Experiment 3 implied more changing states and thus should be more disruptive than the environmental sounds used in

Experiments 1 and 2. In contrast, the sound direction manipulation and, hence, the potential of distractors to exert an attentional capture effect was constant across experiments, and so was its effect on serial recall.

In sum, the present results close a gap in the literature on cross-modal attentional effects; they favor models of attention that posit cross-modal attentional links, and they favor models of working memory in which attention is assumed to play a role in the maintenance of serial recall.

AUTHOR NOTE

REFERENCES

BAAYEN, R. H., PIEPENBROCK, R., & VAN RIJN, H. (1993). The CELEX lexical database (Release 1) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.

BADDELEY, A. D. (1996). The concept of working memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 1-27). Hove, U.K.: Psychology Press.

BADDELEY, A. D., & LOGIE, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge: Cambridge University Press.

BANBURY, S. P., TREMBLAY, S., MACKEN, W. J., & JONES, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors*, **43**, 12-29.

BUCHNER, A., & ERDFELDER, E. (2005). Word frequency of irrelevant speech distractors affects serial recall. *Memory & Cognition*, **33**, 86-97.

BUCHNER, A., IRMEN, L., & ERDFELDER, E. (1996). On the irrelevance of semantic information for the "irrelevant speech" effect. *Quarterly Journal of Experimental Psychology*, **49A**, 765-779.

BUCHNER, A., MEHL, B., ROTHERMUND, K., & WENTURA, D. (2006). Artificially induced valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition*, **34**, 1055-1062.

BUCHNER, A., ROTHERMUND, K., WENTURA, D., & MEHL, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition*, **32**, 722-731.

CAMPBELL, T. (2005). The cognitive neuroscience of auditory distraction. *Trends in Cognitive Sciences*, **9**, 3-5.

CAMPBELL, T., WINKLER, I., KUJALA, T., & NÄÄTÄNEN, R. (2003). The N1 hypothesis and irrelevant sound: Evidence from token set size effects. *Cognitive Brain Research*, **18**, 39-47.

COHEN, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Hillsdale, NJ: Erlbaum.

COLLE, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning & Verbal Behavior*, **19**, 722-735.

COWAN, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

COWAN, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

DOWNING, C. J., & PINKER, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 171-188). Hillsdale, NJ: Erlbaum.

DRIVER, J., & SPENCE, C. (2004). Crossmodal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 179-220). Oxford: Oxford University Press.

EIMER, M., & SCHRÖGER, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, **35**, 313-327.

ELLERMEIER, W., & HELLBRÜCK, J. (1998). Is level irrelevant in "irrelevant speech"? Effects of loudness, signal-to-noise ratio, and binaural unmasking. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1406-1414.

ELLIOTT, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition*, **30**, 478-487.

FAUL, F., ERDFELDER, E., LANG, A.-G., & BUCHNER, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**, 175-191.

HUBLET, C., MORAIS, J., & BERTELSON, P. (1976). Spatial constraints on focused attention: Beyond the right-side advantage. *Perception*, **5**, 3-8.

HUBLET, C., MORAIS, J., & BERTELSON, P. (1977). Spatial effects in speech perception in the absence of spatial competition. *Perception*, **6**, 461-466.

HUGHES, R. W., VACHON, F., & JONES, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 736-749.

JONES, D. M. (1993). Objects, streams, and threads of auditory attention. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention, selection, awareness, and control: A tribute to Donald Broadbent* (pp. 87-104). Oxford: Oxford University Press.

JONES, D. M., ALFORD, D., MACKEN, W. J., BANBURY, S. P., & TREMBLAY, S. (2000). Interference from degraded auditory stimuli: Linear effects of changing-state in the irrelevant sequence. *Journal of the Acoustical Society of America*, **108**, 1082-1088.

JONES, D. M., BEAMAN, C. P., & MACKEN, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209-237). Hove, U.K.: Psychology Press.

JONES, D. M., & MACKEN, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 369-381.

JONES, D. M., MACKEN, W. J., & MOSDELL, N. A. (1997). The role of habituation in the disruption of recall performance by irrelevant sound. *British Journal of Psychology*, **88**, 549-564.

JONES, D. M., MILES, C., & PAGE, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal, or memory? *Applied Cognitive Psychology*, **4**, 89-108.

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

LANGE, E. B. (2005). Disruption of attention by irrelevant stimuli in serial recall. *Journal of Memory & Language*, **53**, 513-531.

LANGENDIJK, E. H. A., & BRONKHORST, A. W. (2002). Contribution of spectral cues to human sound localization. *Journal of the Acoustical Society of America*, **112**, 1583-1596.

LECOMPTE, D. C., & SHAIBE, D. M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *Quarterly Journal of Experimental Psychology*, **50A**, 100-118.

MCDONALD, J. J., TEDER-SÄLEJÄRVI, W. A., DI RUSSO, F., & HILLYARD, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, **15**, 10-19.

MEISER, T., & KLAUER, K. C. (1999). Working memory and changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1272-1299.

NÄÄTÄNEN, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral & Brain Sciences*, **13**, 201-288.

NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.

NEATH, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review*, **7**, 403-423.

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition*, **27**, 37-44.

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior*, **21**, 150-164.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1005-1030.

Spence, C., & Driver, J. (1999). *Spatial location of irrelevant speech does not modulate the the unattended speech effect in seial recall.* Unpublished manuscript.

Spence, C., 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.

Spence, C., & Read, L. (2003). Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *Psychological Science*, **14**, 251-256.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 659-671.

Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics*, **62**, 1258-1264.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America*, **94**, 111-123.

NOTE

1. Directional transfer functions are related to a sound propagation path from a sound source to listener's ears and are defined as the Fourier transforms of the ratio of the sound pressure at the ear entrance to that at the center of the head in a free sound field in the frequency domain.