Memory & Cognition
2007, 35 (1),156-165

Irrelevant speech effects and sequence learning

LISA A. FARLEY AND IAN NEATH
Purdue University, West Lafayette, Indiana

DAVID W. ALLBRITTON
DePaul University, Chicago, Illinois

AND

AIMÉE M. SURPRENANT
Purdue University, West Lafayette, Indiana

The irrelevant speech effect is the finding that performance on serial recall tasks is impaired by the presence of irrelevant background speech. According to the object-oriented episodic record (O-OER) model, this impairment is due to a conflict of order information from two different sources: the seriation of the irrelevant speech and the rehearsal of the order of the to-be-remembered items. We tested the model's prediction that irrelevant speech should impair performance on other tasks that involve seriation. Experiments 1 and 2 verified that both an irrelevant speech effect and a changing state effect would obtain in a between-subjects design in which a standard serial recall measure was used, allowing employment of a between-subjects design in subsequent experiments. Experiment 3 showed that performance on a sequence-learning task was impaired by the presence of irrelevant speech, and Experiment 4 verified that performance is worse when the irrelevant speech changes more (the changing state effect). These findings support the prediction made by the O-OER model that one essential component to the irrelevant speech effect is serial order information.

The irrelevant speech effect is the finding that performance on immediate serial recall tasks is impaired by the presence of background speech, even though the background speech is completely irrelevant to the task. This finding has attracted a lot of attention, since it is a prime example of cross-modal interference: Visual to-be-remembered items are interfered with by irrelevant auditory information. Explanations of this apparently simple result remain controversial (see, e.g., Neath, 2000, and subsequent comments by Baddeley, 2000b, and Jones & Tremblay, 2000). The goal of the present work is (1) to extend the range of tasks that are known to be susceptible to disruption by irrelevant speech and (2) to assess the ability of the three major theories to account for the results.

Empirical Review

In one of the first demonstrations of the irrelevant speech effect, Colle and Welsh (1976) presented eight-item lists of consonants visually. The irrelevant speech, played continuously, was a passage from Franz Kakfa's *Ein Hungerkunstler* in German, a language that none of the subjects reported understanding. The speech was categorized as irrelevant because the subjects were instructed to ignore it and were assured that there would be no subsequent test on it (as, indeed, there was not). Performance was 12% worse in the irrelevant speech condition than in the quiet control condition. A subsequent study (Colle, 1980; see also Ellermeier & Hellbrück, 1998) showed that the amount of impairment caused by irrelevant speech was independent of the intensity of the irrelevant stimuli, at least over the range from 40 to 76 dB(A). The impairment appears to be the same regardless of whether the irrelevant speech accompanies presentation or follows presentation (Miles, Jones, & Madden, 1991), and the magnitude of the effect does not diminish over repeated trials or sessions (Hellbrück, Kuwano, & Namba, 1996; Tremblay & Jones, 1998).

In general, the phonological or semantic relation between the irrelevant stimuli and the to-be-remembered stimuli is not related to the effect (Jones & Macken, 1995). In particular, there are far more studies demonstrating that phonological overlap between the irrelevant speech and the to-be-remembered item has no additional disruptive effect (Bridges & Jones, 1996; Jones & Macken, 1995; LeCompte & Shaibe, 1997) than there are studies showing such an effect (Salamé & Baddeley, 1982).

One of the hallmarks of the irrelevant speech effect is the changing state effect (Jones, Madden, & Miles, 1992): Irrelevant auditory stimuli that change over time produce more of a decrement than do otherwise comparable stimuli that do not change (see also Beaman & Jones, 1997; Jones, Alford, Macken, Banbury, & Tremblay, 2000). For

I. Neath, ineath@mun.ca

example, a single repeated item (e.g., B B B B) will produce less of a decrement than will a sequence of different items (e.g., A B C D).

Another hallmark is that tasks without a serial component are largely insensitive to the disruptive effects of irrelevant speech (Jones & Tremblay, 2000). For example, Baddeley and Salamé (1986) failed to find an irrelevant speech effect on a series of tasks in which subjects were asked to judge whether visually presented items rhymed; Salamé and Baddeley (1990) failed to find an irrelevant speech effect in free recall of 16-item lists; and Beaman and Jones (1997) failed to find an effect of irrelevant speech on a missing item task. Although published research about the effects of irrelevant speech on simple response time (RT) tasks—in which there is no serial order information—is quite sparse, in general, the results suggest that performance is not routinely impaired by the presence of irrelevant speech (e.g., Kjellberg & Sköldström, 1991; Smith, 1989).

One purpose of the present research was to extend the range of tasks that are known to be susceptible to disruption by irrelevant speech. One task that is quite different from immediate serial recall but still involves serial order is sequence learning.[1] In a typical sequence-learning task, targets are presented in one of several locations on a computer screen, and subjects are asked to press a corresponding key as quickly as possible to indicate the location. The targets are shown in a specified pattern for a number of training blocks; then they are changed to a different pattern for one transfer block, and then the original pattern reappears. For example, if there are four target locations, and 1 denotes the location farthest left and 4 denotes the location farthest right, a pattern can be any sequence, such as 124342314321. For the first 9 blocks of trials, the targets will appear accordingly. During the 10th block of trials, the original pattern of the locations will be replaced by a new pattern, or the targets may simply appear at randomly determined locations. For the final block, the original pattern from the first 9 blocks will reappear. Learning will be inferred by comparing data from the transfer block with those from the blocks that immediately precede and follow the transfer block (Hsiao & Reber, 1998). If a subject successfully learns and utilizes the underlying sequence, the pattern change in the transfer block should lead to longer response times and decreased accuracy. When the original sequence returns for the final block, RTs should decrease, and accuracy should increase. Given that subjects who learn the sequence must learn serial information, it is therefore possible that such tasks can be disrupted by irrelevant speech. Our experiments directly tested this idea.

Theoretical Review

There are three major accounts of the irrelevant speech effect: the object-oriented episodic record (O-OER) model (Jones, 1993), the feature model (Neath, 2000), and the phonological store hypothesis (Baddeley, 2000a).

According to the O-OER model (Jones, 1993; Jones, Beaman, & Macken, 1996; Jones & Tremblay, 2000), the irrelevant speech effect is due to a conflict between serial order information from two different sources. In a typical irrelevant speech experiment (e.g., visual to-be-remembered items and auditory irrelevant material), both visual and auditory items are represented using amodal, abstract representations, called *objects*. Objects are created by the same processes as those used in perception, including those dealing with segmentation based on changes in stimulus characteristics. Serial order is encoded by the use of pointers that are associated with individual objects. The different modalities are indicated through streaming, in which items or events are assigned to either the same or different sources. The formation of a pointer is a probabilistic process, and once formed, its strength decays over time. Errors in recall occur when pointers from one stream of objects, such as those representing irrelevant speech, interfere with a different set of pointers, such as those representing the list items.

The changing state effect arises, then, because a repeatedly presented auditory item creates only one object; in contrast, if the auditory input consists of a set of different or varying items, multiple objects are created, along with appropriate pointers. The model accounts for the lack of an effect of phonological or semantic overlap between the to-be-remembered items and the irrelevant information by noting that it is the relationship between items within a sequence that is important.

The O-OER model posits that seriation is a necessary requirement for observing an irrelevant speech effect: If there is no seriation, there can be no interference between the two streams.[2] Our focus is on one particular implication of this: Tasks other than serial recall that emphasize seriation should be susceptible to disruption by irrelevant speech and should also show a changing state effect. In sequence-learning tasks, the subject must know (either consciously or not) the order in which the stimuli appear over trials. Seriation can be plausibly assumed to the extent that the subject has learned the sequence. In such a situation, the two processes of seriation could readily interfere, causing an irrelevant speech effect.

The feature model (Nairne, 1990) offers a quite different account (Neath, 2000). According to this view, the irrelevant speech effect is due to a combination of two factors: feature adoption and attention. Items are represented by sets of modality-independent and modality-dependent features. Feature adoption occurs when some of the features of the irrelevant speech become incorporated into the representation of a to-be-remembered item. This reduces the probability of successfully matching a degraded memory trace to a particular cue, and so an error occurs. An additional element within the feature model is an attentional parameter. Although not well specified, the idea is that the presence of irrelevant speech creates a sort of dual-task situation. One task is the recollection of a list of items, and the second task is ignoring the irrelevant stimuli. The more effort required to ignore or not process the secondary stimuli, the more a subject's performance should be reduced, just as, if more attention is required to perform Task 1, there is a greater effect on the performance of Task 2.

This view accounts for the changing state effect by assuming that it requires more attention to perform the memory task when the secondary task—ignoring the irrelevant speech—is made more difficult by having the irrelevant speech change (see Simulation 3 in Neath, 2000). It explains the lack of an effect of phonological or semantic overlap between the to-be-remembered items and the irrelevant information, because only features are adopted: It takes multiple features to represent one particular phoneme, let alone semantic information about an entire word (see Simulation 2 in Neath, 2000).

This view does not require seriation for an irrelevant speech effect to be observed. It can predict an irrelevant speech effect in a sequence-learning task only to the extent that such tasks require attention. A review of the literature on this topic is beyond the scope of this article; currently, however, it seems unlikely that attention plays a major role in typical sequence-learning tasks (cf. Shanks & Channon, 2002).

The third account of the irrelevant speech effect is the phonological store hypothesis derived from Baddeley's (1986, 2000a) working memory framework. Visually presented verbal items are converted to a phonological code via the articulatory control process and are then deposited in the phonological store. Auditory verbal information automatically enters the store directly. Irrelevant speech affects memory by interfering in some unspecified way with information about the to-be-remembered items in the phonological store.

Because the unit of importance in the phonological store is the phoneme, there should be no effect of semantic similarity between the irrelevant speech and the to-be-remembered items. However, contrary to the data, this view predicts that there should be an effect of phonological similarity, because both the to-be-remembered items and the irrelevant speech items will be temporarily stored together. According to this view, seriation is not necessary. The phonological store hypothesis predicts an effect of irrelevant speech on sequence learning only if the stimuli are stored in the phonological store.

Experiment Overview

Empirically, then, one might expect an irrelevant speech effect in a sequence-learning task because such a task requires learning serial order information and irrelevant speech effects have routinely been found only in those tasks that emphasize serial order. Theoretically, the O-OER model makes the strongest prediction for observing such an effect: Although the feature model and the phonological store hypothesis could accommodate the results, neither makes a particularly strong prediction.

Before examining whether irrelevant speech effects are observable on sequence-learning tasks, however, it is necessary to iron out a methodological wrinkle. Typically, in irrelevant speech effect experiments, immediate serial recall is used in a within-subjects design. That is, each subject experiences both the quiet and the irrelevant speech conditions. The most simple test of irrelevant speech effects in sequence learning would utilize a between-subjects design, in which a particular subject would experience only irrelevant speech or only quiet during the training blocks and the transfer block. However, it is unclear whether irrelevant speech effects would obtain in between-subjects designs, and it is quite possible that they might not. First, to our knowledge, there has been no demonstration of a between-subjects irrelevant speech effect. Hanley and Broadbent (1987) did report an effect of irrelevant speech with a between-subjects design, but they used auditory presentation for the to-be-remembered items, not visual presentation. Thus, they did not show cross-modal interference. Second, there are many memory phenomena that behave differently in between- versus within-subjects (or even blocked vs. random) designs, such as the generation effect (Slamecka & Graf, 1978), the bizarre imagery effect (Einstein & McDaniel, 1987), and the word frequency effect (Watkins, LeCompte, & Kim, 2000), to name only a few. Therefore, in Experiments 1 and 2, we examined whether an irrelevant speech effect would obtain in immediate serial recall in a between-subjects design. Experiments 3 and 4 were then performed to assess whether an irrelevant speech effect would obtain in sequence learning.

EXPERIMENT 1

The purpose of Experiment 1 was to verify that an irrelevant speech effect is observable in a between-subjects design. Half the subjects received the quiet condition first, followed by the irrelevant speech condition, and the remaining subjects received the irrelevant speech condition first, followed by the quiet condition. This design enabled us to compare performance between the two groups on just the first block, to see whether there would be a between-subjects irrelevant speech effect, and also to examine the effects of order of conditions. The basic method was modeled on one used by Neath, Farley, and Surprenant (2003) that produced a reliable within-subjects irrelevant speech effect.

Method

Subjects. Forty DePaul University undergraduates participated in exchange for credit in introductory psychology courses and were arbitrarily assigned to one of two groups. All identified themselves as native speakers of American English.

Design. The presence or absence of irrelevant speech was a within-subjects variable; the order of these conditions—that is, quiet in Block 1 or irrelevant speech in Block 1—was a between-subjects variable.

Stimuli. The to-be-remembered items were the seven consonants F, K, L, M, Q, R, and X, used by Colle and Welsh (1976). The irrelevant speech was a passage from *Die Wilden* by Franz Kafka, spoken in German by a female speaker, previously used by Neath et al. (2003).

Procedure. The subjects were informed that we were interested in how accurately they could remember the order in which a series of letters had been presented. Each letter was shown in uppercase for 1 sec in the middle of the window in 24-point Helvetica. After the final letter had been shown, seven response buttons became active and were labeled with the seven letters in alphabetical order. The subjects were asked to indicate the presentation order by clicking on appropriately labeled buttons, using the mouse. For example, if they thought that the first letter was L, they should click on the button labeled L first. If they thought that the third letter was Q, they should click on the button labeled Q third. The subjects in the irrelevant speech condition were informed that they would hear some German being spoken and that they should ignore it. The irrelevant speech began with the onset of the first list item and ended with the offset of the last list item. An experimenter remained in the room to ensure compliance with the instructions. The subjects received an initial block of 20 lists and were informed they could take rest breaks at any point. Following the first 20 lists, the subjects were given an additional block of 20 lists in the other condition. The subjects were tested individually.

Results

The main result of interest is shown in Figure 1: The proportion of letters correctly recalled in order was greater for the subjects who did not experience irrelevant speech in Block 1 (.625) than for those subjects who did hear irrelevant speech (.497). The data were analyzed with a 2 (orders: quiet first or irrelevant speech first) × 2 (condition: presence or absence of irrelevant speech) ANOVA. The alpha level for this and all the subsequent analyses was set at .05.

There was no effect of order of the conditions [$F(1,38) < 1$], with an overall proportion correct of .592 when quiet came first, as compared with .572 when irrelevant speech came first. There was a main effect of irrelevant speech [$F(1,38) = 30.84$, $MS_e = 0.008$], with a larger proportion of items correctly recalled in order in the quiet condition (.636) than in the irrelevant speech condition (.527). There was also a reliable interaction [$F(1,38) = 4.57$, $MS_e = 0.008$], due to better performance in Block 2 than in Block 1: When the quiet condition was first, the proportion correct was .625, as compared with .648 when quiet came second; when the irrelevant speech condition came first, the proportion correct was .497, as compared with .558 when it came second.

A series of post hoc LSD tests confirmed the presence of a between-subjects irrelevant speech effect regardless of the order (i.e., in both Block 1 and Block 2)—both when quiet came first (.625 vs. .497) and when quiet came second (.648 vs. .558). The tests also confirmed the presence of a within-subjects irrelevant speech effect regardless of the order—both when quiet came first (i.e., .625 vs. .558) and when quiet came second (i.e., .648 vs. .497).

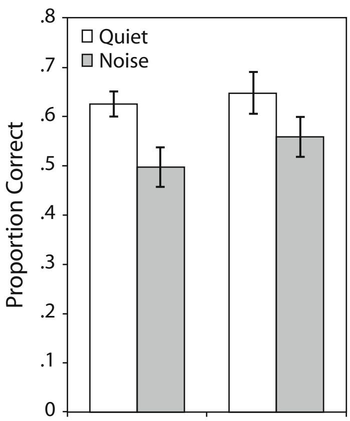

Figure 1. The proportions of letters correctly recalled in order as a function of condition (quiet or irrelevant speech) and presentation order (Block 1 or 2) in Experiment 1. The between-subjects irrelevant speech effect is seen in the two bars for Block 1. Error bars show the standard error of the mean.

Discussion

Experiment 1 demonstrated that irrelevant speech effects can be observed in a between-subjects design, and the magnitude of the effect was similar to that observed in a comparable within-subjects design (Neath et al., 2003). Experiment 1 also showed a reliable interaction between order and condition, suggesting that although the irrelevant speech effect can be observed in between-subjects designs, interpretation of results in blocked designs might be more complex.

EXPERIMENT 2

Although Experiment 1 demonstrated that the presence of irrelevant speech affects memory in a between-subjects design, just as it does in a within-subjects design, it is possible that the effect demonstrated was not the classic irrelevant speech effect. Given that the changing state effect is the empirical signature of the irrelevant speech effect, Experiment 2 was designed to see whether a changing state effect would be observable in a between-subjects design.

Method

Subjects. Seventy-eight DePaul University undergraduates participated in exchange for credit in introductory psychology courses and were arbitrarily assigned to one of three groups. All identified themselves as native speakers of American English.

Design. The one between-subjects variable was condition: quiet, unchanging irrelevant speech, or changing irrelevant speech.

Stimuli. The to-be-remembered items were the same as those in Experiment 1. The irrelevant speech stimuli were digits pronounced by a male speaker. In the unchanging condition, one digit (between 1 and 9) was randomly picked for each subject, and then that digit was played repeatedly every 1,250 msec during presentation of the to-be-remembered items. In the changing condition, the subjects also heard a digit every 1,250 msec during list presentation, but each time, the digit was randomly selected from the entire set, subject to the constraint that a digit never be repeated in immediate succession.

Procedure. The procedure was almost identical to that in Experiment 1, except that there was only one block. The subjects received 25 trials.

Results

Experiment 2 replicated Experiment 1 in that there was again a between-subjects irrelevant speech effect. In addition, there was also a between-subjects changing state effect: The proportion of items correctly recalled in order was best in the quiet condition (.664), intermediate in the unchanging condition (.575), and worst in the changing condition (.518).

An ANOVA on the proportion of letters correctly recalled revealed a main effect of condition [$F(2,75) = 13.91$, $MS_e = 0.010$]. Planned comparisons showed a significant difference in the level of performance between all three conditions: More items were recalled in the quiet than in the unchanging condition [$F(1,75) = 10.28$, $MS_e = 0.010$], and more items were recalled in the unchanging than in the changing condition [$F(1,75) = 4.09$, $MS_e = 0.010$].

Discussion

Experiment 2 demonstrated a changing state effect, using a between-subjects design. This is consistent with the claim that the decrement in memory performance caused by irrelevant speech in Experiment 1 was due to the classic irrelevant speech effect. Having thus demonstrated a between-subjects irrelevant speech effect and a between-subjects changing state effect, we could now use a between-subjects design to see whether irrelevant speech would disrupt performance on a sequence-learning task.

EXPERIMENT 3

Experiment 3 was designed to see whether irrelevant speech disrupts performance on a sequence-learning task. The subjects saw a sequence of symbols appear in the middle of a computer monitor and were asked to identify each symbol by pressing one of four keys.[3] Half of the subjects heard irrelevant speech during presentation, and half did not. Furthermore, half of the subjects saw the symbols occur according to a predetermined pattern, whereas the other half saw the symbols occur in a randomly determined order. Seriation should occur only in the pattern group, since there was no serial order information in the random group. Therefore, if the O-OER model is correct in its assumption that seriation in two streams is critical to observing an irrelevant speech effect, such an effect should occur only when there is a pattern, and not when the symbols appear randomly.

Method

Subjects. One hundred twenty Purdue University undergraduates participated in exchange for credit in introductory psychology courses. All identified themselves as native speakers of American English, and each was randomly assigned to one of the four conditions.

Design. Both the presence or absence of irrelevant speech and the type of sequence (pattern or random) were between-subjects variables. The subjects were exposed either to a repeating pattern or to a random sequence. The pattern that governed which symbol appeared was the same as that used by Destrebecqz and Cleeremans (2001). Using the digits 1–4 to represent the four symbols, the pattern was 342312143241. One block consisted of 96 trials, on which each trial was one keypress, and the pattern was repeated four times per block. On Block 10, the pattern was switched to a different sequence similar in structure (341243142132), and on Block 11, the original pattern was restored. In the random condition, the order of the symbols was determined randomly, subject to the constraints that each target appeared equally often and that no symbol be used twice in succession. The subjects either heard irrelevant speech during the learning task or performed the task in silence. Each block in the pattern condition began at a different randomly determined position within the sequence.

Stimuli. The irrelevant speech was the same as that in Experiment 1. The four symbols were !, @, #, and $ and were shown in 72-point Helvetica in the center of the screen in red against a black background.

Procedure. The subjects were informed that we were interested in how quickly and accurately they could press a key on the keyboard (with a standard U.S. layout) to indicate which symbol had appeared. They were asked to place their left middle finger on the “z” key, their left index finger on the “x” key, their right index finger on the “.” key, and their right middle finger on the “/” key. A graphic illustrated the mapping of keys to symbols: “z” was mapped to “!”, “x” was mapped to “@”, “.” was mapped to “#”, and “/” was mapped to “$”. The subjects were asked to press the key as quickly as they could while still being accurate. Each symbol appeared 500 msec after the previous keypress. If the wrong key was pressed, a message indicated the error during this 500-msec period. The subjects were informed that they could take a short break after every block.

Results

Data analysis. For each of the 120 subjects, all RTs more than three standard deviations greater than the mean were excluded and were counted as incorrect responses. The mean proportion of responses affected by this was .017 in the pattern quiet condition, .014 in the pattern irrelevant speech condition, .017 in the random quiet condition, and .017 in the random irrelevant speech condition.

Figure 2 shows the mean RTs for correct responses for each condition as a function of block, and Table 1 shows the overall mean RTs and accuracy. There are two main results of interest, both of which are apparent in the figure. First, the presence of irrelevant speech caused longer RTs when there was a pattern than when there was no irrelevant speech (the black triangles are well above the white triangles); when there was no pattern, the presence of irrelevant speech had no discernible effect (the black and white circles are roughly equal). Second, although irrelevant speech slowed responding in the pattern conditions, it did not prevent learning: Both the quiet and the irrelevant speech conditions showed slower responding on Block 10 than on Blocks 9 and 11.

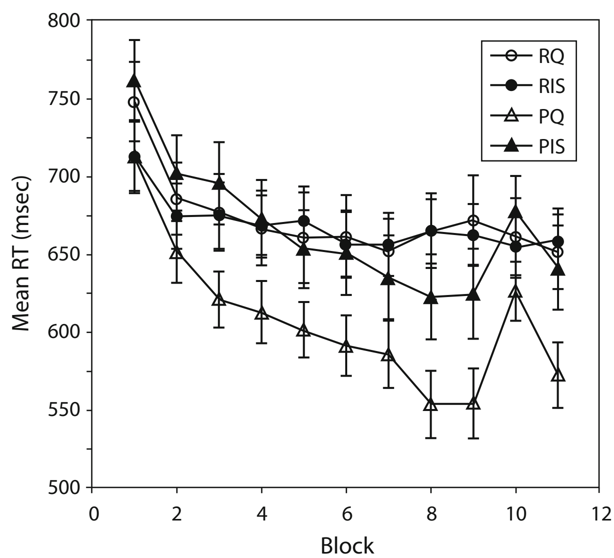

Figure 2. Mean response times (RTs) for correct responses as a function of block number for all four conditions in Experiment 3. R, random; P, pattern; Q, quiet; IS, irrelevant speech.

Table 1
Mean Response Times (RTs, in Milliseconds) for Correct Responses and Proportions Correct for Blocks 1–9 in Experiment 3

	Mean RT		Proportion Correct	
Condition	Pattern	Random	Pattern	Random
Quiet	609.38	676.60	.943	.941
Irrelevant speech	668.88	671.46	.946	.943

The analyses that follow were designed to evaluate the observations above, so the data from the random and the pattern conditions were analyzed separately.[4] There was a highly restricted range for the accuracy data; therefore, detailed analyses involving accuracy will be shown in the Appendix. However, as Table 1 makes clear, there was no suggestion of a speed–accuracy trade-off, since accuracy was approximately equal in all the conditions.

Effect of irrelevant speech. For the pattern conditions, a 2 (background condition: presence or absence of irrelevant speech) × 9 (block) ANOVA on response times for correct responses yielded a main effect of background condition [$F(1,58) = 4.13$, $MS_e = 115{,}611$], with faster responding in the quiet condition (609 msec) than in the irrelevant speech condition (669 msec). There was a main effect of block [$F(8,464) = 36.41$, $MS_e = 3{,}670.7$] but no interaction [$F(8,464) < 1$].

In contrast, for the random groups, there was only an effect of block [$F(8,464) = 17.63$, $MS_e = 1{,}733.6$]. The effect of background condition was not reliable ($F < 1$), and the interaction was not reliable either [$F(8,464) = 1.51$, $MS_e = 1{,}733.6$, $p > .10$].

An additional analysis was performed to determine whether irrelevant speech affected the rate at which the pattern was learned. For each subject, the slope of the best-fitting line for the mean RT for correct responses for each block was calculated. These slopes were then analyzed with a 2 (group: pattern or random) × 2 (background condition: presence or absence of irrelevant speech) ANOVA. There was a main effect of group [$F(1,116) = 22.51$, $MS_e = 144.6$], with steeper slopes in the pattern (−16.26) than in the random (−5.84) condition. However, there was no effect of condition ($F < 1$) and no interaction ($F < 1$). A series of post hoc LSD tests showed that the slopes in the pattern quiet (−17.01) and pattern irrelevant speech (−15.51) conditions did not differ and that the slopes in the random quiet (−7.03) and random irrelevant speech (−4.65) conditions did not differ. However, both of the slopes in the two pattern conditions differed from both of the slopes in the two random conditions.

Evidence of learning. Learning was assessed using a 2 (background condition: presence or absence of irrelevant speech) × 2 (block: mean RT on Transfer Block 10 vs. mean RT averaged over Blocks 9 and 11) ANOVA. For the pattern groups, there was a reliable main effect of background condition [$F(1,58) = 4.47$, $MS_e = 24{,}512$], with faster overall responding in the quiet condition (595 msec) than in the irrelevant speech condition (695 msec). There was also a main effect of block [$F(1,58) = 18.93$, $MS_e = 4{,}581.85$], with slower responding on the transfer block (652 msec) than on surrounding blocks (598 msec). The interaction was not reliable ($F < 1$). The same analysis on data from the random groups yielded F values of less than 1 for both the main effects and the interaction.

Discussion

Experiment 3 clearly shows that irrelevant speech slows responding on a sequence-learning task when there is a pattern to be learned; when there is no serial information, as in the random conditions, irrelevant speech has no effect on response times. The results suggest that it is irrelevant speech that affects performance, rather than learning: Neither the rate of learning, as measured by the slope of the best-fitting line, nor performance on the transfer block differed in the pattern group as a function of whether there was irrelevant speech.

These results are in line with what the O-OER model predicts: Irrelevant speech affects responding on a sequence-learning task when there is a pattern to learn but does not affect responding when there is no serial order information in the task. The model does not specify whether the effect will be on learning of the pattern or on performance. The reason is that serial recall tasks are usually not designed to assess the distinction between learning and performance; typically, results are described in terms of an effect on the memory for an item (e.g., irrelevant speech causes more item or more order errors). The results presented above suggest that learning is not affected; rather, irrelevant speech disrupts performance. Discussion of a possible explanation will be postponed until after the report on Experiment 4, which was designed to see whether a changing state effect would obtain in sequence-learning tasks.

EXPERIMENT 4

It is possible that the decrement observed in Experiment 3 is different from the classic irrelevant speech effect observed in serial recall tasks, such as Experiment 1 and 2. Experiment 4 was designed to test this by determining whether a changing state effect would obtain with sequence learning, using the patterned sequences from Experiment 3 and the two different kinds of irrelevant speech, unchanging and changing, from Experiment 2.

Method

Subjects. Ninety Purdue University undergraduates participated in exchange for credit in introductory psychology courses. All identified themselves as native speakers of American English and were randomly assigned to one of the three conditions.

Stimuli. The stimuli were the same as those in Experiment 3 (!, @, #, and $), but the irrelevant speech was the same as that in Experiment 2 (a single repeating digit vs. a stream of randomly determined digits).

Design and Procedure. The design was similar to that in Experiment 3, with the exceptions that there were no random conditions and that there were two types of irrelevant speech.

Results

Data analysis. As with Experiment 3, for each of the subjects, all RTs more than three standard deviations greater than the mean were excluded and were counted as incorrect responses. The proportion of responses affected by this was .015 in the quiet condition, .013 in the unchanging condition, and .013 in the changing condition. As with Experiment 3, there was a restricted range in the accuracy data, so detailed analyses of the accuracy data will be presented in the Appendix. As Table 2 suggests, however, there was no evidence of a speed–accuracy trade-off.

Figure 3 shows the mean response times for correct responses for each condition as a function of block, and Table 2 shows the mean RT and mean accuracy levels for each condition. There are three main results of interest. First, as in Experiment 3, the presence of irrelevant speech resulted in longer overall response times. Second, the effect was larger for irrelevant speech that changed than for irrelevant speech that did not change. Third, although irrelevant speech slowed responding, it did not prevent learning: All three conditions showed slower responding on Block 10 than on Blocks 9 and 11.

Effect of irrelevant speech. A 3 (condition: quiet, unchanging, or changing) × 9 (block) ANOVA on response times for correct responses yielded a main effect of condition [$F(2,87) = 8.76$, $MS_e = 59{,}723$]. There was also a main effect of block [$F(8,696) = 52.24$, $MS_e = 2{,}263$] but no interaction ($F < 1$). Planned comparisons confirmed that responding was faster in the quiet condition (546 msec) than in the unchanging (590 msec) condition [$F(1,58) = 4.13$, $MS_e = 61{,}187$] and that responding was faster in the unchanging condition (590 msec) than in the changing (635 msec) condition [$F(1,58) = 5.20$, $MS_e = 51{,}993$].

As in Experiment 3, the slope of the best-fitting line for the mean RT for correct responses for each block was calculated. A one-way ANOVA showed no effect of condition ($F < 1$). The mean slope was −15.01 for the quiet group, −11.61 for the unchanging group, and −11.91 for the changing group.

Evidence of learning. Learning was assessed using a 3 (background condition: quiet, unchanging, or changing) × 2 (block: mean RT on Transfer Block 10 vs. mean RT averaged over Blocks 9 and 11) ANOVA. There was a reliable main effect of condition [$F(2,87) = 10.76$, $MS_e = 12{,}081$], with faster overall responding in the quiet condition (533 msec), intermediate responding in the unchanging condition (586 msec), and slowest responding in the changing condition (625 msec). There was also a main effect of block [$F(1,87) = 54.50$, $MS_e = 2{,}699$], with slower responding on the transfer block (610 msec) than on surrounding blocks (553 msec). The interaction was not reliable ($F < 1$).

Table 2
Mean Response Times (RTs, in Milliseconds) for Correct Responses and Proportions Correct for Blocks 1–9 in Experiment 4

Condition	Mean RT	Proportion Correct
Quiet	546.49	.937
Unchanging irrelevant speech	589.76	.928
Changing irrelevant speech	634.53	.915

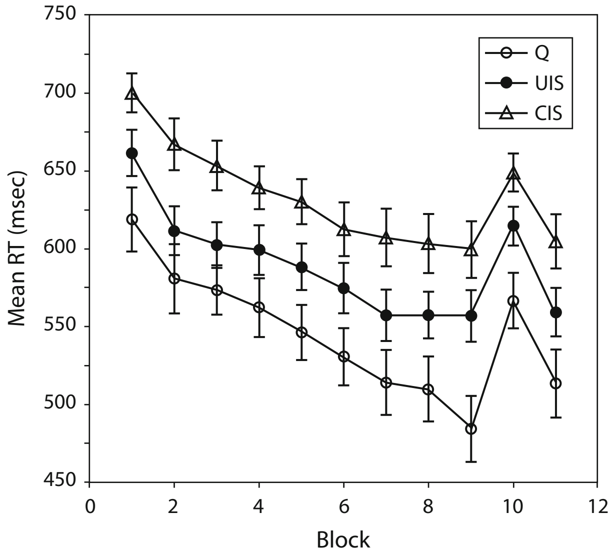

Figure 3. Mean response times (RTs) for correct responses as a function of block number for the quiet (Q), unchanging irrelevant speech (UIS), and changing irrelevant speech (CIS) conditions in Experiment 4.

Discussion

Experiment 4 showed a changing state effect in a sequence-learning task. RTs were longer the more the irrelevant speech changed; these results are consistent with the claim that the disruption observed was due to the classic irrelevant speech effect. As in Experiment 3, irrelevant speech seemed to affect performance, rather than learning: Neither the rate of learning, as measured by the slope of the best-fitting line, nor performance on the transfer block differed between the quiet and the two irrelevant speech conditions.

GENERAL DISCUSSION

The four experiments reported here extend the range of experimental designs and tasks that are known to be susceptible to disruption by irrelevant speech. Experiment 1 demonstrated that with immediate serial recall, irrelevant speech disrupts performance of visually presented items even when manipulated between subjects. Experiment 2 demonstrated a between-subjects changing state effect, suggesting that the effect observed in Experiment 1 really was an irrelevant speech effect.

Experiment 3 showed that irrelevant speech lengthened RTs in a sequence-learning task when there was a pattern to be learned; when there was no pattern—and thus, no serial order component—there was no effect. Irrelevant speech did not affect learning; both the irrelevant speech and the quiet conditions showed typical learning results: slower performance on the transfer block than on the surrounding blocks. Experiment 4 demonstrated a changing state effect in a sequence-learning task, suggesting that the effect observed in Experiment 3 was the irrelevant speech effect. Another similarity between the findings in both Experiments 3 and 4 and those observed with serial recall is that the magnitude of the disruption did not decrease over blocks of trials (e.g., Hellbrück et al., 1996; Tremblay & Jones, 1998).

One potential difference between the irrelevant speech effect in serial recall and that in sequence learning is the locus of the effect. In sequence learning, much is made of the difference between an effect on learning versus an effect on performance. The data presented here are consistent with the claim that irrelevant speech affects performance, rather than learning. In immediate serial recall, the distinction is rarely made explicit, and it is possible that the locus of the effect in the two paradigms is different; further research will be necessary to test this possibility.

Of the three major explanations of the irrelevant speech effect, the O-OER model provides the best account of the data. This view attributes the locus of the effect to a conflict between order information from two sources. Although typically, rehearsal of the to-be-remembered items is one of the seriation processes, it is plausible to view sequence learning as a seriation process within the context of the model. When such a process occurs (as in the pattern conditions in Experiment 3), irrelevant speech disrupts the sequence-learning task. When there is no serial process (as in the random conditions in Experiment 3), there is no disruption.

The model does not say whether irrelevant speech will affect learning or performance, whereas the data suggest an effect only on performance. The effect on performance can be readily explained: The conflict in order information happens at the time the subject must make a decision about which symbol is being shown. Sequence learning might not show an effect of irrelevant speech on learning, due to the nature of the task: Unlike serial recall tasks, in which the order information is presented only once and then tested, sequence-learning tasks present the order information numerous times.

The feature model could easily explain the findings of a between-subjects irrelevant speech and changing state effect. However, the model has more difficulty in explaining why irrelevant speech affects sequence learning only when there is a pattern. One possibility is to assume that attention plays a role in these tasks only when there is something to attend to (i.e., a pattern). However, this explanation may not be tenable if subsequent research supports the idea that attention does not play a substantial role in this type of task.

Similarly, the phonological store hypothesis has no difficulty with the between-subjects irrelevant speech and changing state effects. However, it has even more difficulty than does the feature model in explaining why the effect obtains only in the pattern condition, and not in the random condition. If information about the sequence is in the phonological store for one condition, why would it not be in the store for the other?

The results clearly extend the range of tasks that are known to be susceptible to disruption by irrelevant speech: Such effects, including the changing state effect, are observable in between-subjects designs and in sequence-learning tasks. Given the importance attributed to seriation by the O-OER model, the results also favor its explanation of irrelevant speech effects over those of the feature model and the phonological store hypothesis.

AUTHOR NOTE

I.N. and A.M.S. are now at the Memorial University of Newfoundland. This work is based on the first author's master's thesis. Preliminary results were reported at the XX British Psychological Society Cognitive Section Conference, University of Reading, September 2003. Portions of this project were conducted while I.N. and A.M.S. were Visiting Professors at the Department of Psychology, City University, London. We thank Robert W. Proctor for his extremely useful comments and suggestions. Correspondence may be addressed to I. Neath, Department of Psychology, Memorial University of Newfoundland, St. John's, NL, A1B 3X9 Canada (e-mail: ineath@mun.ca).

Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

REFERENCES

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. (2000a). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-423.

Baddeley, A. D. (2000b). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review*, **7**, 544-549.

Baddeley, A. [D.], & Salamé, P. (1986). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 525-529.

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 459-471.

Bridges, A. M., & Jones, D. M. (1996). Word dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing state? *Quarterly Journal of Experimental Psychology*, **49A**, 919-939.

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.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning & Verbal Behavior*, **15**, 17-31.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, **8**, 343-350.

Einstein, G. O., & McDaniel, M. A. (1987). Distinctiveness and the mnemonic benefits of bizarre imagery. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 78-102). New York: Springer.

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.

Hanley, J. R., & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology*, **78**, 287-297.

Hellbrück, J., Kuwano, S., & Namba, S. (1996). Irrelevant background speech and human performance: Is there long-term habituation? *Journal of the Acoustical Society of Japan*, **17**, 239-247.

Hsiao, A. T., & Reber, A. S. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious.

In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 471-494). Thousand Oaks, CA: Sage.

Jones, D. M. (1993). Objects, streams, and threads of auditory attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (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, 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.: Erlbaum.

Jones, D. M., & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 103-133.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *Quarterly Journal of Experimental Psychology*, **44A**, 645-669.

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

Kjellberg, A., & Sköldström, B. (1991). Noise annoyance during the performance of different nonauditory tasks. *Perceptual & Motor Skills*, **73**, 39-49.

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.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory & Language*, **27**, 382-398.

Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 578-584.

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.

Neath, I., Farley, L. A., & Surprenant, A. M. (2003). Directly assessing the relationship between irrelevant speech and articulatory suppression. *Quarterly Journal of Experimental Psychology*, **56A**, 1269-1278.

Salamé, P., & Baddeley, A. (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.

Salamé, P., & Baddeley, A. (1990). The effects of irrelevant speech on immediate free recall. *Bulletin of the Psychonomic Society*, **28**, 540-542.

Shanks, D. R., & Channon, S. (2002). Effects of a secondary task on "implicit" sequence learning: Learning or performance? *Psychological Research*, **66**, 99-109.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 592-604.

Smith, A. (1989). A review of the effects of noise on human performance. *Scandinavian Journal of Psychology*, **30**, 186-206.

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.

Watkins, M. J., LeCompte, D. C., & Kim, K. (2000). Role of study strategy in recall of mixed lists of common and rare words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 239-245.

NOTES

1. Although these tasks are usually called *implicit sequence learning*, we omit the term *implicit*, since we focus on whether the task is susceptible to disruption by irrelevant speech, rather than on the extent to which the task is truly implicit.

2. Although irrelevant speech can affect performance on tasks that do not emphasize seriation, such as reading, the effects of the irrelevant speech are different from those in serial recall tasks; for example, the semantic features of the irrelevant speech affect its ability to cause disruption (e.g., Martin, Wogalter, & Forlano, 1988).

3. We chose to have the subjects respond to the identity of the targets presented in a common location, rather than to a common stimulus presented at varying locations, because the former task allows the subjects to focus attention consistently on just one part of the visual field.

4. A 2 (group: pattern or random) × 2 (background condition: presence vs. absence of irrelevant speech) × 9 (block) ANOVA on RTs for correct responses yielded only two significant results: a main effect of block [$F(8,928) = 51.25$, $MS_e = 2,702$] and an interaction between block and group [$F(8,928) = 9.52$, $MS_e = 2,702$]. Only the first 9 blocks were included in this and subsequent analyses, because Block 10 was the transfer block.

APPENDIX

Analysis of Accuracy Data From Experiment 3

Effect of irrelevant speech. For the pattern conditions, a 2 (background condition: presence vs. absence of irrelevant speech) × 9 (block) ANOVA on the proportion of correct responses yielded no reliable main effects and no interaction [for the main effect of background condition, $F < 1$; for the main effect of block, $F(8,464) = 1.28$, $MS_e = 0.0011$, $p > .25$; for the interaction, $F(8,464) = 1.42$, $MS_e = 0.0011$, $p > .15$]. The results were identical for the random condition [for the main effect of background condition and the interaction, $F < 1$; for the main effect of block, $F(8,464) = 1.09$, $MS_e = 0.0009$, $p > .30$].

Learning. For the pattern groups, a 2 (background condition: presence vs. absence of irrelevant speech) × 2 (block: accuracy on Block 10 vs. mean accuracy averaged over Blocks 9 and 11) ANOVA resulted in a reliable main effect of block [$F(1,58) = 14.068$, $MS_e = 0.0013$], with less accurate responding in the transfer block (.919) than in the surrounding blocks (.943). The other main effect and the interaction were not reliable (both Fs < 1). For the random groups, there was no effect of irrelevant sound [$F(1,58) = 1.79$, $MS_e = 0.0032$, $p > .15$] and no effect of block [$F(1,58) = 1.79$, $MS_e = 0.0005$, $p > .15$]. The interaction, however, was almost reliable [$F(1,58) = 3.86$, $MS_e = 0.0005$, $p = .054$]. Accuracy on the transfer block was lower in the irrelevant sound condition (.918) than in the other conditions, all of which had accuracy levels between .932 and .940.

Analysis of Accuracy Data From Experiment 4

Effect of irrelevant speech. A 3 (condition: quiet, unchanging, or changing) × 9 (block) ANOVA on the proportion of correct responses yielded a marginal effect of condition [$F(2,87) = 2.88$, $MS_e = 0.012$, $p = .061$]. The proportion of correct responses was .937 in the quiet condition, .928 in the unchanging condition, and .915 in the changing condition. The main effect of block was reliable [$F(8,696) = 3.27$, $MS_e = 0.001$]. The interaction was not reliable ($F < 1$).

Learning. For the pattern groups, a 3 (condition: quiet, unchanging, or changing) × 2 (block: accuracy on Block 10 vs. mean accuracy averaged over Blocks 9 and 11) ANOVA resulted in no effect of condition [$F(2,87) = 2.19$, $MS_e = 0.003$, $p > .10$]. The effect of block was reliable [$F(1,87) = 5.41$, $MS_e = 0.0007$], with less accurate responding in the transfer block (.912) than in the surrounding blocks (.921). The interaction was not reliable [$F(2,87) = 1.16$, $MS_e = 0.0007$, $p > .30$].

(Manuscript received June 4, 2004;
revision accepted for publication October 28, 2005.)