Directly Assessing the Relationship Between Irrelevant Speech and Irrelevant Tapping

Aimée M. Surprenant and Ian Neath
Memorial University of Newfoundland

Tamra J. Bireta
The College of New Jersey

David W. Allbritton
DePaul University

The acoustic confusion effect is the finding that lists of to-be-remembered items that sound similar to one another are recalled worse than otherwise comparable lists of items that sound different. Previous work has shown that concurrent irrelevant speech and concurrent irrelevant tapping both reduce the size of this effect, suggesting similarities between the two manipulations. The authors assessed the relation between irrelevant speech and irrelevant tapping by correlating the disruption each causes to recall of similar- and dissimilar-sounding items. A significant correlation was obtained, indicating a relation between the two. The results indicate that researchers should be sensitive to changes in the magnitude of the effects rather than focusing exclusively on the presence or absence of particular effects. Implications for the 3 major explanations of the irrelevant speech effect are discussed.

The acoustic confusion effect is the finding that lists of similar-sounding items are recalled worse than otherwise comparable lists of dissimilar-sounding items (Baddeley, 1966; Conrad, 1964; Conrad & Hull, 1964). It has long been established that concurrent articulation—repeatedly saying something out loud—eliminates the acoustic confusion effect when the to-be-remembered items are presented visually (Estes, 1973; Murray, 1968; Peterson & Johnson, 1971). Two other manipulations have been claimed to have a similar effect: irrelevant speech (Surprenant, Neath, & LeCompte, 1999) and irrelevant tapping (Saito, 1993). However, the former claim is not without controversy (see, e.g., Larsen & Baddeley, 2003), and the latter is still relatively unexplored (see below).

The purpose of this article is first to review the literature on the effects that irrelevant speech and irrelevant tapping have on the magnitude of the acoustic confusion effect; second, to examine the relation between irrelevant speech and irrelevant tapping; and third, to consider the implications of the literature review and the results of the experiment on the three major theoretical accounts of the irrelevant speech effect. For ease of exposition, we delay consideration of the various theories of irrelevant speech, concurrent articulation, and their relation to the acoustic confusion effect until the General Discussion section.

Aimée M. Surprenant and Ian Neath, Department of Psychology, Memorial University of Newfoundland; Tamra J. Bireta, Psychology Department, The College of New Jersey; David W. Allbritton, Department of Psychology, DePaul University.

Part of this research was conducted while Aimée M. Surprenant and Ian Neath were at Purdue University. We thank Saturo Saito for very helpful comments on an earlier version of this article and for providing the means from Saito (1994).

Correspondence concerning this article should be addressed to Aimée M. Surprenant, Department of Psychology, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X9. E-mail: asurpren@mun.ca

The Irrelevant Speech Effect

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 memory task (Colle & Welsh, 1976). This finding has attracted a great deal of attention as it is an example of cross-modal interference: Irrelevant auditory information interferes with visual to-be-remembered items. However, the effect is also observed with auditory to-be-remembered items (Hanley & Broadbent, 1987; LeCompte, 1996; Surprenant et al., 1999). Another reason that this apparently simple effect has garnered so much research is that there is no consensus regarding its cause (see, e.g., Neath, 2000, and responses by Baddeley, 2000, and Jones & Tremblay, 2000).

In one of the first demonstrations of the irrelevant speech effect, Colle and Welsh (1976) visually presented eight-item lists of consonants. The irrelevant speech was a passage from Franz Kafka's *Ein Hungerkunstler* in German, a language that none of the participants reported understanding. The speech was irrelevant in the sense that participants were instructed to ignore it and were assured there would be no subsequent test on it. In the irrelevant speech condition, performance was 12% worse than in the quiet control condition. Subsequent investigations showed that the degree of impairment is independent of the intensity of the irrelevant speech (Colle, 1980; Ellermeier & Hellbrück, 1998; Salamé & Baddeley, 1987; Tremblay & Jones, 1999) and does not diminish over repeated trials or sessions (Ellermeier & Zimmer, 1997; Hellbrück, Kuwano, & Namba, 1996; Tremblay & Jones, 1998). In general, neither phonological–acoustic similarity nor semantic similarity between the irrelevant stimuli and the to-be-remembered stimuli is needed to produce the effect (Bridges & Jones, 1996; Jones & Macken, 1995; LeCompte & Shaibe, 1997).

A final important characteristic of the irrelevant speech effect is the changing state effect, the finding that irrelevant auditory stim-

TABLE 1
Summary of the effects of irrelevant speech on the acoustic confusion effect (ACE)

Study	ACE-Q	ACE-IS	ISE	TBR	List Length
Colle & Welsh (1976)	0.18	0.03	0.18	L	8
Salamé & Baddeley (1986) Experiment 1	0.15	0.11	0.09	L	5
Experiment 2	0.16	0.13	0.09	L	6
Experiment 3	0.18	0.09	0.16	L	7
Experiment 4	0.05	–0.02	0.12	L	8
Jones & Macken (1995) Experiment 4	0.22	0.19	0.11	L	8
Surprenant et al. (1999) Experiment 1	0.19	0.04	0.10	W	8
Experiment 2	0.20	0.02	0.10	W	8
Experiment 3	0.27	0.15	0.23	L	8
Larsen & Baddeley (2003) Experiment 1	0.27	0.24	0.03	L	6
Experiment 2	0.30	0.19	0.14	L	6
Experiment 3	0.26	0.15	0.17	L	6
M	0.20	0.11	0.13		

Note. ACE-Q = acoustic confusion effect with no irrelevant speech; ACE-IS = acoustic confusion effect with concurrent irrelevant speech; ISE = irrelevant speech effect; TBR = to-be-remembered items, either letters (L) or words (W); Exp. = experiment.

uli that change over time produce more of a decrement than otherwise comparable stimuli that do not change (Beaman & Jones, 1997; Jones & Macken, 1995). For example, a single repeated item (e.g., B B B B) will produce less of a decrement than a sequence of different items (e.g., A B C D). A similar pattern holds for changes within an item (Hughes, Tremblay, & Jones, 2005).

Previous studies involving irrelevant speech and acoustic confusability vary considerably in terms of list lengths, overall level of recall, and the methods for measuring the effects in question. Following Logie, Della Sala, Laiacona, Chalmers, and Wynn (1996) and Neath, Farley, and Surprenant (2003), we present the various effects under consideration in terms of the difference between the two conditions divided by performance in the control or standard condition (see Neath & Surprenant, 2007, for a longer discussion of these types of scores). Thus, the irrelevant speech effect (ISE) is given by

$$ISE = \frac{Q - IS}{Q}, \tag{1}$$

where Q is mean performance in the quiet condition with dissimilar items and IS is mean performance in the irrelevant speech condition with dissimilar items. Similarly, the acoustic confusion effect (ACE) is given by

$$ACE = \frac{Dis - Sim}{Dis}, \tag{2}$$

where Dis is mean performance in the quiet condition with dissimilar items and Sim is mean performance in the quiet condition with similar items. The advantage of such a score over a simple difference score is that it takes overall level of recall into account.[1]

Table 1 shows published studies that include acoustically confusable stimuli and irrelevant speech in the same design. The experiments varied in terms of the number of items in each list (from five to eight) and the type of to-be-remembered materials (letters or words). In addition, many other details such as modality of presentation and response mode were different among the experiments. As can be seen from the table, the general pattern is clear: When both an acoustic confusion effect and an irrelevant speech effect are observed in the appropriate control conditions, the presence of irrelevant speech either eliminates or greatly reduces the magnitude of the ACE.

Two studies reported slightly atypical results. In Experiment 1 of Larsen and Baddeley (2003), there was no irrelevant speech effect ($ISE = 0.03$) and thus no effect of irrelevant speech on ACE. It should be noted, however, that the irrelevant speech in this experiment consisted of a single, repeating item (the word *two*). When unchanging items form the irrelevant auditory stimuli, irrelevant speech often has no disruptive effect (e.g., Jones, Madden, & Miles, 1992).

A second atypical result concerns Experiment 4 of Salamé and Baddeley (1986). Salamé and Baddeley did not observe an ACE in the quiet condition with eight-item lists, even though other researchers have reported such a result (e.g., Colle & Welsh, 1976). If there is no effect of acoustic confusability in the quiet condition, there can be no opportunity for irrelevant speech to have an effect. However, with the seven-item lists, there was an ACE in the quiet condition that dropped in the irrelevant speech condition. Finally, in their five- and six-item lists, the ACE decreased only slightly in the presence of irrelevant speech.

Table 1 clearly shows that the most common pattern of results observed when both an ACE and an ISE are observed in the

[1] Ideally, these scores should be calculated for each participant and then averaged. Because most studies do not report these values, however, we report scores calculated from group means. Scores calculated from group means might differ slightly from scores calculated from participant means. For example, in the experiment reported below, the value of ACE calculated for each participant in the no-tapping condition and then averaged was 0.361 compared with 0.374 when calculated from group means.

appropriate control conditions is the reduction or elimination of the ACE. Thus, irrelevant speech has a similar effect on the ACE as does concurrent articulation.

The Irrelevant Tapping Effect

Saito (1993, 1994) investigated a third type of manipulation that also seems to interact with the ACE. He was interested in looking at which facet of concurrent articulation interferes with the speech motor programs involved in recoding and rehearsal: the articulation itself or a more general motor timing program? He devised a nonverbal motor task that could be varied in complexity: tapping a finger to mimic the rhythm of a sequence of tones. He found that when the tapping task involved a complex rhythm, overall performance was substantially decreased. Saito (1993) argued that the tapping effect was not due purely to distraction because it interfered with the pattern of performance (discussed below), not just overall performance. Instead, Saito (1994) argued that complex tapping interferes with a generic timing mechanism that plays a role in articulation.

One can measure the amount of disruption caused by irrelevant tapping in a manner analogous to ISE and ACE. Thus, the irrelevant tapping effect (ITE) is given by Equation 3:

$$ITE = \frac{No\ Tap - Tap}{No\ Tap} \quad (3)$$

With this effect, it is important to specify the exact nature of the tapping task. For example, Saito (1993) reported one experiment in which participants saw six-item lists of dissimilar- or similar-sounding letters. There were two tapping conditions: simple and complex. In the simple tapping condition, participants pressed a button on a mouse once every 320 ms. In the complex tapping, or syncopated, condition, the same number of presses were required, but there were two rapid beats, or syncopations, in every 1,600-ms block. With simple tapping, the mean group ACE was unaffected, but with complex tapping the ACE was almost entirely eliminated. This turns out to be the general pattern, as shown in Table 2: Simple tapping has no effect on ACE; in contrast, complex tapping reduces or eliminates the ACE.

Saito (1994) expanded on his previous results with two experiments in which participants were asked to recall either dissimilar- or similar-sounding lists of six letters. In Experiment 1, complex tapping decreased the mean group ACE when participants tapped with their right hand but had no effect when participants tapped with their left hand. In Experiment 2, Saito controlled for the amount of practice given to each hand and used vocal rather than written responses. With these changes, the mean ITE was approximately twice as large as in Experiment 1. Moreover, the mean ACE dropped substantially with both left- and right-handed tapping.

Larsen and Baddeley (2003) examined both irrelevant speech and irrelevant tapping. In Experiment 1, they used simple tapping (pressing the same key twice per second), which, replicating Saito's (1993) findings, did not affect ACE. In Experiment 2, the same tapping tasks were used, but the timing was changed from twice a second to a syncopated pattern. Irrelevant complex tapping reduced ACE to zero. In Experiment 3, the participants were asked to tap the keys on the number keypad in a clockwise loop, but at temporally regular intervals. This form of tapping had no effect on the acoustic confusion effect.

Table 2 shows that the most common pattern reported when both an ACE and a complex ITE are observed is the reduction or elimination of the acoustic confusion effect. Thus, irrelevant tapping has a similar effect on the ACE as does concurrent articulation and irrelevant speech.

Summary of Existing Studies

The results shown in Tables 1 and 2 may be described as follows: With visually presented to-be-remembered items, the ACE is either eliminated or greatly reduced by irrelevant speech and by irrelevant complex tapping. The mean group ACE drops from approximately 0.2 to approximately 0.1 with irrelevant speech and from approximately 0.2 to less than 0.1 with complex

TABLE 2
Summary of the effects of irrelevant tapping on the acoustic confusion effect

Study	ACE-Q	ACE-T	ITE	TBR	List Length	Tapping
Saito (1993)	0.19	0.17	0.05	L	6	Simple
Experiment 1						
Experiment 2	0.19	0.03	0.31	L	6	Complex
Saito (1994)	0.17	0.06	0.18	L	6	Complex
Experiment 1						
Right						
Left	0.17	0.17	0.13	L	6	Complex
Experiment 2						
Right	0.22	0.08	0.29	L	6	Complex
Left	0.22	0.05	0.33	L	6	Complex
Larsen & Baddeley (2003)	0.27	0.27	0.12	L	6	Simple
Experiment 1						
Experiment 2	0.30	0.00	0.49	L	6	Complex
Experiment 3	0.26	0.32	0.13	L	6	Sequential
M (complex tapping)	0.22	0.07	0.31			
M (all)	0.22	0.13	0.25			

Note. ACE-Q = acoustic confusion effect with no irrelevant speech; ACE-T = acoustic confusion effect with concurrent irrelevant tapping; ITE = irrelevant tapping effect; TBR = to-be-remembered items, either letters (L) or words (W).

tapping. Empirically, then, it seems likely that the disruption caused by irrelevant speech and irrelevant tapping is due to a common cause. Given the finding of substantial individual differences in the size of the ISE (Ellermeier & Zimmer, 1997), it should be the case that individuals who are susceptible to the effects of irrelevant speech should also be disrupted by the irrelevant tapping task. In other words, the effects should be correlated. Experiment 1 tested this.

Experiment 1

Given that both irrelevant speech and irrelevant tapping have apparently similar effects on the ACE, this experiment was designed to measure the correlation between the two effects and also to compare the amount of disruption with the ACE caused by irrelevant tapping and by irrelevant speech. In Phase 1 of the experiment, the participants saw lists of similar- or dissimilar sounding letters. They were asked to tap a complex pattern during presentation of half of the lists. In Phase 2, the memory task was the same, but half the time participants heard irrelevant background speech during list presentation.[2]

Method

Participants. One hundred Purdue University undergraduates participated in exchange for credit in introductory psychology courses. All identified themselves as native speakers of American English.

Design. There were two phases. In Phase 1, there were three within-subjects variables: serial position, dissimilar- or similar-sounding letters, and the presence or absence of concurrent complex tapping. In Phase 2, there were again three within-subjects variables: serial position, dissimilar- or similar-sounding letters, and the presence or absence of irrelevant speech.

Stimuli. The stimuli in both Phase 1 and Phase 2 were the letters F K L M R Q and B D G P T V, a subset of those used by Colle and Welsh (1976). In Phase 1, a 50-ms 440-Hz tone was used to indicate the rhythm. In Phase 2, the irrelevant speech was a passage from *Die Wilden* by Franz Kafka spoken in German by a female speaker; this passage has previously been used to produce reliable ISEs (e.g., Neath et al., 2003).

Procedure. In Phase 1, participants were asked to recall the order in which they saw six dissimilar-sounding or six similar-sounding letters. Each letter was shown in uppercase for 500 ms in 28-point Helvetica. On half of the trials, a series of tones was played according to the scheme noted in Figure 1 of Saito (1994). A total of 15 tones were played on each trial, 5 of which were before the appearance of the first letter. The intervals between tone onsets were 400 ms, 200 ms, 400 ms, 200 ms, and 400 ms (which then repeated). Participants "tapped" by pressing the space bar on a computer keyboard. At the end of the list, 12 buttons appeared labeled, in alphabetical order, with the 12 letters. Participants were asked to click on the buttons to recreate the order in which the letters were shown (i.e., strict serial reconstruction of order). They were not allowed to change an answer once they had clicked on a button. There were 10 trials in each of the four conditions; the order of the trials was randomly determined for each participant.

Phase 2, which began approximately 5 min after the end of Phase 1, was almost identical to Phase 1 except that instead of tapping on half of the trials, participants heard irrelevant speech via headphones. Participants were informed that they would sometimes hear 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. There were 10 trials in each of the four conditions; the order of the trials was randomly determined for each participant.

Participants were tested individually and were informed they could take rest breaks at any point. An experimenter remained in the room to ensure compliance with the instructions.

Results

For all analyses, an alpha level of .05 was adopted. For correlations with $n = 100$, r values greater than .195 are significant according to a two-tailed test.

Phase 1. The data were analyzed with a 2 (conditions: no tapping vs. complex tapping) $\times$ 2 (similarity: dissimilar vs. similar) $\times$ 6 (serial positions) analysis of variance (ANOVA). There was a main effect of condition, with better performance in the no-tapping (0.569) than in the tapping (0.462) conditions, $F(1, 99) = 104.59$, $MSE = 0.065$. There was also a main effect of similarity, with better recall of dissimilar (0.629) than similar (0.401) items, $F(1, 99) = 300.22$, $MSE = 0.104$. The main effect of serial position, $F(5, 495) = 470.05$, $MSE = 0.030$, was due to a typical serial position curve, as shown in the left panel of Figure 1. **F1**

Of most theoretical importance was the reliable interaction between condition and similarity, $F(1, 99) = 16.18$, $MSE = 0.041$. This was due to a larger ACE in the no-tapping condition (0.699 vs. 0.438, a difference of 0.261) than in the tapping condition (0.560 vs. 0.365, a difference of 0.195).

The interaction between condition and position was not reliable, $F(5, 495) < 1$. The interaction between similarity and position was reliable, $F(5, 495) = 28.68$, $MSE = 0.022$, again because of larger differences at later serial positions. There was a reliable three-way interaction, $F(5, 495) = 4.81$, $MSE = 0.016$.

The ACE was calculated separately for the no-tapping (ACE-NT) and the tapping (ACE-T) conditions (using Equation 2). The mean ACE-NT was 0.361, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.108$). The mean ACE-T was 0.309, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.097$). There was a significantly smaller ACE with tapping than without tapping, $F(1, 99) = 4.36$, $MSE = 0.031$.

The ITE was calculated on the basis of performance with the dissimilar items (ITE-Dissimilar) and similar items (ITE-Similar) separately (using Equation 3). The mean ITE-Dissimilar was 0.196, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.063$). The mean ITE-Similar was 0.131, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.120$).

Phase 2. A similar analysis was conducted on the data from Phase 2. There was a main effect of condition, with better performance in the quiet (0.586) than in the irrelevant speech (0.481)

[2] The irrelevant speech phase always came second to retain compatibility with other correlational studies of irrelevant speech (e.g., Beaman, 2004; Neath et al., 2003).

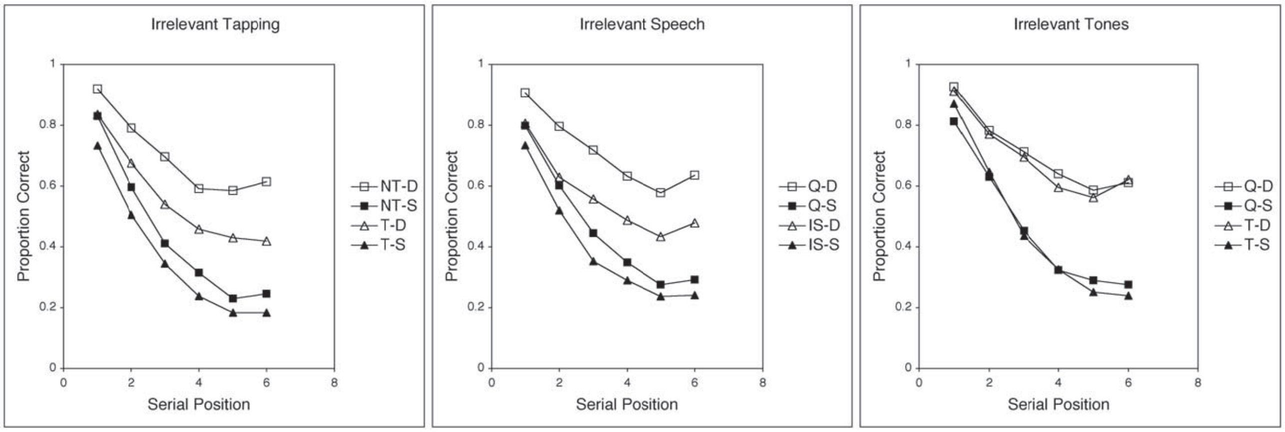

Figure 1. The proportion of dissimilar- and similar-sounding (D and S, respectively) items correctly recalled in Experiment 1 as a function of whether there was concurrent tapping (T) or no tapping (NT; left panel) or irrelevant speech (IS) or quiet (Q) (middle panel). The right panel shows the proportion of dissimilar- and similar-sounding items correctly recalled in Experiment 2 as a function of whether there were concurrent tones (T) or quiet (Q).

conditions, $F(1, 99) = 131.98$, $MSE = 0.050$. There was also a main effect of similarity, with better recall of dissimilar (0.638) than similar (0.428) items, $F(1, 99) = 190.59$, $MSE = 0.139$. The main effect of serial position, $F(5, 495) = 249.16$, $MSE = 0.432$, was due to a typical serial position curve, as shown in the middle panel of Figure 1.

As in Phase 1, there was a significant interaction between condition and similarity, $F(1, 99) = 23.40$, $MSE = 0.042$. This was due to a larger ACE in the quiet condition (0.711 vs. 0.460, a difference of 0.251) than in the irrelevant speech condition (0.566 vs. 0.396, a difference of 0.170).

The interaction between condition and position was marginally significant, $F(5, 495) = 2.18$, $MSE = 0.014$, $p = .055$, and due to a slightly larger difference between conditions at later postitions than at earlier positions. The interaction between similarity and position, $F(5, 495) = 27.11$, $MSE = 0.020$, was also due to larger differences at later positions. The three-way interaction was not reliable, $F(5, 495) = 1.26$, $MSE = 0.013$, $p > .28$.

The ACE was calculated separately for the quiet (ACE-Q) and the irrelevant speech (ACE-IS) conditions. The mean ACE-Q was 0.330, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.089$). The mean ACE-IS was 0.266, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.099$). These two values were significantly different, $F(1, 99) = 7.33$, $MSE = 0.029$.

The ISE was calculated on the basis of performance with the dissimilar items (ISE-Dissimilar). The mean ISE was 0.200, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.106$). ISE-Similar was also calculated: The value was 0.139, and the distribution did not differ from normal (Kolmogorov–Smirnov $d = 0.060$).

Correlations. Table 3 shows the correlation matrix. Of most interest, there was a significant correlation between ISE-Dissimilar and ITE-Dissimilar ($r = .343$). The correlation between ITE-Similar and ISE-Similar was also significant ($r = .201$). Thus, people who are affected by irrelevant tapping tend to be similarly affected by irrelevant speech. It is interesting to note that the magnitude of the irrelevant tapping and irrelevant speech effects were all quite comparable: ITE-Dissimilar = 0.196, ISE-Dissimilar = 0.200, ITE-Similar = 0.131, and ISE-Similar = 0.139. T3

TABLE 3
Correlations among the variables

Variable	1	2	3	4	5	6	7	8
1. ITE-D	—							
2. ISE-D	0.34	—						
3. ITE-S	0.18	0.01	—					
4. ISE-S	0.12	0.22	0.20	—				
5. ACE-NT	0.23	0.26	-0.33	0.21	—			
6. ACE-T	-0.36	-0.03	0.43	0.31	0.38	—		
7. ACE-Q	0.06	0.30	0.07	-0.07	0.52	0.42	—	
8. ACE-IS	-0.10	-0.24	0.16	0.47	0.45	0.60	0.57	—

Note. Values greater than 0.196 are significant at the .05 level by a two-tailed test. ITE = irrelevant tapping effect; ISE = irrelevant speech effect; D = dissimilar; S = similar; ACE = acoustic confusion effect; NT = no tapping; T = tapping; Q = quiet.

The various values involving ACE all correlated with one another, as would be expected. In particular, ACE-NT correlated with ACE-Q in the quiet condition (r = .521), and ACE-T correlated with ACE-IS (r = .603). One suggestion of a difference between ITE and ISE is revealed by correlations with ACE. ITE-Dissimilar (from Phase 1) did not correlate with ACE-Q (from Phase 2; r = .062), whereas ISE-Dissimilar (from Phase 2) did correlate with ACE-NT (from Phase 1; r = .258). Although irrelevant tapping and irrelevant speech may be related and have similar effects on memory, they are not identical and there may very well be differences between them.

The results from the correlational analyses outlined above should be interpreted with caution for two reasons. First, because testing was limited to one relatively short session, we do not have sufficient observations per participant to calculate the reliability of the measures given our data. However, Ellermeier and Zimmer (1997), using a similar method, showed reasonable test–retest reliabilities (r_{tt} = .45) for the irrelevant sound effect. Although one might assume a similar level of intraparticipant reliability for the irrelevant tapping effect, this remains to be demonstrated empirically. A second caution is in the use of difference scores in correlations. Difference scores tend to be less reliable than other measures and, of necessity, are restricted in their variance (Johns, 1981; Peter, Churchill & Brown, 1993). Again, an appropriate psychometric study would settle the issue. In the meantime, we note that the ANOVA results are clear: The same pattern of interactions of condition and similarity were found with the two manipulations, suggesting similar underlying mechanisms.

Experiment 2

One concern with interpreting the results of Experiment 1 is that it is possible that the tones used to indicate the to-be-tapped pattern caused the disruption rather than the tapping per se. This is unlikely for two reasons. First, a single repeating tone generally does not cause disruption (e.g., Jones & Macken, 1993). Second, two earlier studies (Saito, 1993, 1994) have included a control condition in which the tones indicating the tapping pattern were still heard, but there was no tapping. In these conditions, no disruptive effect was observed. Nonetheless, Experiment 2 was run as a control: It was identical to Phase 1 in Experiment 1 except that participants were not asked to tap.

Method

Participants. Fifty-eight undergraduates at The College of New Jersey and DePaul University participated in exchange for credit in introductory psychology courses. All identified themselves as native speakers of American English.

Design, stimuli, and procedure. The experiment was identical to Phase 1 of Experiment 1 except that no mention was made of tapping. Rather, participants were informed that on some trials they would hear beeps and that they should ignore them.

Results and Discussion

The data were analyzed with a 2 (conditions: no tones vs. syncopated tones) × 2 (similarity: dissimilar vs. similar) × 6 (serial positions) ANOVA. As the right panel of Figure 1 shows, there was no effect of the tones: The proportion of letters correctly recalled was 0.587 in the quiet condition compared with 0.577 in the tones condition, $F(1, 57) < 1$. There was a robust ACE, with better recall of the dissimilar-sounding items than the similar-sounding items (0.701 vs. 0.462), $F(1, 57) = 164.45$, MSE = 0.120. There was also the standard effect of position, $F(5, 325) = 202.91$, MSE = 0.037.

The only significant interaction was Similarity × Position, $F(5, 285) = 37.12$, MSE = 0.018, which was due to a larger difference at the end positions than at the earlier positions. Neither the interaction between tones and similarity, $F(1, 57) < 1$, nor the interaction between tones and position, $F(1, 57) = 1.66$, MSE = 0.018, were significant. The three-way interaction was the closest to the adopted significance level, $F(5, 285) = 2.12$, MSE = 0.013, $p < .07$.

Performance in the quiet dissimilar and quiet similar conditions in Experiment 2 was comparable to that observed in both phases of Experiment 1. However, there was no evidence that the irrelevant beeps affected performance at all. Therefore, it seems reasonable to conclude that the correlation between irrelevant speech and irrelevant tapping seen in Experiment 1 was not due to the tones used to mark the to-be-tapped pattern.

General Discussion

In Experiment 1, both irrelevant tapping and irrelevant speech reduced the magnitude of the ACE. This is consistent with the literature review presented in Tables 1 and 2 and similar to the effects of concurrent articulation. In addition, the results of Experiment 1 showed a significant correlation between the disruption caused by irrelevant speech and the disruption caused by irrelevant tapping. Although the correlational data must be interpreted as preliminary, pending confirmation of the reliabilities of the measures, the pattern of interaction between condition and the ACE is the same for both forms of irrelevant information. The data from Experiment 2 confirmed that the tones that delineate the to-be-tapped pattern played no role in the observed correlation.

There are three major theoretical accounts of the ISE: the phonological loop hypothesis (Baddeley, 1986; Larsen & Baddeley, 2003), the feature model (Neath, 2000; Neath et al., 2003), and the changing-state hypothesis (Jones, Beaman, & Macken, 1996; Jones & Tremblay, 2000). How well do these accounts handle these results?

Phonological Loop Hypothesis

The phonological loop hypothesis is based on Baddeley's (1986) working memory framework and rests on the assumption of separate storage and rehearsal processes. According to this view, concurrent articulation and irrelevant tapping are related, but both differ from irrelevant speech. Within the phonological loop, 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 items have automatic (and obligatory) access to the phonological store. Concurrent articulation prevents the participant from using articulatory rehearsal, and so (a) prevents recoding of visually presented information into a phonological form and (b) prevents maintenance rehearsal. Complex irrelevant tapping involves speech production mechanisms (Larsen & Baddeley, 2003), and so it should have similar effects as and be related to concurrent articulation. Both affect the magnitude of the ACE for visually presented items by reducing or blocking conversion of the visual information into a phonological form.

Larsen and Baddeley (2003, p. 1265) described “at least two ways of accounting for [the irrelevant speech effect] within the phonological loop model.” The first is to assume interference at the level of a feature rather than an item, but “such features operate at the phoneme or syllable level.” This idea was not well developed. The second, following Page and Norris (2003), is to assume that irrelevant speech adds noise to the association linking item and serial order information. However, phonological information is represented at a different level in the theory, that of the item. Irrelevant speech should therefore not interact with the ACE because they operate at different levels.

This account readily explains why both concurrent articulation and complex irrelevant tapping interact with the ACE. It does not explain (a) why irrelevant speech interacts with the ACE (see Table 1 and the results of Experiment 1) or (b) why there is a correlation between the magnitude of the disruption caused by irrelevant speech and irrelevant tapping (Experiment 1). Moreover, it does not address why there is a correlation between the magnitude of the disruption caused by irrelevant speech and concurrent articulation (see Neath et al., 2003). Quite substantive changes are needed to the phonological loop account to accommodate the full pattern of results.

The Feature Model

According to the feature model (Neath, 2000; Neath et al., 2003), items are represented in memory as a vector of features. Two factors contribute to the ISE: ignoring the irrelevant stimulus (an attentional factor) and interference between the to-be-remembered items and the irrelevant speech (feature adoption). Any stimulus that takes effort to be ignored (such as a fly buzzing about) will potentially reduce attention to the main task. Attention thus plays a role in a dual-task sense: The more effort required to ignore or not process the secondary stimulus, the more a participant’s performance should be reduced. 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 errors are likely to occur.

Irrelevant speech interacts with the ACE through feature adoption. The features that convey phonological information have a large probability of being replaced by the adopted features. This has the effect of removing the major difference between the similar and the dissimilar items (see Simulation 4v of Neath, 2000, for complete details). Concurrent articulation, according to the feature model, differs from irrelevant speech only in that it takes more attention to repeat a syllable out loud than to hear another voice repeating a syllable out loud (see Simulation 1 of Neath, 2000).

This view predicts that irrelevant speech will interact with the ACE (see Table 1 and the results of Experiment 1) and also predicts the correlation between irrelevant speech and concurrent articulation. It does not predict the interaction between irrelevant tapping and the ACE or the correlation between irrelevant tapping and irrelevant speech because the model has not yet been applied to irrelevant tapping. A relatively straightforward extension could remedy this. To the extent that irrelevant tapping requires attention (like irrelevant speech or concurrent articulation), overall disruption should occur. Moreover, if one makes the assumption that irrelevant complex tapping involves processes similar to those used in subvocal speech, then it might be the case that irrelevant tapping could generate modality independent features. Those features could then be involved in feature adoption and interact with acoustic confusability.

The Object-Oriented Episodic Record Model

According to the object-oriented episodic record model (Jones, 1993; Jones et al., 1996; Jones & Tremblay, 2000), the ISE is due to a conflict between serial order information from two different sources. Both visual and auditory items are represented using amodal, abstract representations, called objects. Serial order is encoded by the use of pointers that are associated with individual objects. 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 objects, such as those representing the list items. The changing state effect arises 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.

To the extent that the irrelevant stream contains dynamic changing items, the amount of interference will be greater. Concurrent articulation, irrelevant speech, and irrelevant complex tapping all provide a second source of order information that conflicts with recalling the order of the to-be-remembered items. Given the assumption that “serial order information is more robustly embodied in representations derived from changing state input than from those derived from nonchanging state input” (Jones & Macken, 1995, p. 114), one would also expect complex tapping to have a large effect but simple tapping to have almost no effect.

Summary

The results, then, most strongly support the object-oriented episodic record model. If one accepts the argument that irrelevant tapping involves some kind of subvocal processing and that the processing generates modality-independent features, then the feature model can accommodate the results. The least well-supported view is the phonological loop hypothesis. According to that view, irrelevant speech is fundamentally different from concurrent articulation and irrelevant tapping, and moreover, irrelevant speech should not interact with acoustic confusability.

The pattern of results both observed in the literature review and found in the experiment suggest that rather than concentrating on the presence or absence of particular effects, perhaps more attention should be placed on changes in the magnitude of the effects. Here, we found differences in both overall magnitude and relative magnitude. Although we do not suggest abandoning Fisherian-type designs that test primarily whether an effect is present or absent, such research should be complemented by studies in which changes in the magnitude of the effects are examined and individual differences variables are studied. Such an approach has led to the conclusion that the three major forms of distraction in short-term working memory tasks—concurrent articulation, irrelevant speech, and irrelevant complex tapping—all have similar effects on memory for acoustically confusable items.

References

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

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

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

Beaman, C. P. (2004). The irrelevant sound phenomenon revisited: What role for working memory capacity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1106–1118.

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, and 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: Human 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 and Verbal Behavior, 19,* 722–735.

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

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55,* 429–432.

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 and Performance, 24,* 1406–1414.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech effect." *Journal of the Acoustical Society of America, 102,* 2191–2199.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior, 12,* 360–372.

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.

Hughes, R. W., Tremblay, S., & Jones, D. M. (2005). Disruption by speech of serial short-term memory: The role of changing-state vowels. *Psychonomic Bulletin & Review, 12,* 886–890.

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance, 27,* 443–463.

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, England: Oxford University Press.

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. 209–238). Hove, England: 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, and Cognition, 19,* 369–381.

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, and Cognition, 21,* 103–115.

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: Human 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,* 544–558.

Larsen, J. D., & Baddeley, A. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56A,* 1249–1268.

LeCompte, D. C. (1996). Irrelevant speech, serial rehearsal, and temporal distinctiveness: A new approach to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1154–1165.

LeCompte, D. C., & Shaibe, D. M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A,* 100–118.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory & Cognition, 24,* 305–321.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78,* 679–684.

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: Human Experimental Psychology, 56A,* 1269–1278.

Neath, I., & Surprenant, A. M. (2007). Accounting for age-related differences in working memory using the feature model. In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory: Behavioural and neural correlates* (pp. 165–179). Oxford, England: Oxford University Press.

Page, M. P. A., & Norris, D. G. (2003). The irrelevant sound effect: What needs modelling, and a tentative model. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56A,* 1289–1300.

Peter, J. P., Churchill, G. A., & Brown, T. J. (1993). Caution in the use of difference scores in consumer research. *Journal of Consumer Research, 19,* 655–662.

Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning and Verbal Behavior, 10,* 346–354.

Saito, S. (1993). The disappearance of phonological similarity effect by complex rhythmic tapping. *Psychologia, 36,* 27–33.

Saito, S. (1994). What effect can rhythmic finger tapping have on the phonological similarity effect? *Memory & Cognition, 22,* 181–187.

Salamé, P., & Baddeley, A. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24,* 263–265.

Salamé, P., & Baddeley, A. (1987). Noise, unattended speech and short-term memory. *Ergonomics, 30,* 1185–1194.

Surprenant, A. M., Neath, I., & LeCompte, D. C. (1999). Irrelevant speech, phonological similarity, and presentation modality. *Memory, 7,* 405–420.

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, and Cognition, 24,* 659–671.

Tremblay, S., & Jones, D. M. (1999). Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1005–1015.