Interference from degraded auditory stimuli: Linear effects of changing-state in the irrelevant sequence

Dylan M. Jones,a) David Alford, William J. Macken, Simon P. Banbury, and Sébastien Tremblay
School of Psychology, Cardiff University, Cardiff, United Kingdom

(Received 24 May 1999; revised 5 October 1999; accepted 9 June 2000)

Cognitive performance, particularly on a number of tasks involving short-term memory for order, is impaired by the mere presence of irrelevant background sound. The current study examines the features of the irrelevant sound that determine its disruptive potency. Previous research suggests that the amount of variability in an irrelevant stream is related to the degree of disruption of memory. The present experiments used a parametric approach to manipulate degree of change more precisely. Increasing levels of degradation, effected either by low-pass filtering (speech) or by digital manipulation (speech and nonspeech), monotonically decreased the degree of interference. The findings support the following propositions: (i) the degree of physical change within an auditory stream is the primary determinant of the degree of disruption; and, (ii) the effects of irrelevant speech and irrelevant nonspeech sounds are functionally similar. © 2000 Acoustical Society of America. [S0001-4966(00)02809-5]

PACS numbers: 43.50.Qp [DWG]

I. INTRODUCTION

Exposure to background sound substantially impairs memory for order (e.g., Colle and Welsh, 1976). The most notable feature of this effect is that it occurs reliably even though individuals are fully aware that the background sound is irrelevant and are instructed to ignore it. In a typical irrelevant sound experiment, a series of digits or letters is presented sequentially on a computer screen for recall at the end of the sequence presentation or after a short (typically 10 sec) delay. Sound, either speech or nonspeech, presented during such tasks appreciably disrupts recall of the sequence. The robustness of disruption is evidenced by the frequency with which the effect has been replicated and by the high proportion of individuals sensitive to the effect (see Ellermeier and Zimmer, 1997). Not only can the effect reach up to a 50% decrement but it is also reliable; typically disruption endures over trials (e.g., Tremblay and Jones, 1998) and is just as marked when tested in experimental sessions separated by up to a week (e.g., Ellermeier and Zimmer, 1997; Hellbrück et al., 1996). However, a substantial amount of research indicates that some sounds are more disruptive than others (e.g., LeCompte et al., 1997), and also that some tasks are more prone to disruption than others (e.g., Beaman and Jones, 1997). The functional character of the task is not the subject of the present study; instead, interest centers on the claim that the degree of change in the irrelevant sequence is the primary source of disruption.

Neither the semantic nor the physical similarity of the irrelevant sound to the to-be-remembered material seems to be a major determinant of the effect. It seems fairly certain that the semantic similarity of the irrelevant material to the items being rehearsed in memory is not an important factor.

Numerous attempts to find effects of meaning have failed (e.g., Buchner et al., 1996; Jones et al., 1990; LeCompte and Schaibe, 1997) and those that have succeeded only found small and marginally significant effects (e.g., Neely and LeCompte, 1999). Also, there is mounting evidence that physical similarities between the stimuli in the irrelevant stream and the rehearsed material are not a major determinant of disruption. Given that the material rehearsed in memory is usually words or syllables, one might expect that disruption would be brought about only by verbal material such as words or syllables. But it has been shown that speech is neither a necessary (e.g., Salamé and Baddeley, 1989) nor a sufficient (e.g., Jones et al., 1992) condition for substantial disruption to occur. The key determinant of disruption seems to be the degree of change within the irrelevant sound sequence. The prime example of this is that presenting a sequence of different tokens such as letters or syllables results in substantially more disruption of serial recall than when a single token is repeatedly played (e.g., Jones et al., 1992; Jones and Macken, 1995a; LeCompte, 1995). This effect of variation applies for nonverbal and less discrete stimuli also. Tones varying either in frequency or timbre are more disruptive than a repeated tone (e.g., Jones and Macken, 1993) and instrumental music with many sharp changes in energy tends to be more disruptive than music with slow changes (Klatte et al., 1995). The notion of changing-state can be defined as the amount of acoustic variation in frequency, timbre, tempo but not in sound level (see Tremblay and Jones, 1999).

In the experiments that follow, we attempted to show how the degree of change in the auditory stream is related to the disruption of serial recall. One way of testing the concept of changing-state is by selecting naturally occurring stimuli and arranging them on a continuum which is supposed to map onto variability. So, if speech tokens are chosen to rhyme with one another, then the degree of disruption is markedly less than if the tokens are phonologically dissimi-
lar (Jones and Macken, 1995a). However, this approach has a number of drawbacks: First, it is far from clear what criteria should be employed in the selection of stimuli to represent the various levels of variability, and, second, it would be difficult to assign a value to each stimulus that could be arranged along a continuum. Overcoming this last drawback would be particularly desirable given that parametrically derived relationships between variables are more informative in analytical terms than a simple contrast between two experimental conditions.

An alternative method is to construct sequences embodying different degrees of change by using a fixed set of stimuli and reducing by degrees the extent to which they change. The general approach adopted in the current study was to systematically degrade the spectral character of changing sequences. By taking a stimulus sequence known to produce an irrelevant sound effect, and systematically degrading it so that the amount of variation within it is gradually reduced, it should be possible to establish a function relating the degree of change to the disruption of memory for order. A similar approach was used by Ellermeier and Hellbrück (1998). They used changes in the signal-to-noise ratio of the irrelevant sequence as the device for degrading the stimulus. That is, they manipulated the degree of change within a sequence by adding more or less noise to a speech sequence. As might be expected on the basis of the changing-state concept, when the signal-to-noise ratio became smaller, so the degree of disruption diminished; moreover, this relationship was monotonic. Abstractly, this is explained by the fact that sequences of noise-masked stimuli have fewer prominent acoustic features; therefore, the degree of variation is reduced gradually as the signal-to-noise ratio diminishes.

However, the technique of using a noise mask is not without its drawbacks (as Ellermeier and Hellbrück, 1998, acknowledge) insofar as the masking effect might be plausibly one primarily of reducing the perceivable number of tokens (distinctive acoustic events) within a stream. This would happen if some of the tokens tended to be more susceptible to masking than others so that, as the signal-to-noise ratio is decreased, they would be reduced to a level below audibility within the noise, with gradually fewer tokens becoming perceptible. Thus, the beneficial effect of adding the noise could be one of reducing the number of tokens, and not one of the relative degrees of change. There is strong evidence to suggest that the degree of disruption is related to the number of tokens per time unit in the irrelevant sequence (referred to as the token-dose effect; see Bridges and Jones, 1996). For example, if the rate at which a changing irrelevant sound sequence is presented is increased, the degree of disruption also increases. However, there is no such effect for a repeated irrelevant sequence.

II. EXPERIMENT 1

In experiment 1, low-pass filtering was used to degrade a stimulus sequence to different degrees (so that some sequences of sound become less distinct). It was predicted that as the degree of filtering increased so the disruption would diminish. The filter is set in such a way as to allow frequencies around the fundamental frequency of the speech sequence to pass unaffected. The degree of degradation of the stimulus by the filter was set by manipulating the rate of roll-off of the filter which served to attenuate progressively the frequencies higher than the fundamental. Steeper rates of roll-off (such as 24 dB per octave) would attenuate many more of these frequencies than shallower degrees of roll-off (say, 6 dB per octave). Phenomenally, the effect of the roll-off becoming steeper would be that the differences between elements in the sound sequence would become less distinct. Importantly, with the settings used in experiment 1, the sound was never completely obliterated even at the steepest rates of roll-off; hence, the token dose remained constant over all levels of roll-off thereby circumventing the difficulty to which the masking technique of Ellermeier and Hellbrück (1998) is prey. The prediction was that, as the roll-off became increasingly steep, the degree of disruption would diminish; that is, the relation between the degree of roll-off and disruption would be inverse and monotonic.

III. METHOD

A. Participants

Twenty-five Cardiff University students each reporting normal hearing and normal or corrected-to-normal vision, received a small honorarium for participating in the experiment.

B. Apparatus and stimuli

1. Visual stimuli for the memory task

Lists of items to be recalled were presented serially on the screen of an Apple Macintosh Performa 6200 microcomputer. Lists were constructed from the random arrangement of the nine letters f, k, l, m, q, r, s, t, and y, with the constraints that a letter could not appear in the same serial position in two consecutive lists and that recognizable words, acronyms, or letter strings were excluded.

2. Auditory stimuli

The words—turn, kilt, band, jaw, fruit, rod, porch, pier, nest, cow, pin, hat, boat, top—were recorded digitally in a male voice, spoken at an even pitch. The person producing the words was asked to produce them at a pitch matching that of a reference pure tone of 103 Hz (A♭2) which was played periodically to him as a reminder throughout the recording session. During the preparation of the materials, any word not matching this pitch was re-recorded. The words were edited using digital signal processing techniques to last 500 msec with a 200-msec inter word interval. The sequence was then band pass filtered using a Kemo Dual Variable Filter Type VBF/14. For each of the auditory conditions, the center frequency of the filter was set at 85 Hz and the lower bound of the filter (the high-pass element) was fixed at a roll-off of 24 dB/octave. Only the rate of attenuation of the frequencies above the center frequency (the low-pass component) was manipulated across auditory conditions within the experiment. The roll-off values for the low-pass component were set at either 0, 6, 12, 18, or 24 dB/octave. The
value of the center frequency of the filter—at 85 Hz, lower
than the fundamental frequency of the speech—was arrived
at following pilot trials to meet the criterion that the lower
degrees of roll-off (6 dB/octave vs 0 roll-off) produced a
perceptible difference in attenuation, at least when judged in
focal attention. The range of values of roll-off used in the
study was constrained by the values available on the Kemo
filter.

The stimuli were recorded on digital audio tape, and
transferred thereafter to Sound Designer digital editing soft-
dware where the amplitude was digitally edited to constant
root-mean-square (rms) value for all stimuli. Generally, this
meant that the stimuli were broadly similar in loudness.
There is a substantial body of evidence showing that both the
overall level of intensity (e.g., Ellermeier and Hellbrück,
1998) and the degree of change in intensity (Salamé and
Baddeley, 1989; Tremblay and Jones, 1999) do not influence
the degree of disruption of serial recall. All samples were
recorded to 16-bit resolution at a sampling rate of 44.1 kHz,
stored digitally, and presented during experimental trials
within a SuperCard environment.

C. Design

A repeated-measures design was used, all participants
undertaking the recall task under all five auditory conditions.
There were 75 trials in all, 15 for each level of roll-off.
Presentation of the conditions was quasirandomized from
trial to trial, such that every condition was presented before
any one was repeated.

D. Procedure

Participants were tested individually, seated in a sound-
proof laboratory approximately 0.5 m from the computer
screen. At the outset, standard instructions were read by par-
ticipants. These informed them of the nature of the recall
task, asked them to ignore any sounds they might hear, and
reassured them that the content of the auditory material
would not be tested. In each trial, the nine letters were dis-
played in random order as described above. Each letter was
displayed for 0.5 sec with an interstimulus interval also of
0.5 sec. Every list was preceded by a warning tone. After
presentation of the last item in the list, the word “wait” was
displayed on the screen for 10 sec during which the subject
was expected to rehearse the list. After this retention interval,
once the word “wait” disappeared, the subject was required
to recall the list in strict serial order. Sound was played via
headphones (Sony CD250 Digital Reference) throughout
both presentation and retention of the lists at a level of ap-
approximately 60 dB(A). Responses were written on a blank
grid comprising rows of nine boxes corresponding in order to
the to-be-remembered sequence. The experimental trials
were preceded by a short practice session, and the experi-
ment lasted approximately 45 min.

IV. RESULTS

The results were scored to a strict serial recall criterion.
Each item in the recalled sequence was scored as correct only if it corresponded exactly to the item at that position in

the to-be-remembered sequence. The main feature of the re-

sults is depicted in Fig. 1, which shows that as the degree of
filtering increased the degree of disruption diminished; more-
over, this relationship is monotonic. The error percentages
were subjected to a two-way repeated measures analysis of
variance, which revealed significant main effects of serial
position (nine positions), $F(8,192) = 30.81, MSE = 16.77, p
<0.001$, of auditory condition (five levels of roll-off),
$F(4,96) = 4.91, MSE = 7.84, p < 0.01$, and also of their in-
teraction, $F(32,768) = 1.86, MSE = 2.50, p < 0.05$. No func-
tional significance is attached to the interaction between se-
rial position and auditory condition since it may simply
reflect a ceiling effect in recency (analysis of serial position
data revealed the same pattern in experiment 2 but are not
reported).

Further scrutiny of the significant effect of auditory con-
ditions shows that, using a test of orthogonal polynomials,
the trend has a strongly linear component, $F(4,96) = 19.34,$
$p < 0.0001$, and, as expected, a nonsignificant quadratic com-
ponent, $F(4,96) = 0.19, p > 0.05$. It could be argued that the
linear function is the result of averaging many nonlinear
functions. Given the relatively great range of roll-offs em-
ployed, the average of nonlinear functions should have led to
more variability in the middle range, however, visual inspec-
tion of the error bars in Fig. 1 (representing standard error)
clearly shows that the variability is similar across the range
of roll-off points.

V. DISCUSSION

The monotonic relation between the degree of degrada-
tion in the stimulus and the degree of disruption points to the
fact that the effect of changing-state is continuous; there is no evidence of a cutoff point above which the disruption becomes more pronounced. That is, there is no evidence here of a signal level which is necessary for the effect to occur, as one might suppose if the effect is dependent upon the categorical perception of the words or their constituent phonemes, for example. This result supports the suggestion that the effect of irrelevant sound is not dependent on the identity of the sounds but that identity semantic, lexical, or phonological.

Although the filtering technique is one that provides clear results for the purposes of supporting the concept of changing-state, its main shortcoming is that it is difficult to apply when attempting to make a comparison across several different classes of stimuli varying greatly in complexity. For example, in making a comparison of the degradation in a speech sequence with that in a sequence of notes from a musical instrument, it would be necessary to adopt vastly different degrees of roll-off in each case (given that the distribution of the energy across the audio spectrum is far less wide in the case of a typical musical instrument). Moreover, the fundamental frequency of each irrelevant stimulus would have to be the same. Given these differences, it might be difficult to be sure that very different ranges of value of roll-off have equivalent effects with each stimulus. Ideally, it would be useful to use a technique that would employ a common metric for the degree of degradation with different classes of stimuli.

We attempted to use just such a technique in experiment 2; instead of low-pass filtering, we employed a method involving a digital signal processing technique in which an intact stimulus (say, a spoken word) is sampled digitally and then the polarity of each of the sample points constituting the stimulus is reversed with a certain probability. By systematically varying the likelihood of reversal, a range of stimuli may be generated that bear different degrees of resemblance to the original token, at one extreme being composed entirely of amplitude-modulated noise, for which the likelihood of polarity reversal is .5, through intermediate stages in which there are approximations to the word (accompained by noise), to the other extreme in which none of the samples is changed and the stimulus is in its original form. Of course, this interference function can in theory be applied to any auditory stimulus with a common metric, namely the probability of a change in polarity. This allows a contrast of the effect of speech and nonspeech stimuli to be made. This is what we attempted in experiment 2, by comparing speech tokens with tokens of a cello note played at different frequencies.

In phenomenal terms the steps in the degradation of the stimulus will give rise to progressive lack of clarity of the stimulus, accompanied by a loss of lexical and phonological identity to the point where only very little of the original stimulus can be discerned. Eventually, when the stimulus is composed entirely of amplitude-modulated noise, a noise burst will be heard. In the case of the cello note, the amplitude of each token is the same, and hence the noisiest version of the transformation will be the same for each token. For the words, there will be small variations in the pattern of noisiness, reflecting minor variations in the amplitude envelopes of the words that served originally as the basis for the transformations.

The purpose of experiment 2 was to explore the shape of the interference function relating degradation of the stimulus sequence and serial recall for different classes of stimuli. It was predicted that the function would be monotonic and general; that is, the degree of disruption would increase as the degree of change increased, and this would be true for both speech and nonspeech (instrumental notes).

VI. EXPERIMENT 2

In experiment 2, interference functions were established for nonspeech (cello notes) and speech (syllable sequences) stimuli. Again, the concept of changing-state makes the prediction that both will have monotonic interference functions, but that, because there are more changing elements in the speech than in the cello note, the slope of the function will be steeper for speech than for cello notes. Six levels of noisiness were compared.

VII. METHOD

A. Participants

Forty-seven participants volunteered to take part in the experiment for a small honorarium. All reported normal hearing and vision.

B. Apparatus and stimuli

1. Visual stimuli

The short-term memory task was the same as that used in experiment 1.

2. Auditory stimuli

Two types of stimuli were employed: cello notes and spoken single-syllable words. Seven words—pier, nest, cow, pin, hat, boat, and top—were recorded digitally in a male voice. Digital samples of a plucked cello note were taken from a compact disk containing musical samples (Digidesign SampleCell). Twelve samples of the plucked cello note covering the octave C4 to B4 served as the basis for the transformations. Words and cello notes were edited to last approximately 500 msec with a 100-msec interstimulus interval. From each of the speech and cello note samples, five additional versions were generated by digital signal processing to form a set of transformed samples along a continuum of noisiness. This was achieved by first representing the sample digitally (with a sampling rate of 22 kHz) and then reversing the polarity of each sample point with different probabilities (Schroeder, 1968; Repp and Frost, 1988). In essence, this changed the proportion of the sample that consisted of noise. As the probability of changing sample points increased, so did the proportion of noise in relation to the signal increase, but the length of the sample and its amplitude envelope remained the same. Six levels of degradation of the original sample were prepared: 0%, 20%, 40%, 60%, 80%, and 100% (corresponding to probabilities of reversing polarity of 0, 0.1, 0.2, 0.3, 0.4, and 0.5). These values en-
The results may be summed up succinctly: As stimulus degradation was reduced, the effect of interference from the auditory stimulus increased, and this effect was significantly linear. Just as in experiment 1, the form of these results is interesting, in that there is no discontinuity in the effect of degradation, as one might expect if there were categorical processes at work, such as the emergence of meaning or the detection of lexical, phonological, or other sublexical units. There is, in other words, no cutoff point that a degraded stimulus must exceed in order to become classified as a word before the disruption becomes marked. This monotonic function can be explained by the concept of changing-state. Simply, as more of the signal becomes available, the amount of variation will be greater and so the degree of disruption will also increase.

IX. DISCUSSION

The two factors—degradation (six levels in the range 0% to 100% noise) and source (cello speech)—were entered into a repeated-measures analysis of variance (ANOVA). The disruption produced by speech was greater than that produced by cello notes, $F(1,46) = 5.63, \text{MSE} = 23.09, p < 0.02$. The main effect of percent noise was highly significant, the level of disruption decreasing markedly as the proportion of noise increased, $F(5,230) = 13.56, \text{MSE} = 0.92, p < 0.0001$. The two factors also interacted significantly, $F(5,230) = 2.37, \text{MSE} = 0.92, p < 0.05$. The results are broadly in line with predictions. Speech showed more disruption than nonspeech, and the degree of disruption rose more rapidly as the percentage of noise decreased. In separate within-participants ANOVAs, each function produced a significant fit to linear polynomials, but not to higher-order polynomials [for cello, linear contrast: $F(1,46) = 13.72, p < 0.001$; quadratic contrast: $F < 1$; for speech, linear contrast $F(1,46) = 56.25, p < 0.001$; quadratic contrast: $F < 1$].

VIII. RESULTS

As in experiment 1, participants’ responses were scored according to the strict serial order criterion. The main features of the results are illustrated in Fig. 2. Errors decreased as the level of noise increased; the slope of this improvement was less steep for the cello notes than for the speech stimuli. The results may be summed up succinctly: As stimulus degradation was reduced, the effect of interference from the auditory stimulus increased, and this effect was significantly linear. Just as in experiment 1, the form of these results is interesting, in that there is no discontinuity in the effect of degradation, as one might expect if there were categorical processes at work, such as the emergence of meaning or the detection of lexical, phonological, or other sublexical units. There is, in other words, no cutoff point that a degraded stimulus must exceed in order to become classified as a word before the disruption becomes marked. This monotonic function can be explained by the concept of changing-state. Simply, as more of the signal becomes available, the amount of variation will be greater and so the degree of disruption will also increase.
ruptive than fully degraded nonspeech, may, however, be taken as evidence to the contrary. One possible explanation of this difference in disruptive potency is that the fully degraded speech sounds still contain some spectral changes from token to token, given the different amplitude envelopes associated with the different speech tokens. While previous studies have shown that amplitude-modulated broadband noise is not disruptive of recall (e.g., Salamé and Baddeley, 1989), the stimuli used in those studies may not have been perceptually segmentable into discrete tokens, a condition necessary for disruption to occur with any signal (Jones and Macken, 1993). However, the 100-ms silent interval between tokens used in experiment 2 is likely to have been sufficient for such segmentation to occur. In any case, although the levels of performance were different with cello notes and with speech, they seemed to be related to the degree of degradation of the irrelevant sequence in the same lawful way.

A similar argument about the functional equivalence of speech and nonspeech sounds can be derived from other factors such as the way in which token set size and disruption are related for tones and speech (Tremblay and Jones, 1998). That is, both speech and tones show the same relation between disruption and token set size: There is an increase in error as the number of distinguishable sounds increases from one to two, but there is very little increase in error as set size increases thereafter. Differences in terms of the relative degree of disruption for speech and nonspeech are sometimes significant (e.g., LeCompte et al., 1997) and in some other studies nonsignificant (e.g., Jones and Macken, 1993). The interpretation of such differences depends in part on how one regards the comparison between simple speech sounds and simple nonspeech sounds. Of course, at one level, this comparison is of doubtful value, given that in acoustic terms speech sounds that are nominally simple are typically more complex than nonspeech sounds.

While the studies reported here clearly demonstrate a linear relationship between the amount of change within an auditory sequence and its power to disrupt recall, the precise mechanism of this effect is perhaps less clearly established. In particular, an as-yet unanswered question is whether change is important as a global property of a sound sequence, or operates on a local, token-to-token mismatch basis. Results from Jones et al. (1992) suggest that global properties are not the critical determinants of disruption, at least insofar as predictable sequences of tokens (e.g., A, B, C, D, A, B, C, D, etc.) are no less disruptive than randomly changing, unpredictable sequences (e.g., A, C, D, B, C, A, B, D, etc.). Further, there is evidence pointing to the fact that change is derived from information on the basis of two immediately successive stimuli: The largest increase in disruption occurs when the set of distinguishable tokens goes from one repeated token to two distinct tokens, with little or no further increase in disruption from sequences containing larger numbers of tokens (see above; Tremblay and Jones, 1998). However, the effects of manipulating the rate of transition between different tokens in the irrelevant sound are not in line with this mismatch hypothesis. A low rate of transition, at which the token was changed every three stimuli (e.g., AAABBBCCC), should produce less disruption than a sequence of irrelevant sound in which the transition is between immediately successive tokens (e.g., ABC). However, these two conditions do not produce significantly different degrees of disruption (see experiment 4 of Tremblay and Jones, 1998). Whether the effect of changing-state is calculated over adjacent tokens remains to be further investigated.

One intuitively appealing and parsimonious account of the irrelevant sound effect is that repeated tokens (or unvarying continuous sound) produce less disruption because by virtue of their repeated presentation they do not produce an attentional response. Based on the concept of the orienting response (see Sokolov, 1963), it was suggested that irrelevant sound recruits attention away from the task at hand but repetitive stimuli cause much less disruption because the attentional response to the sound attenuates as a result of repeated presentation (see Cowan, 1995). However, a number of lines of empirical evidence suggest that habituation plays little or no role in the irrelevant sound effect. First, a key prediction of the habituation framework is that the degree of disruption should increase as the number of different tokens in the irrelevant sound sequence increases. With few tokens in a sequence, each token is repeated more often than if there are many, and therefore habituation should be brought about more rapidly. However, as mentioned earlier, it has been shown that disruption increases sharply when the set of distinct tokens increases from one to two, but beyond two tokens the degree of disruption does not increase significantly (Tremblay and Jones, 1998). Second, a further strong prediction is that the effect of irrelevant sound should diminish as trials are repeated, particularly so if they contain repeated tokens, but empirical evidence shows no hint of diminution in disruption over successive trials or over experimental sessions (Ellermeier and Zimmer, 1997; Hellbrück et al., 1996; see also Jones et al., 1997). A further difficulty is a logical one. The degree of disruption is not the product solely of the nature of the sound; it is less pronounced in tasks not requiring serial order. If the irrelevant sound effect was due wholly to the recruitment of attention from the task, it seems problematic that the nature of the task determines in part the degree of disruption (e.g., Beaman and Jones, 1997, 1998; Salamé and Baddeley, 1990).

A more tenable account of the effect of changing-state is that, generally, the effect of irrelevant sound on serial recall is due to concurrent processes of ordering (or seriation). One source of seriation relates to the maintenance of order during rehearsal of the lists in the memory task (see, e.g., Jones et al., 1996, for a discussion). The other is automatic and reflects the seriation of sound. Generally, it has been assumed that the greater the quantity of changing-state information within the sound sequence, the stronger the cues to seriation (and so the greater the disruption). However, there is some evidence to suggest that the degree of disruption is not predicted solely by a consideration of the degree of change, but it seems the effects of changing-state can be modulated by the rules of auditory organization generally known as “streaming” (Bregman, 1990). For example, if the effect of presenting a fixed sequence of three syllables monaurally (so that a single stream containing a changing sequence is formed) is contrasted with the effect of same se-
quence presented with each syllable assigned to a different location in auditory space (now forming three streams of unchanging syllables), disruption is much less marked in the latter case (Jones and Macken, 1995b). Thus it seems that the effect of degree of change can be understood only in relation to the organization of sounds into streams and to the ensuing consequences of change within such streams. The results of the current series reinforce the generalization that the degree of change within a single coherent stream is related to the magnitude of the irrelevant sound effect.

ACKNOWLEDGMENTS

This work received support from the UK’s Economic and Social Research Council (ESRC) in the form of a project grant. Dylan Jones is also a Visiting Fellow at the Air Systems’ Sector of the Defense Evaluation and Research Agency, Farnborough, Hampshire, UK. David Alford received support from an ESRC CASE studentship. Particular thanks are due to Clive Frankish for his help in undertaking the signal transformation used in experiment 2. Thanks are also due to Karen Anderson Howes and Robert Houghton for a critical reading of an earlier version of the paper.

We used a set of 7 irrelevant words in experiment 2 and a set of 14 words in experiment 1. There is a difference in the set size because the two experiments were run in different laboratories, and the available sound files were employed. Also, 12 cello notes were used in experiment 2 as there are 12 semitones in one octave. As demonstrated in Tremblay and Jones (1998), the degree of disruption by irrelevant sound is independent of the irrelevant token set size.

There are two other prominent accounts of the irrelevant speech effect, the phonological loop hypothesis (Salame and Baddeley, 1982) and the feature model (Neath, 2000), but they do not offer an explanation for the effect of irrelevant nonspeech sounds.


