Contrasting effects of changing rhythm and content on auditory distraction in immediate memory

Article

Accepted Version

Running title: Timing and content of irrelevant speech

Contrasting effects of changing rhythm and content on auditory distraction in immediate memory

Fabrice B. R. Parmentier[1,2,3]

&

C. Philip Beaman[4]

[1] Department of Psychology and Research Institute for Health Sciences (iUNICS), University of the Balearic Islands, Spain
[2] School of Psychology, University of Western Australia, Australia
[3] Instituto de Investigación Sanitaria de Palma (IdISPa)
[4] Centre for Cognition Research and School of Psychology & Clinical Language Sciences, University of Reading, United Kingdom

Corresponding author:
Fabrice Parmentier
Department of Psychology
University of the Balearic Islands
07122 Palma de Mallorca
(Tel) ++34 971 172750
(Fax) ++34 971 172309
Email: fabrice.parmentier@uib.es

Authors' note
Fabrice Parmentier is an Adjunct Senior Lecturer at the University of Western Australia. We wish to thank Gregory Elford for helping collect the data, and Pilar Andrés for her useful comments on an earlier version of this article.

Order of authorship was randomly determined and no priority is implied or should be inferred. Correspondence can be sent to Fabrice B. R. Parmentier, Department of Psychology, Ed. Cientifico-Tecnico (iUNICS), University of the Balearic Islands, Ctra de Valldemosssa, km 7.5, 07122 Palma, Spain or to Philip Beaman, School of Psychology & Clinical Language Sciences, University of Reading, Earley Gate, Whiteknights, Reading, UK, RG6 6AL.

Abstract

Across five experiments, the temporal regularity and content of an irrelevant speech stream were varied and their effects on a serial recall task examined. Variations of the content, but not the rhythm, of the irrelevant speech stimuli reliably disrupted serial recall performance in all experiments. Bayesian analyses supported the null hypothesis over the hypothesis that irregular rhythms would disrupt memory to a greater extent than regular rhythms. Pooling the data in a combined analysis revealed that regular presentation of the irrelevant speech was significantly more disruptive to serial recall than irregular presentation. These results are consistent with the idea that auditory distraction is sensitive to both intra-item and inter-item relations and challenge an orienting-based account of auditory distraction.

A number of key findings have historically been taken as indicators of the involvement of verbal short-term memory storage processes in a verbal serial recall task (a task in which participants encode, maintain and recall the stimuli in their original order of presentation). These include the phonological similarity effect (i.e., lower recall for similar sounding to-be-remembered stimuli, Conrad, 1964), the word-length effect (i.e., lower recall for long compared to short to-be-remembered words, Baddeley, Thomson & Buchanan, 1975), the articulatory suppression/concurrent articulation effect (i.e., reduction of recall performance when participants repeat an irrelevant utterance during the presentation and retention of the to-be-remembered stimuli, Murray, 1968) and the irrelevant speech/sound effect (i.e., reduction of serial recall performance when participants are exposed to a stream of irrelevant verbal stimuli while encoding and maintaining to-be-remembered stimuli, Colle & Welsh, 1976)[1].

These effects have all, in some way, informed the development of the concept of short-term or "working" memory (as reviewed by Baddeley, 1986), but all have also had some aspect of their initial interpretation called into question. The reduction in the phonological similarity effect, for example, or even its absence under some circumstances, rather than signaling the lack of phonological storage or rehearsal (e.g., Larsen & Baddeley, 2003) may be a consequence of scaling effects in the way it has been measured (Beaman, Neath & Surprenant, 2008; Jarrold & Citroën, 2013; Jarrold & Hall, 2013). Similarly, the word-length effect has been shown to be responsive to lexical properties of short- and long-words which co-vary with length (Jalbert, Neath, Bireta,

[1] "Concurrent articulation" is a theoretically neutral description of the articulatory suppression task and will be used here. The "irrelevant sound effect" (Beaman & Jones, 1997) is a term introduced to subsume the "irrelevant speech effect" which acknowledges the effects of non-speech sound but does not necessarily imply that there are no speech-specific impacts under certain circumstances (e.g., Marsh, Hughes & Jones, 2008).

& Surprenant, 2011), and alternative accounts have also been advanced for the concurrent articulation effects (Chein & Fiez, 2010; Nairne, 1990). Of all these immediate memory "effects", however, possibly the one to provoke the most controversy is the irrelevant sound effect.

Originally reported by Colle (1976), the disruption caused to an immediate, ordered recall task by the presence of background speech was first named the "unattended speech effect" (Salamé & Baddeley, 1982). It was later renamed the "irrelevant speech effect" to avoid the theoretical connotations of "unattended" (Jones & Morris, 1992) and was believed to provide evidence for a phonological *input* buffer in immediate memory, rather than the phonemic response (Baddeley & Hitch, 1974, p. 77) or articulatory buffer (Baddeley et al., 1975) previously envisaged (see Shallice & Cooper, 2011, pp. 252-258 for a discussion). The basis for this theoretical shift was Salamé and Baddeley's (1982) finding that irrelevant speech (which participants were specifically asked to ignore) disrupted immediate serial recall of visually-presented verbal items and that speech which was phonologically similar to the items to be recalled in the memory task was the most disruptive. The disruptive effect of speech has been replicated many times, but the effect of phonological similarity between to-be-recalled and irrelevant items has not been replicated (Bridges & Jones, 1996; Jones & Macken, 1995; Larsen, Baddeley & Andrade, 2003; LeCompte & Shaibe, 1996). However, the phonological store's's status as an input buffer (which was inspired by this finding) has nonetheless been retained within models of immediate memory (Baddeley, 2003; Burgess & Hitch, 1999).

Given that confusions within phonological memory cannot be the basis for the irrelevant speech effect, what possibilities remain? Work by Jones and Macken (1993)

established that phonology is neither a necessary, nor a sufficient condition to observe an irrelevant sound effect. According to the currently dominant "changing-state" account of this effect, the disruption occurs because of a conflict between order cues automatically extracted from the irrelevant speech stream and similar cues deliberately maintained in memory (Jones, 1993). The changing-state account is so called because it suggests that pre-attentive order cues in unattended speech or non-speech are given by *changes-in-state* (i.e., abrupt changes in one or more physical dimensions which segment the auditory stream into discrete auditory events or objects). So, for example, a repeated speech utterance or tone played at a constant pitch is "steady-state" and contains no useful order information as the same token occurs repeatedly. Such steady-state stimuli typically produce little or no interference with a concurrent order retention task (Jones & Macken, 1993). In contrast, if multiple tokens are present (different speech utterances or different pitch tones) then the order of presentation of these tokens is automatically registered, interfering with the serial order cues being deliberately held in the immediate memory task (Jones, Macken & Murray, 1993).

The changing-state explanation keeps the effect within the explanatory bounds of immediate memory –, the interference observed is considered to be a consequence of conflict between ordering processes in memory and in perception. However, more recent alternative accounts make no reference to memory per se, but are instead wholly based on attentional processes. There are two forms of attentional account. According to one form, passive listening to sound during a memory task is viewed as a dual-task divided attention cost, where changes in the sound increase the demands on divided

attention (Neath, 2000)[2]. In the second attentional account, hearing changes in speech or sound provokes an orienting response whereas steady-state unchanging or predictable speech or sounds are habituated to (Cowan, 1995; Elliott, 2002; see also Näätänen, 1990). This attentional capture account is the more frequently cited of the two attentional hypotheses (e.g., Röer, Bell & Buchner, 2014) and is broadly applicable to different types of distraction which could be characterized as involving attentional capture and beyond immediate memory to other tasks.

One type of distraction for which attentional capture has a ready explanation is the effect of "deviant" sounds (Escera, Alho, Winkler, Näätänen, 1998; Parmentier, 2008, 2014; Schröger, 1996). These are sounds which violate an expected pattern, such as irregularity in the presentation schedule (so that all the items bar one in the speech are presented in a regular rhythm) which introduces an extra disruptive effect beyond that of existing changes in state within the irrelevant speech (Hughes, Vachon & Jones, 2005, 2007). This effect is beyond the explanatory scope of the changing-state hypothesis, so that a duplex account was proposed (Hughes et al., 2007) in which the changing-state and deviant effects are independent and the deviant effect is incurred by mechanisms broadly equivalent to attentional capture. Although there is independent evidence for this duality (for example the size of the deviant effect is sensitive to the working memory capacity of the experimental participants -Sörqvist, 2010- whereas the changing-state effect is not -Beaman, 2004; Sörqvist, 2010; Sörqvist, Marsh & Nöstl, 2013) a single mechanism account is simpler and more elegant than a duplex account if it is able to cover both situations.

[2] There is also an interference component in this account which applies only to speech stimuli, but this is not considered here.

A difficulty with a general attentional account, however, is a loss of predictive value. Specifically, it is prone to post-hoc data-fitting and logical circularity. The divided-attention account of Neath (2000) for example, simply involves varying an attentional parameter within a mathematical model to allow the model to fit the data. Thus although the model can readily account for existing data, no *a priori* basis is given to determine the value of this parameter in the absence of divided attention (although see Beaman et al., 2008) or to indicate the extent to which different speech/sounds should vary this initial value and why this should be the case. Similarly, the attentional capture approach both assumes attentional capture if a distraction effect occurs and explains that distraction effect in terms of attentional capture, which is logically circular at best and untestable at worst. For example, Buchner and Erdfelder (2005) assume an effect of the word frequency of irrelevant speech "because low-frequency distractor words require more processing resources that could otherwise have been used for keeping the memory representations of the target words active and intact" (Buchner & Erdfelder, 2005, p. 89) and this is contrasted with the null prediction of the changing-state hypothesis. However, this is not a necessary prediction of the attentional account (as Buchner and Erdfelder (2005) are careful to note) – it is equally plausible that high-frequency distractor words might attract more attention than low-frequency words (e.g., if words need to be active above a certain threshold before attention is directed towards them) or that the frequency of the words in question might have no effect on the distraction observed (as seen, for example, in Elliott & Briganti, 2012). Thus, although the finding might be a test of the changing-state account, it is not a good test of the attentional account which could accommodate any pattern of findings in this case. To begin to test an attentional account adequately, it is

therefore necessary to identify a plausible cause of attentional capture *a priori* and then to link this with some known measure of attentional control, as has been done (for example) with the deviant effect (Sörqvist, 2010; Sörqvist et al., 2013).

It seems *a priori* plausible that any purely attentional mechanism will respond to irrelevant sound which is unexpected and unstructured. To appropriately test an account of what materials will cause an irrelevant sound effect therefore it is necessary to examine a situation in which manipulation of the irrelevant speech stream varies one or both of these dimensions, that is the speech is either unexpected, or unstructured, or both. Studies of habituation to irrelevant sound (or its lack) have indicated that stimuli which are simply unexpected are treated equivalently to stimuli which are expected (Jones, Macken & Mosdell, 1997) so expectancy *per se* is unlikely to be a key element in determining attentional capture. The deviant effect, however, shows that stimuli that are both unexpected and irregular do provoke an additional distraction effect, so irregularity may be an informative dimension to examine. One possibility is that there is a general effect of irregular presentation, conceivably encompassing the "deviant" effect of Hughes and colleagues which contrasted regular rhythmic presentation with an introduced irregularity on key trials (and hence is rhythmically irregular as well as unexpected). Studies of the mismatched negativity (MMN) component of event-related potentials have also established that deviations in an otherwise regular temporal schedule automatically capture attention (Alain, Cortese & Picton, 1999; Geiser, Ziegler, Jancke, & Meyer, 2009; Levänen, Ahonen, Hari, McEvoy & Sams, 1996) so there is *a priori* reason for supposing that rhythmic irregularity may act as a general precondition for capturing attention.

Rhythm is defined by a number of features but, in the context of this study, the term is used to refer to the temporal organization of stimuli as a function of the intervals separating them. Explorations of the effect of rhythm in irrelevant sound are surprisingly limited (Hadlington, Bridges & Beaman, 2006, Experiments 3c, 3d; Jones & Macken, 1995, Experiment 2; Larsen & Baddeley, 2003, Experiment 2) and reconsideration of the evidence gives reason to query the reported finding that irregular presentation of sound disrupts immediate memory more than regular presentation (Jones & Macken, 1995). This is of importance for the changing-state hypothesis because rhythm can also be linked to immediate memory function. Saito (2001; see also Saito & Ishio, 1998) reported that rhythm reproduction is impaired by a concurrent articulation task, and that digit span correlates with the ability to memorize rhythmic patterns. Variations in temporal organization (and hence rhythm) have also been shown to influence perceptual grouping in immediate memory (Frankish, 1985; Hitch, Burgess, Towse & Culpin, 1996) and to mediate the implicit learning of sequences, as demonstrated by Bower and Winzenz (1969) in a Hebb repetition paradigm (Hebb, 1961). Hebbian learning was only observed when the repeated sequence had a fixed temporal structure, not when its structure varied from one repetition to the next. All of these results suggest that temporal information influences order memory, and – in line with the changing-state hypothesis – unpredictable changes in the temporal structure of irrelevant speech (an irregular rhythmic presentation) could plausibly disrupt immediate memory for serial order.

However more recent evidence has indicated that order information is encoded in immediate memory in terms of relative or absolute position, rather than timing per se. Several such studies (Lewandowsky, Duncan, & Brown, 2004; Lewandowsky,

Brown, Wright, & Nimmo, 2006; Ng & Maybery, 2005) have pointed to the importance of positional rather than temporal codes in representing order information. For example, Ng and Maybery (2005) found that, with grouped presentation of to-be-recalled items, where items "migrated" across groups at recall these migration errors nonetheless preserved within-group ordinal position. If order information is represented positionally, then a conflict between to-be-recalled positional information and purely temporal information within the irrelevant speech seems less plausible than a more general attentional capture account. A re-examination of the nature of the order information present in irrelevant speech, and specifically the role of temporal information in creating interference, therefore seems necessary.

Of the four studies examining the effects of irregularly-presented irrelevant speech, one (Hadlington et al., 2006, Experiment 3c) found no effect of irregular presentation of irrelevant speech presented binaurally. The evidence for a detrimental effect of irregular presentation of irrelevant sound (over regular presentation) therefore seems to be dependent on the findings of only three experiments. Jones and Macken (1995, Experiment 2) showed similar levels of recall in regular rhythm (steady-state content) and quiet conditions whilst in the irregular rhythm condition immediate memory performance was relatively poor. Larsen and Baddeley (2003, Experiment 2) also reported a detrimental effect of the temporally varying repetition of an irrelevant speech stimulus, as did Hadlington et al. (2006, Experiment 3d). However, Larsen and Baddeley did not contrast temporal variation with a temporally regular condition and Hadlington et al.'s experiment included a pitch shift as well as a temporal irregularity, so the effects of temporal variability *per se* were not demonstrated in either study.

Thus, the assumption that changes in rhythm disturb memory for order effectively rests upon a single experiment by Jones and Macken (1995).

In what follows, we report the results of five experiments – each varying the rhythm and content of irrelevant speech – in an attempt to replicate the basic effect reported by Jones and Macken. In addition to standard inferential analyses to test the null hypothesis, we also make use of Bayesian techniques to determine the relative level of support for the null hypothesis versus a theoretically derived hypothesis. This counters the criticism raised earlier in the introduction, that attentional accounts are (or can be made) flexible enough to address any pattern of data. To do this, it is necessary to be more specific about the size of the effect we might a priori expect, based upon a specific theoretical framework.

Attentional involvement might presumably vary in the extent to which an orienting response is evoked (Cowan, 1995) or attention is divided across tasks (Neath, 2000) so a point-estimate of the effect size is impractical. Nonetheless, taking a Bayesian approach to this problem forces us to be specific about the hypothesis we are setting up as an alternative to the null. Although quantitatively different attentional accounts can be constructed (e.g., the value of the attentional parameters in Neath's (2000) model could vary to an arbitrary amount) the hypothesis we chose was that, if there is an effect of irregularity of presentation which captures attention, then the size of this effect will be no larger than the changing-state effect observed within the same data-set and will, in fact, vary in size between zero and the upper-limit set by the size of the changing-state effect with a smaller effect being more likely. The rationale for this is that attentional capture depends on the presence of changing-state information to provoke an orienting response. Thus, we can reasonably expect the difference

between regular and irregular presentation conditions to be no larger than the difference between steady-state and changing-state content conditions. It will also be no smaller than zero if we anticipate that irregular rhythms should distract more than regular rhythms rather than vice versa.

In the terminology of Dienes' (2014a,b), , this theoretical prediction involves a half-normal distribution with a mode at zero and the standard deviation equal to the likely predicted value (i.e., the mean difference between steady-state and changing-state conditions) because an approximate expected value can be determined and the theory makes a clear directional prediction. The observed sample means on which these predictions are tested against the null are then the mean difference between regular and irregular presentations and the standard error for that sample. This allows us to use Bayes theorem to compare against the null hypothesis the support yielded by the data for this specific prediction of how attentional capture might work and to derive (Bayesian) data indicating which of the two hypotheses is best supported. . To anticipate, no single experiment found a significant effect of change in rhythm (according to conventional significance testing techniques) and the Bayesian evidence, on an experiment by experiment basis, also largely favoured the null hypothesis. All experiments replicated the negative impact of changes in speech content. The results of these experiments are then discussed and new hypotheses suggested, before being pooled in a more powerful combined analysis and the results presented.

General Method for Experiments 1 to 5

Since our five experiments only differed in small ways, we report a single method section highlighting the key differences between experiments. To facilitate their comparison, we also summarize their key specific features in Table 1.

Participants. A total of 120 participants took part in the five experiments (20 each in experiments 1-4, 40 in experiment 5). Participants were undergraduate students from the University of Plymouth who took part in exchange for course credit or young adults from the local community who received a small honorarium. The ratio of males to females participants were 2/18, 6/14, 10/10, 8/12 and 8/32 in Experiments 1 to 5 respectively. The average age of the participants was 20.75 (*SD* = 4.16), 19.95 years (*SD* = 3.12), 25.86 years (*SD* = 7.06), 26.15 (*SD* = 7.17), and 18.83 years (*SD* = 1.32) respectively.

Material. The to-be-remembered stimuli were consonants ('w' excluded) presented visually at the center of a computer screen in Arial font, size 60. Seven monosyllabic words (band, fruit, jaw, kilt, porch, rod, turn; taken from Tremblay & Jones, 1998), were recorded in a male voice to be used as irrelevant speech. The words were spoken in a monotone voice and edited to a 250 ms duration (with a 10 ms rise and fall) and normalized. The experiments were programmed in E-Prime on a PC computer. Participants responded using the computer mouse.

Design and procedure. Participants were required to recall sequences of consonants while, in some trials, ignoring irrelevant speech presented through headphones. A typical trial started with participants clicking on a 'start' button, followed 2 s later by

the presentation of 7 consonants. Each consonant was presented for 750 ms followed by a 250 ms gap. An additional 10 s gap was presented after the last consonant and before the recall screen consisting of the 19 consonants of the alphabet presented in a fixed arrangement across four rows (the same arrangement in every trial). A 'skip' button allowed participants to omit items. When clicked upon, each consonant turned gray and was no longer available for selection. In the quiet condition, no irrelevant speech was presented. The irrelevant speech conditions were characterized by their regular or irregular rhythm of presentation, crossed with steady- or changing-state content, resulting in four conditions. In conditions involving irrelevant speech, its presentation started 1 sec before the to-be-remembered consonants were presented and continued until the time of recall. In regular rhythm conditions, words were presented at a regular rhythm of 2 per second (each word followed by a 250 ms gap), except in Experiment 4 in which it was one per 750 ms. In Experiments 1 to 3, all irrelevant words were used whilst in Experiment 4 and Experiment 5, only six were used ('porch' was omitted). In the steady-state conditions, a single word was used in all trials, randomly picked for each participant. In the changing-state conditions, a sequence of irrelevant words were repeated in a fixed order.

---TABLE ONE ABOUT HERE ---

In the irregular rhythm conditions of Experiments 1 and 2, seven time intervals (10, 90, 170, 250, 330, 410 & 490 ms) between the irrelevant speech stimuli were used. In Experiment 1, these intervals were used in a fixed order (330, 250, 170, 410, 10, 490, &

90 ms) repeated six times, yielding 42 intervals across the course of the retention period.

In Experiment 2, the irregular sequences were made of the above seven intervals repeated six times across 42 intervals randomly ordered. A different random arrangement of intervals was used in each irregular rhythm trial. The same random intervals were used across participants.

In Experiment 3, the intervals separating the irrelevant words were set in order to promote the subjective perception of groups. A single sequence of intervals (250, 0, 0, 500, 0, 250 & 750 ms) was repeated six times.

In Experiment 4 **&** 5**,** in an attempt to emulate the conditions used by Jones and Macken (1995) and Hadlington et al. (2006), we used their set of intervals (0, 100, 200, 800, 900 & 1000 ms). This set of intervals also differed from that used in Experiments 1 to 3 insofar as it contained six values instead of seven and that its intervals were overall longer. In Experiment 4 & 5, the sequence of intervals consisted in a random permutation of the six intervals repeated five times. A single arrangement was used in all changing-rhythm trials for any single participant (while it varied randomly across participants). The retention interval was 12.5 seconds in Experiments 4 and 5.

In Experiment 5, two conditions were employed: A blocked condition in which 20 participants completed a block containing a random mix of quiet, regular, and irregular presentation trials (12 of each type) with the nature of the speech (steady- or changing-state) blocked and counter-balanced. In the mixed condition, a further 20 participants were presented with the same set of stimuli but in a mixed (randomized) presentation order. The retention interval was 12.5 seconds in Experiment 5**.**

Thus, over the course of these experiments the following variations were employed: the rhythmic pattern of the irrelevant stimuli contained a cyclic structure (Experiments 1, 3 & 4), or varied randomly across the course of a trial (Experiment 2). It contained equidistant intervals (Experiments 1 & 2) or intervals chosen to form clear groupings (Experiments 3, 4 & 5).

The task involved 60 test trials (12 per condition) in a mixed random order (except for the blocked condition of Experiment 5). Participants completed 5 practice trials (1 per condition) prior to beginning the test phase. Instructions emphasized the need for accuracy and speed. If unsure of a response, participants were encouraged to try to guess and to press the 'skip' button if they could not.

Results

Only items recalled in their correct serial position were scored as correct. For economy, the analyses reported below do not include serial position as a factor. Analyses not reported here found no evidence for any condition x serial position interaction in any of the experiments and there are no theoretical positions of which we are aware that predict such interactions. However, in order to establish that our data exhibited the typical serial position curves expected from a visual serial recall task (U-shaped curve with strong primary and relatively smaller recency), we report the serial curves for all experiments combined (see Figure 1). The proportion of correct responses was analyzed using a one-way ANOVA with the 5 experimental conditions as independent variables. Follow-up analyses contrasted the effects of content and rhythm in 2 x 2 ANOVAs for repeated measures and comparisons of performance in the quiet condition and irrelevant speech conditions where relevant. Means and standard error

of the mean for all experiments and all conditions are presented in Table 2. Bayes factor calculations were carried out using the online software described in Dienes (2008; 2011; 2014a,b) which assumes as a default a null hypothesis where the true population value is exactly zero.

---FIGURE 1 ABOUT HERE ---

---TABLE TWO ABOUT HERE ---

Experiment 1

As is visible from Table 1, the main effect of condition on performance was significant, $F(4, 76) = 8.694$, $MSE = .007$, $p < .001$, $\eta_p^2 = .314$. A 2 x 2 ANOVA found no evidence that changes in rhythm affected performance, $F(1, 19) = .217$, MSE = 35.933, $p = .646$, $\eta_p^2 = .011$, and there was no interaction between rhythm and content, $F(1, 19) = .956$, MSE = 68.977, p = .341, $\eta_p^2 = .048$. However, performance was better in the steady-state than in the changing-state content condition, $F(1, 19) = 16.007$, $MSE = .75.392$, $p < .001$, $\eta_p^2 = .457$.

A Bayes factor analysis of the regularity manipulation, assuming a half-normal distribution with SD equivalent to the mean difference observed between steady and changing state gives the likelihood of the data given this theory as .0678 and the likelihood of the obtained data given the null as .2668. Hence the Bayes factor for the alternative over the null hypothesis is .254, or strong evidence in favor of the null, where according to these analyses, values under 1 support the null and those above 1 support the alternative hypothesis, and Jeffreys (1961) suggests that values < 1/3 should

be interpreted as strongly supporting the null and values > 3 are strongly supportive of the alternative (specified) hypothesis[3].

Experiment 2

The results for Experiment 2 show a similar pattern to Experiment 1: no effect of rhythm was observed while changing-state content depressed recall levels. The main effect of condition was significant, $F(4, 76) = 12.587$, $MSE = .006$, $p < .001$, $\eta_p^2 = .398$. There was a main effect of state, F(1, 19) = 7.585. MSE = 102.021, p = .013, $\eta_p^2 = .285$ no effect of rhythm, $F(1, 19) = 2.654$, $MSE = .004$, $p = .119$, $\eta_p^2 = .123$ and no interaction between rhythm and content , $F(1, 19) = 1.469$, $MSE = 36.470$, $p = .240$, $\eta_p^2 = .072$.

Bayes factor analysis of the regularity manipulation as in Experiment 1 gives the likelihood of the data given the theory previously outlined as .1107, whereas the likelihood of the data given the null is .0637, resulting in a Bayes factor of 1.734, or weak evidence in favor of the attentional capture by rhythmic irregularity hypothesis.

Experiment 3

The main effect of condition was significant, $F(4, 76) = 11.419$, $MSE = .007$, $p < .001$, $\eta_p^2 = .375$. There was a main effect of changing-state content, $F(1,19) = 20.728$, $MSE = 93.071$, $p < .001$, $\eta_p^2 = .522$, no effect of rhythm, $F(1, 19) = 1.324$, $MSE = 65.547$,

[3] Note that some programmes (e.g., Rouder et al., 2009) provide Bayes factors in favour of the null over the alternative hypothesis and all comments about evidence < 1/3 or >3 being strong evidence for one hypothesis over the other must be inverted. Likelihoods are provided here for readers who wish to make these inversions for themselves without re-running the full Bayes factor analyses.

p = .264, η_p^2 = .065 and no rhythm by content interaction, F(1, 19) < .001, MSE = 40.279, p >.999, η_p^2 < .001.

A Bayes factor analysis of the regularity manipulation as before gives the likelihood of the data given the attentional capture hypothesis as .0679 and the likelihood of the obtained data given the null as .1136 hence the Bayes factor for alternative over null is .598, or weak (inconclusive) evidence in favor of the null

Experiment 4

The effect of condition was significant, F(4, 76) = 11.083, MSE = .007, p < .001, η_p^2 =.368. There was a main effect of content, F(1, 19) = 27.033, MSE = 90.199, p < .001, η_p^2 = .587no effect of rhythm, F(1, 19) = 2.499, MSE = 75.198, p = .13, η_p^2 = .116, or rhythm by content interaction, F(1, 19) = 2.333, MSE = 68.545, p = .109, η_p^2 = .109.

A Bayes factor analysis of the regularity manipulation gives the likelihood of the data given attentional capture by irregular rhythms as .0042 and the likelihood of the obtained data given the null as .059 hence the Bayes factor for alternative over null is .071, or strong evidence in favor of the null.

Experiment 5.

A 2 (blocked vs mixed) x 5 (sound condition) ANOVA on the number correct reveals a significant effect of sound condition F(4, 152) = 9.164, MSE = 74.932, p < .001, η_p^2 = .194 but no main effect of blocking vs mixing, F (1, 38) = .042, MSE = 1260. 604, p

= .839, η_p^2 = .001 and no significant interactions F (4, 152) = .600, MSE = 74.932, p = .663, η_p^2 = .016.

When the mixed and blocked conditions are combined in a 2 x 2 ANOVA, the changing-state effect is significant, in line with the four previous experiments, and the existing literature, $F(1, 39) = 9.236$, $MSE = 112.982$, $p = .004$, $\eta_p^2 = .191$. The effect of rhythm approached significance $F(1, 39) = 3.846$, $MSE = 63.487$, $p=.057$, $\eta_p^2 = .090$ in the opposite direction to predictions, but there was no interaction, F(1, 39) = .086, MSE = 69.073, p =.770, η_p^2 =. 002. Thus, in a condition replicating Jones and Macken (1995, Experiment 2), with participants exposed to blocks of quiet, steady-state regular and irregular speech, no disruptive effect of irregular timing is observed.

A Bayes factor analysis of the regularity manipulation gives the likelihood of the data given attentional capture by irregular rhythms as .0038 and the likelihood of the obtained data given the null as .0463 hence the Bayes factor for alternative over null is .082, or strong evidence in favor of the null.

The observation, in Experiments 2, 4 and 5, of a numerical (but non-significant) performance *advantage* for conditions with irregular rhythms over those with regular rhythms (see Table 2) suggests the intriguing possibility that a small difference in the opposite direction to that originally hypothesized might be present. This possibility is consistent with the observation that, whilst all but one of the Bayes factor analyses report evidence in favor of the null over the alternative hypothesis, the likelihoods of the null given the data reported above are not, of themselves, of a confidence-inspiring magnitude in any of the analyses conducted, A combined analysis of all our data, in order to overcome the potentially low statistical power of each individual experiment, is reported below. The analysis investigates the above observation, and pursues the

suggestion of a negative effect of steady-state information relative to quiet (a result previously considered inconsistent with the changing-state hypothesis, LeCompte, 1995).

Combined analysis.

A combined 2 (timing) x 2 (content) analysis across all experiments confirmed the large main effect of varying content, $F(1, 119) = 67.808$, $MSE = .99.841$, $p < .001$, $\eta_p^2 = .363$, and a small but statistically significant effect of varying timing , $F(1, 119) = 4.422$, $MSE = 58.116$, $p = .038$, $\eta_p^2 = .036$, reflecting the overall lower performance in regular rhythm conditions. No significant interaction was observed between the two, $F < 1$. Paired-sample t-tests also confirmed that both steady-state conditions impaired performance relative to quiet, $t(119) = 5.336$, $p < .001$ (regular), $t(119) = 3.341$, $p = .001$ (irregular). These results are shown in Figure 2.

--- FIGURE TWO ABOUT HERE ---

A final Bayes factor analysis of the regularity manipulation was carried out on the data from Experiments 1-5 combined in a meta-analysis using the mean and SE from Study 1 as the prior mean and prior SD and the mean and SE from Study 2 as the likelihood's mean and likelihood's SD to calculate the posterior mean and posterior SD. The posterior mean and SD thus obtained were then entered as the new prior mean and SD and the mean and SE from Study 3 used as the likelihood's mean and SD and so on until the data from all five experiments were combined (Dienes, 2008). This then gives the likelihood of the combined data given the hypothesis that rhythmic

irregularity will capture attention as .0314 and the probability of the data given the null as .5189, yielding BF = .061 or strong evidence for the null over the alternative, attentional capture, hypothesis .

General discussion

The present study measured the effect of variations in the timing and content of irrelevant speech stimuli on verbal serial recall. Five experiments consistently found that serial memory was affected by variations in the content but not timing of the irrelevant speech with Bayesian evidence strongly favoring the null in three out of the five experiments (see Table 2) and in a combined cross-experiment analysis. The negative effect of changing-state content is in line with previous empirical reports (e.g., Jones & Macken, 1993). In contrast, temporal variations of the irrelevant speech failed to impair serial recall performance, regardless of whether the rhythmic pattern of the irrelevant stimuli varied randomly across the course of a trial (Experiment 2[4]) or contained a cyclic structure (Experiments 1, 3 & 4); whether it contained equidistant intervals (Experiments 1 & 2) or intervals chosen to induce groupings (Experiments 3 & 4). Experiments 1-4 mixed variations in rhythm and content whilst Jones and Macken (1995) included variations of rhythm only (steady-state content). The possibility that in the presence of variations of content, variations in rhythm are less salient and not pre-attentively processed by the cognitive system was tested in Experiment 5 but there was no sign that the presence of changing-content irrelevant speech in a repeated-measures design masked an interference effect specific to changing rhythm (Experiment 5).

[4] Although this was the only experiment where Bayesian evidence favoured the alternative hypothesis, this evidence was inconclusive (< 3; Jeffreys, 1961).

These results contrast with the detrimental effect of rhythmic variations reported by Jones and Macken (1995, Experiment 2) but are in line with Hadlington et al.'s (2006, Experiment 3c) failure to find an effect of variation in rhythm alone when irrelevant stimuli were presented to both ears. The reasons for Jones and Macken's original observation that irregular rhythm impairs memory more than regular presentation is unclear, but it is clear that – in the light of the current study – the weight of evidence tells against such an outcome.

The null results of Experiments 1-5 challenge the changing-state hypothesis but they also provide an opportunity to refine that hypothesis. Novelty, complexity, and unpredictability in various forms are key to the disruptive effects of speech (see Beaman, 2005 for a review) but, although abrupt change in many physical dimensions (e.g., frequency) is a defining feature of the changing-state effect, not all physical or perceptual dimensions are equally potent. For example, abrupt changes in intensity do not produce an effect (Ellermeier & Hellbrück, 1998; Tremblay & Jones, 1999). Our study further suggests that the temporal characteristics of the irrelevant speech stimuli do not convey order cues per se. One strong possibility is that the temporal characteristics of speech stimuli cease to be important once order cues have been stripped from the initial representation (Maybery, Parmentier, & Jones, 2002; Parmentier & Maybery, 2008) and in line with evidence that serial order is encoded positionally, not temporally (Lewandowsky et al., 2004, 2006), except in conditions where temporal information is predictable and correlates with serial position (e.g., Geiger & Lewandowsky, 2008; Surprenant, Neath, & Brown, 2006).

These results also challenge an attentional account of the irrelevant sound effect (Cowan, 1995), which postulates that auditory distracters automatically capture

attention and that habituation mediates distraction. The assumption is that auditory novelty invokes an orienting response that causes interference with ongoing cognition until such time as the auditory stimuli cease to be novel. This view predicts that the unpredictable presentation of stimuli should provoke an orienting response and hence distraction. Regular presentation of the auditory distracters should, on the other hand, be less distracting. This attentional account has widespread generality and offers an appealing account of the deviant effect and the changing-state effect (the difference between steady- and changing-state material) as well as the basic irrelevant speech/sound effect, but it does not explain the task-specificity of the irrelevant sound effect such that disruption is more readily observed primary tasks requiring the maintenance of order information (Beaman & Jones, 1997, 1998; Farley, Neath, Allbritton & Surprenant, 2007; Henson, Hartley, Burgess, Hitch & Flude, 2003; Jones & Macken, 1993). Nor does the attentional capture account explain why there is little or no indication of habituation to irrelevant speech or sound over the longer term (Hellbrück, Kuwano, & Namba, 1996; Jones, Macken & Mosdell, 1997; Tremblay & Jones, 1998) unless participants are pre-exposed to the stimuli prior to the experimental task (Röer, Dentale & Buchner, 2012, but see also Röer, Bell & Buchner, 2014, for contrary evidence). Critically, the account is not well-specified. If the difference between the effects of changing-state and steady-state content is simply a matter of habituation to the steady-state stimuli, then a substantial degree of habituation should be readily observable over very short time periods (within the space of a single trial). Although we have not directly addressed the issue of habituation here (see Beaman, Campbell & Marsh, 2014, for an attempt to do so) we have gone some way to making the predictions of an attentional account more explicit, in order to provide a more

satisfactory test of the underlying theory in the context of irregular vs. regular presentations. The strengths of the Bayesian approach are that it requires us to do this, and in quantifying (albeit approximately) the expected disruption due to attentional capture we were able to satisfy ourselves that no such effect was observed. The weaknesses of the approach are that we only directly compared the null hypothesis to the alternative we have specified – it is, however, open to others to formulate quantitative predictions from an attentional capture perspective and test them in a similar manner. The Bayesian approach we have taken is, at least, transparent in its assumptions and these are therefore open to challenge in a way that less well-specified accounts are not.

Overall, the results of the current study fail to provide support for the idea that regular presentation schedules of irrelevant speech result in less disruption to serial recall than the same speech presented irregularly. As such, the data are inconsistent with accounts of the irrelevant sound effect which assume that attention capture by unpredictable stimuli underpins distraction, in line with empirical work distinguishing between novelty distraction (Parmentier, 2014) and the irrelevant sound effect (Hughes, Vachon & Jones, 2005, 2007). Additionally, the results are inconsistent with the idea that a change in rhythm constitutes a "change-in-state" within the auditory stream, and that such changes in rhythm provoke irrelevant speech disruption. Furthermore, we found that irregular irrelevant speech produced less disruption than regular speech. Finally, the study also replicates the findings of LeCompte (1995) that steady-state speech can also produce an "irrelevant speech effect". The steady-state effect, it seems, is small but reliably non-zero if studies are of sufficient statistical power.

How might these unexpected findings be accommodated? Although the data we obtained support the null over the alternative hypothesis that irregular presentation might capture attention, close examination of the probability of any of these data given the null hypothesis shows that – on an experiment-by-experiment basis – there is in fact little evidence for the null, even though it is to be preferred over the attentional capture alternative. This conclusion is reinforced by the findings of the combined analysis, that regular presentation is significantly *more* disruptive than irregular presentation. Although any explanation of this is necessarily *post hoc*, we suggest that, whereas rhythmic presentation unambiguously results in a single, uniform stream of elements, irregular presentation, in contrast, results in a relatively uneven stream. Presenting items in temporal irregular fashion thus yields irregular groupings of the irrelevant tokens, potentially resulting in fewer between-unit transitions over the same time period (a lower "dose" effect; Bridges & Jones, 1996) as shown below:

(1) [A] [B] [C] [D] [E] [F] [A] [B] [C] [D] [E] [F] [A] [B] [C] [D] [E] [F]

(2) [AB] [C] [DEF] [AB] [C] [DEF] [AB] [C] [DEF]

This dose account assumes that the parsing of the stream into discrete utterances is not all-or-none but allows for a modest amount of changing-state information, insufficient to be greatly disruptive, to be accommodated within a given utterance. Changes-in-state are usually considered between auditory objects (Jones & Macken, 1993) but, at both physical and perceptual levels, changes also occur within objects, so the difference between changing-state and steady-state is a matter of degree rather

than kind. At a physical level, even a "steady-state" stimulus such as a frequency modulation or pitch glide (Jones et al., 1992), consists of discrete changes when analyzed across different frequency bands (such as occurs within the cochlea). At a perceptual level, individual speech utterances contain phonetic as well as physical transitions and the parsing of a speech element to a single utterance is subject to top-down as well as bottom-up influences (e.g., "birth" and "day" constitute two discrete utterances and so represent a changing-state sequence whereas "birthday" is a single utterance and, arguably, steady-state). If stream segmentation is more ambiguous with irregular than with regular rhythms, then regular rhythms will result in more coherent streams and a greater word "dose" (Bridges & Jones, 1996) producing, as here, a more disruptive irrelevant speech effect. Thus, variable -content stimuli will always produce more disruption than unvaried-content, but variable-presentation rates may result in less disruption than regular, rhythmic presentation rates. *Ceteris paribus*, one would expect any dose effect induced by irregular presentation rates to be less than or equal to the size of the deliberately created dose effect documented by Bridges and Jones (1996). As such, when looking for such an effect in future studies, the Bayesian approach adopted here is recommended as a means of providing positive evidence for any null effects of dose.

References

Alain, C., Cortese, F., & Picton, T. W. (1998). Event-related brain activity associated with auditory pattern processing. *Neuroreport, 9*, 3537-3541.

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

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews, 4*, 829-839.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior, 14*, 575-589.

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Applied Cognitive Psychology, 19*, 1041-1064.

Beaman, C. P., Campbell, T., & Marsh, J. E. (2014). *Orienting and habituation to irrelevant sound: A Bayesian perspective*. Manuscript submitted for publication.

Beaman, C. P., & Jones, D. M. (1997). The role of serial order in the irrelevant speech effect: tests of the changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory & Cognition, 5*, 315-321.

Beaman, C. P., Neath, I., & Surprenant, A. M. (2008). Modeling distributions of immediate memory effects: No strategies needed? *Journal of Experimental Psychology: Learning, Memory & Cognition, 34*, 219-229.

Beaman, C. P. & Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *Quarterly Journal of Experimental Psychology, 51A, 615-636*.

Buchner, A., & Erdfelder, E. (2005). Word frequency of irrelevant speech distractors affects serial recall. *Memory & Cognition, 33*, 86–97.

Burgess, N., & Hitch, G. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digits series. *Journal of Experimental Psychology Monograph Supplement, 80*, 1-17.

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, 49*, 919-939.

Chein, J. M., & Fiez, J. A. (2010). Evaluating models of working memory through the effects of concurrent irrelevant information. *Journal of Experimental Psychology: General, 139*, 117-137.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15*, 17-32.

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

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

Dienes, Z. (in press). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*.

Dienes, Z (in press). How Bayesian statistics are needed to determine whether mental states are unconscious. In M. Overgaard (Ed.), *Behavioural methods in consciousness research*. New York: Oxford University Press.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Sciences, 6*, 274-290.

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference*. New York: Palgrave Macmillan

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

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

Elliott, E. M., & Briganti, A. M. (2012). Investigating the role of attentional resources in the irrelevant speech effect. *Acta Psychologia, 140*, 64-74.

Escera, C., Alho, K., Winkler, I. & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience, 10*, 590-604.

Farley, L. A., Neath, I., Allbritton, D. W., & Surprenant, A. M. (2007). Irrelevant speech effects and sequence learning. *Memory & Cognition, 35*, 156-165.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory & Language, 24*, 381-389.

Geiger, S. M., & Lewandowsky, S. (2008). Temporal isolation does not facilitate forward serial recall – or does it? *Memory & Cognition, 36*, 957-967.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex, 45*, 93-102.

Hadlington, L., Bridges, A. M., & Beaman, C. P. (2006). A left-ear disadvantage for the presentation of irrelevant sound: Manipulations of task requirements and changing state. *Brain & Cognition, 61*, 159-171.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37-51). Oxford University Press: London, UK.

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 (E), 17*, 239-247.

Henson R.N.A., Hartley T., Burgess N., Hitch G., & Flude B. (2003). Selective interference with verbal short-term memory for serial order information: a new paradigm and tests of a timing signal hypothesis. *Quarterly Journal of Experimental Psychology, 56A*, 1307-1334.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology, 49A*, 116-139.

Hughes, R. W., Vachon, F., & Jones, D. M. (2005) Auditory Attentional Capture During Serial Recall - Violations at Encoding of an Algorithm-Based Neural Model. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*, 736–749.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds - Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory & Cognition, 33*, 1050–1061.

Jalbert, A., Neath, I., Bireta, T. J., & Surprenant, A. M. (2011). When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory & Cognition, 37*, 338-353.

Jarrold, C., & Citroën, R. (2013). Re-evaluating key evidence for the development of rehearsal: phonological similarity effects in children are subject to proportional scaling artefacts. *Developmental Psychology, 49*, 837-847.

Jarrold, C., & Hall, D. (2013). The development of rehearsal in verbal short-term memory. *Child development perspectives, 7*, 182-186.

Jeffreys, H. (1961). *Theory of probability*. New York: Oxford University Press.

Jones, D. M. (1993). Objects, streams and threads of auditory attention. In: A. D. Baddeley & L. Weiskrantz (Ed.s). *Attention: Selection, awareness and control.* (pp.87-104). Oxford: Clarendon press.

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

Jones, D. M., & Macken, W. J. (1995). Organizational factors in the effect or irrelevant speech: The role of spatial location and timing. *Memory & Cognition, 23*, 192-200.

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

Jones, D. M., Macken, W. J., & Murray, A. C. (1993). Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation. *Memory & Cognition, 21*, 318-328.

Jones, D.M. & Morris, N. (1992). Irrelevant speech and serial recall: implications for theories of attention and working memory. *Scandinavian Journal of Psychology, 33*, 212–229

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

Larsen, J. D., Baddeley, A. D., & Andrade, J. (2000). Phonological similarity and the irrelevant speech effect: Implications for models of short-term verbal memory. *Memory, 8*, 145-157.

LeCompte, D. C. (1995). An irrelevant speech effect with repeated and continuous background speech. *Psychonomic Bulletin & Review, 2*, 391-397.

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.

Levänen, S., Ahonen, A., Hari, R., McEvoy, L., & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cerebral Cortex, 6*, 288-296.

Lewandowsky, S., & Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory & Language, 54*, 20-38.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*, 771-790.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory & Language, 58*, 682-700.

Maybery, M. T., Parmentier, F. B. R., & Jones, D. M. (2002) Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory & Language, 47*, 360-385.

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

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

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

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

Ng, H. L. H., & Maybery, M. T. (2005). Grouping in short-term memory: Do oscillators code the positions of items? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*, 175-181.

Parmentier, F. B. R. (2008). Towards a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition, 109*, 345-362.

Parmentier, F. B. R. (2014). The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. *Psychological Research, 78*, 321-338.

Parmentier, F. B. R., & Maybery, M. T. (2008). Equivalent effects of grouping by time, voice and location on response timing. *Journal of Experimental Psychology: Learning, Memory & Cognition, 34*, 1349-1355.

Röer, J. P., Bell, R., & Buchner, A. (2014). Evidence for habituation of the irrelevant-sound effect on serial recall. *Memory & Cognition*.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225-237.

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

Saito, S. (2001). The phonological loop and memory for rhythms: An individual differences approach. *Memory, 9*, 313-322.

Saito, S., & Ishio, A. (1998). Rhythmic information in working memory: Effects of concurrent articulation on reproduction of rhythms. *Japanese Psychological Research, 40*, 10–18.

Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience, 8*, 527-539.

Shallice, T., & Cooper, R. P. (2011). *The organization of mind*. London: Oxford University Press.

Sörqvist, P. (2010). High working memory capacity attenuates the deviation effect but not the changing-state effect : Further support for the duplex-mechanism account of auditory distraction. *Memory & Cognition*, *38*, 651-658.

Sörqvist, P., Marsh, J. & Nöstl, A. (2013). High working memory capacity does not always attenuate distraction : Bayesian evidence in support of the null hypothesis. *Psychonomic Bulletin & Review, 20,* 897-904.

Surprenant, A. M., Neath, I., & Brown, G. D. A. (2006). Modeling the effects of age-related differences in immediate memory using SIMPLE. *Journal of Memory and Language, 55*, 572-586.

Tremblay, S., & Jones, D. M. (1998). Habituation versus changing state explanations of the irrelevant sound effect: Evidence from the effects of token set size. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*, 659-671.

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

Experiment	General Design	IS Rhythm	IS Content
1	5 conditions in different random order for each participant: Quiet + steady/changing-state content x steady/changing-state rhythm	Steady-state: 2 words per second (250 ms gap between words) Changing-state: Fixed sequence of time intervals (330, 250, 170, 410, 10, 490, & 90 ms) repeated six times.	Steady-state: single word randomly picked for each participant from set of seven. Changing-state: fixed sequence ("turn, kilt, band, jaw, fruit, rod, porch") repeated six times.
2	Same as Experiment 1	Steady-state: Same as Experiment 1 Changing-state: Random arrangement of time intervals (330, 250, 170, 410, 10, 490, & 90 ms) in every trial, same set of arrangements for all participants	Same as Experiment 1
3	Same as Experiment 1	Steady-state: Same as Experiment 1 Changing-state: A single, fixed, sequence of intervals (250, 0, 0, 500, 0, 250 & 750 ms) repeated six times.	Same as Experiment 1
4	Same as Experiment 1	Steady-state: One word every 750 ms. Changing-state: Random permutations of the six intervals (0, 100, 200, 800, 900 & 1000 ms) repeated five times. Single arrangement used in all trials for any single participant but varying randomly across participants.	Steady-state: single word randomly picked for each participant from set of six ("porch" excluded). Changing-state: fixed sequence ("turn, kilt, band, jaw, fruit, rod") repeated six times.
5	Five conditions from Experiment 1 but with the IS content conditions either blocked or randomized (between-subjects).	Same as in Experiment 4	Same as in Experiment 4

Table 1. Main design features characterizing Experiments 1 to 5.

	Sound Condition					
Experiment	Q	SSR	SSIR	CSR	CSIR	Bayes factor
1	62.44 (2.84)	58.39 (3.96)	57.20 (4.23)	48.8* (4.41)	41.25* (3.48)	0.25
2	61.37 (3.13)	52.38* (2.38)	52.98* (2.32)	44.52* (2.65)	48.39* (2.85)	1.74
3	65.36 (3.63)	62.62* (3.69)	60.54 (3.17)	52.80* (3.77)	50.71* (3.76)	0.60
4	70.95 (3.27)	67.98* (4.09)	73.87 (3.44)	59.76* (4.48)	60.00* (3.80)	0.08
5: Total	61.30 (2.40)	55.68 *(2.80)	58.54 (2.79)	50.96* (2.93)	53.05* (2.87)	0.82
Blocked presentation	60.87 (3.3.2)	55.77* (3.95)	58.93 (4.03)	53.30* (3.94)	53.24* (3.64)	
Random presentation	61.72 (3.54)	55.60* (4.06)	58.15* (4.00)	48.63* (4.37)	52.86* (4.53)	

Table 2. Performance according to irrelevant speech conditions in Experiments 1 to 5. Mean accuracy levels in the quiet condition (Q), steady-state content regular rhythm (SSR), steady-state content irregular rhythm (SSIR), changing-state content regular rhythm (CSR), and changing-state content irregular rhythm (CSIR) conditions are given for Experiments 1 to 5, and for the group condition of Experiment 5. Asterisks indicate conditions that differed significantly from the Quiet condition based on a two-tailed t-test for dependent samples (smallest $t(19) = 2.11$, $p = .048$). Bayes factors are for the effect of irregular presentation. All values are given to two d.p.s. Standard errors are given in parentheses.

Captions

Figure 1. Proportion correct serial recall as a function of serial position, averaged across all five experiments.

Figure 2. Mean percentage correct serial recall in the Quiet, Steady-State Regular (SSR), Steady-State Irregular (SSIR), Changing-State Regular (CSR), and Changing-State Irregular (CSIR) conditions pooled across Experiments 1-5. Error bars represent one standard error of the mean.

Figure 1

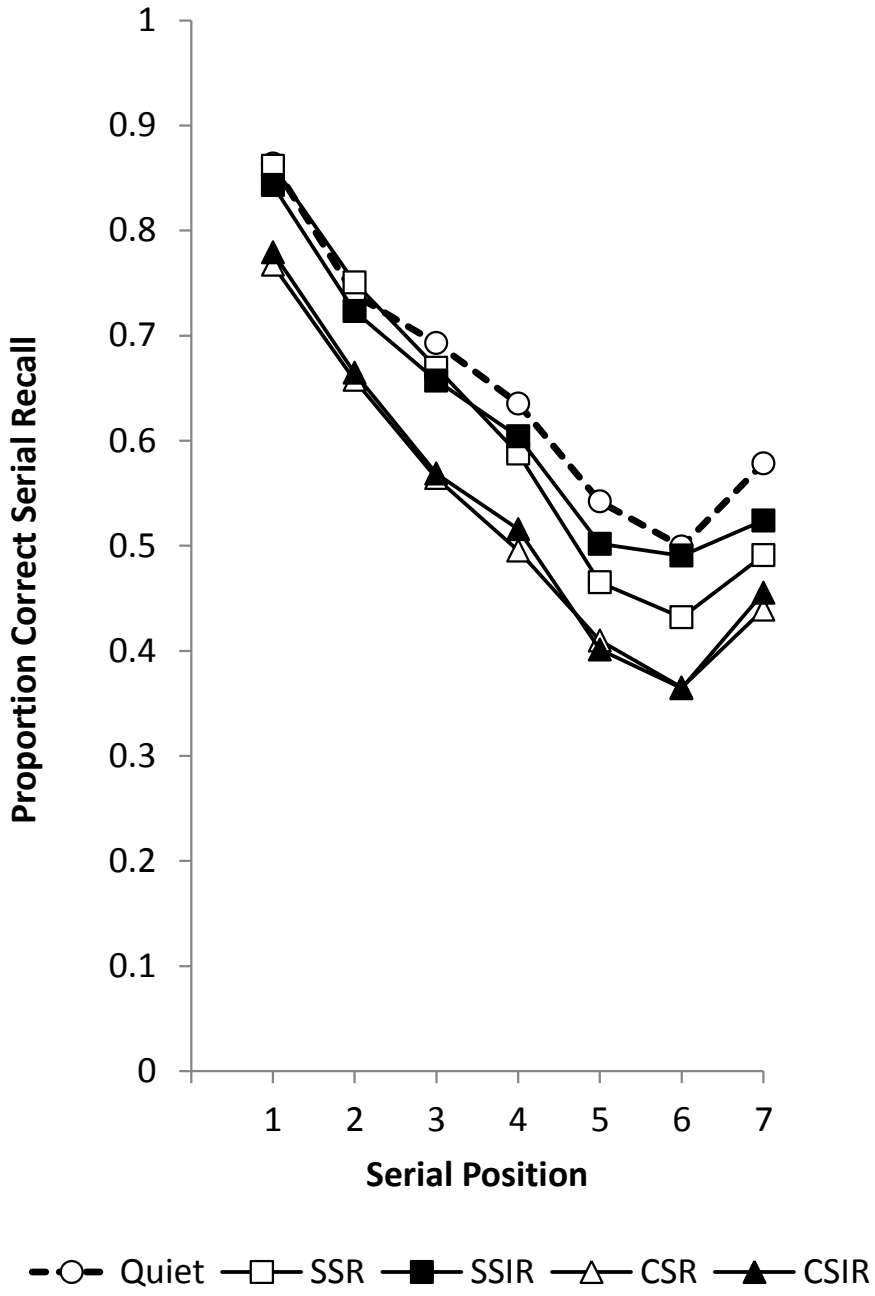
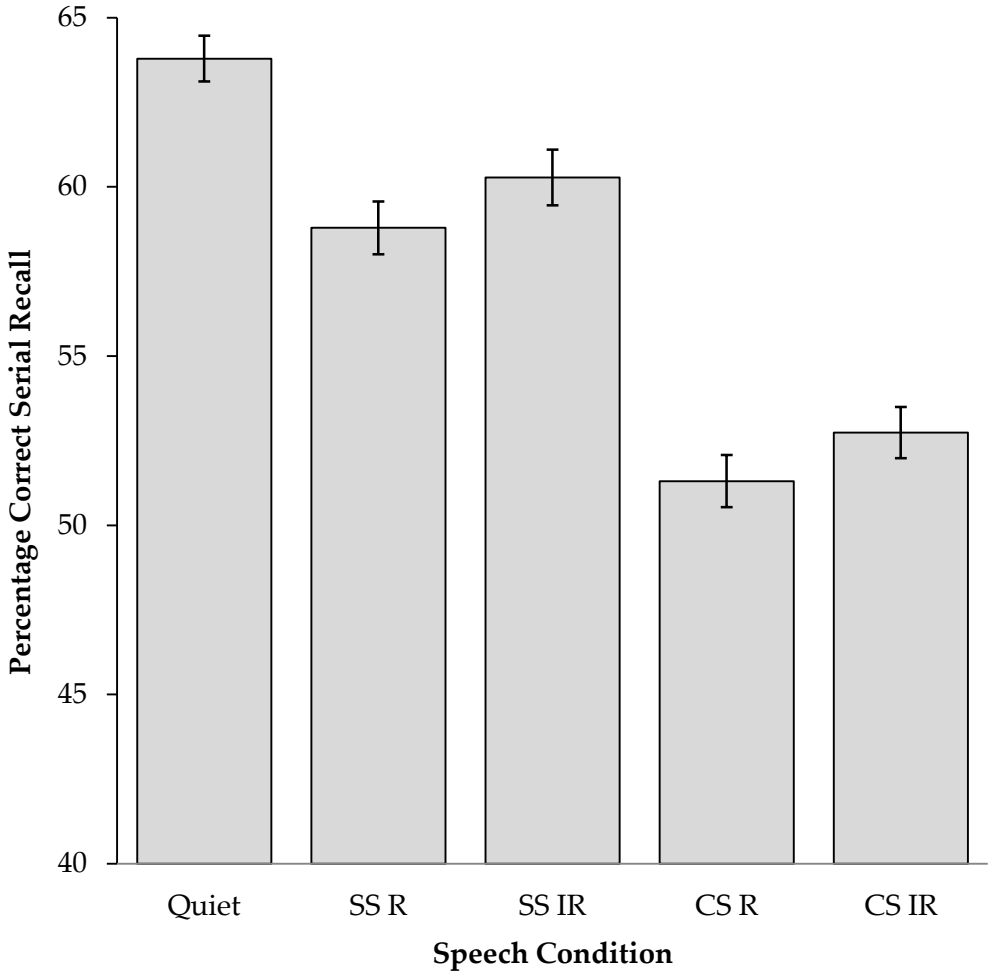

Proportion Correct Serial Recall
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
Serial Position
Quiet
SSR
SSIR
CSR
CSIR

Figure 2


65
60
55
50
45
40
Percentage Correct Serial Recall
Quiet
SS R
SS IR
CS R
CS IR
Speech Condition