When is Forewarned Forearmed? Predicting Auditory Distraction in Short-Term Memory

Robert W. Hughes[1] & John E. Marsh[2,3]

[1]Department of Psychology, Royal Holloway, University of London, Egham, UK

[2]School of Psychology, University of Central Lancashire, Preston, UK

[3]Department of Building, Energy, and Environmental Engineering, University of Gävle, Gävle, Sweden

RUNNING HEAD:

When is Forewarned Forearmed?

Correspondence: Robert W. Hughes, Department of Psychology,

Royal Holloway, University of London, Egham, Surrey, United Kingdom,

TW20 0EX; Phone: (+44) 01784 443280

email: Rob.Hughes@rhul.ac.uk

Abstract

Two experiments critically examined a predictive-coding based account of the vulnerability of short-term memory to auditory distraction, particularly the disruptive effect of changing-state sound on verbal serial recall. Experiment 1 showed that providing participants with the opportunity to predict the contents of an imminent spoken distractor sentence via a forewarning reduced its particularly disruptive effect but only to the same level of disruption as that produced by 'simpler' changing-state sequences (a sequence of letter-names). Moreover, a post-categorically unpredictable changing-state sequence (e.g., "F, B, H, E …") was no more disruptive than a post-categorically predictable sequence ("A, B, C, D …"). Experiment 2 showed that a sentence distractor was disruptive regardless of whether participants reported adopting a serial rehearsal strategy to perform the focal task (in this case, a missing-item task) whereas, critically, the disruptive effect of simpler changing-state sequences was only found in participants who reported using a serial rehearsal strategy. Moreover, when serial rehearsal was not used to perform the focal task, the disruptive effect of sentences was completely abolished by a forewarning. These results indicate that predictability plays no role in the classical changing-state irrelevant sound effect and that foreknowledge selectively attenuates a functionally distinct stimulus-specific attentional-diversion effect. As such, the results are at odds with a unitary, attentional, account of auditory distraction in short-term memory and instead strongly support a duplex-mechanism account.

KEYWORDS: Auditory distraction; Serial recall; Predictive coding; Short-term memory; Interference-by-process; Duplex-mechanism account.

There has in recent decades been a great deal of interest in the idea that a key function of the brain is to generate a predictive model of the environment and one's actions upon it (for reviews, see Clark, 2013; Huang & Rao, 2011; Winkler, Denham, & Nelken, 2009). In the present paper, we scrutinize the explanatory value of this predictive coding theory in the context of its recent application to the marked disruptive effect of task-irrelevant sound on short-term serial recall, a phenomenon that has long played a prominent role in the study of short-term memory, selective attention, and their inter-relationship (e.g., Baddeley, 1983, 2007; Beaman & Jones, 1997; Buchner, Irmen, & Erdfelder, 1996; Colle & Welsh, 1976; Ellermeier & Zimmer, 1997; Elliott, 2002; Elliott, & Briganti, 2012; Hanley & Hayes, 2012; Hughes & Jones, 2001; Jones & Macken, 1993; Jones, Macken, & Nicholls, 2004; Klatte, Lachmann, Schlittmeier, & Hellbrück, 2010). The predictability-based account claims that the disruption is due to the fact that the involuntary formation of a model of a relatively unpredictable sound sequence draws upon attentional resources required to perform the focal memory task (Bell, Röer, Marsh, Storch, & Buchner, 2017; Cowan, 1995; Elliott, 2002; Röer, Bell, & Buchner, 2015). We will argue here, however, that whereas predictive coding plays a role in some cases of auditory distraction, it does not underpin the classical effect of changing-state sound on serial short-term memory (e.g., Jones, Madden, & Miles, 1992; Jones & Macken, 1993).

It is well established that serial recall is impaired appreciably by the mere presence of task-irrelevant sound (e.g., Colle & Welsh, 1976; Elliott, 2002; Hanley & Bakopolou, 2003; Hughes, Vachon, & Jones, 2007; Jones & Macken, 1993; Röer et al., 2015; Salamé & Baddeley, 1982). In a typical study in this paradigm, a sequence of around 6-8 verbal stimuli (e.g., seven or eight digits) is presented on a screen, one item at a time, at the rate of about 1 or 2 per second, which must then be recalled in strict serial order (i.e., visual-verbal serial recall). During presentation of the to-be-remembered items—or/and, in some studies, during a retention interval between the last item and a recall cue—sound is presented at a moderate intensity [up to around 65dB(A)] that is irrelevant to the recall task and which is to be ignored. Of particular interest for present purposes is that the key precondition for reliable and marked disruption of serial recall, as far as the sound is concerned, is that it must contain acoustical variation. This has been dubbed the ch*anging-state effect,* the necessary and sufficient condition for which is an acoustic change between immediately successive sounds that form a single perceptual stream (cf. Bregman, 1990). Thus, a sequence of changing speech stimuli (e.g., "A, Q, J, G...") presented in a single voice, or a sequence of pure tones changing to a relatively modest degree in frequency from one to the next, impairs serial recall appreciably. In contrast, a repeated stimulus (e.g., "A, A, A, A…", or a repeating tone) produces much less disruption, if any (e.g., Campbell, Beaman, & Berry, 2002; Elliott, 2002; Hughes et al., 2007; Jones et al., 1992, 1995; Jones & Macken, 1993; LeCompte, 1996; Parmentier & Beaman, 2014).

The predictability-based account posits that changing-state sound is more disruptive than steady-state sound because the former draws more attentional resources away from the focal memory task (Bell et al., 2017; Cowan, 1995; Elliott, 2002; Röer, Bell, & Buchner, 2014, 2015; Röer, Bell, Dentale, & Buchner, 2011). This account appeals to the well-accepted notion embedded in predictive coding theory that a predictive model of the ongoing auditory scene is generated, one function of which is to alert the organism to potentially important environmental changes (e.g., Hughes & Jones, 2003; Näätänen, 1990; Schröger, 1997; for a recent review, see Winkler & Schröger, 2015). Two key tenets of the predictability-based account of the changing-state effect are that: i) the process of predictive model-building demands limited attentional resources; and ii) attentional resources are demanded to the extent that the successive elements in the sequence are difficult to predict, that is, to the extent that the model needs to be updated in light of prediction-errors. Thus, an accurate predictive model of a steady-state sequence can be fashioned quickly because of the high degree of predictability of each successive sound. Hence, generating such a model takes relatively few resources away from the task at hand, leaving performance relatively unscathed. In contrast, the attempt to build a model of a changing-state sequence draws relatively heavily on attentional resources due to the relative unpredictability of its elements, thereby impairing serial recall appreciably. In this view, then, “the changing-state effect is just another example of the general principle that distraction increases with the unpredictability of the distractor material” (Röer et al., 2015, p. 693).

In the present article, our argument is not with the assumption that a predictive model of the auditory scene is generated nor with the idea that such model-building plays some role in auditory distraction during short-term memory performance. Rather, our argument is that attentional diversion based on unpredictability (or rather, we have argued, the violation of predictions; cf. Vachon et al., 2012; see also Parmentier, Elsley, Andrés, & Barceló, 2011) only accounts for one of two distinct forms of auditory distraction. In particular, we argue that it does not underpin the classic changing-state effect. According to the duplex-mechanism account of auditory distraction (Hughes, 2014; Hughes, Hurlstone, Marsh, & Jones, 2013; Hughes, Vachon, & Jones, 2005, 2007), whereas some kinds of auditory input can indeed disrupt serial recall (and other tasks) via predictability-based attentional diversion (see below), the changing-state effect is instead due to *interference-by-process* (e.g., Hughes & Marsh, 2017; Jones & Tremblay, 2000). On this account, any sequential predictiveness (or invariance) within the auditory scene is conceptualized as but one cue out of several that the perceptual system can exploit in the task of determining whether or not successive sounds have emanated from the same environmental event. That is, the brain must perceptually organize the initially undifferentiated mixture of soundwaves reaching the ears into temporally-extended objects or ‘streams’ corresponding to the various sound-emitting events that created that mixture (i.e., *sequential streaming*; cf. Bregman, 1990; Winkler & Schröger, 2015). Thus, whilst our account, like the attentional diversion account, incorporates the concept of predictive coding, the changing-state effect is a by-product of streaming processes, not predictive coding per se. In line with Bregman (1990), we assume that streaming, including the use of predictability as a cue to streaming, is a preattentive process; it does not draw some limited attentional resource away from ongoing task-relevant processing (e.g., Macken et al., 2003). Rather, the changing-state effect arises because a by-product of preattentive streaming interferes specifically with a similar process involved in the focal serial recall task. Specifically, if there are spectral differences between successive, segmentable, sounds but those sounds are nevertheless similar enough to be integrated into a single stream (i.e., attributed to the same environmental source), the order of those sounds is preattentively registered (Jones et al., 1999; Macken et al., 2003). These involuntarily registered order cues interfere with the similar, but deliberate, process of assembling and maintaining the serial order of the to-be-remembered items in the form of a vocal-motor plan (e.g., Hughes & Marsh, 2017; Jones & Macken, 1993).

According to the duplex-mechanism account (Hughes, 2014), distraction due to attentional diversion is distinct from distraction caused by interference-by-process, though both are related ultimately to sequential streaming (Hughes, 2014; Macken, 2014). On this account, there are two main sources of attentional diversion[1]: In *stimulus-aspecific attentional diversion* (cf. Eimer, Nattkemper, Schröger, & Prinz, 1996), the auditory input diverts attention from an ongoing task due to the fact that it does not "fit" within the prevailing auditory scene; its attention-diverting power does not derive from any feature specific to the stimulus itself. For example, a B tone in the context of A tones ("AAAAABAA") could

[1] To avoid potential confusion, it may be worth emphasizing that these are two sub-types *within* the second main type of distraction (attentional diversion) in the duplex-mechanism account.

divert attention but, equally, so could an A tone in the context of B tones (“BBBBBABB”; Schröger, 1997). Thus, in the context of serial recall, we have shown that a single sound presented in a different voice (Hughes et al., 2007, 2013) or with a different timing (Hughes et al., 2005) from the remainder of the tokens within an irrelevant sound sequence disrupts performance (the *deviation effect*; for a review, see Hughes, 2014). We have argued that such stimulus-aspecific attentional diversion occurs to the extent that a sound is perceived as not belonging to an already established stream; that is, it diverts attention because it violates the current preattentively-generated model of the auditory scene (e.g., Hughes, 2014; Hughes & Jones, 2003; Hughes et al., 2005; Vachon, Hughes, & Jones, 2012; see also Sussman, Horváth, Winkler, & Orr, 2007; Winkler & Schröger, 2015). This conceptualization of aspecific attentional diversion differs in two important ways from that in the attentional diversion account of the changing-state effect (Bell et al., 2017; Röer et al., 2015): First, as noted, we argue that the process of integrating sounds into a stream—which forms the context against which a sound can divert attention—is preattentive; it does not, therefore, consume attentional resources away from focal task processing. Second, stimulus-aspecific attentional diversion occurs when a sound is perceived that cannot be integrated into an already established stream whereas the changing-state effect arises due to the perception of changes *within* an already established stream.

The second, non-mutually exclusive, sub-type of attentional diversion within the duplex-mechanism account of auditory distraction is *stimulus-specific attentional diversion* (cf. Eimer et al., 1996; Hughes, 2014). Here, the specific content or quality of the distractor is critical to its attention-diverting power. Attention is diverted in this case because the stimulus, in and of itself, has some kind of relevance or interest for the organism. The most famous example of such stimulus-specific attentional diversion in the cognitive-experimental literature is the so-called “cocktail party effect” in which hearing one’s own name is especially likely to divert attention from ongoing goal-relevant processing (Moray, 1959). But other examples would include a crying baby (particularly for its main caregiver(s)) or the word "food" for a hungry person, taboo words (Röer, Körner, Bell, & Buchner, 2017a) and other emotionally valent words (Buchner, Mehl, Rothermund, & Wentura, 2006; Buchner, Rothermund, Wentura, & Mehl, 2004; Marsh et al., 2018).

Predicting Distraction: The Effects of Foreknowledge

One means of examining the extent to which predictability-based attentional diversion underpins the disruptive effect of task-irrelevant auditory stimuli on serial recall is to examine whether giving participants the power to predict the sequence via a forewarning about its content attenuates that disruptive effect. For example, in Hughes et al. (2013), we showed that the disruptive effect on serial recall of a single deviation in the voice conveying a sequence of irrelevant speech tokens was eliminated when participants were informed a few seconds before a trial that the irrelevant sequence would contain a deviant (compared to not being informed)[2]. In contrast, forewarning participants that an upcoming sequence would be a changing-state sequence (as opposed to a steady-state sequence) had no effect on its disruptive impact. These results are consistent with the duplex-mechanism account: A forewarning eliminated the effect of a deviant because this effect is underpinned by aspecific attentional diversion caused by the violation of a predictive model. That is, the forewarning

[2] Bell et al. (2017) reported that the effect of a deviant was not reduced by foreknowledge and concluded that this finding "disconfirms the prediction of the duplex model that the deviation effect is more amenable to cognitive control than the changing-state effect" (p. 366). However, the deviant in Bell et al.'s study only produced a very small performance decrement (around 2-3%) whereas the decrement in the relevant experiment of Hughes et al. (2013, Experiment 2) was around 10%. As the authors themselves have pointed out (Bell et al., 2017; Röer et al., 2015), it would be difficult to observe a benefit of foreknowledge on distraction if the amount of distraction is very small in the first place. This discrepancy in the size of the deviation effect may be due to the use of a (probably more salient) deviation in voice in Hughes et al. (2013) as opposed to a word-deviation (e.g., "fall, fall, fall, dog...") in Bell et al. (2017). In line with this interpretation, the voice-deviation effect has been found to have an effect size (Cohen's *d*) in the range of .82 – 1.67 (based on four experiments reported by Hughes et al., 2007, 2013) whereas the word-deviation effect (based on the three experiments reported in Bell et al. 2017) is consistently much smaller (Cohen's *d* in the range .32 – .49).

afforded the opportunity to include, in a top-down manner, the deviant as part of the predictive model, hence stripping it of its usual disruptive power. In contrast, the finding that a forewarning had no influence on the changing-state effect is in line with the view that this effect is not caused by attentional diversion but by some other mechanism (e.g., interference-by-process).

Interest in the present article centres on the recent claim that the changing-state effect is indeed reduced by foreknowledge, suggesting, contrary to the duplex-mechanism account, that the changing-state effect is, like the deviation effect, due to predictability-based attentional diversion. Röer et al. (2015) suggested that foreknowledge about whether an imminent distractor sequence was a changing- or a steady-state sequence (cf. Hughes et al., 2013) may not have been specific enough to produce a foreknowledge effect. When they instead presented a forewarning that specifically informed participants about what the changing-state sequence would contain—e.g., a written transcript of an upcoming speech sequence—its disruptive effect was indeed attenuated. This result was taken as supporting the view that the changing-state effect is due to attentional diversion (Röer et al., 2015; see also Bell et al., 2017).

However, the central argument we make in the present article is that the attenuation of the changing-state effect by foreknowledge reported by Röer et al. (2015) was not an attenuation of the changing-state effect but the attenuation of an attentional diversion effect superimposed upon a real changing-state effect. The starting point for our argument is the observation that the critical “changing-state sequence” in Röer et al. (2015) was a naturally spoken sentence (with a different sentence presented on each changing-state trial)—e.g., “Pour water, lemon juice, and sugar in a cooking pot, then stir gently until it boils and gradually fold in beaten egg white”. In contrast, the steady-state condition comprised a single word repeated in a regular rhythm. Röer et al.’s (2015) rationale for using what they called a ‘complex changing-state sequence’ such as a sentence was that, according to the authors, a forewarning may only attenuate the effect of a changing-state sequence when the following three preconditions are met: i) the disruptive effect of the sound sequence must be relatively large, otherwise the attenuating impact of foreknowledge on that disruptive effect may not be empirically detectable; ii) related to the first precondition, the changing-state sequence must not already be highly predictable (e.g., “ABABAB…”) because providing a forewarning could not in that case effectively increase its predictability; and iii) participants are able, based on the forewarning, to form a stable representation of the sequence; thus, when, instead of sentences, a sequence of unrelated words was used as the changing-state sequence (what they called a ‘simple changing-state sequence’), no attenuation of its effect by a forewarning was found. This was because, the authors argued, a stable representation of a random word-sequence could not be readily formed.

We contend, and will go on to show, that Röer et al.’s (2015) conclusions are unsafe. To use a sentence as the changing-state sequence and contrast this with a regularly presented steady-state word is problematic because this putative manipulation of changing- vs. steady-state sound is multiply confounded: Whilst a sentence is undoubtedly a changing-state sequence, it also has several other attributes that differentiate it from the temporally-regular repetition of a single word, including sentential meaning and grammatical and syntactical structure. We suggest that any or some combination of the properties of a natural sentence that are absent from a single repeating word could have caused an additional, functionally distinct, attentional diversion effect over and above an underlying (pure) changing-state effect and that only this attentional diversion-caused portion of the disruption was attenuated by foreknowledge. More specifically, we suggest that the inherent interest or relevance of a coherent, meaningful, spoken sentence to a language-using human participant (at least or especially a sentence that is also unfamiliar) causes some degree of stimulus-specific attentional diversion over and above the classic, acoustic-based, changing-state effect. In this view, foreknowledge attenuates a sentence's disruptive power by reducing its interestingess/relevance. If this is the case, the results of Röer et al. (2015; see also Bell et al., 2017) leave the interference-by-process account of the changing-state effect—and the duplex-mechanism account of distraction more broadly—unscathed.

The general goal of the present study, then, was to demonstrate that the changing-state effect is not related to the predictability of the sound sequence, contrary to the account of Röer and colleagues. We sought to meet this objective in part by testing the hypothesis that the attenuation by foreknowledge of the disruptive effect of a 'complex' changing-state sequence on serial recall (Bell et al., 2017; Röer et al., 2015) is the attenuation of an attentional diversion effect that is functionally separate from the 'true' changing-state effect.

Experiment 1

In Experiment 1, we first sought to replicate the finding that what Röer and colleagues called a 'complex' changing-state sequence—a coherent sentence—is more disruptive than a sequence of unrelated verbal tokens (Bell et al., 2017; Röer et al., 2015). Whilst the authors did not dwell on this aspect of their results, we suggest that this finding is in itself difficult to reconcile with a key part of their reasoning. As described earlier, they reasoned that the greater attenuating effect of foreknowledge on distraction by a sentence compared to the effect of foreknowledge on distraction by a sequence of unrelated tokens is due to the fact that it is easier to build a mental representation of the former than the latter kind of sequence: "*meaningful, syntactically coherent* speech can be easily processed when being attended before the trial, and can, therefore, be effectively converted into a predictive representation of the upcoming distractor sequence" (Bell et al., 2017, p. 366; emphasis added). But it seems to us that the reason why 'meaningful, syntactically coherent speech can be easily processed' is that these properties endow the successive tokens in a sentence with higher post-categorical transitional probabilities than is the case for a sequence of unrelated tokens. In other words, it is the relatively high degree of post-categorical *predictability* of the successive elements in a sentence that makes the building of a stable mental representation of it relatively easy. Indeed, for the same reason, a coherent sentence is far better recalled in a short-term serial recall task than is a list made up of the same words in a random order: "…syntactic and semantic factors increase the predictability of words in a sentence and may enhance memory in a process akin to chunking (Miller, 1956)" (Jefferies, Ralph, & Baddeley, 2004, p. 624; see also Brener, 1940). If so, the finding that a sentence is more, not less, disruptive than a sequence of unrelated (and hence less predictable) tokens is the opposite result to that predicted by the predictability-based account of the changing-state effect. In contrast, our alternative hypothesis about why the effect of a sentence is reduced by foreknowledge provides a straightforward explanation of the fact that a sentence is more disruptive than a sequence of unrelated tokens: It seems reasonable to suppose that an unfamiliar meaningful sentence would be instrinsically more relevant or interesting to the participant—and hence would be more likely to cause (stimulus-specific) attentional diversion—than a sequence of unrelated tokens.

A second approach to testing the predictability-based account of the changing-state effect is to systematically manipulate the predictability of the changing-state sequence itself. Indeed, there is already evidence using such an approach that appears to be at odds with the predictability-based account: A relatively predictable speech sequence such as "CHJUCHJUCHJU . . ." is no less disruptive than a relatively unpredictable one such as "HJUCUCJHCUHJ . . ." (Jones et al., 1992; see also Tremblay & Jones 1998). However, proponents of the predictability-based account could perhaps counter that the predictability of the 'predictable' sequence in Jones et al. (1992) could only have emerged over the course of a trial or/and of the experimental session and hence its actual predictability may have remained relatively low. We therefore took a different approach that capitalized on the fact that the predictability-based account clearly places a great deal of importance on the *post-categorical* predictability of a changing-state sequence. This is evident from the fact that the account supposes that the post-categorical features of a forewarning (meaning, syntax) play a key role in its attenuating effect on a sentence's disruptive effect (Bell et al., 2017). Accordingly, a sentence's disruptive potency (in the absence of foreknowledge) must derive, in part at least, from the relative unpredictability of its post-categorical features. This leads to a simple prediction: Regardless of foreknowledge, if the sequence itself is already highly post-categorically predictable then it should be less disruptive than a sequence that is post-categorically unpredictable. Of course, we have just argued that this has already been disconfirmed by Röer et al.'s (2015) own data (see also Bell et al., 2017) in the finding that a (relatively predictable) sentence was more not less disruptive than an unrelated (and hence unpredictable) sequence of tokens. Nonetheless, we also tested the prediction more systematically in the present experiment by contrasting two 'simple' changing-state (CS) conditions: a *Simple CS-predictable* condition in which the irrelevant sequence comprised the first 18 letters of the alphabet recited in alphabetical—and hence post-categorically predictable—order (e.g., "A, B, C, D…") and a *Simple CS-unpredictable* sequence in which the exact same tokens were presented in a random—and hence relatively post-categorically unpredictable—order (e.g., "D, M, J, F …").[3] Moreover, a different random order was used for each trial in this latter condition, thereby making the sequence not only relatively unpredictable within a trial but also across trials (cf. Vachon et al., 2012). This should, from the standpoint of the predictability-based account, increase further the likelihood of observing greater disruption in this condition compared to the Simple CS-predictable condition (in

[3] Marsh et al. (2014) contrasted the effect of these two types of sequence but did not include a steady-state condition and so, strictly speaking, they did not have a measure of the changing-state effect.

which the exact same sequence was presented on every trial). In contrast, the interference-by-process account denies any role for either post-categorical features (unpredictable or otherwise) or predictability in the changing-state effect (e.g., Hughes et al., 2013; Jones et al., 1992; Marsh et al., 2018). As such, we predict no difference in the degree of disruption between these two conditions.

Finally, we manipulated whether or not participants received a written specific forewarning about the content of the upcoming irrelevant spoken sequence. We predict that foreknowledge will, in a highly selective way, reduce that portion of the disruptive effect of a sentence that goes over and above that caused by the simpler (or, from our perspective, 'purer') changing-state sequences. That is, it will selectively reduce the proportion of the sentence effect that we ascribe to its power to produce stimulus-specific attentional diversion and not affect the proportion of its effect that we ascribe to a separate, underlying, acoustic-based changing-state effect.

Method

Participants. For this experiment, we aimed initially to recruit twenty-four participants, the same number recruited to test and demonstrate that attentional diversion by a deviant sound is eliminated by a forewarning (Hughes et al., 2013, Experiment 1). We therefore advertised 32 timeslots for testing (to accommodate a few no-shows or other difficulties in data collection). In the event, 28 participants were recruited and successfully tested. They were students from Royal Holloway, University of London, who took part in the experiment in return for being entered into a raffle for a £50 gift voucher. All reported normal or corrected-to-normal vision and normal hearing. Ethical approval for the study was granted by the Ethics committee of Royal Holloway, University of London.

Apparatus and materials. *Memory lists*. The lists to be recalled were visually presented and comprised eight digits taken pseudo-randomly without replacement from the set 1-9 with the constraints that there were no ascending or descending runs of more than two digits and the list did not start with the digit 1. The digits were presented one at a time, in the central position of the computer monitor in black Arial Font size 72, with a duration of 800 ms and an inter-stimulus interval (offset-to-onset) of 200 ms.

To-be-ignored auditory sequences. All auditory stimuli were pre-recorded in the same male voice using a Sennheiser ME 65 microphone to 16-bit resolution at 22 kHz sampling rate with *Sony Sound Forge Pro 10* software (Sony Creative Software). To generate the steady-state (SS) sequence and the two types of Simple changing-state (CS) sequences, the first 18 letters of the English alphabet were individually recorded and then each was edited to last 250 ms using the time-stretch function in *Sound Forge Pro 10* without altering pitch or compromising intelligibility. For the SS sequence, the letter “A” was repeated 18 times with an inter-stimulus interval of 206 ms (there were 17 such intervals) and so the total duration of the SS sequence was 8 s. The Simple CS-predictable sequence comprised the first 18 letters of the English alphabet presented in canonical order and always starting with *A* (i.e., *A, B, C*...through to *R*) with the same stimulus-timing as the SS sequence. The Simple CS-unpredictable sequence comprised the same 18 letters as for the Simple CS-predictable sequence but in a random order (a different random order for each Simple CS-unpredictable trial within a block), again lasting 8 s. For the Complex CS sequences, 20 sentences were recorded. The sentences were drawn from various categories in line with Röer et al. (2015; although in the present experiment they were spoken in English—and hence were understandable by all our participants—rather than German). The sentence categories were: weather forecast, prose text, cooking recipe, scientific textbook, poetry, operating manual, road message, and aphorism (see Appendix 1). Each sentence was spoken in a manner appropriate for the sentence category and was recorded and re-recorded (if necessary) by the first author until it lasted approximately 8 s. It was then edited to last exactly 8 s (the same as the other sequence-types) by eliminating a few ms of silence at the end of the recording or artificially reducing or increasing the length of the sentence via the timestretch function in *Soundforge 10* (with care taken not to affect intelligibility). A different sentence was presented on each Complex CS(Sentence) trial (again as in Bell et al., 2017; Röer et al., 2015). Regardless of auditory condition, the onset of the first word/letter coincided with the onset of the first to-be-remembered digit. All sounds were presented at approximately 65dB(A). Finally, the number of words in each sentence equaled that in the other conditions (i.e., 18). This differs from Röer et al.'s (2015) method insofar as the number of words in their sentences varied, though it averaged 18.

Forewarning. The experiment was split into two blocks, a 'with-forewarning block' and a 'no-forewarning block' (see Design). On every trial in the with-forewarning block, a written transcription of the auditory sequence to be presented on the upcoming trial was shown on the screen for 16 s, followed by the text 'Get ready' for 3 s, then a fixation cross for 1500 ms, followed by the first to-be-remembered item and the onset of the auditory sequence. The 'no-forewarning' block was identical except instead of a written transcription of the upcoming auditory sequence, the text 'No information' was displayed for 16 s. The experiment was run on a PC using an E-Prime 2.0 program (*Psychology Software Tools*) that controlled stimulus presentation.

Design

The experiment had a 2 (Foreknowledge: with-foreknowledge, no-foreknowledge) × 4 [Auditory condition: SS, Simple CS-unpredictable, Simple CS-predictable, Complex CS(sentence)] within-participant design. There were 80 trials divided into two blocks: a with-forewarning block (40 trials) and a no-forewarning block (40 trials), the order of which was counterbalanced across participants. Each of the two blocks comprised 10 trials in each of the four auditory conditions. The order of the auditory conditions within a block was pseudorandom with the constraint that each condition was presented once every four trials and the immediate repetition of the same condition was restricted to once per auditory condition across a block. A unique sentence was used for each of the 20 Complex CS(sentence) trials encountered across the two blocks and a unique ordering of the 18 letters was used for each of the 20 Simple CS-unpredictable trials across the two blocks.

The 20 sentences were divided into 2 sets of 10 (see Appendix 1). Half the participants within each block-order received Set 1 during the no-forewarning block and Set 2 during the with-forewarning block while the reverse was the case for the other half of participants. At the end of the first block, an on-screen instruction appeared explaining that the second block would either involve the presentation of a written transcript of the upcoming auditory sequence or the presentation of the words 'No information' (depending on the block-order for the given participant).

Procedure

Participants were first provided with a general set of instructions in which it was explained that they would be presented with auditory sequences during the to-be-remembered lists but that they were to ignore the sound the best they could. It also explained that in one block they would be provided with foreknowledge of the content of the upcoming auditory sequence and that in another block they would not. At the start of the first block, participants were given two practice trials (without an irrelevant sound sequence) to familiarize them with the serial recall task and the response interface. During the 16 s period during the practice trials at which, during the experimental trials, the written transcript of the upcoming auditory sequence or 'No information' text would appear, instructions were provided as to what they would see at that point in the experimental trials (e.g., 'At this point in the experimental trials, you will see the text 'No information'). They were also given two practice trials before the second block which again included instructions about what would be presented during that 16 s period of the trial.

Following the final to-be-remembered item in each trial, the digits 1-9 were presented at random positions (different random positions for each trial) within a circular array on the screen. Participants were required to mouse-click on eight of the digits in the order in which they believed the digits had been presented on that trial. Beneath the circular array of digits, there were eight horizontally-arranged empty boxes. As the first digit was clicked within the circular array, it immediately appeared in the left-most box, as the second was clicked it appeared in the second left-most box and so on until all 8 boxes were filled. Participants were instructed that they could make a guess if they were unsure of any response or they could click on a '?' that appeared in the centre of the circular array to record a 'don't know' response. Participants were free to take as long as they wished to respond. After making 8 responses, a cross appeared at the centre of the screen prompting participants to click the mouse to continue to the next trial. The experiment took approximately 1 hr.

Results

As standard, recalled items were only scored as correct if they appeared in the same absolute serial position as that in which they were presented. Figure 1 shows the proportion of correctly recalled items, collapsed over serial positions, in the four auditory conditions in the No-forewarning condition and the With-forewarning condition. The pattern of results is clear-cut and is precisely as predicted by the duplex-mechanism account: With no forewarning, disruption was produced in all CS conditions compared to the SS condition and this disruption was particularly marked in the Complex CS(sentence) condition. Moreover, there was little difference in the disruption caused by the Simple CS-unpredictable compared to the Simple CS-predictable sequence. The provision of forewarning had a clear but highly selective effect: Forewarning brought performance in the complex CS(sentence) condition up to broadly the same level as that in the two Simple CS conditions, while having no effect on the disruptive potency of those simple CS sequences.

A 4 (Auditory condition) × 2 (Forewarning) repeated-measures Analysis of Variance (ANOVA) corroborated the foregoing impression of the data: There was a main effect of

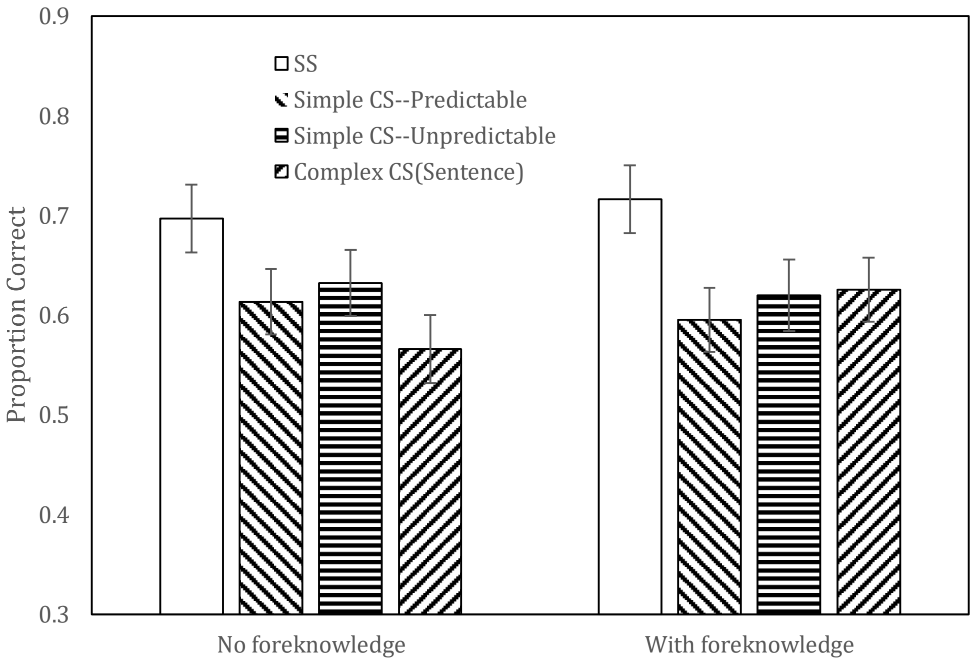

Figure 1. Proportion correct recall in the four auditory conditions as a function of foreknowledge in Experiment 1. Error bars represent the standard error of the mean.

Auditory condition, *F*(3, 81) = 22.36, *MSE* = .006, *p* < .001, η_P^2 = .45, but not of Foreknowledge, *F* < 1. Crucially, the interaction between Auditory condition and Foreknowledge was significant, *F*(3, 81) = 3.21, *MSE* = .006, *p* = .027, η_P^2 = .11. A simple effects analysis of this interaction confirmed the following pattern of effects: In the No-foreknowledge condition, Simple CS-predictable and Simple CS-unpredictable sequences both disrupted performance compared to the SS sequence (*p* < .001 for both contrasts) but, consistent with the duplex-mechanism account and at odds with the predictability-based account, there was no reliable difference between the Simple CS-predictable and Simple CS-unpredictable conditions (p = .34). A supplementary Bayes factors analysis of this null effect (following Masson, 2011) showed that the posterior probability that the data favored the null hypothesis (that there is no difference between the two conditions) [p_{BIC} (H_0|D)], over the alternative hypothesis (that there is a difference) was .71. Furthermore, when we combined our data with those from the same two conditions from Experiment 1 of Marsh et al. (2014), then the p_{BIC} (H_0|D) was .88. Note also that the slight numerical difference between the post-categorically predictable and post-categorically unpredictable conditions in our experiment was, in any case, in the opposite direction to that predicted by the predictability-based account.

Performance was poorer in the Complex CS(Sentence) condition than any other condition ($p < .02$ for all contrasts). In the With-foreknowledge condition, in contrast, performance in the Complex CS(Sentence) condition rose to the same level as that found in the Simple CS conditions such that there was now no longer any difference between any of the CS conditions (all $p > .05$). But another theoretically important observation was that even with foreknowledge all CS sequences remained markedly disruptive compared to the SS sequence (all contrasts $p < .001$).

Discussion

The results of Experiment 1 confirm the predictions of the duplex-mechanism account and are problematic for the predictability-based account. First, as already reported by Bell et al. (2017; see also Röer et al., 2015), a sentence produced more disruption than a sequence of unrelated verbal tokens (Simple CS-unpredictable condition). This is consistent with our hypothesis based on the duplex-mechanism account that a sentence produces an additional stimulus-specific attentional diversion effect over and above the (pure) acoustic-based changing-state effect that would be expected to be produced by both a sentence and a sequence of unrelated tokens. At the same time, this result is problematic for the predictability-based account: The higher degree of predictability of successive tokens in a sentence means that a sentence should be less, not more, disruptive than a sequence of unrelated (and hence relatively unpredictable) tokens. Second, a more systematic and controlled manipulation of the post-categorical predictability of a sequence—a contrast between letter-names presented in alphabetical order compared to the same letters presented in a random order—had no effect (see also Marsh et al., 2014). Third, a written forewarning about the upcoming distractor sequence only reduced the disruption caused by the complex (sentence) sequence. Importantly, it reduced a sentence's disruptive power only to the extent that a sentence then produced as much disruption as that produced by the purer changing-state sequences (Simple CS-predictable and Simple CS-unpredictable).

The finding that foreknowledge only brings performance with a sentence distractor to the same level as with simpler changing-state sequences is as predicted by the duplex-mechanism account: The additional disruption caused by a sentence (and that which is reduced by foreknowledge) is due to stimulus-specific attentional diversion while the remaining disruption is due to its acoustic changing-state quality. Indeed, in line with this two-component account of the sentence effect, the disruptive impact of a single deviant sound (e.g., Hughes et al., 2007)—universally attributed to attentional diversion (and only attentional diversion)—is abolished by foreknowledge (Hughes et al., 2013).

The limited nature of the reduction of a sentence's disruptive potency produced by foreknowledge was also observed by Röer and colleagues. They argued that the remaining power of a sentence to disrupt performance despite foreknowledge, and the power of simpler changing-state sequences to do so regardless of foreknowledge, may be attributable to a 'basic call for attention' that cannot be over-ridden by top-down knowledge. That is, the mere onset of each sound consumes attentional resources because each sound must be evaluated with regard to whether a full attention switch to that sound is warranted. Thus, very much like our duplex-mechanism account, the predictability-based account invokes one form of distraction that is amenable to top-down control (e.g., via foreknowledge)—actual attentional diversion—and another that is not: a 'basic-call-for-attention'-based distraction. At first glance, this seems to render the two theoretical accounts difficult to tease apart. Fortunately, however, the postulation that the residual effect of a sentence (with foreknowledge) and the effect of simple, even highly predictable, changing-state sequences is due to a 'basic call for attention' in fact affords a definitive test of the two accounts, which we carried out in Experiment 2.

Experiment 2

The interference-by-process account of the changing-state effect—in which the effect reflects a conflict between the processing of order information in the sound and the processing of order in the focal serial recall task—predicts, by definition, that there should be little or no interference if no order processing is being carried out to perform the focal task. This prediction has been confirmed numerous times by various research groups (Beaman & Jones, 1997; Elliott et al., 2016; Henson, Hartley, Burgess, Hitch, & Flude, 2003; Hughes et al., 2007; Jones & Macken, 1993; Joseph, Hughes, Sörqvist, & Marsh, 2018; see also Farley, Neath, Allbritton, & Surprenant, 2003; Neath, Guérard, Jalbert, Bireta, & Surprenant, 2009). This is in itself problematic for an attentional-diversion account of the changing-state effect because there is nothing in the account that would lead one to expect that the changing-state effect should be restricted to particular kinds of focal processes. Indeed, effects that all theorists agree are due to attentional diversion, such as that caused by a single deviant sound, are indeed not restricted to tasks that require or encourage order processing (Hughes et al., 2007; Vachon, Labonté, & Marsh, 2017).

We capitalize on this key distinction between the interference-by-process and attentional-diversion accounts of the changing-state effect in Experiment 2 to provide a convergent and arguably definitive test of the two accounts in relation to the effects of complex sentential distracting material and its modulation by foreknowledge. If the residual effect of sentences (with foreknowledge) and the effect of simpler changing-state sequences (regardless of foreknowledge) is due to an indomitable 'basic-call-for-attention', then these effects should be found also even when order-processing is not used to perform the focal task. In contrast, on the duplex-mechanism account, a sentence (with foreknowledge) and relatively pure changing-state sequences (regardless of foreknowledge) should no longer produce much if any disruption. That is, on this account, once the stimulus-specific attentional diversion effect of a sentence has been eliminated by foreknowledge, it should, like purer changing-state sequences, have little or no disruptive effect compared to a control (steady-state) condition. This would show quite emphatically that it is precisely that portion of the disruptive effect of a sentence that is due to attentional diversion—and not that portion that is due to its changing-state quality—that is reduced by foreknowledge.

In Experiment 2, then, we changed the focal task from serial recall to one that does not necessitate the retention of serial order (the missing-item task; e.g., Beaman & Jones, 1997; Buschke, 1963; Hughes, Marsh, & Jones, 2011; Klapp, Marshburn, & Lester, 1983; Marsh et al., 2018). In the missing-item task, all but one of a familiar closed set of items are presented (e.g., eight digits from the set 1-9) in a random order and the task is to report which item was missing (e.g., *6* is missing from the list *31784952*). Crucially, this task does not require the order of the presented items to be retained or reproduced. As predicted by the interference-by-process account, little or no changing-state effect is found in this task because there is little or no focal order processing to be interfered with by the involuntary processing of the order of the changing-state sequence (Beaman & Jones, 1997; Elliott et al., 2016;

Hughes et al., 2007; Jones & Macken, 1993; Joseph et al., 2018). Notably, effects that are universally attributed to attentional diversion such as that of a deviant sound (e.g., Bell et al., 2017; Hughes et al., 2007) are indeed observed in the missing-item task and to a similar degree as they are in serial recall (Hughes et al., 2007; Vachon et al., 2017).

The current experiment was a close replication of an experiment reported in Bell et al. (2017; see their 'Replication in Sweden' experiment, p. 365; hereafter, we refer to this as Bell et al.'s 'Swedish experiment') in which, like the present Experiment 1, foreknowledge reduced the disruptive effect of a sentence on serial recall. The critical difference in the present experiment was that the focal task was the missing-item task rather than serial recall. Thus, we contrasted the effect of three kinds of irrelevant auditory sequence—a steady-state word (SS), a simple changing-state sequence (a sequence of unrelated words; Simple CS), and a complex CS sequence (a sentence)[4]—on missing-item identification, with and without the provision of a forewarning about the content of the upcoming irrelevant sequence. As in Bell et al.'s (2017) Swedish experiment, the forewarning on this occasion involved not only a written transcript of the upcoming spoken distractor sequence but presenting the spoken sequence itself (that is, the spoken sequence was presented twice; once as part of the forewarning and once as the task-irrelevant auditory sequence accompanying the memory list).

A further novel aspect of the present experiment was that, for the first time in the context of the study of auditory distraction, we systematically collected subjective report data regarding the strategy participants adopted to perform the missing-item task, using a questionnaire developed by Morrison, Rosenbaum, Fair, and Chein (2016). This was done in light of speculations that some participants may adopt a serial rehearsal strategy to perform

[4] Bell et al. (2017, Experiment 2) also included a condition with a single deviant sound. This condition was omitted here as it is not essential for the purposes of the present experiment (see also Footnote 3).

the missing-item task even though the task does not necessitate the processing of the order of the items and that this may explain the fact that, on occasion, changing-state sound has been found to disrupt missing-item task performance to some extent (Jones & Macken, 1993; Marsh et al., 2018). Indeed, Morrison et al. (2016) reported that 28.6% of their participants reported using a serial rehearsal or grouping strategy in the missing-item task.

Thus, the predictions, based on our duplex-mechanism account, differed according to the focal-task strategy participants reported: For those reporting using some strategy other than serial rehearsal to perform the missing-item task—which we expected would be the majority—the predictions were as follows: The Simple CS sequence should produce little or no disruption compared to the SS condition, that is, there should be little or no changing-state effect. The Complex CS(sentence) sequence should disrupt performance compared to the Simple-CS and SS sequence because we argue that a sentence produces a stimulus-specific attentional diversion effect over and above any changing-state effect it may produce. Finally, foreknowledge about the upcoming irrelevant sequence should eliminate the disruption caused by the Complex CS(sentence) sequence, raising the level of performance in that condition to the same as that in the Simple-CS and SS conditions. In other words, in the presence of foreknowledge, there should no longer be any differences among any of the auditory conditions.

The predictions of the duplex-mechanism account in relation to any participants reporting the use of a serial rehearsal strategy to perform the missing-item task are the same as those for Experiment 1 (and Bell et al.'s, 2017, Swedish experiment) using serial recall. This is because it is the engagement in serial rehearsal, not the nominal nature of the task per se, that renders performance susceptible to a changing-state effect (see, e.g., Beaman & Jones, 1998). That is, there should for these participants (and only these participants) be a residual disruptive effect of sentences (under foreknowledge) and an effect of the Simple-CS sequence (regardless of foreknowledge) compared to the SS condition because the use of a serial rehearsal strategy will render them vulnerable to interference-by-process.

In contrast, on the predictability-based account of the changing-state effect, the pattern of performance should replicate that reported in Bell et al. (2017) in the context of serial recall regardless of self-reported focal-task strategy, because this account does not posit any role for the qualitative nature of the processes used to perform the focal task. First, the Simple-CS sequence should, like in serial recall, disrupt performance compared to the SS sequence because of an indomitable basic-call-for-attention. Second, a sentence should disrupt performance more than the Simple-CS sequence, even though, as argued earlier, this prediction in fact appears to be opposite to that which flows from the predictability-based account of the changing-state effect. Third, foreknowledge should attenuate the disruptive effect of a sentence but a sentence should, like the Simple-CS sequence, still disrupt performance compared to the SS condition due, again, to a basic-call-for-attention.

Participants. This experiment was conducted in the same laboratory and using the same participant pool (though different participants) as Bell et al.'s (2017) Swedish experiment. We aimed to recruit 72 participants on the basis that 28% of participants reported using a serial rehearsal/grouping strategy in Morrison et al. (2016) which would, in theory, have given us 20 participants for the 'serial rehearsal' group. While this is not a large number, we were confident that it would be sufficient given the robustness of the changing-state effect. In the event, we were only able to recruit 62 participants in a timely fashion. Fortuitously, however, as described later, the percentage of participants reporting a serial rehearsal/grouping strategy turned out to be 40% in our experiment, hence we had 25 in the serial rehearsal group and 37 in the non-serial-rehearsal group.

The 62 participants were students at the University of Gävle, Sweden, took part in the experiment in return for two cinema tickets. All reported normal hearing and normal or corrected-to-normal vision. The study was conducted in accordance with the declaration of Helsinki and the ethical guidelines given by the *American Psychological Association*. All participants were adults and participated with informed consent. The participants signed an information agreement form. The study did not involve sensitive personal data, did not entail a physical intervention or methods with the purpose of affecting a research person, the data collection did not include apparent risk of injury, and no data could or can be traced to individual persons. Because of this, there was no external ethical review, in accordance with Swedish law.

Apparatus and materials. *Memory lists.* The memory lists were, like, Experiment 1 (and Bell et al.'s 2017, Swedish experiment), drawn from the digit set *1-9*. Eight of these digits appeared sequentially in the central position of a computer monitor in black 100 pt equidistant Monaco font on a white background for 1 s each with no inter-stimulus interval. The digit missing from the list (and hence the digit to be identified) was determined randomly for each trial.

Auditory sequences. There were three types of irrelevant auditory sequence: SS, Simple CS and Complex CS(Sentence) which were identical to the corresponding conditions from Bell et al.'s (2017) Swedish experiment. In the SS condition, a one-syllable word was randomly selected and repeated 8 times (e.g., "child, child, child, child, child, child, child, child"). In the Simple CS condition, 8 different monosyllabic words were presented in random order (e.g., "road, song, day, hear, went, man, hand, mine"; note that the actual words used for this and the SS condition were the Swedish monosyllabic translations of these words). The words for these conditions were sampled randomly without replacement from a pool of 128 of the most common monosyllabic words in the Swedish language. Finally, for the Complex CS(sentence) condition, coherent sentences were presented (e.g., the Swedish translation of: "Put water, lemon juice, and sugar into a pot. Bring it to boil, stirring continuously, and progressively fold in the egg white"). Note that unlike the present Experiment 1, Bell et al. (2017) did not control for the number of words across the sentences nor the number of words in the sentence condition compared to the other two conditions; we did not either therefore in this experiment because the goal was to show that the same modulation of the sentence effect by foreknowledge found by Bell et al. (2017) using serial recall can also be found in the missing-item task under identical stimulus conditions. All auditory distractor sequences lasted 8 s each and were presented binaurally at approximately 65 dB(A) through Sennheiser HD-202 headphones that participants wore throughout the experiment. All spoken material was recorded in the same voice and sampled with a 16-bit resolution at a sampling rate of 44.1 kHz using *SoundForge 8* and edited with Audacity software (Audacity Development Team, 2015).

Forewarning. In the foreknowledge condition, a transcript of the to-be-ignored sequence was visually-presented in 32 pt Monaco font for 16 s prior to the trial. Moreover, the auditory sequence that would be subsequently presented as the distractor sequence on that given trial was presented during the first 8 s of this pre-trial phase. For the no-foreknowledge block, the words "no information" were presented in 32 pt Monaco font during the 16 s pre-trial phase, and no auditory sequence was played. Following the 16 s pre-trial phase, a blank white screen was presented for 1 s followed by the presentation of the first of the eight digits and the start of the accompanying auditory distractor sequence. The experiment was executed on a PC running an E-Prime 2.0 program (Psychology Software Tools) that controlled stimulus presentation and recorded participant responses.

Design. The experiment had a 2 × 3 × 2 mixed-measures design, with Foreknowledge (with-foreknowledge, no-foreknowledge) and Auditory condition (SS, Simple CS, and Complex CS(Sentence) as within-participant factors and Task-strategy as a between-participants factor (serial rehearsal, non-serial-rehearsal). The assignment of participants to strategy-group was determined following data collection on the basis of their self-reported strategy-use (see Procedure). The dependent variable was whether or not, for each trial, the missing item was identified. The with-foreknowledge and no-foreknowledge trials were blocked and each block consisted of 8 trials per auditory condition, presented in a random order with the constraint that no two trials from the same condition were presented in immediate succession. The order of the two blocks was counterbalanced across participants.

Procedure. Two practice trials were undertaken before the experimental trials, one with foreknowledge and one without for each of the three auditory conditions. Following the offset of the last memory item of each trial, the digits 1 to 9 appeared in a horizontal array. Beneath the array there was a single response box. Participants were required to mouse-click on the digit corresponding to the digit that was missing from the just-presented list. Once a digit was selected, a copy of it appeared in the response box for 1 s before the program prompted the participant to click on a "Begin Trial" button to initiate the next trial.

Task-Strategy Questionnaire. Following completion of the last block of trials, a strategy questionnaire (Morrison et al., 2016) was administered onscreen to participants. We asked our participants to identify the strategy that they primarily used when undertaking the missing-item task. Following Morrison et al. (2016), participants were asked to endorse one of the following 10 strategies: *I expected certain items to appear and mentally **checked them off** as they arrived* (Checklist); *I silently **repeated** the items* (Rehearsal), *I remembered the items in **groups*** (Grouping); *I thought about the way the items **sounded*** (Sound); *I answered based on what items seemed recent or **familiar*** (Familiarity); *I simply **concentrated** on the items* (Concentrate); *I created a visual **image** based on the meaning of the items* (Imagery); *I pictured the way the items **looked** on the screen* (Look); *I thought about **other things** that could relate to the items* (Association); *I used the **meaning** of the items to remember or connect them* (Semantic). Participants could also indicate if they had used a strategy not identified by the questionnaire ('Another unidentified strategy'), or whether they used no particular strategy. Furthermore, there was also an option to indicate if they had not comprehended the demands of the missing-item task so that we were able to assert whether their selection of strategy appeared legitimate. The words presented in bold were adopted from Morrison et al. (2016) and presented to participants with the intention of highlighting the key characteristic of each strategy. However, the labels in parentheses represent the nomenclature Morrison et al. adopted for different strategies and were not presented to the participants. The order in which the strategies were presented was counterbalanced across participants to avoid response bias for locations within the list.

Results

The results of the task-strategy questionnaire showed that 25 of the 62 (40%) participants reported using a 'serial rehearsal' strategy or 'grouping' strategy (14 of the 25 reporting the former, 11 the latter) to perform the missing-item task in this experiment despite the fact that the task does not nominally require the retention of order information. As grouping can be considered a process within the broader category of serial rehearsal (e.g., "Whatever else a grouping method is, it is a method of rehearsal", Wickelgren, 1964, p. 414; see also Taylor, Macken, & Jones, 2015), we treated participants that reported either of these strategies as 'serial rehearsers'. While this is a larger proportion than that reported by Morrison et al. (28.6%), for present purposes it is fortuitous that it is a relatively large number as it suits our objective of contrasting the pattern of auditory distraction effects as a function of task-strategy. The majority of participants (the remaining 37) reported using some kind of non-serial-rehearsal strategy (usually 'checklist' or an 'Another unidentified strategy'[5]); these were combined to form a 'non-serial-strategy' group.

[5] Specifically, 18 of these 37 participants reported using a 'checklist' strategy, 9 reported using 'another unidentified strategy', 3 a 'familiarity' strategy, 2 reported 'no particular strategy', while 'look', 'imagery'

Figure 2 shows the data for the non-serial-strategy group (panel A) and for the serial rehearsers (panel B) according to auditory condition and foreknowledge. A 3(Auditory condition) × 2(Foreknowledge) × 2(Task-strategy) mixed ANOVA revealed a main effect of Auditory condition, $F(2, 120) = 12.73$, $MSE = .025$, $p < .001$, $\eta_p^2 = .18$, and, importantly, reliable interactions between Auditory condition and Foreknowledge, $F(2, 120) = 7.41$, $MSE = .016$, $p < .005$, $\eta_p^2 = .11$, and, of particular interest in the present experiment, between Auditory condition and Task-strategy, $F(2, 120) = 3.18$, $MSE = .025$, $p < .05$, $\eta_p^2 = .05$. The three-way interaction was not significant. Simple effects analyses of the reliable interactions showed the following pattern of effects: First, regardless of foreknowledge, there was a disruptive effect of the Simple-CS sequence compared to the SS sequence but, critically, only for the serial rehearsal group, $p < .001$, and not for the non-serial-strategy group, $p > .05$. In other words, supporting the duplex-mechanism account, only those who reported using a serial rehearsal strategy exhibited a changing-state effect and this effect, as in Experiment 1, was unaffected by foreknowledge. A supplementary Bayes factors analysis of the absence of a changing-state effect for the non-serial-strategy group showed that the posterior probability that the data favored the null hypothesis (that there was no difference between the two conditions), pBIC (H0|D), over the alternative hypothesis (that there was a difference) was .83.

Second, collapsing across strategy-groups, in the absence of foreknowledge, sentences disrupted performance significantly more than did simple-CS sequences (again as in

---

and ‘sound’ were each reported by 1 participant. An anonymous reviewer queried the extent to which the most popular of these (‘checklist’) might also entail a form of ‘rehearsal’. To clarify, therefore, our view is that whilst the checklist strategy may well involve a form of rehearsal, it does not entail the motor-based serial rehearsal of the list as presented (see, e.g., Jones et al., 2004).

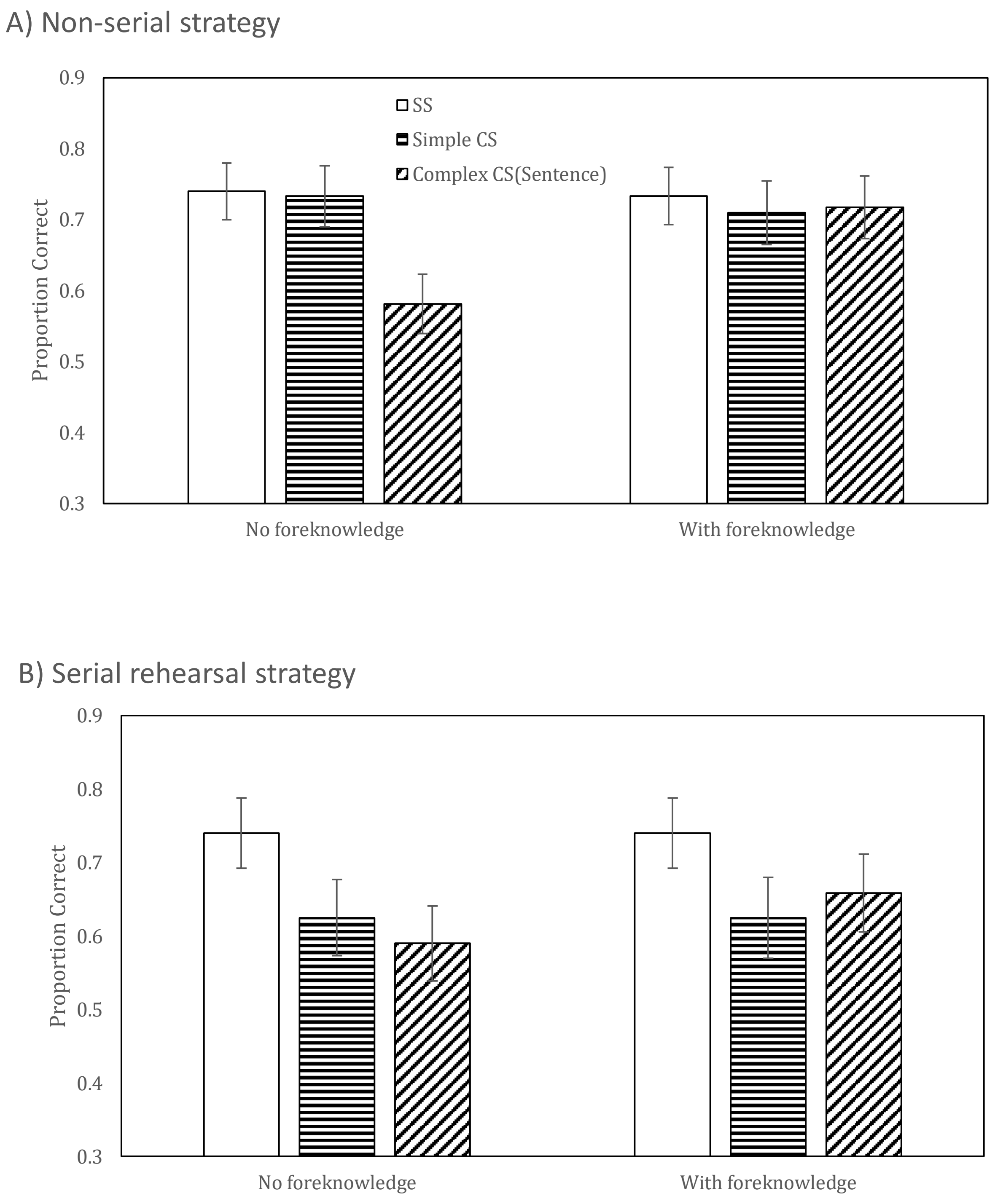

Figure 2. Proportion of lists for which the missing item was correctly identified as a function of auditory condition and foreknowledge for those participants who reported using a strategy other than serial rehearsal (Panel A) and those who reported using a serial rehearsal strategy (Panel B) in Experiment 2.

Experiment 1), $p < .01$, but, third, foreknowledge eliminated this disruptive sentence effect [Complex CS(Sentence) *vs.* Simple-CS, $p > .05$]. Thus, we observed a sentence effect in a context in which there is no changing-state effect (i.e., as evident from the pattern of distraction found for the non-serial-strategy group), and this sentence effect, unlike the changing-state effect, is eliminated by foreknowledge.

Discussion

The pattern of results from Experiment 2 provides strong support for the duplex-mechanism account (e.g., Hughes, 2014; Hughes et al., 2005, 2013) and goes against the predictability-based account of distraction by changing-state sound (Bell et al., 2017; Röer et al., 2015). The latter account invokes a basic-call-for-attention mechanism in order to explain the fact that, in serial recall, a simple-CS sequence disrupts performance compared to a steady-state sequence despite foreknowledge (Bell et al., 2017; current Experiment 1). Because the account does not make different predictions as a function of the nature of the focal-task processing, it cannot explain the fact that, regardless of foreknowledge, a simple-CS sequence is no longer disruptive when serial processing is not engaged to perform the focal task. The basic-call-for-attention construct was required also by the predictability-based account to explain the fact that, in serial recall, a sentence distractor still produces some degree of disruption even with foreknowledge (Bell et al., 2017; Röer et al., 2015; present Experiment 1). Thus, the account is also seriously challenged by our finding that the effect of a sentence distractor is indeed completely abolished by foreknowledge in the context of a missing-item task so long as participants do not adopt a serial rehearsal strategy. In short, there is no evidence for an indomitable basic-call-for-attention mechanism, the inclusion of which was critical to the predictability-based account of the findings of Röer and colleagues.

In contrast, the pattern of results confirms the predictions of the duplex-mechanism account: First, in line with the interference-by-process explanation of the changing-state effect incorporated within this account, there is only a changing-state effect when participants adopt a serial rehearsal strategy (such as in serial recall; Elliott et al., 2016; Hughes et al., 2007 Jones & Macken, 1993; present Experiment 1) or when they happen to adopt a serial rehearsal strategy even when the task does not nominally call for order processing (present experiment). Indeed, the present experiment is the first to our knowledge to show the dependence of the changing-state effect on the serial nature of focal-task processing without manipulating the explicit demands of the task (cf. e.g., Jones & Macken, 1993) but by examining individual differences in spontaneous strategy-adoption within the context of the same task.[6]

Second, the results of this experiment strongly support our central argument that a distractor sentence produces two functionally distinct forms of distraction: an attentional diversion effect and a changing-state effect, where only the former is attenuated by foreknowledge. This was shown by the finding that there is a sentence effect in a context in which there is no changing-state effect (as evident from the data from the non-serial-strategy group) and, moreover, this sentence effect—which we argue must be caused by attentional diversion—is abolished with foreknowledge.

General Discussion

The results of the present study support the duplex-mechanism account of auditory distraction (e.g., Hughes, 2014; Hughes et al., 2005) over the attentional account of Röer and

[6] While subjective reports are of course always to be treated with caution, there is good consensus that the kind of strategies in question in the present study involve explicit, effortful, processes, the kind that are most likely to be cognitively accessible and hence reportable (e.g., Smith & Miller, 1978; Schiffrin & Schneider, 1977). There are aspects of Morrison et al.'s (2016) data that also point to the validity of the questionnaire. The fact that, for example, 25% of their participants reported using a 'checklist' strategy for the missing-item task (while only 1.8% participants reported such a strategy across all other tasks) and 76.7% of participants reported using serial rehearsal or grouping to complete a serial recall task was clearly in line with apriori expectations based on convergent empirical evidence (e.g., that serial recall is particularly vulnerable to the effects of articulatory suppression, e.g., Baddeley et al., 1984).

colleagues. To summarize the key impacts of the present experiments, the predictability-based account of the changing-state effect (Röer et al., 2015) was undermined by the finding that a post-categorically predictable changing-state sequence ("A, B, C, D…") is no less disruptive than a post-categorically unpredictable changing-state sequence (e.g., "B, F, C, K…"). We also replicated the finding that a sentence is more disruptive than an unrelated sequence of tokens (Bell et al., 2017; Röer et al., 2015) even though, again, the elements in the former are more predictable than in the latter. Experiment 1 also replicated the finding that, in serial recall, foreknowledge of the nature of an upcoming auditory distractor sequence selectively attenuates the disruptive effect of a sentence and does not affect the impact of 'simple changing-state speech' (a sequence of unrelated verbal tokens) as compared to a steady-state verbal token. The two theoretical accounts offer different explanations for this pattern: From the standpoint of the duplex-mechanism account, we hypothesized that a sentence is disruptive in part due to a (stimulus-specific) attentional diversion effect that is functionally separate from an underlying, acoustic-based, changing-state effect caused by interference-by-process. On this account, it is the attentional-diversion portion of the sentence effect, not that attributable to interference-by-process, that is, like other attentional-diversion effects (e.g., the deviation effect; Hughes et al., 2007), attenuated by top-down control factors such as foreknowledge. Thus, the underlying changing-state effect is left unaltered by foreknowledge (cf. Hughes et al., 2013). The predictability-based account also attributes that portion of the sentence effect that is attenuated by foreknowledge to attentional diversion but attributes that portion of the sentence effect that is immune to foreknowledge (as well as the effect of a simple-CS sequence) to an indomitable basic-call-for-attention whereby the mere onsets of stimuli draw some attentional resources from a focal task. The results of Experiment 2 supported the duplex-mechanism explanation: When participants are not engaged in serial processing within the focal task, there is no changing-state effect (regardless of foreknowledge) and only the portion of a sentence effect that we attribute to attentional diversion remains, as evident from the fact that the effect, like the deviation effect (cf. Hughes et al., 2013), is eliminated by foreknowledge.

The present results, then, indicate that the findings of Bell et al. (2017) and Röer et al. (2015) have no bearing on the causal mechanism underlying the changing-state effect nor, therefore, on the interference-by-process explanation of that effect. As described in the Introduction, this explanation posits that the acoustic changes that underpin the changing-state effect must be carried on a common stream, unlike, for example, the deviation effect (e.g., a single change of voice within the irrelevant sound sequence; Hughes et al., 2007) which we have argued causes (aspecific) attentional diversion because the change in this case marks the onset of a stimulus that does not fit any current stream. Evidence for the interference-by-process account comes from the finding that the relationship between the degree of change between successive sounds and their disruptive effect is non-monotonic. For instance, when the magnitude of change between successive tones is so large (e.g., ACACA…, where A and C differ by 10 semi-tones) that it is likely to cause the sequential partitioning of those tones into two, interleaved, steady-state streams (i.e., *A***C***A***C***A*...), the disruption is *less* marked than with a sequence containing more modest changes and hence ones likely to be perceived as variation within a single stream (e.g., ABABA…; where A and B differ by only 5 semi-tones; Jones et al., 1999; see also Macken et al., 2003). As well as providing compelling support for the intimate link between streaming and the changing-state effect, such findings are difficult to reconcile with an attentional-diversion based account; it is far from clear how an attentional account could explain the non-monotonic relation between the amount of change and the degree to which serial recall is disrupted.

The value of the findings of Röer and colleagues (Bell et al., 2017; Röer et al., 2015), then, lies instead in being the first to show that complex speech such as sentences have the power to cause not only a changing-state effect but also attentional diversion. It is important to highlight, however, that the purpose of the present experiments was not to determine which of the properties of a sentence, or combination thereof, causes this additional attentional-diversion effect; they were designed only to show that they have a functionally different disruptive effect over and above the changing-state effect. Nonetheless, it may be worth elaborating a little here on the *stimulus-specific relevance* hypothesis alluded to earlier. We speculate that whilst the elements in an unfamiliar sentence are indeed relatively unpredictable—the key construct invoked in Röer and colleagues' account—this in itself is not the critical attention-diverting property of a sentence. Indeed, Experiment 1 showed quite clearly that unpredictability per se does not have disruptive power, with no difference being found between a post-categorically unpredictable sequence and a post-categorically predictable one. We suggest instead that the suprasegmental meaning and hence 'relevance' or 'interest' of a sentence to a human participant is key to its attention-diverting power. It is already known that such suprasegmental features are processed even though the speech is to be ignored (Röer, Bell, Körner, & Buchner, 2019). In this view, then, the key action of foreknowledge is to make the sentence more familiar and hence less interesting. That is, whilst foreknowledge would indeed increase the predictability of the successive elements in the sentence (part of what increased familiarity would entail), we speculate that it is the fact that the increased familiarity renders the sentence less interesting that is instrumental in attenuating its disruptive effect. Thus, we suggest that the additional disruptive effect of a sentence (over and above the changing-state effect) is mainly if not entirely driven by its post-categorical content, not by physical-level violations of expectancies (as in the acoustic deviation effect, Hughes et al., 2013). Note that this stimulus-specific based account can also explain why a sentence tends to be more disruptive than what could reasonably be characterized as the less meaningful and hence less interesting/relevant sequences used for the simple-CS sequences in the present experiments ("F, H, B, E…" or "road, song, day…"; see also Bell et al., 2017; Röer et al., 2015) despite the fact the successive elements in these simple-CS sequences are less predictable.

A potential difficulty for the foregoing stimulus-specific relevance hypothesis is that it would seem to predict that forward narrative speech should be more disruptive than speech that is foreign to the participant or rendered meaningless by being presented backwards, when several studies suggest that this is not the case (Jones, Miles, & Page, 1990; Marsh, Hughes, & Jones, 2009; Röer, Körner, Buchner, & Bell, 2017b). However, we note that the meaningful distractor material in these studies was always either a sequence of individual words (which, as already noted, may have relatively low relevance/interest value in any case) or presented in the form of a narrative from a single literary source that continued throughout the experimental session (Jones et al., 1990) in contrast to the discretely-presented (i.e., trial-by-trial), variously-themed, sentences as used here (and in Bell et al., 2017; Röer et al., 2015). Indeed, we have emerging evidence that discrete sentences presented in a forward direction and in a language the participant understands are indeed more disruptive than sentences presented in reverse or in a language foreign to the participant (Marsh, Kershaw, Vachon, & Hughes, in preparation). One possibility, therefore, is that continuous narrative speech (when presented as task-irrelevant material) soon loses its meaning or interest for the participant (and hence its attention-diverting power) whereas discrete novel sentences varying in content and intonation appear to have more sustained attention-diverting power. At first glance, however, the results of a study by Röer et al. (2014) seem to contradict those of Marsh et al. (in prep.) and hence the conclusions we have drawn from them. Röer et al. (2014) found that discretely-presented sentences played forwards were no more disruptive of serial recall than the same sentences played backwards. The first thing to note is that this null effect, if taken at face value, poses a problem not only for our stimulus-specific relevance hypothesis but also for the predictability-based account because a sentence is clearly more predictable than the same sentence presented backwards. That is, the predictability-based account predicts that a backward sentence should be more disruptive than a forward sentence. The second point, however, is that a feature of Röer et al.'s (2014) method means that their results do not cause a difficulty for the stimulus-specific relevance hypothesis after all. Specifically, for a given participant, the same sentence (whether forward or backward) was presented for all trials in the given condition (moreover, unlike the present experiments, auditory condition was blocked). From the standpoint of the stimulus-specific relevance account, such repeated presentation of the same sentence across a block of trials may soon strip that sentence of its relevance and hence attention-diverting power; indeed, such repetition would, in effect, be akin to having foreknowledge of the auditory sequence. The difficulty that Roer et al.'s (2014) results pose for the predictability-based account remains however: the fact that the sentence was repeated does not alter this account's (incorrect) prediction that a backward sentence should be more disruptive than a forward sentence. Indeed, if anything, the difference between conditions should be even greater as memory for (and hence foreknowledge of) a forward sentence that is repeated from trial-to-trial is likely to be better than for a repeating backward sentence.

The same stimulus-specific attentional diversion account as offered here for the effect of sentences has been applied recently to account for other post-categorical auditory distraction effects during serial recall that also appeared, otherwise, to undermine the interference-by-process account of disruption by changing-state sound (Marsh et al., 2018). Specifically, serial recall is disrupted to a greater degree by a sequence of emotionally valent compared to neutral words. This has been taken as evidence against the notion that the disruptive impact of sound on serial recall is attributable purely to the sound's acoustic, pre-categorical, characteristics (Buchner et al., 2004, 2006). In Marsh et al. (2018), however, we tested the hypothesis that this additional effect of valence is, like the sentence effect studied here, due to stimulus-specific attentional diversion that is unrelated to an underlying changing-state effect. In line with our hypothesis, we showed that promoting an increase in focal-task engagement by making the visual to-be-remembered items more difficult to encode eliminates the valence effect—just as it does the deviation effect (Hughes et al., 2013)—but has no impact on the disruptive effect of words per se (compared to quiet). Moreover, similar to the present Experiment 2 in terms of its logic, Experiment 2 of Marsh et al. (2018) showed that a valence effect is also found in the missing-item task but that the general effect of words is significantly attenuated, providing convergent evidence that the valence effect, too, is unrelated to the classical changing-state effect.

To conclude, the present results reinforce our contention that all post-categorical auditory distraction effects in serial recall—including the particularly disruptive effect of sentences, valent words (Buchner et al., 2006, 2004), one's own name (Röer et al., 2013), taboo words (Röer et al., 2017a), and so on—are epiphenomenal; they are attentional diversion effects that have, unlike the changing-state effect, relatively little relevance to the understanding of serial recall or short-term memory more generally because they are found outside these domains. Thus, the present findings (see also Elliott et al., 2016; Joseph et al., 2018; Marsh et al., 2018) underscore the importance of determining, whenever serial short-term memory is found to be impaired by task-irrelevant sound, which of two distinct mechanisms of distraction—attentional diversion and interference-by-process—is at play or the degree to which each might be contributing to the total amount of disruption.

References

Baddeley, A. D. (1983). Working Memory. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 311-324.

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, *36*(2), 233-252.

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, and Cognition*, *23*, 459-471.

Bell, R., Röer, J. P., Marsh, J. E., Storch, D., & Buchner, A. (2017). The effect of cognitive control on different types of auditory distraction: A preregistered study. *Experimental Psychology*, *64*(5), 359-368.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organisation of sound*. Cambridge, MA: MIT Press.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*, 467–482.

Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the irrelevance of semantic information for the irrelevant speech effect. *Quarterly Journal of Experimental Psychology A*, *49*(3), 765-779.

Buchner, A., Mehl, B., Rothermund, K., & Wentura, D. (2006). Artificially induced valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition*, *34*, 1055-1062.

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory &*

Cognition, 32, 722-731.

Buschke, H. (1963). Relative retention in immediate memory determined by the missing scan method. *Nature, 200,* 1129–1130.

Campbell, T., Beaman, C. P., & Berry, D. C. (2002). Auditory memory and the irrelevant sound effect: Further evidence for changing-state disruption. *Memory*, *10*(3), 199-214.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204.

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

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

Eimer, M., Nattkemper, D., Schröger, E., & Prinz, W. (1996). Involuntary attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action* (Vol. 3, pp. 389–446). London, UK: Academic Press.

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

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

Elliott, E. M., & Briganti, A. M. (2012). Investigating the role of attentional resources in the irrelevant speech effect. *Acta psychologica*, *140*(1), 64–74.

Elliott, E. M., Hughes, R. W., Briganti, A., Joseph, T. N., Marsh, J. E., & Macken, W. J. (2016). Distraction in verbal short-term memory: Insights from developmental differences. *Journal of Memory & Language*, *88,* 39-50.

Farley, L. A., Neath, I., Allbritton, D. W., & Surprenant, A. M. (2007). Irrelevant speech effects and sequence learning. *Memory & Cognition*, *35*(1), 156-165.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review, 10*(2), 435-444. doi: 10.3758/BF03196503

Hanley, J. R., & Hayes, A. (2012). The irrelevant sound effect under articulatory suppression: Is it a suffix effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 482-487.

Henson, R., 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. *The Quarterly Journal of Experimental Psychology, 56*(8), 1307-1334.

Huang, Y., & Rao, R. (2011). Predictive coding. Wiley Interdisciplinary Reviews: *Cognitive Science, 2,* 580–593.

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal, 3,* 30-41.

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., & Jones, D. M. (2013). Cognitive control of auditory distraction : Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 539–553.

Hughes, R., & Jones, D. M. (2001). The intrusiveness of sound: Laboratory findings and their implications for noise abatement. *Noise & Health, 4*(13), 51-70.

Hughes, R. W., & Jones, D. M. (2003). Indispensable benefits and unavoidable costs of unattended sound for cognitive functioning. *Noise & Health, 6,* 63-76.

Hughes, R. W., & Marsh, J. E. (2017). The functional determinants of short-term memory:

Evidence from perceptual-motor interference in verbal serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 43,* 537-551. http://dx.doi.org/10.1037/xlm0000325

Hughes, R. W., Marsh, J. E., & Jones, D. M. (2011). Role of serial order in the impact of talker variability on short-term memory: Testing a perceptual organization-based account. *Memory & Cognition, 39*(8), 1435-1447.

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.

Jefferies, E., Ralph, M. A. L., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*(4), 623-643.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 464-473.

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

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological, and is it a store? *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *30,* 656-674.

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

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology, 4*(2), 89-108.

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

Joseph, T. N., Hughes, R. W., Sörqvist, P., & Marsh, J. E. (2018). Differences in auditory distraction between adults and children: A duplex-mechanism approach. *Journal of Cognition*, *1*(13), 1-11.

Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve The "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, *112*, 240-264.h

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology*, *22*(8), 1168–1191.

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

Macken, B. (2014). Auditory distraction and perceptual organization: streams of unconscious processing. *PsyCh Journal*, *3*(1), 4-16.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology. Human Perception and Performance, 29*(1), 43-51.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110*, 23-38.

Marsh, J. E., Kershaw, M. B. A., Vachon, F., & Hughes, R. W. (2018). *The Attention-Diverting Power of a Meaningful Spoken Sentence*. Manuscript in preparation.

Marsh, J. E., Röer, J. P., Bell, R., & Buchner, A. (2014). Predictability and distraction: Does the neural model represent postcategorical features? *PsyCh Journal*, *3*(1), 58-71.

Marsh, J. E., Yang, J., Qualter, P., Richardson, C., Perham, N., Vachon, F., & Hughes, R. W. (2018). Postcategorical auditory distraction in short-term memory: Insights from increased task load and task type. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *44*(6), 882-897.

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679-690.

Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & Cognition*, *44*(6), 922-936.

Moray, N. P. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

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

Neath, I., Guérard, K., Jalbert, A., Bireta, T. J., & Surprenant, A. M. (2009). Irrelevant speech effects and statistical learning. *The Quarterly Journal of Experimental Psychology, 62*(8), 1551-1559.

Parmentier, F. B., & Beaman, C. P. (2014). Contrasting effects of changing rhythm and content on auditory distraction in immediate memory. *Canadian Journal of Experimental Psychology*, *69*(1), 28-38.

Parmentier, F. B., Elsley, J. V., Andrés, P., & Barceló, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119*(3), 374-380.

Röer, J. P., Bell, R., & Buchner, A. (2013). Self-relevance increases the irrelevant speech effect: Attentional disruption by one's own name. *Journal of Cognitive Psychology*, *25*, 925-931.

Röer, J. P., Bell, R., & Buchner, A. (2014). Evidence for habituation of the irrelevant-sound effect on serial recall. *Memory & Cognition*, *42*, 609–621. http://dx.doi.org/10.3758/s13421-013-0381-y

Röer, J. P., Bell, R., & Buchner, A. (2015). Specific foreknowledge reduces auditory distraction by irrelevant speech. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 692-702.

Röer, J. P., Bell, R., Dentale, S., & Buchner, A. (2011). The role of habituation and attentional orienting in the disruption of short-term memory performance. *Memory & Cognition*, *39*, 839–850. http://dx.doi.org/10.3758/s13421-010-0070-z

Röer, J., Bell, R., Körner, U., & Buchner, A. (2019). A semantic mismatch effect on serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 515-525.

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017a). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion*, *17*(4), 740-750.

Röer, J. P., Körner, U., Buchner, A. & Bell, R. (2017b). Semantic priming by irrelevant speech. *Psychonomic Bulletin & Review*, *24*(4), 1205-1210.

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

Schneider, W., & Schiffrin, R. M. (1977). Automatic vs controlled processing. *Psychological Review*, *84*, 1-64.

Schröger, E. (1997). On the detection of auditory deviants: A preattentive activation model. *Psychophysiology*, *34,* 245–257. doi:10.1111/j.1469-8986.1997.tb02395.

Smith, E. R., & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. *Psychological Review*, *85*(4), 355-362.

Sussman, E. S., Horváth, J., Winkler, I., & Orr, M. (2007). The role of attention in the formation of auditory streams. *Perception & Psychophysics*, *69*(1), 136-152.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 659-671.

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken expec tations: Violation of expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 38,* 164–177. DOI: 10.1037/a0025054.

Vachon, F., Labonté, K., & Marsh, J. E. (2017). Attentional capture by deviant sounds: A noncontingent form of auditory distraction? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(4), 622-634.

Wickelgren, W. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, *68*(4), 413–419.

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences, 13,* 532–540. doi:10.1016/j.tics.2009.09.003

Winkler, I., & Schröger, E. (2015). Auditory perceptual objects as generative models: Setting the stage for communication by sound. *Brain and Language*, *148,* 1-22.

Author note

Robert W. Hughes, Department of Psychology, Royal Holloway, University of London. John E. Marsh, School of Psychology, University of Central Lancashire and the Department of Building, Energy, and Environmental Engineering, University of Gävle. We would like to thank Samuel Thomas, Hannah McCaldin, and Matthew Talbot for data collection and Eshim Shahid for sourcing the sentences used in Experiment 1. We thank Sebastian Arnström for recording the auditory stimuli and collecting the data for Experiment 2. John E. Marsh's contribution to this article was supported by a grant from the Swedish Research Council (2015-01116) awarded to Patrik Sörqvist and John Marsh. Correspondence may be addressed to Robert W. Hughes at Rob.Hughes@rhul.ac.uk.

Appendix 1: Sentences Used in Experiment 1

Set 1

1. (Weather forecast). A widespread frost at first, perhaps the odd mist or fog patch, and then bright with sunny spells.
2. (Weather forecast). A cold and raw day with a chilly north-easterly wind, showery rain, chiefly to the south-east.
3. (Prose). Let them have everything - health, food, a place to live, entertainment - they are and remain unhappy and low.
4. (Prose). I object to violence because when it appears to do good; the good is temporary, evil is permanent.
5. (Cooking recipe). Put over a medium heat and warm, whilst stirring constantly, until it is very thick and just bubbling.
6. (Scientific text). No obvious distinction was found for the lower critical solution temperature and swelling response to a temperature jump.
7. (Poem). “That I scarce was sure I heard you"- here I opened wide the door, darkness there, nothing more.
8. (Operating manual). Old taps may seize in the closed position and thus prevent the water from reaching into the machine.
9. (Road message). Congestion caused by vehicles restricts emergency, recovery or winter maintenance vehicles from providing assistance or from clearing roads.
10. (Aphorism). Millions long for immortality but do not know what to do with themselves on a rainy Sunday afternoon.

Set 2

1. (Weather forecast). A cold start for many, then most places dry with sunny spells, cloudier in the southeast with rain.
2. (Weather forecast). Mainly dry with the best of the cloud breaks around the North Sea coast, clouds thickening at times.
3. (Prose). All things are subject to interpretation whichever interpretation prevails at a given time is a function of power.
4. (Prose). Humanity is an ocean; if a few drops of the ocean are dirty, the ocean doesnt become dirty.
5. (Cooking recipe). Whisk in the mashed date mixture, then fold into the egg white mixture until it is well combined.
6. (Scientific textbook). The decrease of crosslinker dosage or increase of chain length thereof could enhance swelling capacities of the hydrogels
7. (Poem). Christmas is coming. The goose is getting very fat, please put a penny in the old man’s hat.
8. (Operating manual). Remove the four protective screws and the rubber bush with the respective spacer, situated on the appliance’s rear
9. (Road message). Travel conditions are dangerous and you should avoid the specified roads, if you travel, you may experience disruption.
10. (Aphorism). The optimist proclaims we live in the best of possible worlds; and the pessimist fears this is true.