The benefit of foreknowledge in auditory distraction depends on the intelligibility of pre-exposed speech

Florian Kattner[1*], Beth H. Richardson[2], and John E. Marsh[2,3]

[1]*Health and Medical University, Germany*

[2]*University of Central Lancashire, United Kingdom*

[3]*Lulea University of Technology, Sweden*

Author Note

[*]*Corresponding author's address:*

Florian Kattner
Health and Medical University
Olympischer Weg 1
14471 Potsdam
Germany
florian.kattner@health-and-medical-university.de

Abstract

Introduction: Providing participants with an opportunity to listen to a forthcoming distracter sentence has been shown to attenuate its disruptive effect on short-term memory. On the stimulus-specific attentional diversion account, foreknowledge selectively reduces any potential diversion produced by interest in the post-categorical (e.g., semantic or syntactical) properties of a discrete sentence. This account assumes that the beneficial effect of foreknowledge depends crucially on the intelligibility of pre-exposed sentential speech.

Method: During a visual-verbal serial recall paradigm, participants undertook two counterbalanced blocks of trials wherein they were either pre-exposed to impending auditory distracter sentences (foreknowledge) or not (no foreknowledge). Pre-exposed sentences were intelligible, partially intelligible or unintelligible while sentences accompanying serial recall were all intelligible. Participants were instructed to attend to the sentences during pre-exposure and ignore them when they accompanied the serial recall task.

Results: Foreknowledge of an impending distracter sentence attenuated its later distractive power in serial recall, but only when the foreknowledge was at least partially intelligible.

Discussion: Consistent with the stimulus-specific attentional diversion account, the intelligibility of speech presented during a foreknowledge period is a key requirement for attenuation of auditory distraction by sentential speech. This suggests that intelligible foreknowledge increases familiarity of the material thereby reducing attentional diversion due to interest. These results reinforce the view that foreknowledge reduces disruption produced by the semantic/syntactical properties of discrete sentences but has little effect on that produced by its acoustic properties.

Keywords: foreknowledge; speech intelligibility; irrelevant speech; auditory distraction

The benefit of foreknowledge in auditory distraction depends on the intelligibility of pre-exposed speech

A central role of the cognitive system is to attenuate or block the undesired processing of material that is currently irrelevant to efficient goal-directed behavior. When untempered, such processing of irrelevant material can disrupt cognitive activity. For optimal functionality two opposing requirements must be satisfied by the selective attention system. On the one hand, the system must ensure that, from the array of inputs impinging on various senses, only the fraction that currently supports task-requirements reaches perceptual awareness. On the other hand, the system must remain receptive to the continued processing of currently irrelevant information such that it can, upon evaluation, compete for and assume control of action if it indicates a threat or potential opportunity (a focusability-distractibility dilemma; see e.g., Allport, 1989; Johnston & Strayer, 2001). While distractibility is thus important, a key role for the cognitive system is ensuring, through various means, that the impact of undesired processing of irrelevant material on concurrent cognitive activity is minimized. The focus of the current study is on top-down factors – specifically foreknowledge concerning potential distraction or distracters (Hughes et al., 2013; Röer et al., 2015; Sussman et al., 2003; Vachon et al., 2012) – that modulate the focusability-distractibility balance through preventing a diversion of attention to the task-irrelevant sound (Hughes et al., 2013).

Irrelevant Sound Effect Paradigm

Over the last decade, the irrelevant sound effect paradigm has leant itself well to studying the controllability of auditory distraction via top-down cognitive factors. This paradigm involves the (typically visual) sequential presentation of six to eight items (usually digits or letters) at a rate of one or two items per second. Immediately following presentation of the last item, or in

some studies after a short retention interval, participants are required to recall the items in absolute serial order. During this serial recall task, task-irrelevant sequences of sound are presented usually accompanying visual presentation of items, and/or sometimes during their retention. Of interest to studies of selective attention is the observation that the presence of auditory distractors during the task appreciably impairs serial recall performance. Compelling is that this disruption happens despite participants being told to ignore the sound and reassured that they will not be tested on its content (for a review, see Hughes & Jones, 2001; Jones, 1999; Jones et al., 2010).

To produce significant disruption to serial recall the sound must demonstrate acoustical variation. A changing-state sound sequence wherein acoustic changes occur between immediately successive sounds (e.g., "Q, G, A, T" or a sequence of changing tones) within a single perceptual stream (e.g., Bregman, 1990) produces marked disruption of serial recall compared to a steady-state sequence wherein relatively modest changes occur between sounds (e.g., "T, T, T, T", or a repeated tone; Elliott, 2002; Hughes et al., 2007; Jones et al., 1992, 1995; Jones & Macken, 1993; LeCompte, 1996). There are two explanations for this changing-state effect. According to the *duplex-mechanism account*, the changing-state effect is attributable to interference-by-process (e.g., Hughes & Marsh, 2017; Jones & Tremblay, 2000). The pre-attentive process of organizing auditory items into temporally-extended auditory objects or streams (sequential streaming; Bregman, 1990) interferes with the deliberate process of forming and maintaining a vocal-motor plan representing the serial order of the visual-verbal to-be-remembered items.

An alternative, attentional account (Bell et al., 2017; Cowan, 1995; Elliott, 2002; Röer et al., 2011, 2014, 2015) proposes that exposure to sound results in the involuntary formation of a

model of the sound sequence (a neural or predictive model) that usurps attentional resources to the extent that the sequence of sounds is unpredictable thereby reducing their availability for efficient performance in an ongoing focal task. Within the attentional framework, a predictability-based account can be distinguished from a stimulus-specific account of attentional diversion. According to the *predictability-based account* (cf. Eimer et al., 1996), steady-state sequences produce little disruption because each successive sound has a high degree of predictability (based on the previous stimulation) and hence an accurate predictive model of the sequence can be quickly generated. Thus, relatively few attentional resources are required to build the predictive model and focal task performance will not suffer much. In contrast, changing-state sequences are more disruptive because they are (usually) less predictable. When successive elements within a sequence are difficult to predict, more attentional resources are required to update the predictive model in the event of prediction errors. Thus, building a predictive model of an unpredictable changing-state sequence usurps attentional resources that could otherwise be dedicated to focal task performance, due to the unpredictable nature of its elements. In contrast, according to the *stimulus-specific account* (e.g., Hughes, 2014; Hughes & Marsh, 2020), attentional capture is elicited by certain properties of the irrelevant stimulus itself (e.g., semantic aspects) rather than by its predictability based on the previous stimuli, and such an account has been suggested to explain why meaningful sentences (e.g., “Christmas is coming, the goose is getting very fat, please put a penny in the old man’s hat”) sometimes produce greater disruption of serial recall than sequences of unrelated words (e.g., stimulus-specific self relevance hypothesis; Hughes & Marsh, 2020). The stimulus-specific attentional diversion account of the additional disruption produced by a meaningful discrete sentence conflicts with the predictability account. To fully grasp the intricacies of the arguments as to why this is the

case, it is useful to introduce a means of examining the extent to which giving participants the ability to predict an auditory sequence (by providing foreknowledge) attenuates the disruption it produces to serial recall.

Foreknowledge Effects

Research has suggested that providing participants with the ability to predict the content of the irrelevant sequences through forewarning, reduces their disruptive power. For example, informing participants a few seconds before a trial that the impending irrelevant sequence will contain a deviant has been shown to reduce the disruptive effect on serial recall produced by a single change voice shift in a stream of spoken letters (compared to no prior information; Hughes et al., 2013). Furthermore, specific foreknowledge also reduces the disruption produced by discrete meaningful sentences. For example, Röer et al. (2015) presented participants with a written transcript of an upcoming to-be-ignored speech sequence (e.g., “Put the printer cartridge in the printer so that the tape covers the ink supply port, and start cleaning the print head”) and observed an attenuation of its disruptive effect on serial recall. This finding has been conceptually replicated with auditory presentations of sentences during the foreknowledge period (e.g., Bell et al., 2017). In stark contrast, providing participants with foreknowledge that an impending sequence comprises changing-state letters (as compared to a repeated letter; steady-state) does not attenuate its disruptive power (Hughes et al., 2013).

The predictability-based account assumes that the benefit of foreknowledge arises because it allows participants to form a stable representation of the sequence to the extent that the sequence is predictable. According to this logic it is difficult to create a stable representation of a random-word or random-letter sequence. However, it is relatively easy to build a mental representation of a meaningful sentence because semantic and syntactic features permit easy

translation into a predictive model of the impending sentential sequence (Bell et al., 2017). The problem with this predictability-based account is that, through its assumed importance of post-categorical predictability, it predicts that changing-state sequences of sounds that are predictable (e.g., “A, B, C, D”) should be less disruptive to serial recall than changing-state sequences of sounds that are unpredictable (e.g., “J, M, F, D”). This is because it should be relatively easy to form a stable model of the predictable sound sequence and thus fewer attentional resources should be diverted away from the ongoing focal task. However, several studies have demonstrated that post-categorical predictability within a sound sequence is unrelated to the magnitude of disruption it produces to visual-verbal serial recall (Hughes & Marsh, 2020; Jones et al., 1992; Marsh et al., 2014).

Another prediction that flows from the predictability-based account is that post-categorically predictable discrete sentences should, as standard (i.e., without foreknowledge) produce less disruption than sequences comprising random words or letters that are relatively unpredictable. This is because a predictive representation of the upcoming distracter sequence should be relatively easy to generate for meaningful, coherent speech as compared with sequences of random words (Bell et al., 2017). However, discrete sentences are in fact more, not less, disruptive than sequences of random words (Bell et al., 2017; Röer et al., 2015). Furthermore, recent findings from our laboratories indicate that discrete meaningful sentences produce more disruption than meaningless (foreign-language or reversed) sentences (Marsh et al., 2021; Ueda et al., 2019) and intelligible sentences produce more disruption than unintelligible sentences providing the participants understand the language within which the material is presented (Ueda et al., 2019). A further finding that undermines the predictability-based account is that specific foreknowledge (a written transcript of the impending sequence)

only removes the additional disruption a sentence produces relative to a changing-state sequence of unpredictable (random) verbal items (e.g., letter names; Hughes & Marsh, 2020). The view that the higher predictability of the elements within a meaningful sentence does not lead to less disruption is further reinforced by the observations that 1) the disruption a sentence produces relative to a steady-state sequence is not completely abolished by foreknowledge, and 2) the attenuation by foreknowledge is greater for disruption produced from a sentence than from a changing-state sequence of unpredictable items.

The duplex-mechanism account (e.g., Hughes, 2014; Hughes et al., 2005) appears to offer a straightforward account of the overall patterns of performance as a function of type of distracter sequence and foreknowledge. It eschews the notion that predictability underpins the changing-state effect, the (additional) disruption produced by a meaningful sentence and its attenuation via the presentation of specific foreknowledge. Instead, the duplex account proposes that the disruption produced by simple changing-state sequences of items such as letter-names or random words, relative to a steady-state sequence comprising a repeated item, is produced by an interference-by-process. This occurs between the pre-attentive seriation process automatically applied to the irrelevant sound and deliberate serial rehearsal process applied to the visual-verbal to-be-remembered items. The additional disruption produced by a discrete meaningful sentence is attributable to a qualitatively distinct mechanism: stimulus-specific attentional diversion. The intrinsic relevance or interest of the unfamiliar meaningful sentence — attributable to its semantic or syntactic/grammatical properties — endows it with the capacity to divert attention from the ongoing focal task. Thus, the additional disruption produced by a discrete meaningful sentence relative to a changing-state sequence of random items reflects disruption attributable to stimulus-specific attentional diversion that is distinct from, but additive to, the changing-state

effect. Further, it is this additional effect — attributable to interest or the relevance of a discrete meaningful sentence — that can be attenuated by foreknowledge (Hughes & Marsh, 2020).

Recently emerging evidence gels with the duplex-mechanism account's interpretation of the existing pattern of results (Ellermeier et al., 2015; Marsh et al., 2021; Ueda et al., 2019). For example, Ueda et al. (2019) manipulated the intelligibility of speech that was in the participants' native or non-native language using locally time reversal. Compared to a continuous pink noise control condition, normal speech and locally time-reversed speech of short segment duration — preserving intelligibility — disrupted serial recall to a greater extent than locally time-reversed speech with longer segment durations and a globally inverted (completely reversed and thus unintelligible) speech signal. However, such a pattern was observed only when the speech was presented in the participants' native language. This suggests that any or some combination of attributes of the natural sentences used by Ueda et al. (2019) — including grammatical or syntactic structure, and sentential meaning in a language understood by the participant — produces an additional stimulus-specific attentional diversion effect that is additive with the changing-state effect that arguably represents a pre-categorical acoustic and pure form of distraction observed regardless of whether the language is understood.

Current Research

On the stimulus-specific attentional diversion view within the duplex-mechanism account (e.g., Hughes, 2014), foreknowledge reduces the disruptive impact of a discrete meaningful sentence because it reduces its interestingness or relevance (see Hughes & Marsh, 2020). The general goal of the present study was to demonstrate that the intelligibility of the pre-exposed speech information is therefore integral to the attenuation by foreknowledge of the disruptive effect of sentential speech on serial recall. To this end, we used time reversal to degrade local phonetics in

spoken sentences. This technique divides the speech signal into segments of short (usually constant) duration, and each segment is then played backwards while the order of segments (and thus the global structure of the sentence) is preserved. The benefit of pre-exposure to normal spoken sentences (100% intelligibility rate) is compared with locally-reversed sentences of short reversal frames (70 ms) for which intelligibility is typically moderate (50% intelligibility rate), and for sentences with relatively long reversal frames (140 ms) which is typically unintelligible (see Saberi & Perrott, 1999; Ueda et al., 2017). Greater disruption of serial recall has already been observed from normal speech and locally time-reversed speech of short segment duration as compared to long segment duration (Ueda et al., 2019). Assuming the additional disruption from intelligible over unintelligible speech reflects a stimulus-specific attentional diversion, we expected that pre-exposure to intelligible as compared with unintelligible speech of the same global temporal structure should attenuate its attention-diverting effect. This is because participants will process intelligible speech only and thereafter become familiar with and less interested in the suprasegmental (e.g., semantic, syntactic) properties of speech driving the attentional diversion. We did not predict any attenuation of disruption via pre-exposure to unintelligible speech because here, although the global temporal structure of the impending irrelevant sequence is pre-exposed, the disruption produced by such acoustic properties of irrelevant sound — that arguably underpin the changing-state effect — is not attributable to attentional diversion. Therefore, serial recall should not benefit from such pre-exposure.

Method

Participants

Ninety-six participants were recruited via Prolific Academic (57 women, 38 men, 1 other). Ages ranged between 18 and 30 years (M = 23.4; SD = 3.6). Prior to starting the task, all participants provided informed consent and indicated that they were not diagnosed with hearing loss. Participants were compensated with £10 through Prolific Academic.

Apparatus and Stimuli

The experiment was designed as an online study and programmed in PsyToolkit (Stoet, 2010, 2017). The study could be run either on a desktop (28% of participants) or laptop computer (71.1% of participants), but not on mobile phones or tablets. Any web browser could be used except Safari. Participants were required to use headphones for the study. Most participants reported to have used in-ear (48.5%) or over-ear headphones (43.3%).

To make sure that participants were not using loudspeakers, a headphone-screening test (Woods et al., 2017) was to be completed successfully prior to the main serial recall task. In this test, participants were first asked to adjust the volume of their headphones (or computer settings) to a comfortable level while continuous pink noise was presented. Then three 200-Hz tones (1000 ms each, separated by 500-ms intervals) were presented successively on each trial, and participants had to indicate which tone was “softer” than the other two by clicking on a numbered box (“1”, “2”, or “3”) that was presented together with the particular tone. The sound pressure level of one tone was 6 dB lower than the level of the two other tones, but one of the two tones with the higher level had the phase reversed between the left and right channels (due to acoustical interference, this should reduce the level that reaches the ear when using loudspeakers, thus making it difficult to detect the softer tone; see Woods et al., 2017). The headphone test was

passed if at least five responses were correct in a block of six trials. If a block was not successful, the test continued with a new block until either five correct responses were made in a block or ten blocks were completed, whichever came first. If the headphone test was not passed, a message was shown on the screen, informing participant that the study could not be continued due to an insufficient audio equipment (participants were allowed to restart the experiment).

The to-be-ignored speech sounds comprised recordings of fifty unique English sentences from different categories (aphorisms, cooking recipes, poems, prose, operating manuals, road traffic messages, scientific reports, weather forecast). Each sentence was spoken by Alexa (www.alexa.com) and had a duration of 8 s. In addition to the original speech sample, two types of locally reversed speech were created (using Python) by dividing the speech recording into segments of either 70 ms or 140 ms and temporally reversing the signal within each segment while preserving the order of segments (including 4-ms Hann-shaped rise/fall times in each segment). It has been found previously that intelligibility of locally-reversed speech increases with shorter segment durations, with 140 ms being unintelligible and 70 ms being partially (about 50%) intelligible (see Saberi & Perrott, 1999; Ueda et al., 2017).

Design and Procedure

Each participant completed two separate blocks of 32 trials each (24 unique sentences and 8 silent trials), with foreknowledge being provided in one block, but not in the other. The order of blocks was counterbalanced across participants. In the foreknowledge block, either silence, original speech, 70-ms reversed speech, or 140-ms reversed speech was presented in a foreperiod prior to the presentation of to-be-remembered items and the irrelevant sound (which was either silence or the respective original speech sample). Each foreknowledge condition was repeated eight times throughout the block (with unique sentences), and the order of trials was

fully randomized. In the no-foreknowledge block, no sound was presented in the fore period, but the number of trials was identical. Participants were instructed to listen to the foreknowledge information (in the foreknowledge block), but to ignore the fully intelligible sentences that were presented later during serial recall. Prior to each block, participants completed three practice trials on which serial recall was performed either in silence or with a distracter sentence that was not presented during the main experiment (and different foreknowledge versions).

Participants could start each trial of the serial recall task by clicking on a yellow box with the text "BEGIN" on a black background screen. Then, in the foreknowledge block (except on silent trials), the message "Listen to the information!" was presented in yellow, while the respective speech sound (original or reversed speech with 70 ms or 140 ms segment durations) was presented as foreknowledge for 8s. After the foreknowledge period, there was no sound and a white "+" sign was presented in the center of the screen for 5 s, followed by the message "Get Ready" for 1 s (and followed by a 1-s interval). Then a random sequence of eight digits (from 1-9, drawn without replacement) was presented on the screen, each digit for 800 ms followed by an inter-stimulus interval of 200 ms. Except on silent trials, the original speech recording was presented as to-be-ignored sound during the presentation of the to-be-remembered digits (starting with the second digit). Following a 1-s delay after the last digit disappeared, the response matrix showing the digits 1-9 was presented and participants were asked to click the digits in the order they were presented (see Figure 1 for an illustration of the design of a typical trial with foreknowledge). Feedback was presented after the last response, indicating the number of digits that were recalled in the correct serial position. The next trial started after a 1.5-s inter-trial interval. A trial in the block without foreknowledge (and the silent trials) were identical except

that no sound was presented in the fore-period, while the message "(no information)" was shown on the screen for 8 s.

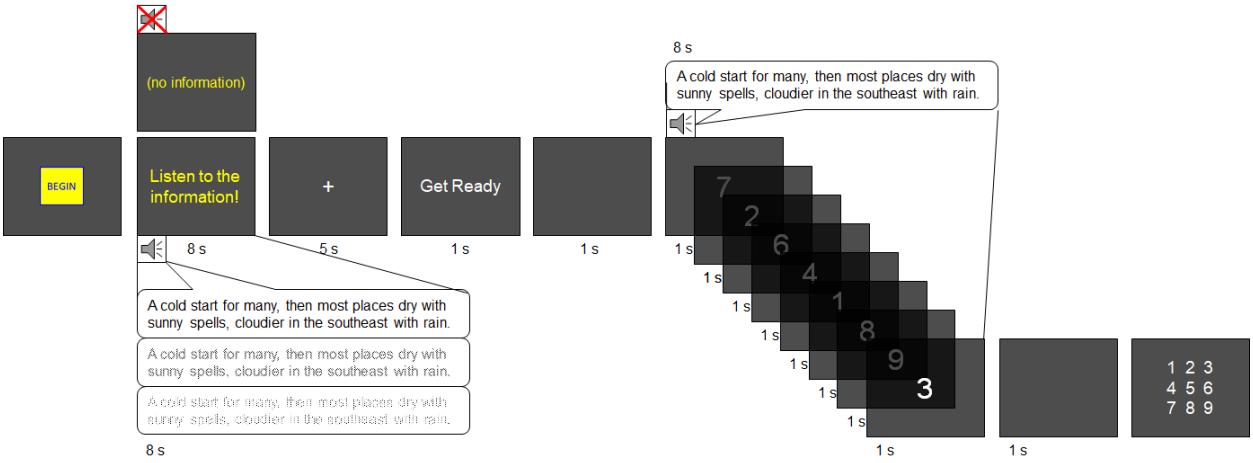

Figure 1. Illustration of the experimental task used in the present study with original or degraded speech presented as foreknowledge prior to the serial recall task during which the original sentence was presented as irrelevant sound. No foreknowledge information was presented on control trials without irrelevant speech during serial recall (not illustrated) as well as in the otherwise identical block without foreknowledge.

After the serial recall task, participants were asked to indicate whether they used external help (paper and pencil or from another person) when remembering the digits, whether they said the digits aloud, whether the volume was turned off during the task, whether the headphones were removed or unplugged, whether there were external sources of visual or auditory distraction, whether they switched between different tasks or browsers during the task, and whether they experienced technical problems. In addition, participants were asked to indicate how motivated they were to obtain the best test-score possible and how concentrated they were (both on a scale from 1 to 5).

Results

Commitment of Participants and Data Processing

Motivation and concentration were reported to be high in general on a 5-point scale (M = 4.09; SD = 0.85 and M = 4.22; SD = 0.79, respectively). Most participants reported to have said the digits aloud when trying to remember them (61.9%), and several participants reported external distraction (11.3%) such as “dog barking”, “people coming into the room”, or “office background noise”. Only a few participants reported to have switched between different tasks or browsers (3.1%).

Four participants reported removing or unplugging their headphones during the task, and two reported turning off the volume during the task (we note that it is unclear whether this was during the trial or in the breaks between trials). Eight other participants reported to have used external help (e.g., paper and pencil) or help from another person. The data of these participants were not included in the analysis, which was then based on N = 82 participants (48 women, 33 men, 1 other; age: M = 23.1; SD = 3.4 years). For n = 43 participants, the foreknowledge block came first, and for n = 39 participants, the no-foreknowledge block came first.

Visual-Verbal Serial Recall

The average proportion of digits that were recalled in the correct serial position as a function of the sound and foreknowledge condition is illustrated in Figure 2. First of all, performance during silence in the present online study (M = 6.06 digits / 75% recalled in the correct serial position; SEM = 1.7%) was comparable to the non-distracting control conditions in many previous studies using similar procedures in a laboratory setting (e.g., 73-77% or 6.53-6.95 of 9 digits in silence, Ellermeier et al., 2015; 71% or 5.66 of 8 digits during white noise, Kattner & Ellermeier, 2020; or 76% of 5-7 digits during white noise, Kattner & Meinhardt, 2020), thus

underlining the general quality of the data obtained in the present online study. In addition, it can be seen that performance was lower when speech was presented as irrelevant sound (and foreknowledge in the respective block), compared with the silent trials. This observation was confirmed by a 2 (block: foreknowledge vs. no foreknowledge) × 4 (foreknowledge sound[1]: silence, speech, 70-ms reversed speech, 140-ms reversed speech) × 2 (block order) mixed-factors ANOVA with block and sound as repeated-measures factors, revealing a significant main effect of sound, $F(3,240) = 27.36$; $p < .001$; $\eta_G^2 = 0.054$ (contrast between speech and silence: $F(1,80) = 43.09$; $p < .001$; $\eta_G^2 = 0.063$). In addition, recall performance was significantly worse in the block without foreknowledge ($M = .70$; $SD = .15$) than with foreknowledge ($M = .73$; $SD = .15$); $F(1,80) = 5.42$; $p = .022$; $\eta_G^2 = 0.008$. Most importantly, the interaction between foreknowledge sound and block was also significant, $F(3,240) = 5.10$; $p = .002$; $\eta_G^2 = 0.006$, indicating that the disruptive effect of irrelevant speech depends on the intelligibility of the foreknowledge that was presented prior to the task. There was no main effect of the order of the two blocks, $F(1,80) = 0.82$; $p = .37$; $\eta_G^2 = 0.006$, and no interaction of block order with foreknowledge sound, $F(3,240) = 0.38$; $p = .77$; $\eta_G^2 < 0.001$, or block, $F(1,80) = 0.37$; $p = .546$; $\eta_G^2 = 0.001$. There was also no three-way interaction, $F(3,240) = 1.64$; $p = .186$; $\eta_G^2 = 0.002$, indicating that the foreknowledge effect on auditory distraction did not depend on which block was presented first.

Most importantly, a follow-up analysis within the foreknowledge block confirmed that recall accuracy differed significantly as a function of the type of foreknowledge sound presented prior to the to-be-remembered digits, $F(3,240) = 18.63$; $p < .001$; $\eta_G^2 = 0.062$. A planned-contrasts analysis (conducted with the {emmeans} package in R) revealed that performance

[1] Note that this factor was dummy-coded for the no-foreknowledge block, which had the same trial structure as the foreknowledge block, but the foreknowledge period was always silent. That is, any difference as a function of foreknowledge sound in the no-foreknowledge block must be due to the irrelevant sound presented later during the encoding phase (silence or speech).

differed significantly between silence and all types of foreknowledge; $t(240) > 4.60$; $p < .001$ (adjusted for multiple comparisons according to Benjamini & Hochberg, 1995). In addition, the presentation of unintelligible 140-ms locally reversed speech as foreknowledge resulted in lower serial recall accuracy during irrelevant speech (i.e., more distraction) compared to both original speech, $t(240) = 2.68$; $p = .012$; $d_z = 0.173$, and partially intelligible 70-ms reversed speech, $t(240) = 2.30$; $p = .027$; $d_z = 0.148$, whereas there was no significant difference between the original speech and 70-ms reversed speech conditions, $t(240) = 0.38$; $p = .702$. This indicates that the release of auditory distraction depended on the (partial) intelligibility of the foreknowledge sentences.

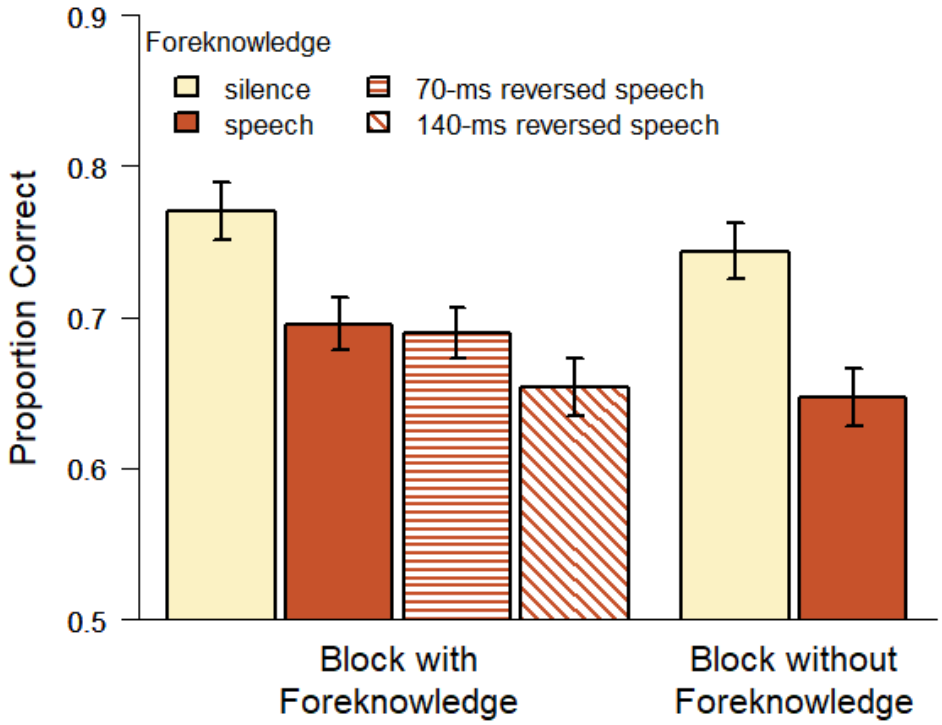

Figure 2. Serial recall accuracy as a function of the irrelevant sound presented during the task (either silence or original speech) and the type of foreknowledge presented prior to the task or not depending on the block. Error bars depict standard errors of the means.

Discussion

The intention of this study was to understand the role of foreknowledge in attenuating distraction produced by task-irrelevant auditory sentential material during the serial recall task. This was addressed by presenting participants with foreknowledge comprising intelligible, partially intelligible or unintelligible sentential material. Thereafter, participants were presented with intelligible speech when undertaking serial recall. The extent to which foreknowledge attenuated the disruption produced by sentential speech was related to the intelligibility of the foreknowledge: Pre-exposure to intelligible and partially intelligible speech reduced subsequent disruption of serial recall produced by intelligible speech. However, pre-exposure to unintelligible speech of the same global temporal structure (and changing-state information) led to no foreknowledge benefit. While the effect sizes of these foreknowledge-related reductions of auditory distraction are relatively small, the results also suggest that full intelligibility of foreknowledge is not necessary in order to attenuate the attentional diversion that is produced by task-irrelevant speech.

The results of the present study offer support to the duplex-mechanism account of auditory distraction (e.g., Hughes, 2014; Hughes et al., 2005; Hughes & Marsh, 2019, 2020; Marsh et al., 2020). According to this account, the disruption produced by a discrete meaningful sentence has two components: The changing-state effect that is attributable to acoustic variation within the to-be-ignored sequence, and stimulus-specific attentional diversion attributable to diversion of attention to its unfamiliar and interesting/relevant semantic or syntactical/grammatical properties. The top-down control benefits of foreknowledge arise because they selectively attenuate the attentional diversion component of the disruption produced by a sentence but leaves the changing-state effect unchecked. This is because the changing-state

effect is produced via pre-attentive processing and is thus outside the reach of top-down control. On this view, foreknowledge renders the sentence more familiar and less interesting thereby reducing its disruptive effect. Since such a mechanism requires processing of semantic and syntactical/grammatical sentence properties during the foreknowledge period, it follows that speech intelligibility during foreknowledge is a pre-requisite for the benefits of foreknowledge to manifest. Similarly, given that the disruption produced via the global changing-state properties of to-be-ignored sequences is produced via pre-attentive processing – and thus not amenable to top-down control – it follows that pre-exposure to such information will not attenuate any disruption subsequently produced by that sequence. Thus, the lack of a foreknowledge benefit from unintelligible sentences (predicting the global temporal envelope of the to-be-ignored speech) is also predicted by the duplex mechanism account.

Based on the data reported here, an alternative predictability-based attentional account (e.g., Bell et al., 2017; Röer et al., 2015) might seem to offer a perfectly adequate explanation of the data. For example, the generation of a predictive model – representing semantic, syntactical/grammatical information – would be possible only for intelligible speech during the foreknowledge period. Thus, the failure to find a foreknowledge benefit for unintelligible speech would be readily predicted by that account. However, the predictability account is at odds with the findings that the post-categorical predictability of items (e.g., words, letters, or digit-names) within a sequence does not attenuate the magnitude of disruption to visual-verbal serial recall (Hughes & Marsh, 2019; Marsh et al., 2014) and that a discrete meaningful sentence, possessing higher post-categorical predictability compared with a relatively unpredictable sequence of random items (e.g., words or letter-names), produces greater, not less, disruption of serial recall. Further, the predictability account does not offer a ready explanation for why discrete sentences

in a participant's own-language produce greater disruption than those presented in a different language or globally time-reversed (Marsh et al., 2021; Ueda et al., 2019).

The predictability account has been questioned also in relation to the disruption produced by sentential (so-called "complex changing-state") distracters. A meaningful discrete sentence produces greater disruption than a sequence of unrelated words or letters (e.g., Hughes & Marsh, 2020), even though the predictability of its elements should be higher. There is some debate as to the source of the additional disruption produced by a discrete sentence over a sequence of unrelated items (Hughes & Marsh, 2020). One possibility is that it represents an additional form of disruption superimposed on the acoustically driven changing-state effect which is attributable to semantic or syntactical/grammatical properties of sentential speech that elicit stimulus-specific attentional diversion (Hughes & Marsh, 2020). The argument is that either particular meaning or specific syntactical/grammatical regularities of discrete sentences produce disruption because their inherent interest or relevance causes a diversion of attention away from the ongoing focal task (stimulus-specific relevance hypothesis; Hughes & Marsh, 2020). In this way, the mechanistic underpinning of the disruptive effects of a discrete meaningful sentence is similar to the disruption produced by a single, irregular item (a deviant) that differs from prevailing items within the auditory scene (e.g., a letter or tone A within the context of a sequence of repeated B letters or tones; Schröger, 1997). In the context of serial recall, deviant events including a change of voice (Hughes et al., 2007, 2013) or timing (Hughes et al., 2005) of a single item within a sequence substantially disrupt performance (the deviation effect). However, this deviation effect arguably represents a *stimulus-aspecific attentional diversion* (e.g., Eimer et al., 1996) according to which the poor fit of an auditory input within a prevailing auditory scene causes attention to be diverted from a focal task (i.e., similar to the predictability account). Unlike stimulus-specific

attentional diversion, the capability of a stimulus to divert attention from an ongoing focal task is not attributable to a specific feature of that stimulus.While we have argued that the intelligibility of pre-exposed speech is a prerequisite for obtaining a foreknowledge benefit, it is perhaps surprising that the benefit was not more pronounced following pre-exposure to intelligible as compared to partially intelligible sentences. There are several possible interpretations of this effect. First, it is possible that listening effort was increased for partially-intelligible, as compared with intelligible, sentences and that such processing increased familiarity (and thus reduced intrigue) with the elements of the sound that could be understood, thereby reducing its attentional diverting power. Second, increased familiarity with the sentential material, and thus maximally reduced attentional diversion via foreknowledge, may require only the comprehension of a few sentence elements or chunks that can be apprehended even in cases wherein sentential material is partially intelligible. These potential interpretations also speak to the theoretical mechanism underpinning the foreknowledge benefit.

In the context of the current study, we consider familiarity with the sentential material as compared with the voice delivering the stimuli. On the face of it, one might expect familiarity with the speaker to also give rise to stimulus specific attentional diversion. However, the current evidence is mixed. For example, in the context of visual-verbal serial recall, Barker and Elliott (2019, Experiment 1) found no more distraction from spoken sentences for a group of participants who received the speaker as their course instructor as compared to a group of participants who did not. Furthermore, no additional disruption was found for a group of participants who were specifically instructed that they would be ignoring sentences spoken by their course instructor as compared to a group who also received the course instructor, but which were not informed about the identity of the voice. Further, Barker and Elliott (2019, Experiment

2) reported no more disruption from spoken distractors for a group of participants who were familiarized with the speaker's voice for four days prior to test as compared with a control group who were not. In contrast, in the case of an auditory-verbal serial recall task wherein both to-be-remembered and to-be-ignored material are presented in the auditory domain, Kreitewolf et al. (2019) found greater disruption from a familiar as compared to an unfamiliar voice. At odds with this finding, Johnsrude et al. (2013) found that familiarity with an irrelevant voice (e.g., the participants' spouse's voice) helps participants to ignore distractor speech in order to comprehend an unfamiliar target speaker. Clearly further investigation is required to examine the discrepancies between these findings and whether modality of to-be-remembered information is a moderating factor.

Notwithstanding its difficulties in accounting for the lack of predictability effects, the attentional account (Bell et al., 2017; Röer et al., 2015) might assume that the benefit of foreknowledge results from a predictive model that is a faithful representation of the lexical-semantic and syntactical structure of the pre-exposed sentence. If so, then one might expect a benefit of foreknowledge to the rate that the perceivable lexical-semantic and syntactical structure of a pre-exposed sentence maps onto the subsequent distracter sentence. The fact that this was not observed in the current study appears to be at odds with the attentional account. On the other hand, the stimulus specific attentional diversion account of the sentential distraction and its attenuation via foreknowledge, necessarily tolerates some discrepancy between the pre-exposed sentence and the distracter sentence. For example, current experiments from our laboratories are exploring whether foreknowledge comprising a paraphrase of the impending distracter sentence offers the same magnitude of benefit as a verbatim repetition. Further, we are investigating whether a change from one language at pre-exposure (e.g., written or spoken

foreknowledge) to another during test results in equal foreknowledge benefits compared to no switch conditions. Such a finding would arguably be accommodated better within a stimulus-specific attentional diversion account than an attentional, predictability-based account. This is because the unfamiliarity or intrigue driving attentional capture should be attenuated even by non-verbatim (or cross-language) repetitions of sentential meaning.

It is important to note here that the level of interest for the participant that a discrete, isolated meaningful sentence has is likely to be much less in comparison to the same sentence embedded with numerous others in a single literary source presented throughout a block of serial recall trials. In such experimental settings (Jones et al., 1990; Marsh et al., 2009) a narrative that is meaningful to participants (e.g., presented in their primary language) produces no more disruption compared with a meaningless version (e.g., presented in a language that they do not understand, or reversed). This suggests that contra to single discrete meaningful sentences covering a variety of topics and presented in isolation, narrative speech soon loses its relevance or interest for participants (Hughes & Marsh, 2020).

The stimulus-specific attentional diversion account underpinning one of the two components of the distracter sentence effect suggests that the disruption produced by a discrete, meaningful sentence, should be attenuated by other top-down factors known to modulate the disruption produced by sound events that are irregular with respect to a preceding pattern of sound stimulation (the deviation effect). For example, individual differences studies show a reduced susceptibility to the deviation effect among participants of high working memory capacity, but no difference in the susceptibility to the changing-state effect (Hughes et al., 2013; Labonté et al., 2022/this issue; Marsh et al., 2017; Sörqvist, 2010)(but see Körner et al., 2017). The stimulus-specific attentional diversion account suggests that individuals with higher working

memory capacity may be less susceptible to disruption produced by meaningful (e.g., intelligible) as compared with meaningless (e.g., unintelligible) to-be-ignored speech without foreknowledge and thus facilitated less by the presence of foreknowledge. Further, presenting to-be-remembered digits in a difficult to encode format (e.g., Parmentier et al., 2008; Yi et al., 2004) to promote task-engagement reduced the disruptive impact of an irrelevant auditory deviant on visual-verbal serial recall performance but failed to modulate the changing-state effect (Hughes et al., 2013; but see Bell et al., 2020; Kattner & Bryce, 2021). Analogous effects have also been observed using Navon letters whereby a large letter is composed of smaller letters of a different identity (Marsh et al., 2020). Here, the deviation effect emerges when participants are oriented to attend and serially recall large letters (low encoding-load) but not the list of small letters (high encoding-load; see also Hughes & Marsh, 2019). The stimulus-specific attentional diversion account predicts that such task-engagement manipulations should also attenuate the component of sentential distraction specifically attributable to attentional diversion, but leave the component produced by changing-state unscathed.

Limitations

One limitation of the current study is that we did not fully cross the semantic/syntactical properties and acoustic properties of the distractor sentences. Consequently, acoustic foreknowledge was provided by speech at all levels of intelligibility, but the semantic/syntactic properties were only provided by intelligible and partially intelligible sentences. Our design did not fully cross semantic/syntactical information with acoustical information due to a difficulty in generating a condition within which a semantic/syntactical overlap can occur in the absence of acoustical overlap. For example, at first glance, one suggestion to this problem might be to present a visual transcript of the impending auditory sentence during the foreknowledge period

(e.g., Röer et al., 2015) which comprises semantic/syntactical information in the absence of acoustic information. However, in such a condition it is highly probable that participants use subvocal motor-processing to encode various acoustic cues (e.g., prosody) during the foreknowledge period that would overlap with the impending distractor sentence. Another solution might be to make use of cross-language repetition (as outlined earlier) since providing the lexical-semantic and syntactical properties can be held constant between languages, the acoustic information can be manipulated independently. However, an experiment deploying cross-language repetition in addition to manipulations of speech intelligibility (e.g., via locally reversed speech) requires numerous further conditions and requires proficient bilingual participants.

On the face of it, a further limitation of the current study is that the degree of match between the foreknowledge and distractor sentences differs as a function of the intelligibility of the foreknowledge sentences (e.g., there is a 1:1 match for fully intelligible sentences). However, the degree of match was clearly not important for the benefit of foreknowledge to manifest, as it was observed with equal magnitude when fully intelligible or partially intelligible sentences were presented during the foreknowledge period. Furthermore, results from our laboratories demonstrate that a perfect acoustical match between foreknowledge and distractor sentences is not sufficient to produce a benefit of foreknowledge when the sentences are spoken in a language foreign to participants (i.e., a benefit of foreknowledge was only observed if participants understood the language; Marsh et al., 2018).

Conclusion

To conclude, the present results provide further evidence suggesting that disruptive effects attributable to the post-categorical properties of task-irrelevant sound are qualitatively distinct

from those attributable to its pre-categorical and purely acoustic changing-state properties. The former are attentional diversion effects that are amenable to top-down cognitive control while the latter is beyond the reach of cognitive control because it stems from a clash between the pre-attentive processing of sound and the deliberate processing involved in performing the focal task.

References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). MIT Press. https://psycnet.apa.org/record/1990-97026-016

Barker, B. A., & Elliott, E. M. (2019). The Role of Talker Familiarity in Auditory Distraction. *Experimental Psychology*, *66*(1), 1–11. https://doi.org/10.1027/1618-3169/a000425

Bell, R., Mieth, L., Buchner, A., & Röer, J. P. (2020). Monetary incentives have only limited effects on auditory distraction: evidence for the automaticity of cross-modal attention capture. *Psychological Research*. https://doi.org/10.1007/s00426-020-01455-5

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. https://doi.org/10.1027/1618-3169/a000372

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*(1), 289–300. https://doi.org/10.2307/2346101

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound* (4th ed.). MIT Press.

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

Eimer, M., Nattkemper, D., Schröger, E., & Prinz, W. (1996). Involuntary attention. In O. Neumann & F. Sanders (Eds.), *Handbook of Perception and Action* (pp. 389–446). Academic Press. https://doi.org/10.1016/S1874-5822(96)80022-3

Ellermeier, W., Kattner, F., Ueda, K., Doumoto, K., & Nakajima, Y. (2015). Memory disruption by irrelevant noise-vocoded speech: Effects of native language and the number of frequency

bands. *The Journal of the Acoustical Society of America*, *138*(3), 1561–1569. https://doi.org/10.1121/1.4928954

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory and Cognition*, *30*(3), 478–487. https://doi.org/10.3758/BF03194948

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30–41. https://doi.org/10.1002/pchj.44

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & 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. https://doi.org/10.1037/a0029064

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

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 and Cognition*, *43*(4), 537–551. https://doi.org/10.1037/xlm0000325

Hughes, R. W., & Marsh, J. E. (2019). Dissociating two forms of auditory distraction in a novel Stroop serial recall experiment. *Auditory Perception and Cognition*, *2*(3), 129–142.

Hughes, R. W., & Marsh, J. E. (2020). When is forewarned forearmed? Predicting auditory distraction in short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(3), 427–442. https://doi.org/10.1037/xlm0000736

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, and Cognition*, *31*(4), 736–749. https://doi.org/10.1037/0278-7393.31.4.736

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, and Cognition*, *33*(6), 1050–1061. https://doi.org/10.1037/0278-7393.33.6.1050

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., & Carlyon, R. P. (2013). Swinging at a Cocktail Party: Voice Familiarity Aids Speech Perception in the Presence of a Competing Voice. *Psychological Science*, *24*(10), 1995–2004. https://doi.org/10.1177/0956797613482467

Johnston, W. A., & Strayer, D. L. (2001). A dynamic, evolutionary perspective on attention capture. In C. L. Folk & B. S. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 375–397). Elsevier Science. https://doi.org/10.1016/S0166-4115(01)80017-0

Jones, D. M. (1999). The cognitive psychology of auditory distraction: The 1997 BPS Broadbent Lecture. *British Journal of Psychology*, *90*(2), 167–187. https://doi.org/10.1348/000712699161314

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional Equivalence of Verbal and Spatial Information in Serial Short-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 1008–1018. https://doi.org/10.1037/0278-7393.21.4.1008

Jones, D. M., Hughes, R. W., & Macken, W. J. (2010). Auditory distraction and serial memory: The avoidable and the ineluctable. *Noise and Health*, *12*, 201–209. https://doi.org/10.4103/1463-1741.70497

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. https://doi.org/10.1037/0278-7393.19.2.369

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The 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*, 89–108. https://doi.org/10.1002/acp.2350040203

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin and Review*, *7*(3), 550–558. https://doi.org/10.3758/BF03214370

Kattner, F., & Bryce, D. (2021). Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/https://doi.org/10.1037/xhp0000982

Kattner, F., & Ellermeier, W. (2020). Distraction at the cocktail party: Attenuation of the irrelevant speech effect after a training of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(1), 10–20.

https://doi.org/10.1037/xhp0000695

Kattner, F., & Meinhardt, H. (2020). Dissociating the Disruptive Effects of Irrelevant Music and Speech on Serial Recall of Tonal and Verbal Sequences. *Frontiers in Psychology*, *11*, 346. https://doi.org/10.3389/fpsyg.2020.00346

Körner, U., Röer, J. P., Buchner, A., & Bell, R. (2017). Working memory capacity is equally unrelated to auditory distraction by changing-state and deviant sounds. *Journal of Memory and Language*, *96*, 122–137. https://doi.org/10.1016/j.jml.2017.05.005

Kreitewolf, J., Wöstmann, M., Tune, S., Plöchl, M., & Obleser, J. (2019). Working-memory disruption by task-irrelevant talkers depends on degree of talker familiarity. *Attention, Perception, and Psychophysics*, *81*, 1108–1118. https://doi.org/10.3758/s13414-019-01727-2

Labonté, K., Marsh, J. E., & Vachon, F. (2022). Distraction by auditory semantic deviations is unrelated to working memory capacity: Further evidence of a distinction between acoustic and categorical deviation effects. *Auditory Perception & Cognition*.

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*(5), 1154–1165. https://doi.org/10.1037//0278-7393.22.5.1154

Marsh, J. E., Campbell, T., Vachon, F., Taylor, P., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, and Psychophysics*, *82*, 350–362. https://doi.org/10.3758/s13414-019-01800-w

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110*(1), 23–38. https://doi.org/10.1016/j.cognition.2008.08.003

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. https://doi.org/10.1002/pchj.50

Marsh, J. E., Threadgold, E., Barker, M. E., Litchfield, D., Degno, F., & Ball, L. J. (2021). The susceptibility of compound remote associate problems to disruption by irrelevant sound: a Window onto the component processes underpinning creative cognition? *Journal of Cognitive Psychology*, *33*(6–7), 793–822. https://doi.org/10.1080/20445911.2021.1900201

Marsh, J. E., Vachon, F., & Sörqvist, P. (2017). Increased distractibility in schizotypy: Independent of individual differences in working memory capacity? *Quarterly Journal of Experimental Psychology*, *70*(3), 565–578. https://doi.org/10.1080/17470218.2016.1172094

Parmentier, F. B. R., Elford, G., Escera, C., Andrés, P., & Miguel, I. S. (2008). The cognitive locus of distraction by acoustic novelty in the cross-modal oddball task. *Cognition*, *106*(1), 408–432. https://doi.org/10.1016/j.cognition.2007.03.008

Röer, J. P., Bell, R., & Buchner, A. (2014). Evidence for habituation of the irrelevant-sound effect on serial recall. *Memory and Cognition*, *42*, 609–621. https://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*(3), 692–702. https://doi.org/10.1037/xhp0000028

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 and Cognition*, *39*,

839–850. https://doi.org/10.3758/s13421-010-0070-z

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, *398*, 760. https://doi.org/10.1038/19652

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*, 245–257. https://doi.org/10.1111/j.1469-8986.1997.tb02395.x

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: A review. *Noise and Health*, *12*, 217–224. https://doi.org/10.4103/1463-1741.70500

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin and Review*, *10*, 630–637. https://doi.org/10.3758/BF03196525

Ueda, K., Nakajima, Y., Ellermeier, W., & Kattner, F. (2017). Intelligibility of locally time-reversed speech: A multilingual comparison. *Scientific Reports*, *7*, 1782. https://doi.org/10.1038/s41598-017-01831-z

Ueda, K., Nakajima, Y., Kattner, F., & Ellermeier, W. (2019). Irrelevant speech effects with locally time-reversed speech: Native vs non-native language. *The Journal of the Acoustical Society of America*, *145*(6), 3686. https://doi.org/10.1121/1.5112774

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken expectations: Violation of

expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*, 164–177. https://doi.org/10.1037/a0025054

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*, 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Yi, D. J., Woodman, G. F., Widders, D., Marois, P., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*(9), 992–996. https://doi.org/10.1038/nn1294